Rhotic variation in the Spanish spoken by Puerto Ricans in Puerto Rico and Western Massachusetts

Alba Arias Alvarez

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

Part of the Spanish Linguistics Commons

Recommended Citation

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
Rhotic variation in the Spanish spoken by Puerto Ricans in Puerto Rico and Western Massachusetts

A Dissertation Presented

by

ALBA ARIAS ALVAREZ

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

SEPTEMBER 2018

Hispanic Literatures and Linguistics
Rhotic variation in the Spanish spoken by Puerto Ricans in Puerto Rico and Western Massachusetts

A Dissertation Presented

by

ALBA ARIAS ALVAREZ

Approved as to style and content by:

Meghan Armstrong-Abrami, Chair

Patricia Gubitosi, Internal Member

Marcos Rohena-Madrazo, External Member

Luiz Amaral, Unit Director
Spanish and Portuguese Program

William Moebius, Department Chair
Dept. of Languages, Literatures, and Cultures
DEDICATION

Para ti, Lola de Vera.
Estrela de ollos verdes, confidente y, ante todo, abuela.
ACKNOWLEDGMENTS

I would like to formally express my thanks for the enormous individual and institutional support that I received throughout this life’s chapter that is about to end. Without assistance, generosity, patience and financial support this dissertation would not have been possible.

First of all, I would like to express my deepest gratitude to my advisor and thesis director, Meghan Armstrong-Abrami. Without her inestimable guidance, help and encouragement, I would not have been able to face the challenges of graduate school. My project has definitely benefited from her high standards for methods and analyses. I am truly honored to have been your advisee. Your support helped me to become the scholar I am today. Thank you, Meghan.

I would also like to thank Patricia Gubitosi for guiding me with her knowledge, wisdom and expertise. Her advice helped me to become a better sociolinguistic thinker. Everything I know about sociolinguistics is thanks to you and I am truly lucky for having been your student for the last six years.

I am also immensely grateful to Marcos Rohena-Madrazo for agreeing to be part of this project and for all his advice, comments and support. Not only do I appreciate his valuable suggestions but also his help in recruiting participants during my trip to Puerto Rico.

I want to express my heart-felt gratitude to the Puerto Rican Spanish speakers of Holyoke, Salinas and San Juan, who welcomed in their homes and let me record them. I appreciate you all graciously putting up with my tiresome project demands. Without your generosity, this research would have been totally impossible to complete. GRACIAS.
I am thankful for the generous funding I received to make this research project possible. In particular, the Doctoral Dissertation Research Improvement Grant from the National Science Foundation (Id: 1651050), which covered all my recording materials, research assistants, travel expenses to Puerto Rico and Holyoke and participants’ compensation. This project was also made possible by the financial support of the Graduate School of the University of Massachusetts Amherst, which awarded me with The Dissertation Grant and Pre-Dissertation Grant. I have been extremely fortunate to be a part of the UMass Graduate School, which helped me to prepare my grant materials and supported me in becoming a better scholar through their workshop offerings, writing retreats and individual assistance. Go, UMass!

Data collection would not have been possible with the enormous support I received from Juliana Ivette Cruz Martinez in Puerto Rico, and Jharikem Borrero in Holyoke. In addition, many thanks to Alicia LeClair, Patrick Murphy, Claire Laing and Mallory Fuller, who helped me with the arduous work of coding.

I am grateful for the constant statistical advice of Andrea Cataldo, from the Institute for Social Science Research of the University of Massachusetts Amherst. Moreover, I want to express my gratitude to Covadonga Sánchez, who provided me assistance in the analysis of the quantitative results as well as to Christian DiCanio, who, in addition to helping me with the terrific R, provided me with the necessary Praat script to measure fricatives. Andie Faber, thank you so much for helping me with my non-native English skills. Your help is invaluable.

At various stages in the completion of this dissertation, several linguists provided useful materials or answered my (sometimes embarrassingly basic) questions. For this, I
am grateful to Christian DiCanio, Iraida Galarza, Fiona Dixon and Cova Sánchez. I would also like to thank professors from the Universidad Río Piedras for making their corpus accessible to me and for help with participant recruitment.

I feel honored and extremely fortunate to be surrounded by such wonderful people at the University of Massachusetts Amherst. Cova and Fiona, we are las tres patas pa un banco. Cova, tú eres la siempre sensata, práctica e inteligente amiga con los pies en el suelo, capaz de resolver cualquier problema en el horizonte. Fiona, tú eres la decidida aventurera, apasionada de la vida y del ginger. I love you so much. Since the first day I arrived in Amherst, you have been my unconditional friends and you will keep in my heart for the rest of my life. Without you, I would have never felt at home in the U.S. Christina, thanks for the constant support, help, and tardes de nachos talking about life, love and new challenges on the horizon. Eres powerful, amazing feminista con una fuerza todopoderosa. Claudia, thanks for being the best and most amazing listener in the world; Ana, thanks for opening your heart and letting me open mine; Eli, I truly appreciate your library –and life in general- check-ins (although I haven’t tried your gluten-free tiramisú); Dayana, thanks for reminding me that I am not alone; Sheryl, thanks for guiding me and making me alert for the next step. María Turrero, Edu, Andie, Odalis, Carla, Elena, Rosa, Isaaquito, Abril, Pilar, Santi and to the rest of UMass SpanPort community, as well as to Brian, who reminded me to eat healthy and be happy during the stressful writing period, thank you for all the support and encouragement that kept me alive to complete this dissertation. THANK YOU!
And last, but not least, my family.

Mamá, gracias por apoyarme de manera incondicional desde el minuto cero. Sé lo difícil que es para ti la otra cara de la disertación. A pesar de los 5273.82 km que nos separan, agradezco que nunca se agoten las palabras de ánimo. Gracias por las lentejas de los lunes, por los deberes extra de matemáticas, por celebrar mi único suspenso y por el resto de un infinito etcétera. Todo, absolutamente todo lo que has hecho por mí, ha impactado de alguna u otra forma mi manera de ser y me ha hecho ser mejor persona.

Papá, gracias por tus palabras de sosiego y calma. Gracias por las llamadas matutinas en las que me recuerdas que paso a paso todo se logra. Admiro tu prudencia y tus silencios. Estoy muy orgullosa de mi identidad asturiana y, en gran parte, te lo debo a ti.

Marta, eres bella, por dentro y por fuera. Fuerte, decidida. Gracias por todo el apoyo y ayuda brindada para afrontar mis miedos, mis debilidades a lo largo de mi etapa como estudiante graduada y, sobre todo, en el tramo final de la disertación.

Mamá, papá, Marta, os quiero con toda mi alma.

Finalmente, no quiero finalizar esta sección sin hacer hincapié en que esta disertación está dedicada a la memoria de mi abuela, Lola de Vera. En ella puedo ejemplificar algunos de los conceptos fundamentales de la sociolingüística y, en concreto, del contacto de lenguas. Espero que me hayas perdonado las muchas veces en las que no
dejaba que me hablas en gallego. Ahora lo extraño, como todo lo que me hace recordarte. Sigue brillando y guiándome allá alto. Te quiero, abuela.
ABSTRACT

RHOTIC VARIATION IN THE SPANISH SPOKEN BY PUERTO RICANS IN PUERTO RICO AND WESTERN MASSACHUSETTS

SEPTEMBER 2018

ALBA ARIAS ÁLVAREZ, B.A., UNIVERSITY OF OVIEDO, SPAIN
M.A., UNIVERSITY OF OVIEDO, SPAIN
M.A., UNIVERSITY OF MASSACHUSETTS AMHERST
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Meghan Armstrong-Abrami

The Spanish trill is known to present a wide range of phonetic variation in Puerto Rican Spanish (PRS), attested not only on the island but in the diaspora. Combining auditory and acoustic analysis, this research project studies acoustic data on onset /r/ in Holyoke, MA, the city with the largest per capita population of Puerto Ricans living outside the island. The aim of this dissertation is to analyze whether there is trill variation in the PR community in Holyoke, and, whether it mirrors the variation found in Puerto Rico. Special attention is paid to glottal, velar, or uvular /r/ realizations. Recent work suggests that the phonemes /ɾ/ and /h/ have been contextually neutralized in perception, which would result in the loss of a phonemic contrast. Therefore, this project also investigates whether there is evidence for this production neutralization using measurements that had never been acoustically examined for this dialect (center of gravity, skewness, and kurtosis). Three experimental production tasks were designed and employed: a picture description task, a map task and a reading task. Forty-five participants performed the experimental tasks: 21 were recorded on the island and 24 in Holyoke. As a result, a total
of 4,393 phonemic /r/ and /h/ were analyzed. Results indicate that there are substantial similarities in rhotic variation as well as some variation between Puerto Rico and Holyoke: (1) the same trill realizations are found and (2) the means of center of gravity, skewness and kurtosis are significantly different between /h/ and /r/ in both locations, suggesting an absence of neutralization. However, different linguistic and sociolinguistic variables affect (1) the use of the backed /r/ and (2) phonemic /h/-/r/ distinction. Findings suggest that the PR community in Holyoke tries to maintain their language, one of the most noticeable signs of immigrants’ origin, to strengthen authenticity in the same way that they keep other PR cultural experiences. The differences found suggest that, although Holyoke maintains a close bond with Puerto Rico due to the back and forth migration waves, diasporas are still changing communities which create sites of super-diversity, with different patterns as a result of these new dialect contact situations.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>INTRODUCTION</td>
</tr>
<tr>
<td>1.1</td>
<td>Theoretical issues</td>
</tr>
<tr>
<td>1.2</td>
<td>The present study</td>
</tr>
<tr>
<td>2.</td>
<td>THE PHENOMENON: RHOTIC VARIATION IN PRS</td>
</tr>
<tr>
<td>2.1</td>
<td>Phonetic description and distribution of /r/</td>
</tr>
<tr>
<td>2.2</td>
<td>Variation of /r/: Puerto Rican Spanish</td>
</tr>
<tr>
<td>2.3</td>
<td>The Backed /r/</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Phonetic description and socioindexical meaning</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Main studies on backed /r/ in PRS</td>
</tr>
<tr>
<td>2.4</td>
<td>Neutralization of /r/ and /h/ in PRS</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Acoustic analysis among fricatives: continuous measurements</td>
</tr>
<tr>
<td>3.</td>
<td>SOCIOLINGUISTIC CONTEXT: PUERTO RICAN DIASPORA AND DESCRIPTION OF THE COMMUNITIES UNDER STUDY</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.2</td>
<td>The Puerto Rican Diaspora</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Puerto Ricans on the Mainland United States</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Puerto Ricans in Massachusetts</td>
</tr>
<tr>
<td>3.2.2.1</td>
<td>Demographic and social aspects of the Puerto Rican population in Massachusetts</td>
</tr>
<tr>
<td>3.3</td>
<td>The communities under study</td>
</tr>
</tbody>
</table>
3.3.1 Holyoke, Massachusetts

3.3.1.1 Overall information
3.3.1.2 The Puerto Rican community in Holyoke

3.3.2 Puerto Rico: Salinas & San Juan

3.2.2.1 Salinas
3.2.2.2 San Juan

3.3.3 Motivation for choosing Holyoke, MA and Salinas, PR

4. METHODOLOGY

4.1 Introduction
4.2 Research questions and goals
4.3 Participants

4.3.1 Holyoke Puerto Rican speakers
4.3.2 Puerto Rican Islanders

4.4 Materials

4.4.1 Picture description task
4.4.2 Map task
4.4.3 Reading task
4.4.4 Sociolinguistic questionnaire

4.5 Procedures
4.6 Acoustic analysis and coding

4.6.1 Linguistic variables
4.6.2 Sociolinguistic variables

4.7 Data analysis

5. PRODUCTION EXPERIMENT

5.1 Qualitative results

5.1.1 Phonetic realizations
5.1.2 Frequency distribution
5.1.3 Average durations

5.2 Quantitative results: Statistical analysis
5.2.1 Results for presence vs absence of backed /r/ as dependent variable

5.2.1.1 General results

5.2.1.2 Linguistic variables

5.2.1.2.1 Results for Holyoke

5.2.1.2.2 Results for Puerto Rico

5.2.1.2.3 Summary – the effect of linguistic variables on the production of backed /r/

5.2.1.3 Sociolinguistic variables

5.2.1.3.1 Results for Holyoke

5.2.1.3.2 Results for Puerto Rico

5.2.1.3.3 Summary – the effect of sociolinguistic variables on the production of backed /r/

5.2.2 Results for the acoustic analysis: center of gravity, skewness and kurtosis

5.2.2.1 General results

5.2.2.2 Linguistic variables

5.2.2.2.1 Results for Holyoke

5.2.2.2.2 Results for Puerto Rico

5.2.2.2.3 Summary – the effect of linguistic variables on acoustic aspects of backed /r/

5.2.2.3 Sociolinguistic variables

5.2.2.3.1 Results for Holyoke

5.2.2.3.2 Results for Puerto Rico

5.2.2.3.3 Summary – the effect of sociolinguistic variables on acoustic aspects of backed /r/

5.2.3 Results for the possible /h/-/r/ neutralization

5.2.3.1 General results

5.2.3.2 Linguistic variables

5.2.3.2.1 Results for Holyoke

5.2.3.2.2 Results for Puerto Rico

5.2.3.2.3 Summary – the effect of linguistic variables on acoustic aspects of /h/ and /r/
5.2.3.3 Sociolinguistic variables .................................................. 202

5.2.3.3.1 Results for Holyoke .................................................. 203
5.2.3.3.2 Results for Puerto Rico ............................................. 209
5.2.3.3.3 Summary – the effect of sociolinguistic variables on acoustic aspects of /h/ and /r/ ................................................. 214

6. CONCLUSIONS ........................................................................ 217

6.1 Summary of the results .......................................................... 217
6.2 Intellectual merit .................................................................... 240
6.2 Limitations of the study and future research ............................ 243

APPENDICES

A. PICTURE DESCRIPTION TASK .............................................. 248
B. MAP TASK ........................................................................... 249
C. READING TASK .................................................................... 252
D. HOLYOKE QUESTIONNAIRE IN SPANISH ......................... 257
E. HOLYOKE QUESTIONNAIRE IN ENGLISH .......................... 263
F. PUERTO RICO QUESTIONNAIRE IN SPANISH .................. 269
G. PUERTO RICO QUESTIONNAIRE IN ENGLISH .................. 274

BIBLIOGRAPHY ......................................................................... 279
LIST OF TABLES

Table | Page
--- | ---
1. Trill variation found in previous studies on PRS | 20
2. Negative attitudes towards the velarization and lateralization of /r/ | 26
3. Hispanic or Latino population in Massachusetts | 56
4. States with the highest Puerto Rican population | 56
5. Puerto Rican Population in Massachusetts from 1970 to 2016 | 57
6. Puerto Rican Population in Hampden County and selected cities and towns | 58
7. Hispanic or Latino and Race in Holyoke, MA | 64
8. Holyoke sample based on life stage | 92
9. Holyoke sample based on linguistic proficiency | 93
10. Holyoke sample based on language use | 94
11. Puerto Rico sample based on life stage | 95
12. Puerto Rico sample based on linguistic proficiency | 95
13. Holyoke and Puerto Rico samples based on Socioeconomic status | 110
14. Hierarchy of /r/ variation by place | 130
15. Frequency of backed /r/ and /h/ realizations considering POSITION | 131
16. Overall average durations in milliseconds | 133
17. Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and SPEAKER as a random effect, for the all the data combined | 136
18. Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and SPEAKER as a random effect for the linguistic variables among the Holyoke data | 138
19. Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data | 139
20. Summary of the linear mixed-effect regression model with presence of backed /r/ as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Holyoke data | 142
21. Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data | 149
22. Frequency of backed /r/ realizations by age and generation | 155
23. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect, for the all the data combined | 162
24. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect, for the all the data combined | 164
25. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect, for the all the data combined | 165
26. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among Holyoke data | 166
27. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among Holyoke data | 168
28. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the linguistic variables among Holyoke data ........................................................................................................ 169
29. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data ........................................................................................................ 170
30. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data ........................................................................................................ 172
31. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data ........................................................................................................ 173
32. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Holyoke data ........................................................................................................ 176
33. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among Holyoke data ........................................................................................................ 178
34. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among Holyoke data ........................................................................................................ 179
35. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data ........................................................................................................ 180
36. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data ........................................................................................................ 181
37. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data ........................................................................................................ 183
38. Means for each acoustic measurement for phonemic /r/ and phonemic /h/............. 186
39. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect, for the overall /h/-/r/ data........................................ 186
40. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect, for the overall /h/-/r/ data........................................ 188
41. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect, for the overall /h/-/r/ data........................................ 189
42. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among the Massachusetts data. Only interactions are included.................................. 194
43. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among the Massachusetts data. Only interactions are included.................................. 196
44. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the linguistic variables among the Massachusetts data. Only interactions are included.................................. 196
45. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Interactions are only included .......................................................... 197
46. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Interactions are only included .................................................................................. 199
47. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Interactions are only included .................................................................................. 200
48. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Holyoke data. Interactions are only included .................................................................................. 204
49. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Holyoke data. Interactions are only included .................................................................................. 208
50. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Holyoke data. Interactions are only included .................................................................................. 209
51. Summary of the linear mixed-effect regression model with Center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Only interactions are included .................................................................................. 210
52. Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Only interactions are included .................................................................................. 213
53. Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Only interactions are included .................................................................................. 214
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Word-medial tap production with one apical occlusion for <em>perro</em> ‘dog’</td>
<td>14</td>
</tr>
<tr>
<td>2. Articulation of the tap</td>
<td>14</td>
</tr>
<tr>
<td>3. Word-medial trill production with three apical occlusions for <em>barra</em> ‘bar’</td>
<td>15</td>
</tr>
<tr>
<td>4. Articulation of the trill</td>
<td>15</td>
</tr>
<tr>
<td>5. General distribution of Spanish rhotics by Hualde</td>
<td>18</td>
</tr>
<tr>
<td>6. Articulation of the alveolar trill and the backed variants</td>
<td>24</td>
</tr>
<tr>
<td>7. Geographic distribution of the most common realizations of the backed /r/ in 1928</td>
<td>27</td>
</tr>
<tr>
<td>8. Municipios considered in Graml’s research project</td>
<td>31</td>
</tr>
<tr>
<td>9. Articulation of the glottal [h] and velar [x] fricatives</td>
<td>39</td>
</tr>
<tr>
<td>10. Puerto Ricans in mainland U.S. and in Puerto Rico from 2006 to 2017</td>
<td>52</td>
</tr>
<tr>
<td>11. Puerto Rican relocation Post-Maria (February 2018)</td>
<td>54</td>
</tr>
<tr>
<td>12. Massachusetts Counties</td>
<td>58</td>
</tr>
<tr>
<td>13. Holyoke’s location in Hampden County</td>
<td>61</td>
</tr>
<tr>
<td>14. Population density by ethnicity in Holyoke, MA</td>
<td>63</td>
</tr>
<tr>
<td>15. Income by Location in Holyoke, MA</td>
<td>63</td>
</tr>
<tr>
<td>16. Holyoke’s neighborhoods</td>
<td>66</td>
</tr>
<tr>
<td>17. Map of Puerto Rico showing Salinas</td>
<td>71</td>
</tr>
<tr>
<td>18. Salinas’ Barrios</td>
<td>72</td>
</tr>
<tr>
<td>19. Median Household Income in Salinas</td>
<td>73</td>
</tr>
<tr>
<td>20. Location of San Juan and Salinas in Puerto Rico</td>
<td>74</td>
</tr>
<tr>
<td>21. “El Rincón Boricua” Restaurant in Holyoke, MA</td>
<td>76</td>
</tr>
<tr>
<td>22. Public Art Mural “Isla del Encanto” in Holyoke, MA</td>
<td>76</td>
</tr>
<tr>
<td>23. Nueva Esperanza Inc., a community based program in South Holyoke</td>
<td>76</td>
</tr>
<tr>
<td>24. “Nuestras Raíces”, urban agriculture organization in South Holyoke</td>
<td>76</td>
</tr>
<tr>
<td>25. Normative trill in the utterance <em>perro</em> ‘dog’</td>
<td>103</td>
</tr>
<tr>
<td>26. Word-medial trill production with two apical occlusions for <em>narrar</em> ‘narrate’</td>
<td>116</td>
</tr>
<tr>
<td>27. Approximant trill production for <em>rotito</em> ‘little hole’</td>
<td>117</td>
</tr>
<tr>
<td>28. Fricative trill production for <em>recoestado</em> ‘leaned back’</td>
<td>118</td>
</tr>
<tr>
<td>29. Tap production for <em>rincón</em> ‘corner’</td>
<td>119</td>
</tr>
<tr>
<td>30. Pre-breathy segmental portion (voiced) followed by three occlusions in <em>mirra</em> ‘myrrh’</td>
<td>120</td>
</tr>
<tr>
<td>31. Pre-breathy segmental portion (voiced) followed by one occlusion in <em>Israel</em> ‘Israel’</td>
<td>121</td>
</tr>
<tr>
<td>32. Pre-aspirated segmental portion (voiceless) followed by one occlusion in <em>riñón</em> ‘kidney’</td>
<td>122</td>
</tr>
<tr>
<td>33. Backed /ɾ/ realization in <em>perrito</em> ‘little dog’</td>
<td>123</td>
</tr>
<tr>
<td>34. Backed /ɾ/ realization followed by a tap in <em>corriendo</em> ‘running’</td>
<td>124</td>
</tr>
<tr>
<td>35. Post-tap/trill frication/approximant realization in <em>honra</em> ‘honor’</td>
<td>125</td>
</tr>
<tr>
<td>36. Voiced /h/ realization in <em>ajo</em> ‘garlic’</td>
<td>126</td>
</tr>
<tr>
<td>37. Voiceless /h/ realization in <em>jamon</em> ‘ham’</td>
<td>127</td>
</tr>
<tr>
<td>38. Overall frequency distribution of /ɾ/ variation</td>
<td>128</td>
</tr>
<tr>
<td>39. Frequency distribution of /ɾ/ variation by place</td>
<td>129</td>
</tr>
<tr>
<td>40. Overall frequency distribution of /h/ variation</td>
<td>131</td>
</tr>
</tbody>
</table>
78. Boxplots for center of gravity values for PHONEMIC SOUND and FOLLOWING SOUND in Puerto Rico ................................................................. 198
79. Boxplots for center of gravity values for PHONEMIC SOUND and ORIGIN in Holyoke ........................................................................................................ 205
80. Boxplots for center of gravity values for PHONEMIC SOUND and GENERATION in Holyoke ........................................................................................................ 206
81. Boxplots for center of gravity values for PHONEMIC SOUND and SOCIOECONOMIC STATUS in Holyoke ........................................................................................................ 207
82. Boxplots for center of gravity values for PHONEMIC SOUND and GROUP AGE in Puerto Rico ........................................................................................................ 211
83. Boxplots for center of gravity values for PHONEMIC SOUND and TASK in Puerto Rico ........................................................................................................ 212
CHAPTER 1
INTRODUCTION

1.1 Theoretical issues

Variability is a defining attribute of human speech. This observation is the keystone of sociolinguistic research (Hay & Drager, 2007). Since the nineteenth century, many researchers (Paul, 1880; Pinget, 2015) studied the variation found in speech considering different factors, such as the articulatory constraints and laws of aerodynamics that operate within the vocal tract (Stevens, 1998, in Pinget, 2015), idiosyncratic features (style, speech rate, etc.), sociolinguistic features (origin, age, generation, social identities, etc.) or linguistic contextual features, such as prosody (Hay & Drager, 2007; Pinget, 2015). Variation happens at all levels of linguistic representation, but as Hay and Drager (2007) state, the study of socially conditioned phonetic variation in speech (or Sociophonetics) has been more predominant than in any other linguistic discipline. The ultimate goal of this field of study is to show that variation is systematic. Speakers produce sounds differently from one another and such variation can be sociolinguistically constrained. In fact, the most recent use of instrumental or acoustic phonetic analysis “is beginning to reveal that even extremely fine phonetic details are learned and transmitted for social means” (Hay & Drager, 2007, p. 90).

Beginning with Labov’s work in New York City (1966) and Martha’s Vineyard (1963), many sociophonetic studies have examined the systematicity of phonetic variation based on social factors within the variationist framework (Guy, 1981; Milroy, 1987; Romaine, 1978; Trudgill, 1974). One of the most critical Labov’s motivations in
those early studies was to show, against previous research on American urban dialects (Hubbell, 1950), that “linguistic variation is not random, inexplicable, or theoretically irrelevant” (Foulkes, Scobbie, & Watt, 2010, p. 707). Such framework helps to better understand speakers’ linguistic choices and the way that those variants spread throughout a given community. An example of the importance of sociolinguistic factors in affecting and directing change is the use of coda /r/ in some varieties of English (North America, Scothland, and Ireland), which has been proven to index social class (Labov, 1972; Romaine, 1978; Stuart-Smith, 2007). Speakers with higher socioeconomic status use more consonantal rhoticity than those speakers with lower socioeconomic status (Foulkes et al., 2010, p. 705). The relationship between class and rhoticity is arbitrary since in other areas, (e.g. England), coda /r/ conveys the opposite social evaluation: rhoticity is seen as a sign of low social status (Wells, 1982). As Hay and Drager (2007) mention, the variationist methodology usually invokes the apparent-time hypothesis, in which speakers’ phonological systems persist balanced during their adulthood. Therefore, any differences between older and younger speakers examined at the same are considered to be changes in progress (Bailey, 2001; Hay & Drager, 2007). In addition to focusing on the socio-indexical differences in speech, sociophonetic research has examined stylistic variation, focusing on changes in formality of setting (Labov, 1972) or in audience composition (Bell, 1984). For example, Hay et al. (1999) demonstrate that the speech of Oprah Winfrey could be anticipated by the ethnicity of the referee. The monophthongization of /ay/ is promoted by African-American referees, while such variable is disfavored when the referee is non-African American.
The majority of these pioneering sociophonetic studies consist of “auditory analysis of variables into (often binary) phonetic categories” (Hay & Drager, 2007, p. 91), where data is often divided in terms of gender, age, speaker sex, ethnicity, and social class (Labov, 1966, 1972). As Hay and Drager (2007) explain, there is a historical reason behind this: the only equipment necessary to conduct an auditory analysis is a tape recorder, whereas more specialized equipment is required for an acoustic analysis. Moreover, the most common statistical techniques for sociolinguistics take categorical variables as input (Tagliamonte, 2006). Such auditory analysis fails, then, to capture variation in a multidimensional continuum. As Podesva (2006, p. iv) argues: “an approach taking into account the phonetic details of variation can bring to the surface a rich palate of meanings that cannot be accessed by categorical investigations alone”.

Instrumental analyses of vowel systems have been done in order to corroborate theories of sound change in structuralist linguistics (Hockett, 1965; Martinet, 1955) but also to examine sound changes in progress (Foulkes et al., 2010). Consonant variables, however, have generally been examined auditorily rather than acoustically in sociolinguistic research, although recent analysis have begun to apply more sophisticated analytic techniques (Kissine et al., 2003; Podesva, 2006; Stuart-Smith, 2004). Podesva (2006) examines the speech of three gay men, focusing on the intensity and duration of /t/ release. Findings reveal that phonetic detail is used to create social meaning. The /t/ release is more frequent in professional contexts than social contexts. Such feature is employed in the production of the gay identity, which contrasts with the stereotypically gay style (high pitch, prolonged /l/ and wide pitch ranges), revealing that linguistic styles are “as diverse as the individuals and communities producing them” (2006, p. 187).
Similarly, Stuart-Smith (2004) showed that the spectral energy of the fricative /s/ is manipulated by speakers in Glasgow to convey speaker gender. In addition, Kissine et al. (2003) took into account similar procedures by examining devoicing of Dutch fricatives with the goal of comparing heterosexual and gay speakers. In line with those studies, the current research combines both, auditory and acoustic analyses to better understand the variable of interest: rhotic variation.

Based on variationist sociolinguistics, a new sociolinguistic approach has emerged in recent years which focuses on diasporic communities. This discipline has different denominations depending on the researcher: sociolinguistics of globalization (Blommaert, 2010), sociolinguistics of diaspora (Márquez Reiter & Martín Rojo, 2015) or diaspora linguistics (Bodomo, 2012) (Gubitosi and De Oliveira, in press). Initially, diaspora\(^1\) and migration obtained theoretical importance in sociology, anthropology and cultural studies, as well as in the context of globalization processes, with a focus on social practices, social identities or migration status. In fact, transidiomaticity, deterritorialization or transcultural flows became a core focus of social science (Appadurai, 1996). However, the mobility of communities also implies the mobility of sociolinguistic and linguistic resources, where “sedentary” patterns of language use are complemented by ‘translocal’ forms of language use, and [that] the combination of both often accounts for unexpected sociolinguistic effects” (Blommaert & Dong, 2007, p. 4). Diasporas are, hence, changing communities (Hall, 1990) open to the “lateral

\(^1\) Diaspora refers to the processes of dispersion and displacement where migrants move to a foreign country where they constitute a new imagined community (Anderson 1991).
connections” (Clifford, 1994, p. 306) that are negotiated through language. Previous diaspora research state that such minority communities are threatened by the majority population of their new country where they settle (Ladilova, 2015), allowing for processes of assimilation and acculturation as well as language change, hybridizations or even creolization (Gubitosi and De Oliveira, in press). For example, in the new land or community, the separation from the country of origin could carry pressure towards assimilation in the new host society. Conversely, the use of social media or other electronic contact with the country of origin can contribute to language maintenance in the diaspora (Blommaert & Dong, 2007). Mobility implies different dimensions that can collapse in specific spaces where speech communities live and interact with one another. For that reason, “the structure of people’s repertoires and the patterns of multilingual language use… become less predictable and significantly more complex” (Blommaert, Collins & Slembrouck 2005; Collins, 2007, in Blommaert & Dong, 2007, p. 4). As Vertovec (2007) states, diasporic communities create sites of super-diversity, with different patterns, as a result of these new dialect contact situations.

As Blommaert and Dong (2007) describe, new migrants normally settle in older immigrants’ areas where the older immigrants are the ones renting spaces to the new ones. Those new communities live in legally and economically insecure conditions, dependent upon solidarity networks such as religious institutions (Blommaert et al., 2005; Blommaert, Collins & Slembrouck, 2005). With regards to linguistic practices, these new areas generate multilingual repertoires in which the lingua franca and migrant languages are used. The immigrants maintain intensive contact with different networks, including involving their community of origin. Hence, their spatial organization is “local as well as
translocal, real as well as virtual”, having effects “on the structure and development of language repertoires and patterns of language use” (Blommaert & Dong, 2007, p. 8).

Moreover, Blommaert and Dong (2007) bring attention to the fact that such neighborhoods often embroil collaborative work to assist or translate when necessary in order to accomplish communication. The diasporic community under study in this dissertation belongs to this description of globalized neighborhoods, as described in Chapter 3.

The concepts of space and mobility get primary importance in this new sociolinguistic discipline. Migration implies the change in the spatial organization of someone’s life, “emigrating and immigrating”, leaving and settling (Blommaert & Dong, 2007). “A sociolinguistics of globalization is necessarily a sociolinguistics of mobility” (Blommaert, 2003, p. 611): language and any other cultural manifestations travel across space and time. Such movement across space does not happen over empty spaces. They are home to certain communities, filled with conceptions, rules, expectations of what is proper or not in language use.

“Mobility, sociolinguistically speaking, is therefore a trajectory through different stratified, controlled and monitored spaces in which language ‘gives you away’.

Big and small differences in language use locate the speaker in particular indexical –that is, identity role, ascriptive categories (…) (Blommaert & Dong, 2007, p. 6).

Moreover:

“What is globalized is not an abstract language, but specific speech forms, genres, styles, and forms of literacy practice. And the way in which such globalized
varieties enter into local environments is by a reordering of the locally available repertoires and the relative hierarchical relations between ingredients in the hierarchy. Sociolinguistic globalization results in a reorganization of the sociolinguistic stratigraphy, a process which does not necessarily lead to a new solid and lasting hierarchy but may be seen as an ongoing, highly volatile process cross-cut, again, by matters of scale” (Blommaert, 2003, p. 608).

As suggested in the first quote, this new approach addresses the language-ideological level. On the basis of such discipline, language identities, ideologies and attitudes affect language change (Gal & Woolard, 2001; Schieffelin, Woolard & Kroskrity, 1998). “The key to understanding the processes of ‘globalized’ insertion of varieties into newly stratified orders of indexicality, is to discover what such reordering of repertoires actually mean, and represent, to people” (Blommaert, 2003, p. 609). As an example, due to reterritorialization practices (Rosa 2015), Gubitosi and De Oliveira (in press) explain that segregation and isolation could happen among diasporic communities, strengthening speakers’ identities and “remap[ping] the linguistic and social landscape of their new home”. In fact, the Portuguese diaspora in Massachusetts have restructured and re-territorialized the new homeland to identify it with the motherland, the Azorean archipelago (Gubitosi and De Oliveira, in press). Language, hence, is the central factor of this process, allowing the speakers to negotiate their identities through it.

In sum, this approach questions how any domain of language use must be examined as a piece of a global social system (Coupland, 2003, p. 466) where the hierarchical difference between global and local indexicalities is lost: they happen
simultaneously (Blommaert, 2003, p. 601).

The present study follows the variationist methodology from previous sociophonetic studies, incorporating the main theoretical issues from the sociolinguistics of globalization. Therefore, we focus “not on language-in-place but on language-in-motion” (Blommaert & Dong, 2007), giving prevalence to the concept of mobility among diasporic communities and being aware of the different possible spatiotemporal frames or scales that could interact with one another.

1.2 The present study

“The Puerto Rican migration has brought to the American melting pot many intriguing issues” (Ramos-Pellicia, 2004, p. 1). Since Puerto Ricans are U.S. citizens, their migration patterns are similar to any migration movement within mainland U.S. (Ramos-Pellicia, 2004). However, they maintain significant differences in cultural heritage, “their culture and language are as foreign to general American culture as the cultures and languages of many other ethnic groups that migrate from Europe, Africa, the Middle East and Asia” (Ramos-Pellicia, 2004, p. 1).

The aim of the present sociophonetic study is to examine how language variation plays out in the Puerto Rican diaspora in Western Massachusetts in the U.S., where Puerto Rican migration waves have been in constant increase since 1950. Concretely, research on diasporic speakers takes place in the City of Holyoke, Massachusetts. In spite of presenting the largest per capita population of Puerto Ricans living outside the U.S. (U.S. Census, 2010), this area has not been studied in this regard.
Instead of losing their lifestyle and ethnic identity through the Americanization process, there is a continued affirmation of Puerto Rican identity in Holyoke, MA, where Puerto Rican migrants shaped a new environment similar to that of Puerto Rico. They maintain their self-identity and cultural traditions, which is evidenced through language (Spanish), music, dance, religion and the Puerto Rican flag. All of them serve as a vehicle for cultural survival, affirming a Puerto Rican identity (Rodríguez, Sánchez Korrol, & Alers, 1984, p. 7). Consequently, without leaving aside the “lateral connections” (Clifford, 1994, p. 306) that are negotiated through language, in this dissertation we argue that in Holyoke, MA the back and forth migration waves with the land of origin as well as the use of social media or other electronic contact within the Puerto Rican community contribute to language maintenance in the diaspora (Blommaert & Dong, 2007).

The specific variable of interest is the Puerto Rican Spanish (henceforth PRS) rhotic trill, a sound known to show a great deal of variation in Puerto Rico: [ɾ], [r], [h], [hr], [hr], [ɾ], [xr], [xr], [x], [R], [χ] (Graml, 2009; Navarro Tomás, 1948). Using experimental designs and acoustic measurements, the first step of this dissertation is to provide a comprehensive description of the PRS onset trill variation in Western Massachusetts and show whether there are differences among those realizations found in the diaspora and in the regions where those communities come from on the island (Salinas and San Juan). Special attention is paid to the backed /r/, considered a stigmatized sound in PRS, as well as a very strong marker of local Puerto Rican identity.

In light of this trill variation, recent work suggests that the phonemes /ɾ/ and /h/²

---
² In Caribbean Spanish, the glottal allophone [h] is produced for the phoneme /x/ (Lipski, 2008). Throughout this dissertation focused on PRS, it is refered to [h] as an allophone of /h/.
have been perceptually neutralized in PRS (Delgado-Díaz & Galarza, 2015). This implies a sound change in progress on the island of Puerto Rico, which would result in the loss of a phonemic contrast in a minimal pair such as *Ramón* (‘Raymond’) vs. *jamón* (‘ham’).

The second step of this study is to continue this research, understanding whether there is evidence for this production neutralization in the Western Massachusetts diaspora using measurements such as center of gravity, skewness, and kurtosis (Haley, Seelinger, Callahan Mandulak & Zajac, 2010; Jongman, Wayland & Wong, 2000). Such measurements have not been used to study this variable in Puerto Rican Spanish in the past.

Previous sociolinguistic research has shown that social diversity and change can result in fluctuations in the linguistic variation present in a specific speech community, suggesting that both phonetic and social processes cooperate in a very complex manner (Foulkes et al., 2010). In line with this belief, linguistic and sociolinguistic factors such as generation, voicing pattern and phonological context (Delgado-Díaz & Galarza, 2015; Ramos-Pellicia, 2004, 2007; Willis, Delgado-Diaz, & Galarza, 2015) will be examined in order to determine whether they affect the /r/-/h/ merger as well as the backed /r/ variation. It was postulated that given the constant back and forth movement in this community, PRS speakers in Western Massachusetts will mirror the patterns found on the island. Therefore, this project will contribute to the literature on PRS in the United States, showing interesting implications for the body of research on language variation and change in diasporic communities.

As mentioned in the theoretical section, the majority of previous sociolinguistic research dealing with phonological variables is based on auditory analysis (Foulkes et al.,
Drawing from the more recent subfield of sociophonetics, this dissertation combines auditory and acoustic analyses in order to better understand the phonetic details of variation that cannot be examined by categorical analyses alone. Therefore, in order to differentiate perceived identical segments by their acoustic properties, the present study takes into account acoustic measurements that have been shown to be useful in differentiating place of articulation among fricatives: center of gravity, skewness, and kurtosis (Erker, 2010; Gordon, 2002; Gradoville, 2011; Haley et al., 2010). Following Erker (2010), who used duration and center of gravity to study coda /s/ in Dominican Spanish, continuous measurements would more adequately show the relationship between the variation found in the acoustic signal and its conditioning factors and to better identify patterns of variation.

By providing a detailed analysis of trill variation for different generations of PRS speakers in Western Massachusetts and analyzing whether there is continuity between linguistic and sociolinguistic patterns and those of Puerto Rico, this research project will add to the growing body of literature in the field of Spanish sociophonetics. Additionally, this study would be the first analysis to extract center of gravity, skewness and kurtosis measurements (Haley et al., 2010) to better understand the points of articulation involved in fricative realizations of /r/ (Ramón, Raymond [xa.món]) comparing them to those allophones produced by PRS speakers for /h/ (jamón, ham [ha.món]). In this manner, the present research project will contribute not only to the body of research on changes in progress in the diaspora, but also to the literature on the phonology of Spanish spoken in the United States.
In addition to the theoretical goals of this study, there are also practical objectives. Due to the large Puerto Rican population in the mainland U.S., the consideration of PRS trill variation as well as the potential /h/-/r/ merger are also relevant for instructional purposes. We believe that the linguistic situation in a community involving Latino immigrants needs to be studied and documented first in order to promote the incorporation of its educational policy to the classroom and inform instructors working in the design of teaching materials for the Latino community. This study has the potential to help instructors and speakers in general to understand that the backed /r/ realization is a legitimate variation that exists in Puerto Rico as well as in the diaspora, avoiding teaching corrections and the stigmatization of its use.

This dissertation consists of six chapters. Following this introduction, Chapter 2 discusses the phenomenon under study, rhotic variation in Puerto Rican Spanish, focusing on the backed /r/ realization and the evidence for the possible merge of /r/ and /h/ in this variety. A demographic, historical, social and linguistic description of the Puerto Rican Diaspora in Western Massachusetts is presented in Chapter 3. It also describes the communities under study in Mainland U.S. (Holyoke, MA) and on the island of Puerto Rico (Salinas and San Juan). Goals, research questions and hypothesis are detailed in Chapter 4. The methods used in this study are also described. Qualitative and quantitative findings are presented and preliminary discussed in Chapter 5. Chapter 6 summarizes the findings and concludes the dissertation.
CHAPTER 2

THE PHENOMENON: RHOTIC VARIATION IN PRS

2.1 Phonetic description and distribution of /r/

According to prescriptive descriptions, there are two contrastive rhotic sounds in Spanish, a trill /ɾ/ and a tap /ɾ/. Although both phonemes tend to be produced as [+voiced] and [+alveolar], the phonetic realizations for the phonemic trill and tap vary greatly not only articulatorily, but also acoustically (Bradley & Willis, 2012; Martínez Celdrán, 1998; Quilis, 1993). On the one hand, for its normative realization, the Spanish tap /ɾ/ is produced with the tongue apex touching the alveolar ridge once (Figures 1 and 2). The variable of interest in this dissertation, the Spanish trill /ɾ/, is characterized as having two or more brief occlusions between the alveolar ridge and the tongue apex for its normative realization (Hualde, 2005; Martínez Celdrán, 1998; Quilis, 1993). (Figures 3 and 4).
Figure 1: Word-medial tap production with one apical occlusion for *perro* ‘dog’

<table>
<thead>
<tr>
<th>r</th>
<th>tap</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>r</td>
</tr>
</tbody>
</table>

| perro |

Figure 2: Articulation of the tap (screenshot taken from *Sounds of Speech*, University of Iowa, http://soundsofspeech.uiowa.edu/index.html#spanish #spanish)

[\textsuperscript{r}] La punta de la lengua se adelgaza para elevarse rápidamente y tocar con sus bordes los alvéolos de los incisivos superiores en un movimiento que es a la vez ascendente y hacia dentro.
Figure 3: Word-medial trill production with three apical occlusions for *barra* ‘bar’

Figure 4: Articulation of the trill (picture taken from *Sounds of Speech*, University of Iowa, [http://soundsofspeech.uiowa.edu/index.html#spanish](http://soundsofspeech.uiowa.edu/index.html#spanish))
Solé (2002) points out that, in first language acquisition, trills are sounds characterized for being mastered very late. In fact, they are not even present in the babbling stage, the moment in which infants start exploring the different possibilities that the vocal tract entails (Solé, 2002). Along with sibilants, trills are, in fact, the last sounds to be acquired (Jiménez, 1987; Solé, 2002, Vihman, 1996). Therefore, not only second language learners, but also native speakers can have difficulties in order to produce the articulatory conditions for multiple occlusions. All this suggests that, as Recasens and Pallarès (1999) describe, rhotic trills are articulatorily difficult to produce, requiring a specific control of aperture and airflow with a very small deviation in subglottal and oropharyngeal pressure (Henriksen, 2014; Solé, 2002). Thus, a minimal imbalance in tension needed can motivate the lack of vibration (Solé, 2002).

However, it is noteworthy to mention that, despite this complex production mechanism, which requires not only a critical positioning of the articulators but also specific aerodynamic conditions, trills are common realizations in phonological systems (Solé, 2002). In fact, Maddieson (1984) reveals that 36.4% of all languages of the world present a trill. Therefore, as Solé (2002, p. 656) states: “trills are not rare segments, one in every three languages has an apical trill, but have a relatively low frequency in phonological systems as compared with other segment classes, presumably because the production mechanism is quite complex”.

Due to these precise articulatory and aerodynamic requirements, beyond normative Spanish, previous studies have shown a vast amount of trill variation. As Hammond (2000) states, among Spanish consonants, rhotics present the most allophonic variation. Such realizations can differ in terms of manner of articulation (approximants,
fricatives, taps, flaps and vocoids), place of articulation (coronal, velar, and uvular), or laryngeal setting (voiced, voiceless, breathy voiced) (Blecua, 2001; Bradley, 2006; Bradley & Willis, 2012; Colantoni, 2001; Díaz-Campos, 2008; Henriksen, 2014; Henriksen & Willis, 2010; Willis, 2006, 2007; Willis & Bradley, 2008). Due to this amount of variation, it has been argued that there is not a specific phonemic trill realization in Spanish varieties that can be characterized as the prototypical variant (Bradley & Willis, 2012; Henriksen, 2014; Henriksen & Willis, 2010).

In sum, the Spanish phonemic trill is not always produced as the Real Academia Española (1992) states, as *sonido apicoalveolar vibrante multiple*, ‘apico-alveolar trill’. Its prescriptive accuracy is affected by different external factors such as speech style, dialectal variation, the articulatory quality and aerodynamic conditions of the adjacent sounds and other social factors (Henriksen, 2014).

With respect to rhotic distribution (see Figure 5 below for examples with IPA transcription), the trill contrasts with the tap in intervocalic position. Consequently, there are minimal pairs such as *caro* ‘expensive’ vs. *carro* ‘car’. Moreover, due to coarticulation, a word-internal tap might contrast with a word-initial trill following a vowel in connected speech (e.g. *aroma* ‘aroma’ vs. *a Roma* ‘to Roma’) (Hualde, 2005, p. 183). Outside of those contexts, trill and taps can be produced in complementary distribution: trills occur in word-initial position (e.g. *roca* ‘rock’) or after a heterosyllabic consonant (/n,l,s/) (e.g. *Israelita* ‘Israeli’) and taps appear after tautosyllabic consonant in an onset cluster (e.g. *broma* ‘joke’) or word-final position before a vowel-initial word (e.g. *ser amigos* ‘to be friends’). Finally, trills and taps can be produced in free variation
in word-medial position before a consonant (e.g. *parte* ‘part’) and in word-final position before a consonant or a pause (e.g. *ser poeta* ‘to be a poet’, *ser o no ser* ‘to be or not to be’).

In this dissertation project, trill variation will be analyzed in three positions: intervocalic position, word-initial position and after /n, l, s/.

(1) a. Contrast tap /ɾ/ vs. V__V Intervocalic trill /ɾ/ /karo / ‘expensive’ vs. /karo / ‘cart; car’

b. Only trill /ɾ/ #__ Word-initial /roka / ‘rock’

C. After a heterosyllabic consonant /alrededor/ ‘around’, /enredo / ‘mess’, /israelita / ‘Israeli’

c. Only tap /ɾ/ C__ After a tautosyllabic consonant /broma / ‘joke’, /gramo/ ‘gram’

V__#V Word-final before a vowel /ser amigos / ‘to be friends’

d. Variable rhotic V__C Before a consonant /parte / [parte ] ~ [parte ] ‘part’

V__#C Word-final before a consonant /ser poeta / ‘to be a poet’

V__## Word-final before a pause /ser o no ser / ‘to be or not to be’

Figure 5: General distribution of Spanish rhotics by Hualde (2005, p. 183)

2.2 Variation of /ɾ/: Puerto Rican Spanish
As stated above, phonemic /r/ variation has served as a defining feature of Spanish dialectal variation, in both Peninsular and Latin American varieties (Bradley, 1999; Colantoni, 2001; Lipski, 1994). There are a number of distinct dialectal realizations of the Spanish phonemic trill which, aside from the normative realization (or ‘standard’, in consonance with the Real Academia Española), include: an assibilated trill (Central Mexico, parts of Central America, in the Andean region, Paraguay and northern Argentina) (Hualde, 2005); a velar or uvular fricative, which is of particular interest in the present study (both at the level of individual idiosyncrasy and as a dialectal variant in Puerto Rico (Graml, 2009; Hualde, 2005; Lipski, 1994); a pre-breathy trill followed by either a tap or trill (in Puerto Rico, Cuba and the Dominican Republic) (Hualde, 2005; Willis, 2006, 2007); a retroflex trill or approximant (Costa Rica, Guatemala, Nicaragua and in some varieties of Spanish in the United States) (Hualde, 2005; Lipski, 1994; Ramos-Pellicia, 2004); a tap (some Mexican regions such as Yucatán) and strengthening of rhotics in codas and onset clusters (northern and western Spain) (Hualde, 2005). This variation has not only been analyzed in monolingual speech communities such as Cibao and Santo Domingo, Dominican Republic (Willis, 2006, 2007; Willis & Bradley, 2008); Jerez de la Frontera, Spain (Henriksen & Willis, 2010) or Veracruz, Mexico- (Bradley & Willis, 2012), but also in diasporic and bilingual speech communities where Spanish is spoken as a minority language, such as the contact situation of the United States: Lorain, Ohio (Ramos-Pellicia, 2004, 2007), Grand Rapids, Michigan (Valentin-Márquez, 2007) or Chicago, Illinois (Henriksen, 2014).

While PRS is certainly not the only dialect that presents trill variation, Hammond (2000) found that among all countries considered (Puerto Rico, Mexico, Peru, Chile,
Costa Rica, Venezuela, Argentina, Ecuador, and Colombia), Puerto Rico presents the greatest amount of trill variation. In fact, as many as twenty-two realizations have been shown in previous studies: [ɾ], [r], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ], [ɾ]
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>[xɾ]</td>
<td>Velar + alveolar Fricative velar + tap</td>
<td>Graml (2009)</td>
</tr>
<tr>
<td>[z] [g]</td>
<td>Uvular Uvular fricative</td>
<td>Valentin-Márquez (2007)</td>
</tr>
<tr>
<td>[h]</td>
<td>Glottal Glottal</td>
<td>Graml (2009)</td>
</tr>
</tbody>
</table>
Table 1 presents the allophonic variation for the PRS trill among 5 different studies: Navarro Tomás (1948), Vaquero & Quilis (1989), Hammond (2000), Valentín-Márquez (2007) and Graml (2009). The main exhaustive analysis on trill variation in PRS on the island of Puerto Rico come from Vaquero and Quilis (1989) and Graml (2009). Vaquero and Quilis (1989) were the first researchers who presented a detailed spectrographic analysis of trill and tap variation, revealing 8 alveolar realizations and 4 backed or velar realizations for the phonemic trill. They show a detailed description of each realization in terms of place and manner of articulation, and provide other information such as duration, number of occlusions or sonority. However, aside from describing the rhotic allophonic variation, they do not analyze the linguistic variables affecting their use. Graml (2009) distinguishes up to 11 different variants: trill [r], tap [ɾ], an approximant post-alveolar [ɻ], a fricative post-alveolar [ɬ], a pre-aspirated trill [hr], a pre-aspirated tap [hr], a fricative velar [x], a fricative velar & trill [xr], a fricative velar & tap [xɾ], an uvular trill [ʀ] and a glottal variant [h]. Graml’s study is examined in the following section along with Valentín-Márquez’s (2007) dissertation work, a pivotal comparative study between Puerto Rico and the Mainland U.S.

2.3 The Backed /r/

2.3.1 Phonetic description and socioindexical meaning

Although “[t]he most prototypical member of the class of rhotics are trills made with the tip or blade of the tongue” (Ladefoged and Maddieson 1996, p. 215) some European languages present a dorsal trill realization, such as German or Dutch. This backed realization is also found in different Romanic languages such as French, the Sicilian
variety of Italian (Haden, 1955) or some varieties of Portuguese (Rogers, 1948). In Spanish, while some researchers have claimed that the backed /r/ is a unique phenomenon of Puerto Rican Spanish, others have shown that it is found sporadically in other Spanish varieties (Graml, 2009), such as in Dominican Spanish (Henríquez Ureña, 1940; Terrel, 1980), Cuban Spanish (Cuéllar, 1971; López Morales, 1971), and in Key West, Florida due to Cuban migration (Graml, 2009), as well as Mexico (Alvar, 1969), Panama, Colombia, Venezuela, and Trinidad (Graml, 2009).

Despite finding this fricative realization in other Spanish varieties, the backed /r/ is considered a salient feature of PRS along with coda /s/ weakening [eh.tá] for estás or liquid neutralization [a.mól] for amor (Potowski, 2015). This backed /r/ has different variants itself, as can be seen in Figure 6. Of these backed /r/ variations, [x] is the only allophone whose conditioning has been described in the literature on PRS (Graml, 2009; Valentín-Márquez, 2007). Its origin remains uncertain; however, it has been postulated that it is the result of an exterior influence from the French, African slaves, or the indigenous Taíno population (Valentín-Márquez, 2007).

---

3 Previous studies that analyzed the backed /r/ (Graml, 2009; Valentín-Márquez, 2007) use the term ‘velarization’ for any realization with a posterior feature, without distinguishing among velar or uvular realizations. In this dissertation project, we use the term backed /r/ is used for the same purpose, to include both velar or uvular realizations.
Figure 6: Articulation of the alveolar trill and the backed variants (picture taken from Sounds of Speech, University of Iowa, http://soundsofspeech.uiowa.edu/index.html#spanish)

With respect to the socioindexical meaning of backed articulations of /r/, members of the Puerto Rican speech community in Puerto Rico have been shown to associate it with rural origin or low sociocultural class (Dillard, 1962; Graml, 2009; Medina-Rivera, 1997; Navarro Tomás, 1948). Moreover, it is not only stigmatized
throughout Latin America, but also considered ‘incorrect’ by academic institutions in Puerto Rico (Valentín-Márquez, 2007). In fact, Navarro Tomás (1948, p. 93), one of the main Hispanic phoneticians of the twentieth century, describes this realization in the following terms: “Al forastero le queda el recuerdo de la rr velar como uno de los rasgos más salientes y menos favorables de la pronunciación puertorriqueña” (“velarized rr remains in the outsider’s memory as one of the most salient and least favorable features of Puerto Rican pronunciation”) (Valentín-Márquez’s (2007, p. 73) translation).

Likewise, Tío (2001, p. 49) claims that the backed /r/ along with the lateralization of /r/ are “vicios que apartan el español de Puerto Rico de la lengua general y que además de afean la lengua ponen un sello de inferioridad social o cultural en el hablante” (“vices that set apart Puerto Rican Spanish from the general language and that, besides making the language ugly, put a seal of social or cultural inferiority on the speaker”) (Valentín-Márquez’s translation, 2007, p. 73). Similarly, Matluck (1961, p. 334) states: “para quien visite la Isla, [la rr velar] es el más desconcertante de todos los fenómenos lingüísticos puertorriqueños” (for whoever that visits the island, [the backed ‘r’] is the most desoncertant Puerto Rican linguistic feature).

Those stigmatized opinions are also present among participant’s judgements about language attitude. The following table (Table 2) is taken from Valentín-Márquez (2007, p. 74) who summarizes the negative attitudes towards liquid features in PRS. The column on the left refers to the backed /r/.
Table 2: Negative attitudes towards the velarization and lateralization of /r/. (Source: copied from Valentín-Márquez, 2007, p. 74).

<table>
<thead>
<tr>
<th>Studies</th>
<th>Percentages of informants that expressed negative judgments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>velarization of (rr)</td>
</tr>
<tr>
<td>López Morales 1979</td>
<td>66.5</td>
</tr>
<tr>
<td>Matta de Fiol 1981</td>
<td>85.0</td>
</tr>
<tr>
<td>Emmanuelli 1986</td>
<td>32.8</td>
</tr>
<tr>
<td>Medina Rivera 1997</td>
<td>66.5</td>
</tr>
</tbody>
</table>

With the exception of Emmanuelli (1986), a study conducted in Ponce, Puerto Rico (South) that found positive/neutral attitudes regarding the backed /r/ realization, the majority of studies have found primarily negative attitudes around this linguistic feature (López Morales, 1979a; Matta de Fiol, 1981; Medina-Rivera, 1997). López Morales (1979a) was the first researcher who studied the attitudes associated to the backed realization in Puerto Rico. Since the researcher distributed a survey only among university students in San Juan, this study is not able to offer an exhaustive analysis on regional differences regarding attitudes related to backed /r/. Matta de Fiol (1981) published a similar quantitative study after analyzing the answers of 60 speakers in San Juan.

Medina-Rivera’s (1997) analysis as well as other studies (Lamboy, 2004; López Morales, 1983) suggest that positive attitudes toward backed realizations may be seen as a sign of Puerto-Ricanness, not only in Puerto Rico but also in the diaspora.

### 2.3.2 Main studies on backed /r/ in PRS

In the following paragraphs, an overview of the noteworthy studies analyzing PRS backed /r/ variation is provided. Both studies on the sociolinguistic distribution of backed /r/ realizations on the island of Puerto Rico and in the Puerto Rican communities in the
U.S. diaspora are considered. It is important to mention that, most of the time, these studies refer to any realization with a posterior feature as “velarization”, without distinguishing among velar or uvular realizations.

Although it is not quantitative, Navarro Tomás’ (1948) linguistic atlas stands out due to its detailed comprehensive analysis of Puerto Rican Spanish. His study is the first description of the PRS consonantal inventory, being considered the most pioneering research project in Hispano-American dialectology (Graml, 2009). He combined his impressionistic observations with the results from a questionnaire based on 445 questions related to phonetic, morphologic, syntactic and lexical features. Furthermore, he is the first researcher to describe the backed /ɾ/ realization, showing its geographic extension through the island and finding that it is more common in the southeastern and northwestern regions of Puerto Rico (see Figure 7 below).

Figure 7: Geographic distribution of the most common realizations of the backed /ɾ/ in 1928 (Navarro Tomás, 1948) (Source: figure taken from Valentín-Márquez, 2007, p. 50)

As Valentín-Márquez (2007) points out, it is important to mention that Navarro

Navarro Tomás’s fieldwork followed methods for dialectology from the 19th century, and therefore the majority of the participants were rural men with low education. Furthermore, only one speaker was considered for each community of the study. Three
decades after Navarro Tomás’s study, different linguistic and sociolinguistic analysis on the backed /r/ emerged, focusing on different areas of the island of Puerto Rico.

Among the studies from the island, specifically the metropolitan area, Terrel (1980), Matta de Fiol (1981), and López Morales (1983) provide comprehensive analyses. Terrel (1980) presents an analysis of liquid variables among educated speakers in San Juan. With regards to the distribution of the backed /r/ realization, findings reveal that this sound is not more common among males than females. Matta de Fiol (1981) analyzed the posterior realizations of /r/ in three different contexts (formal, semi-formal and informal) among 60 speakers of San Juan. The researcher also considered socioeconomic level and sex among the sociolinguistic variables. Results indicate that backed realizations were more common among men, middle-age adults, and lower-socioeconomic status speakers while they were involved in an informal conversation, and middle-lower class participants when they were involved in formal and semiformal styles. López Morales (1983) analyzed rhotic variation among 105 speakers in San Juan. Participants were categorized in terms of sex, age, social class and origin (participants born in San Juan or anywhere else but who emigrated afterwards to San Juan).

Quantitative results showed that the backed /r/ is more common among men, lower socioeconomic participants, individuals from rural origin who emigrated to San Juan as adults, and speakers over 55 years of age. González Vargas provides qualitative data from a survey regarding the frequency of backed /r/ among PRS speakers in San Juan. Since the backed /r/ is not produced among the 10 participants of the island, the researcher concludes that the backed /r/ is disappearing from PRS (Graml, 2009).
Among the analyses focused in the non-metropolitan area, that is to say, the rural municipalities outside San Juan, following a quantitative approach, Hammond (1991) compared rhotic variation between three different groups: participants living in the interior of the island, university students living in the western area and professionals living in San Juan. This categorization allowed him to distinguish among habla jíbara, habla popular and habla culta (or, as Valentín-Márquez, 2007, p. 55 translates, ‘rural speech’, ‘popular speech’ and ‘educated speech’ respectively). In line with López Morales’s (1983) study, men produced more backed /r/ realizations than women. Furthermore, this sound was more common among participants who are identified as speakers of habla jíbara. As Valentín-Márquez (2007) points out, two of the methodological problems of the study is that it was not controlled by age and, since young speakers were predominant in the habla jíbara group and older speakers were more common in the habla culta group, the age factor may account for (some of) those differences.

Medina-Rivera (1997) also investigates Spanish in the non-metropolitan area of Puerto Rico, analyzing the interaction of realizations of the tap and the trill with social and stylistic factors in Caguas, a municipality located in the eastern central region of the island of Puerto Rico. The researcher considered sex, parents’ region of origin, and level of education. In opposition to López Morales (1983) findings, women favored the backed /r/ realization. The author explains this result taking into account level of education; since it is similar across all participants, this variable could have influenced the expected result, that is to say, that men typically produce more stigmatized variants than women. Considering previous sociolinguistic research, men are more indifferent with regards to
the stigma that sounds carry, while women are more aware of linguistic prestige and stigmatization (Labov, 1994). Regarding the stylistic factors, this study considers the type of discourse, formality of situation and the degree of familiarity with the interviewer, suggesting that participants who are familiar with the interviewer produced more velar realizations than those participants who did not know the researcher before the experimental task. This study also revealed that the stigmatized /r/ articulation is more common in group conversations than in one-on-one interviews, making it the first to consider diaphasic variation in the analysis of the backed /r/.

Alers-Valentín (1999) analyzed the distribution of /r/ in the northwestern of Puerto Rico (from Isabela to Mayagüez). Age, sex and education were some of the variables considered. The results show, in line with López Morales (1983), that male speakers produced more backed realizations than women. Moreover, this realization emerged more among older and less educated speakers. Later, Holmquist (2003, 2004) incorporated the influence of participants’ ties to the community in order to analyze the linguistic behavior in Castañer, a rural setting in the central-western area of Puerto Rico. Findings suggest that those participants who have stronger ties or spend more time in the community were more likely to preserve the features of that specific community. In particular, it is found that the backed /r/ is more common among speakers who are involved in farming activities. Age was not a variable affecting the production of the backed /r/ since both old and young speakers produce this realization.

With the exception of Navarro Tomás (1948), all the previous studies on backed /r/ have focused on a specific area of Puerto Rico. Moreover, as Graml (2009) states, since there are methodological limitations (e.g., some studies focus on only one area,
some only on students in San Juan, or other only on illiterate participants) it is difficult to analyze the role of the social aspects affecting the use of the backed /r/. Graml (2009) provides the first analysis focusing on both linguistic and sociolinguistic factors regarding trill variation along different geographical areas on the island of Puerto Rico (see Figure 8). Moreover, she investigates the influence of language attitudes on the use of the backed /r/.

![Figure 8: Municipalities from Graml (2009). (Figure taken from Graml, 2009, p. 101). Graml looks at speakers from western (circles), central (squares), metropolitan area (rhombuses) and eastern (triangles) Puerto Rico.](image)

As presented above, findings revealed 11 different phonetic variants for phonemic /r/: trill [ɾ], tap [ɾ], an approximant post-alveolar [ɹ], a fricative post-alveolar [ʃ], a pre-aspirated trill [hr], a pre-aspirated tap [hr], a fricative velar [x], a fricative velar & trill [xr], a fricative velar & tap [xr], an uvular trill [ʀ] and a glottal variant [h]. Graml demonstrates that the backed /r/ is more common in the central and southern areas of Puerto Rico (30.8%), followed by the western (22.8) and the metropolitan area (17.3%). The non-metropolitan area (i.e. outside of San Juan), thus, is where the backed /r/ is most produced. In terms of the linguistic factors conditioning [x] variation, Graml (2009)
illustrates that the backed /r/ is more common in intervocalic position (39.6%), before /a/ (35.4%) and in unstressed syllables (52.4%). Regarding sociolinguistic factors, although the difference is minimal (1.1%), male speakers produced slightly more backed /r/ than females. In terms of age, Graml divided her participants in 5 groups: group 1 (0-19 years old), group 2 (20-39 years old), group 3 (40-59 years old), group 4 (60-79 years old) and group 5 (80-99 years old). The older groups (4, 3 and 5) produced more backed /r/ than the younger groups (1 and 2). In fact, only 1.1% of the backed /r/ realizations was produced by participants from group 1, while 34.5% produced it from group 5. In terms of socioeconomic status, the lower socioeconomic group produced more backed /r/ than the other two groups. As for education, those participants whose highest level of education was elementary used the backed /r/ more than those with a secondary education or higher. Moreover, findings show that when participants read a list of words or a poem backed /r/ is produced less (15.8%) than when the task involved more spontaneous speech, in line with previous sociolinguistic research (Dillard, 1962; Guitart, 1981).

Finally, it is noteworthy to mention that Graml provides evidence that the backed /r/ is one of the stigmatized sounds being corrected in schools and only accepted at home environments in Puerto Rico.

With respect to the studies focused on the Puerto Rican communities in the Continental U.S., it is important to mention that although migratory waves have existed since before 1900 (Sánchez Korrol, 1994; Valentín-Márquez, 2007), academic researchers did not focus on diasporic Puerto Rican communities in the U.S. until 1972

---

4 Graml uses income to make distinctions with regards to the socioeconomic status.
(Valentín-Márquez, 2007). Moreover, Northeastern cities such as New York or Philadelphia have been the most analyzed areas for Puerto Rican sociolinguistic studies in the U.S. (Fishman Cooper, & Ma, 1972; Flores Ferrán, 2004; Lamboy, 2004; Ma & Herasimchuck, 1972; Otheguy & Zentella, 1990; Poplack, 1979).

Fishman, Cooper, & Ma (1972) described the sociolinguistic norms of 431 PRS speakers in a lower-class Puerto Rican community of Jersey City, New Jersey. As part of this project, Ma and Herasimchuck (1972) analyzed the stylistic and social variation of rhotics (trill and tap) and other non-liquid sounds. The researchers showed five different realizations for the phonemic trill: [r], [χ], [hr], [ɾ], [ʁ] which divided in either standard realizations ([r], [hr], [ɾ]) or non-standard realizations ([ʁ], [χ]). Five styles were considered: casual speech, careful speech, list recitation, text reading and list reading. Results show that normative [r] was predominant among young women with the highest levels of education and from the coastal areas of the island of Puerto Rico; whereas the backed /r/ was more common when the task involved less formality.

Lamboy (2004) compared first generation Puerto Rican immigrants with second generation speakers living in the New York City area, not finding significant differences between those groups. However, some patterns emerged: first generation speakers over age 55 produced backed /r/ at high rates (92% of the time) while for the second generation, young adults produced more backed /ɾ/ than middle-aged or elderly speakers.

With regard to comparative studies between the continental U.S. and the island of Puerto Rico, Jany (2000) focuses on PRS in New York and different regions of Puerto Rico, analyzing the pronunciation of English borrowings. The author’s observations are
limited regarding the backed /r/, mentioning simply that it is a realization found in each group of speakers, but is less common than other phonetic features, such as /s/ weakening. Valentín-Márquez (2007) compares liquids in the PRS spoken in Grand Rapids, Michigan to the PRS spoken in Cabo Rojo, Puerto Rico. For phonemic /r/, he found a similar overall distribution of the variants in both settings under study: two alveolar and three posterior realizations [r], [ɾ], [x], [ʁ], [ʁ], being [r], [ɾ], [x] the most common articulations and, thus, main variables considered for the statistical analysis. In both samples, the most produced articulation was the standard variant [r], reaching up to 60% of all occurrences, followed by [ɾ] (20% of the cases) and [x] (frequencies of around 15%). In fact, given the similar frequencies of each articulation in both settings, the differences between the Grand Rapids and Cabo Rojo samples in terms of /r/ variation were not significant. As to the linguistic factors conditioning this variation, there was a statistically significant arrangement of the realizations in all factors analyzed. In both communities, [r] emerged in stressed syllables, after the fricative [h] and in postpausal contexts. The backed realization, [x] was most common in unstressed syllables and in word-initial position. Moreover, it emerged in postnasal position, which the researcher explains could be related to the articulation of word-final /n/ as a velar articulation, a common realization in the Hispanic world. Furthermore, Valentín Márquez (2007) found one difference in the phonolexical variants considered in his study. The percentages of backed realizations among nationality words (Puerto Rico/puertorriqueño) were higher in Grand Rapids than in Cabo Rojo, implying a relationship between the affirmation of national identity and the use of stigmatized realizations. Another explanation would be
that PRS speakers in Michigan are trying to distinguish themselves from other Spanish speaking communities, using the language creatively.

With regard to the social factors conditioning [x] variation, Valentín-Márquez considers socioeconomic level, sex, life stage, and integration into the Puerto Rican community. For socioeconomic level, the researcher examines occupation and educational attainment. Although it was not statistically analyzed, data show a pattern of higher frequencies of [x] when the speaker pertains to a lower social class. However, in Puerto Rico it is possible to hear the backed /r/ among speakers with graduate degrees, implying the necessity to analyze the interplay of class and style as well as ‘the indexicality of social values associated with the use of [x] to signify in-group membership, perhaps based on perceptions of core Puerto Ricanness” (2007, p. 240). In terms of sex, although the analysis did not yield statistical significance, in Cabo Rojo [x] was more frequently among male speakers while in Grand Rapids, contrary to expectations, men and women produced this feature equally. Similarly, in terms of life stage, differences emerged among generational groups in Cabo Rojo but not in Grand Rapids. Thus, for Cabo Rojo, middle-aged speakers produced more [x] than the adolescents, a result that is consistent with the general opinion about the use of the backed /r/ in Puerto Rico: “[it] is perceived as more typical of older speakers” (2007, p. 241). The author also explains that it is possible that the backed /r/ is gradually diminishing in Cabo Rojo since adolescents produced 7% of [x] occurrences, middle-aged adults articulate 19% and the elderly speakers produce 45% of the overall occurrences. Valentín-Márquez also studies the relationship between language use and national identification, looking at the level of integration into the Puerto Rican
community. Speakers in Puerto Rico presented a higher social network index (SNI) score, implying a higher level of community integration than speakers on Mainland U.S. However, the quantitative distribution of /r/ variants is similar in both settings. These findings reveal that differences between Grand Rapids and Cabo Rojo related to their speakers’ networks do not contribute to differences in the use of the backed /r/.

Ramos-Pellicia (2004) examines rhotics in syllable-final and word-final positions across three generations of PRS speakers in Lorain, Ohio. She also explores this variation on the island of Puerto Rico to determine whether PRS in Lorain displays different or similar patterns in the production of rhotics for the aforementioned phonological positions. For both Lorain and Puerto Rico, lateralization, trill, retroflection and deletion are reported. Results reveal that PRS speakers in Lorain and on the island of Puerto Rico follow a similar pattern in the production of /r-ɾ/. However, Ramos-Pellicia reports different frequency distributions. For example, Lorain Puerto Ricans produce more retroflection, which is almost never used on the island. An interesting finding is that at least for the first generation, speakers in Lorain are aware of the negative prestige that lateralization has in Puerto Rico, so they use other /r/ realizations, such as the retroflex trill in order to avoid the production of [l]. Moreover, the preference for the normative trill declines across the three generations in Lorain, with the first generation favoring normative /r/, and the third generation using more lateralization as well as retroflex /r/ than second and first generations. Although this study does not focus on the backed /r/, results regarding differences among generations are pertinent to the present study,
suggesting that generation could have an effect on backed /r/ variation among the current data.

In conclusion, since Navarro Tomás’ investigation on the island of Puerto Rico in 1948, a prosperous body of research has developed related to the backed /r/ in PRS. These analyses typically examine both linguistic and sociolinguistic factors affecting the production of this stigmatized sound. While linguistic factors account for considerably more variation, sociolinguistic factors play a role as well, including speakers’ origin (more frequent among rural speakers), sex (men produce it more often than women), and age (most commonly present among middle-aged adults) (López Morales, 1983; Matta de Fiol, 1981). Furthermore, the backed /r/ has been shown to be produced more in informal interviews, such as narratives or group conversations than in more controlled experiments, such as reading tasks (Graml, 2009; Medina-Rivera, 1997).

The current dissertation analyzes data collected in Salinas, PR and Holyoke, MA to contribute to the sociodialectal description of the backed /r/ in Puerto Rican Spanish, both in the Northeastern United States and in Southeastern Puerto Rico. Previous studies have made categorical distinctions with regard to the backed /r/; however, this study employs continuous acoustic measurements to better understand backed /r/ variation and determine whether the allophones produced are found on a continuum of backedness that could result in a /h/-/r/ neutralization due to their similarities in terms of place of articulation. This potential merge as well as the importance of using continuous measurements for these kinds of acoustic analyses are discussed in the following section.
2.4 Neutralization of /r/ and /h/ in PRS

Spanish varieties vary with regards to the inventory of fricative phonemes as well as to the phonetic realizations of such phonemes. In this same vein, Hualde (2005) explains that in Standard Peninsular Spanish there are four voiceless fricative phonemes: /f θ s x/, where /x/ is realized as postvelar or uvular. However, in Standard Latin American Spanish there are three voiceless fricatives: /f s x/, where /x/ is medio-velar or laryngeal. Most Caribbean Spanish varieties, including PRS, also use the glottal or laryngeal allophone [h] for the phoneme /x/ ([ha.món] for *jamón*, ‘ham’). In fact, it could be argued that the phoneme in these varieties is /h/ instead of /x/. As mentioned above, following such claim, this dissertation refers to [h] as a phonetic realization of /h/.

As mentioned in the previous section, /r/ has a posterior fricative realization in PRS. Accordingly, there are claims suggesting the neutralization of the /r/ and /h/ phonemes in PRS (Delgado-Díaz & Galarza, 2015; Dillard 1962; Lipski, 1994) due to the similarities in the allophonic realizations of the backed /r/ and /h/: /r/ can be realized as [ɾ], [x], [xr], [xr], [hr], [h] (Graml, 2009; Navarro Tomás, 1948) while /h/ can be produced as [h], [ɦ], or [x] (Hualde, 2005; Willis et al., 2015). Both the /r/ and /h/ phonemes, thus, share [h] and [x] as allophones, favoring the neutralization in those contexts, which could imply a sound change in progress on the island (Figure 9).
The neutralization of a phonemic contrast is not odd within the phonology of Spanish. Many neutralized contrasts have been studied in the previous literature. For example, nasal consonants assimilate the place of articulation of the following consonant. As a result, the contrast among the bilabial, alveolar and palatal phonemes is neutralized in coda position. Regarding plosives, the contrast between voiced and voiceless pairs is more robust in the onset than in the coda, implying a neutralization in voice. Furthermore, the tap and trill can also be neutralized in coda position, as previously mentioned. Besides neutralization, there are other phonological processes that affect the Spanish phonological system, such as assimilation, dissimilation, weakening and deletion, strengthening, epenthesis or metathesis (Hualde, 2005). Many of those processes affect the phonological system in specific Spanish varieties. For example, the potential neutralization analyzed in this dissertation (backed /r/ and /h/) only occurs in PRS. Another process of neutralization well known in the literature is the neutralization of liquids, found not only in the Caribbean but also in the south of Spain. In such areas, there is no contrast between liquids and rhotics in coda position. In Cuban and Chilean
Spanish, liquids also assimilate to the following consonant in coda position, as in pulga [púgga] ‘flea’ (example taken from Hualde, 2005, p. 108). In Andalusian Spanish, /s/ assimilate in place and manner of articulation to the following segment but keeping some of its features such as voicelessness or aspiration, as in costa [kð̃usta] ‘coast’ (example taken from Hualde, 2005, p. 108). In Chilean Spanish, the velar fricative /x/ becomes a palatal fricative [ç] before a glide or front vowel (Flores, 2016; Hualde, 2005). In Eastern Andalusian, final mid vowels have a more open quality in words where a final /s/ is not produced and such open quality could be transmitted to other mid vowels in the same word (Hualde, 2005).

The neutralization of /r/ and /h/ in PRS would imply a phonological process where those two phonemes that are generally found in phonemic opposition, would fail to contrast in some specific environment: onset position. Its study is crucial in order to comprehend the phonological processes that affect the PRS phonological system as well as the fricative Spanish system in general. Thus, a secondary goal of this dissertation is to analyze whether there is such neutralization in production and, if so, what are the linguistic and sociolinguistic factors affecting such neutralization.

With regards to the previous research focused on the /r/-/h/ neutralization, Delgado-Díaz and Galarza (2015) show in a perception experiment that there is an effect of phonological context on the possible neutralization between /r/ and /h/: Puerto Ricans are more accurate when identifying the backed /r/ ([x]) in intervocalic position (70% correct identifications) than in word initial position (38% correct). That is to say, the contrast between the backed /r/ and /h/ tends to be maintained in intervocalic position and lost in post-pausal position. Therefore, there is no evidence to justify a case of complete
neutralization since the contrast is maintained in some contexts. The authors suggest that this tendency can be explained in terms of allophonic distribution and its relation with voicing since Willis et al.’s (2015) findings show that /h/ in intervocalic position is voiced 90.4% of the time, but in initial position is voiceless 86.5% of the time. It is possible, then, that it is more difficult for Puerto Rican listeners to distinguish backed realizations of /r/ from /h/ in word initial position only, where this context does not provide listeners with cues to the phonemic distinction. Delgado-Díaz and Galarza also show that neutralization occurs in the direction /r/→/h/ and not in the opposite direction, since PRS listeners had a higher rate of /h/ identifications, implying that PRS listeners do not perceive /h/ as a backed /r/ ([x]). That is to say, PRS listeners do not perceive Ramón, Raymond [xa.móŋ] as jamón, ham [ha.móŋ]). In contrast, non-PRS-speaking listeners (from Bolivia, Mexico and Spain), mapped the posterior realization of /r/ ([x]) with /x/ in both contexts (intervocalic and post-pausal position), revealing that neutralization among those speakers could be explained as a lack of contact with this variety.

Since this investigation shows that the neutralization is incomplete and possibly subject to sociolinguistic factors, Delgado-Díaz and Galarza (2016) conducted a second study, which analyzed the sociolinguistic implications on the perception of the backed /r/ as well as the interplay of those sociolinguistic factors and the phonetic factors found in the study presented above. Using a lexical identification task, they found that backed /r/ perception is influenced not only by phonological context, but also by speaker’s sex and age, in line with the production patterns and attitudes related with the backed /r/ found in previous research (Graml, 2009). Puerto Rican listeners in Caguas and San Juan are more accurate in perceiving the backed /r/ as a realization of /r/ in intervocalic position (75.9%)
than in initial position (60.6%), in line with their previous study (Delgado-Díaz & Galarza, 2015). Moreover, their results suggest that when the speaker is an older male, listeners do not neutralize the backed /r/ with /h/. Those findings are not surprising if we assume that listeners mold their perception based on the expectations of the speaker’s production. In this regard, previous research has proved that the backed /r/ is more common among older men (Delforge, 2013; Graml, 2009; López Morales, 1983; Medina-Rivera, 1999; Valentín-Márquez, 2007). Therefore, perceptual results reflect the same backed /r/ production and attitude patterns that had been found in previous research.

With respect to the interplay between the role of the phonological context and the social factors, results show that listeners had a high accuracy rate in distinguishing the backed /r/ from /h/ in intervocalic position and a low accuracy rate in initial position, with the only exception of the 70-year-old male. This finding suggests that the analysis of social cues becomes more relevant in those cases in which there is greater phonetic overlap, in this case, initial position (Delgado-Díaz & Galarza, 2015; Willis et al., 2015). Further steps need to be taken in order to better understand /r/-/h/ neutralization research in production and corroborate (or not) the perceptual patterns presented above.

2.4.1 Acoustic analysis among fricatives: continuous measurements

Previous studies have described /r/ as well as /h/ variants but none have used continuous measures to analyze such variation. Considering that research on fricative variation among other varieties has benefitted from continuous analysis, the present study takes into account acoustic measurements that have been shown to be useful in differentiating place of articulation among fricatives in order to better understand the points of
articulation involved not only in the production of the backed /r/ but also in the production of /h/ (Erker, 2010; Gordon, 2002; Gradoville, 2011; Haley et al., 2010). Since categorical descriptions have obscured decisive facts about phonological variation and the factors that favor it (Erker, 2010), those continuous measurements will allow a description of the variation present in speech and determine whether the allophones produced for /r/ and /h/ are classified under one of those phonological categories (/r/ vs. /h/), or conversely, whether there is a continuum among backed realizations which might be favoring the neutralization in PRS. Hence, the analysis of acoustic cues allows the classification or identification of place of articulation in backed fricatives with a higher degree of accuracy. Moreover, they help to combine both auditory and acoustic analysis.

As Jongman, Wayland, and Wong (2000, p. 1252) define, fricatives are known to be produced “with a very narrow constriction in the oral cavity. A rapid flow of air through the constriction (the position of which depends on the particular fricative) creates turbulence in the flow, and the random velocity fluctuations in the flow act as a source of sound (e.g., Stevens, 1971, 1998; Shadle, 1990).” The spectral shape of fricatives is resolved by the shape and size of the oral cavity in front of the constriction (Jongman et al., 2000). That is to say, the shorter the anterior cavity is, the less defined the spectrum is and vice versa, the longer the oral cavity, the more defined the spectrum (Jongman et al., 2000; Stevens, 1998). Subsequently, if we are comparing alveolar and palato-alveolar fricatives, the first ones would be characterized by a flat spectrum while the second ones would display a well-defined spectral shape (Jongman et al., 2000). Moreover, since palato-alveolar fricatives are produced with a longer anterior cavity in comparison to
alveolar fricatives, they would display a spectral peak at lower frequencies (Jongman et al., 2000).

Forrest, Weismer, Milenkovic, & Dougall (1988) created a method to measure the place of articulation among voiceless obstruents. The spectrum is considered as a random probability distribution and its central tendency, dispersion, asymmetry, and shape are calculated (Forrest, Weismer, Milenkovic, & Dougall, 1988). Those measurements, the first four spectral moments, mean, variance, skewness, and kurtosis, have been used to define and measure other consonants and have been demonstrated to be effective with both voiceless and voiced fricatives (File-Muriel & Brown, 2011; Flores, 2016; Jongman et al., 2000; Silbert & de Jong, 2008). The center of gravity of a spectrum is the frequency at which sonic energy is most concentrated (Erker, 2010). Following the previous literature (Flores, 2016; Jongman et al., 2000; Silbert & de Jong, 2008), lower center of gravity values reflect a place of articulation further back along the vocal tract. Skewness refers to the distribution’s asymmetry (spectral tilt or slant of the energy distribution). As Jongman et al. (2000, p. 1253) claim, “positive skewness suggests a negative tilt with a concentration of energy in the lower frequencies”. Therefore, we expect that the higher the value, the more backed the realization. “Negative skewness is associated with a positive tilt and a predominance of energy in the higher frequencies” (2000, p. 1253). Finally, zero values are associated to symmetric distribution of lower and higher frequencies, that is to say, “a symmetrical distribution around the mean” (Jongman et al., 2000, p. 1253). Kurtosis captures the distribution’s peakedness.

---

5 In previous literature, “spectral means” is interchangeable with “center of gravity”.
“Positive values indicate a relatively high peakedness (the higher the value, the more peaked the distribution), while negative values indicate a relatively flat distribution. Positive kurtosis thus suggests a clearly defined spectrum with well-resolved peaks, while negative kurtosis indicates a flat spectrum without clearly defined peaks” (Jongman et al., 2000, p. 1253).

It is noteworthy to mention that previous studies have not been able differentiate place of articulation in a speaker-invariant manner. For instance, for the palato-alveolar (/ʃ/) and alveolar (/s/) fricatives in English, inter-speaker overlap obscures place of articulation, although statistically significant group differences have been shown for the spectral mean (Jongman et al., 2000; Nittrouer, 1995). As Haley et al. (2010) point out, to obtain a more reliable and complete analysis of the spectrum and reduce inter-speaker variability, some studies have combined two or more spectral moments, particularly mean and skewness, to distinguish between sibilant fricatives (Forrest et al., 1988; Jongman et al., 2000).

Three studies using continuous measurements contribute to the design of the present project: Jongman et al. (2000), Erker (2010) and Flores (2016). Jongman et al. (2000) present a comparative analysis of acoustic characteristics of English fricatives. In order to classify fricative place of articulation, the authors considered spectral (peak location, spectral moments, locus equations, F2 onset), amplitudinal (overall and relative noise amplitude) and temporal measurements (noise durations). Their findings reveal that spectral peak location, spectral moments, and amplitude consistently serve to differentiate among four places of articulation: labiodental /f,v/, dental /θ, ð/, alveolar /s,z/ an palato-alveolar /ʃ, ʒ/, contributing to the body of research on the mapping between phonetic
categories and acoustic properties among English fricatives. Moreover, with regard to those spectral properties that differentiate place of articulation, a main effect is found not only for linguistic factors (voicing) but also for sociolinguistic ones. In fact, with regards to gender, findings reveal that female speakers produce higher values for spectral mean, variance and kurtosis, as well as lower values for skewness in comparison to male speakers. For the purposes of this dissertation, it is important to mention that their study demonstrates that spectral means (or center of gravity) are highly correlated with place of articulation: the more backed the constriction of air flow is in the vocal tract, the higher the spectral mean.

Erker’s (2010) analysis also present the necessity of using continuous measurements to better understand linguistic variation. Since categorical descriptions (generative model or the Labovian variationist paradigm) have obscured decisive facts about phonological variation and the factors that favor it, Erker justifies the use of continuous measurements (frication duration and center of gravity) to analyze coda /s/ in the speech of Dominican speakers in NYC. The author claimed that an instrumental analysis shows a reliable and objective approach “of uncovering, capturing, and quantitatively expressing differences in acoustic properties that are otherwise obscured at the outset of segmental analysis” (2010, p. 13). That is to say, purely segmental analysis obscures the actual systematic acoustic variation of speech. The author found significant differences in center of gravity as well as in frication duration. Moreover, those acoustic differences were correlated with some of the independent variables considered in the study. As a result, findings showed that subsegmental analysis of /s/ weakening in Dominican Spanish is more detailed and adequate than segmental analysis. It offers a
multidimensional acoustic alternative for a more accurate explanatory mechanism and richer descriptive adequacy in the research of Spanish. Therefore, an acoustic phenomenon could be better understood along a continuum than in discrete units typical of segmental analysis.

Flores (2016) analyzed velar palatalization in Chilean before front vowels and the glide /ie/ following a variationist and speech accommodation framework. This author used center of gravity to analyze the linguistic and sociolinguistic factors affecting the fronting of /x/ productions. In line with Jongman et al., (2000) and Silbert and de Jong (2008), higher center of gravity values entail a place of articulation further back along the vocal tract. Therefore, the palatalized realization has a more anterior place of articulation than the velar fricative variant. Findings showed that among the linguistic factor groups, preceding phonetic context, syllable stress and word position favor palatalization. In particular, the palatalized variant [ç] is favored when the /xe/ variable is produced after a front vowel ([i] or [e]) or an alveolar consonant ([s] or [n]), when it is produced in the tonic and post-tonic syllables and in word-initial position. Therefore, this study reveals that center of gravity is a useful measurement to acoustically analyze the palatalization of fricative /x/ in Chilean Spanish.

Just like Erker (2010), who realizes the difficulty to assign /s/ tokens to one of three segmental phonetic categories: [-s], [-h] or deletion, the same arduous task results in trying to ascertain the different realizations of the backed /r/. Following previous studies, the use of continuous acoustic variables to better understand the point of articulation of the backed /r/ as well as the possible neutralization between /r/ and /h/ in production is necessary.
CHAPTER 3

SOCIOLINGUISTIC CONTEXT: PUERTO RICAN DIASPORA AND
DESCRIPTION OF THE COMMUNITIES UNDER STUDY

3.1 Introduction

The United States has the second largest Spanish-speaking population in the world, reaching more than 46 million speakers (Instituto Cervantes, 2017). The vitality of U.S. Spanish stems from the large number of immigration rates from other Spanish-speaking countries. The majority of the Latino population comes from Mexico (65% of all U.S. Latinos), followed by Puerto Ricans (9%), Cubans (4%), and Dominicans (2.8%) (Pew Hispanic Center, 2012; Potowski, 2015). This demographic circumstance has recently attracted a considerable deal of interest in the field of sociolinguistics, where numerous researchers and scholars have analyzed the Spanish-English contact situation of the U.S., in which Spanish has the status of a minority language (Potowski, 2015; Ramos-Pellicia, 2004; Torres, 1997). This context has significant consequences for the transmission of Spanish among different generations, as well as for the identity factors or attitudes of its speakers (Potowski, 2015). Language certainly influences the social life of members of Latin American communities in the diaspora, and they negotiate their identities through their interactions with the culture of the receiving country as well as with their own culture (Potowski, 2015).

3.2 The Puerto Rican diaspora

3.2.1 Puerto Ricans on the Mainland United States
Puerto Ricans make up the second largest Hispanic population on the mainland United States, reaching approximately 5.4 million inhabitants (U.S. Census, American Community Survey 1-year population estimates, 2016). Puerto Rican migration is of particular interest due to the distinctive geopolitical and sociocultural relationship that Puerto Rico retains with the United States. Since 1917, Puerto Ricans have been U.S. citizens by birth. However, they maintain significant differences in cultural heritage; “their culture and language are as foreign to general American culture as the cultures and languages of many other ethnic groups that migrate from Europe, Africa, the Middle East and Asia” (Ramos-Pellicia, 2004, p. 1).

Although Puerto Ricans have been coming to the Mainland United States since mid-1800s, it was not until the end of World War II when significant numbers of Puerto Ricans migrated looking for a better quality of life and employment opportunities. By then, a demand for agricultural workers had developed in the U.S., especially in the northeast (Carlos Vega Collection). Since islanders worked for low wages, they were convenient laborers in agriculture, from tobacco picking in Connecticut to farming in New Jersey (Rivera, 2001). The causes of this migration have been evaluated from different perspectives besides the lack of job opportunities on the island, such as the initial overpopulation in Puerto Rico, U.S. citizenship, Puerto Rican participation in the armed forces, economic transformation in Puerto Rico during Operation Bootstrap\(^6\) or support and help from relatives and friends already settled on the mainland. Nevertheless, Puerto Rican migration to Mainland U.S. has continued throughout the decades as a

\(^6\) Operation Bootstrap refers to a number of ambitious economical projects that had the goal of transforming the economy of Puerto Rico, developing it into an industrial one.
result of the social, political and economic relations between Puerto Rico and the United States. It is important to note that recent Puerto Rican migration is different from the first waves. In fact, since 1990, most Puerto Rican migrants have been technical workers and college-educated professionals. They are, therefore, more educated and more skilled than the migrants from previous decades (Rivera, 2001). This wave moved to the Mainland U.S. in order to find better opportunities and incomes (Rivera-Batiz and Santiago, 1996, in Rivera, 2001). As a result, the population of Puerto Ricans on the U.S. mainland more than doubled since 1980, rising up to 4.9 million Puerto Ricans in 2012 from 2 million. By 2016, it reached 5.4 million people (U.S. Census, American Community Survey 1-year population estimates, 2016)\(^7\).

It is also important to consider that returning to the island is also a common occurrence. As Garcia Passalaqua claimed: “Thousands of Puerto Ricans live literally in the air, coming and going, between the metropolis and the colony, between one society and another. We are people in transience” (Garcia Passalaqua, 1994, p. 103, in Rivera 2001, p. 21). Migratory waves between the island of Puerto Rico and Mainland U.S. are described, hence, as circular migration or back and forth movement (Center for Puerto Rican Studies, 2016), which is defined as the process in which people who have been living in the United States return to Puerto Rico sometime within a given period of time (Rivera, 2001). Moreover, this *va y ven* (back and forth) migratory movement “often propelled by economic pressures, (…) reinforces many links to the island, although it

---

\(^7\) Census data can be hard to interpret since the numbers reported of Puerto Ricans does not necessarily entail the same number of Puerto Rican Spanish speakers in the mainland.
also reflects the repeated ruptures and renewals of ties, dismantlings and reconstructions of familiar and communal networks in old and new settings” (Rodríguez et al., 1984).

To date, this pattern of migration continues. What is more, it has been accelerated since 2017 due to a catastrophe that devastated Puerto Rico: Hurricane Maria. In September 2017, a Category 4 storm hit Puerto Rico. It is considered the worst storm on the island since San Felipe Segundo in 1928 (Center for Puerto Rican studies, 2018). The storm caused more than a thousand deaths, billions of dollars in damages, and left around 3.4 million Puerto Ricans without potable water, power, and telecommunications (Center for Puerto Rican studies, 2018). Recovery is estimated to take years. In March 2018, six months after Maria devastated the island, the early stage of emergency disaster relief was completed and Puerto Rico entered into a new phase of recovering and rebuilding, where 15% of Puerto Ricans were still without electricity and 12% of residents in the western area lacked potable water (Center for Puerto Rican studies, 2018). Initial estimates for the economic impact of the storm suggest a 20% recession in economic activity.

Another consequence of Hurricane Maria is the massive exodus to Mainland U.S., which is redefining the diasporic community. According to the Center for Puerto Rican studies (2018), Hurricane Maria has accelerated Puerto Rican migration to the United States to a point where depopulation is highly affecting the island in multiple aspects. During the Millennial migration period (from 2006 to 2016), a decade of economic stagnation, Puerto Rico lost 525,769 residents (10% of the island’s population). In fact, while at the beginning of the crisis in 2006 there were the same amount of Puerto Rican residents in the United States than on the island, by 2016 the balance had substantially shifted, leaving 3.4 million residents on the island and 5.5 million Puerto Ricans living on
the mainland U.S. (Center for Puerto Rican studies, 2018). However, due to the direct consequences of Hurricane Maria, from 2017 to 2019, the island could lose more than 470,000 people, which would imply up to 14% of the island’s population (see Figure 10). Consequently, although Puerto Rican migration to the United States has been described as “circular” migration, “the post-Hurricane Maria exodus has changed the prevailing narrative of Puerto Rican migration” (Center for Puerto Rican studies, 2018, p. 4).


Based on school enrollment data, the Center for Puerto Rican Studies (2018) shows that more than 135,000 Puerto Ricans from the island have been relocated in different states after the hurricane, following existing population patterns and nodes. In fact, Florida was the preferred destination for Puerto Rican migrants in the post-economic crisis exodus or “Millennial Migration” (the decades prior to Maria), and continues to
present the largest number of Puerto Ricans on the Mainland post-Maria. Between December 1 and February 22, 2017, “11,553 Puerto Rican students enrolled in Florida’s school districts, 2,874 in Pennsylvania, 2,556 in Massachusetts, 2,218 in New York, 1,827 in Connecticut, 886 in New Jersey, and 607 in Illinois” (Center for Puerto Rican studies, 2018, p. 4)\(^8\). Therefore, Florida presents the largest share of relocations, reaching up to 56,477 people (42% of the total flow), followed by the Northeastern States included in the study. In particular, Massachusetts is the second preferred destination of Puerto Rican migrants, with 15,208 relocations (11% of the total flow), followed by Connecticut (10%, n=13,292), New York, 11,217 (8%, n=11,217), Pennsylvania (7%, n=9,963), and New Jersey (4%, n=5,027).

FEMA (Federal Emergency Management Agency) records indicate that 40,013 household members and 19,271 households relocated to Mainland U.S. as a consequence of Hurricane Maria. According to this agency, traditional settlement locations for Puerto Ricans, such as Massachusetts, New York, Pennsylvania, Connecticut, California, Ohio and Illinois are presenting population growth as well as more dispersed settlement patterns. As an example, New York State presents higher enrollment rates in the upstate area than in New York City (Center for Puerto Rican studies, 2018). Following FEMA data, Florida received a higher number of Puerto Rican evacuees (18,013, 45% of the total FEMA Evacuees), followed by Massachusetts (8%, n=3,399). Figure 11 shows Puerto Rican Stateside Relocation Post-Maria (as of February 2018), comparing data

---

\(^8\) The Center for Puerto Rican studies present those numbers based on school enrollment data from some states’ Department of Education (Florida, New York, Connecticut, Massachusetts, Pennsylvania, and New Jersey school districts). These states represent 80% of the stateside Puerto Rican community (Center for Puerto Rican studies, 2018)
from FEMA and survey of selected Department of Education (Center for Puerto Rican studies, 2018). In general terms, data from both sources is consistent since most of the States present a similar distribution of settlement patterns.


This brief introduction about Hurricane Maria is included since it also affected the trajectory of this dissertation. As it is mentioned in the Methodology section, only data pre-Hurricane Maria was considered. No Puerto-Rican post-Hurricane Maria migrants in Massachusetts were considered for this dissertation.

Considering the Puerto Rican demographic situation described above, from a linguistic standpoint, since migratory waves from Puerto Rico to the Mainland U.S. started in the mid-19th century, PRS (or “el español de aquí y de allá” –the Spanish of here and there-, Valentín-Márquez, 2007, p. XII) has attracted a considerable deal of
interest to scholars, with two principal positions: those who make a distinction between the Spanish spoken on the island of Puerto Rico and the Spanish spoken by the Puerto Rican community on the U.S. mainland; and those who consider PRS to be a single variety whose speakers double the island’s population, including the Puerto Ricans that live on the U.S. mainland (Valentín-Márquez, 2007, p. XII). Regardless of position, the back-and-forth migratory movement of Puerto Ricans has led to changes in the linguistic practices of Puerto Rican speakers in both areas, as has been shown for the case of Lorain, OH (Ramos-Pellicia, 2004, 2007) Grand Rapids, MI (Valentín-Márquez, 2007), Chicago, IL (Potowski, 2015; Torres & Potowski, 2008) and New York, NY (Lamboy, 2004; Raña Risso, 2013; Torres, 1997). Similar back-and-forth migratory movements are found in the Puerto Rican community in Western Massachusetts. This research considers what the linguistic consequences of these back and forth movements might be. As stated above, with the exception of the impressionistic work of Shouse de Vivas (1978), the variety of Spanish spoken in the urban areas such as Holyoke or Springfield in Western Massachusetts remains unexplored, in spite of the rich history of Puerto Ricans in this geographical area, where often cultural heritage is often well-maintained (Rivera, 2001).

3.2.2 Puerto Ricans in Massachusetts

The Latino population in the state of Massachusetts does not reflect the numbers presented above with regards to the United States. As stated before, Mexicans are the largest Latino group in the Mainland U.S. However, in Massachusetts, Puerto Rican residents are the dominant group, reaching 45% of its overall Latino population, followed
by Mexicans (7%) and Cubans (2%) (U.S. Census, 2012-2016 American Community Survey 5-Year Estimates).

Table 3: Hispanic or Latino population in Massachusetts. Source: U.S. Census, 2012-2016 American Community Survey 5-Year Estimates.

<table>
<thead>
<tr>
<th>Hispanic or Latino</th>
<th>Puerto Rican population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total population</td>
<td>10.9% (n=731,739)</td>
</tr>
<tr>
<td>Mexican</td>
<td>0.7% (n=45,175)</td>
</tr>
<tr>
<td>Puerto Rican</td>
<td>4.5% (n=304,770)</td>
</tr>
<tr>
<td>Cuban</td>
<td>0.2% (n=12,825)</td>
</tr>
<tr>
<td>Other Hispanic or Latino</td>
<td>5.5% (n=368,969)</td>
</tr>
</tbody>
</table>

Massachusetts has, in fact, the fifth highest Puerto Rican population in Mainland United States (see Table 4), with a higher Puerto Rican representation than in the United States as a whole (U.S. Census, 2012-2016 American Community Survey 5-year data).

Table 4: States with the highest Puerto Rican population. Source: U.S. Census, 2012-2016 American Community Survey 5-Year Estimates.

<table>
<thead>
<tr>
<th>State</th>
<th>Puerto Rican population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connecticut</td>
<td>7.9% (n=283,278)</td>
</tr>
<tr>
<td>New York</td>
<td>5.6% (n=1,104,443)</td>
</tr>
<tr>
<td>New Jersey</td>
<td>5.3% (n=470,954)</td>
</tr>
<tr>
<td>Florida</td>
<td>5.1% (n=1,014,340)</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>4.5% (n=304,770)</td>
</tr>
<tr>
<td>UNITED STATES</td>
<td>1.7% (n=5,275,008)</td>
</tr>
</tbody>
</table>

There has been a historical presence of Puerto Ricans in urban centers in Massachusetts, with migration occurring since 1950. Puerto Ricans moved directly from Puerto Rico to Western Massachusetts to work in tobacco fields. They also migrated north from urban areas in the United States such as nearby Hartford, CT but also New York City, NY, to find employment in blue-collar industries or in seasonal agriculture.
(Our plural history, 2008). As a consequence, the Puerto Rican population in Massachusetts has increased throughout the decades, reaching (as mentioned above) 304,770 habitants in 2016 (see Table 5) (US Census, 2010; U.S. Census, 2012-2016 American Community Survey 5-Year Estimates).

Table 5: Puerto Rican Population in Massachusetts from 1970 to 2016 (US Census, 2010; Our plural story, 2008; Pew Hispanic Center, 2012; U.S. Census, 2012-2016 American Community Survey 5-Year Estimate)

<table>
<thead>
<tr>
<th>Year</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>23,332</td>
</tr>
<tr>
<td>1980</td>
<td>76,450</td>
</tr>
<tr>
<td>1990</td>
<td>151,193</td>
</tr>
<tr>
<td>2000</td>
<td>199,207</td>
</tr>
<tr>
<td>2010</td>
<td>266,125</td>
</tr>
<tr>
<td>2016</td>
<td>304,770</td>
</tr>
</tbody>
</table>

From 2006 to 2016, 728,486 Puerto Ricans left the island of Puerto Rico to settle in Mainland U.S. Of these Puerto Rican migrants, 6% (n=44,377) moved to Massachusetts, where 36% settled in Hampden County, 17% headed to Essex County, 15% arrived in Worcester County and, finally 12% moved to Suffolk County (See Figure 12) (Boston Planning and Development Agency Research Division, 2017).
Thus, Hampden County, located in the Pioneer Valley in Western Massachusetts, has the largest Puerto Rican population of the state, reaching 90,000 Puerto Rican residents (20% of the overall county’s population) (see Table 6). Holyoke, the city under study in this dissertation, is located in said county (Boston Planning and Development Agency Research Division, 2017).


<table>
<thead>
<tr>
<th>County/City</th>
<th>Puerto Rican Population</th>
<th>% of Puerto Rican Population</th>
<th>% of Latino Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hampden County</td>
<td>92,252</td>
<td>19.7%</td>
<td>23.3%</td>
</tr>
<tr>
<td>Springfield</td>
<td>55,815</td>
<td>36.2%</td>
<td>43.2%</td>
</tr>
<tr>
<td>Holyoke</td>
<td>18,478</td>
<td>45.9%</td>
<td>50%</td>
</tr>
<tr>
<td>Chicopee</td>
<td>9,185</td>
<td>16.4%</td>
<td>18.7%</td>
</tr>
</tbody>
</table>

It is noteworthy to mention that the numbers presented above have increased drastically since October 2017 due to the massive exodus resulting from Hurricane
Maria. As stated in the previous section, following FEMA data Massachusetts received the second higher number of Puerto Rican evacuees (8% of the total FEMA Evacuees, n=3,399).

3.2.2.1 Demographic and social aspects of the Puerto Rican population in Massachusetts

With regard to age, the Puerto Rican community in Massachusetts is significantly younger than the overall Massachusetts residents. Furthermore, with a median age of 26, they are also younger compared to Puerto Ricans in Puerto Rico, who have a median age of 40. While 35% of Puerto Ricans in Massachusetts are children, only 21% of the island population are children. Consequently, 18.9% of the overall population in Puerto Rico is elderly, whereas the elderly constitute 5.7% in Massachusetts (Boston Planning and Development Agency Research Division, 2017).

In terms of educational attainment, on the one hand, only 12.1% of Puerto Rican residents in Massachusetts have a bachelor’s degree in comparison to Puerto Ricans in the rest of the United States (18.4%) and Puerto Rico (25.1%). Moreover, about 31.1% of Puerto Ricans in Massachusetts have less than a high school education, while this demographic constitutes 21.2% of Puerto Ricans in the U.S. and 24.2% in Puerto Rico (Boston Planning and Development Agency Research Division, 2017).

With respect to languages spoken at home, 67% of Puerto Rican residents in Massachusetts speak a language other than English, which is, most of the time, Spanish. Furthermore, 21.7% of Puerto Ricans in Massachusetts say that they speak English less than “very well”, which is higher than the national percentage (17.2%) of Puerto Ricans
in the United States (Boston Planning and Development Agency Research Division, 2017).

With regards to income, Puerto Rican households in MA have an average income of $27,000, lower than in other parts of the United States ($42,856) but higher than in Puerto Rico ($19,977). It explains that as to housing, only 19.4% of Puerto Rican residents in Massachusetts live in owner-occupied housing, in comparison to the 68.6% in Puerto Rico and 36.9% in the rest of the U.S. (Boston Planning and Development Agency Research Division, 2017).

Finally, as to poverty rate, considering that poverty status is regulated by income, family size and the age of workers, 31.3% of Puerto Rican families live in poverty in Massachusetts. This percentage is higher in Puerto Rico (39%) but lower elsewhere in the U.S. (20%) among Puerto Ricans (Boston Planning and Development Agency Research Division, 2017). Additionally, more than half of all Puerto Rican children live in poverty in Massachusetts. This number is interesting considering that it implies more than double the child poverty rate in the overall United States and three times the overall child poverty rate in Massachusetts (UMass Donahue Institute, 2017). Therefore, Puerto Ricans in Massachusetts face higher rates of poverty and lower levels of educational attainment.

### 3.3 The communities under study

In addition to Springfield, MA, one of the principal locations where Puerto Rican settlements were established throughout the decades in Western Massachusetts is the city of Holyoke. As mentioned before, with the exception of the impressionistic work of
Shouse de Vivas (1978) in Springfield, the Spanish varieties spoken in urban centers of Massachusetts remain unexplored.

In the following section, the communities under study are described socially and linguistically. Holyoke, MA is presented first, followed by the two locations on the island of Puerto Rico: San Juan and Salinas.

3.3.1 Holyoke, Massachusetts

3.3.1.1 Overall information

Holyoke is located in the heart of the Pioneer Valley, along the Interstate-91 of Hampden County, Massachusetts. Sitting on the western bank of the Connecticut River, it borders Chicopee, MA; Easthampton Town, MA; Westfield, MA; and West Springfield Town, MA (see Figure 13). It is situated only eight miles north of Springfield, one of the principle metropolitan areas of Massachusetts. Holyoke encompasses an area of approximately 22.8 square miles and has an estimated population of 40,280 habitants (U.S. Census, 2016 Population Estimate). This medium-size urban city has the largest population per capita of Puerto Ricans outside the island of Puerto Rico, while Springfield occupies the 7th place in that ranking (U.S. Census, 2010).

![Figure 13: Holyoke’s location in Hampden County. (Source: Wikipedia Commons, Justin H. Petrosek, in https://commons.wikimedia.org/wiki/File:Holyoke_ma_highlight.png)](https://commons.wikimedia.org/wiki/File:Holyoke_ma_highlight.png)
Holyoke is considered an “immigrant, industrial city” which was founded in the 19th century (Holyoke History Collection). Due to its huge potential for waterpower, in 1847 merchant investors from Boston planned to build a dam and canal system along the Connecticut River, creating a textile manufacturing center on this site that was the most productive in New England. After 1850, Holyoke became famous for high-quality paper, cotton, silk, nylon, thread, and industrial machinery. Therefore, Holyoke’s population increased from 4,600 in 1855 to 35,600 in 1890. The labor force in those industries was largely made up of immigrants who came from different countries, such as Ireland, Canada, England, Scotland, Germany and Poland. As a result, in 1889, Holyoke had a 48% foreign born population. In the 1900s Italians, Jews and Russians also migrated to the town and by 1960, Puerto Ricans constituted the most recent influx of new residents. Today, Puerto Ricans compose 91.8% of the Latino population and almost 50% of the Holyoke’s overall population.

The most recent Census data shows that 46.4% of Holyoke’s population speak a language other than English at home (U.S. Census, 2012-2016 American Community Survey 5-Year Estimates), a higher percentage than the national average of 21%. Spanish is, as expected due to the higher rate of Puerto Rican population, the most common foreign language. In fact, in 2015, 39.4% of the overall population of Holyoke, MA were native Spanish speakers. The next two most common foreign languages are Polish (1.12 %) and French (0.63%) (Data USA).
In terms of education, 79.2% of the Holyoke population have a high school degree or higher and 25.6% hold a bachelor’s degree or higher (U.S. Census, 2012-2016 American Community Survey 5-Year Estimates).

With regards to income and poverty, the median household income (U.S. Census, 2012-2016 American Community Survey 5-Year Estimates) is 38,829, which is less than the median income in the United States ($55,322) or in Massachusetts ($70,954). The number of people per household in the period between 2012 and 2016 is 2.6.

Figure 14: Population density by ethnicity in Holyoke, MA. Latino population in red, White population in blue. (Source: Prison Policy Initiative https://www.prisonpolicy.org/graphs/holyoke_latino_white.html).

Figure 15: Income by Location in Holyoke, MA (Source: Data USA https://datausa.io/profile/geo/holyoke-ma/, taking data from the census tract (2016) U.S. Census, 2012-2016 American Community Survey 5-Year Estimates)
Figures 14 and 15 are of particular interest since they show how the areas where the Latino population is most concentrated those areas correspond to the Flats, Downtown, and South Holyoke (as seen in Figure 16 below), which are also the zones with the lowest median household income in Holyoke. The Latino population and, ultimately, Puerto Ricans are the ethnic group with the lowest income in the city.

3.3.1.2 The Puerto Rican community in Holyoke

According to the 2012-2016 American Community Survey 5-Year Estimates, Holyoke is inhabited by 40,280 people of whom 49.97% (20,130) are Hispanic or Latino, 43.16% (17,387) are White, Not Hispanic or Latino and 4.41% (1,780) are Black or African American non-Hispanic. These numbers indicate that the Latino or Hispanic population is the largest ethnic demographic in Holyoke. As stated previously, among the Hispanic and Latino group, Puerto Ricans are the dominant group in Holyoke, reaching up to 50% of the total population, followed by Mexicans (0.5%) and Cubans (0.1%) (See Table 7). In this regard, it is important to reiterate that Holyoke presents the largest population per capita of Puerto Ricans outside the island of Puerto Rico (US Census, 2010).

Table 7: Hispanic or Latino and Race in Holyoke, MA. Source: U.S. Census, 2012-2016 American Community Survey 5-Year Estimates.

<table>
<thead>
<tr>
<th>Hispanic or Latino and race (Holyoke, MA)</th>
<th>Estimate</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total population</td>
<td>40,280</td>
<td>40,280</td>
</tr>
<tr>
<td>Hispanic or Latino (of any race)</td>
<td>20,130</td>
<td>50%</td>
</tr>
<tr>
<td>Mexican</td>
<td>221</td>
<td>0.5%</td>
</tr>
<tr>
<td>Puerto Rican</td>
<td>18,478</td>
<td>45.9%</td>
</tr>
<tr>
<td>Cuban</td>
<td>58</td>
<td>0.1%</td>
</tr>
<tr>
<td>Other Hispanic or Latino</td>
<td>1,373</td>
<td>3.4%</td>
</tr>
</tbody>
</table>
As mentioned above, Puerto Rican migration to Holyoke began around 1950, although the largest influx arrived in the 1960s. Besides political and economic pressures on the island of Puerto Rico, there were other factors that made Western Massachusetts, and Holyoke in particular, an optimal destination for Puerto Ricans. The Northeast was in need of seasonal farm laborers (in tobacco fields and for other seasonal agriculture). Therefore, many workers moved from other urban areas in the U.S. or from the island to Springfield, MA, largest city in Western New England, where cheap housing and low-skill jobs were available. Springfield, indeed, is considered the first Puerto Rican settlement in the state of Massachusetts (Williams, 1992, in Rivera, 2001). Soon, surrounding cities and towns such as Westfield, Chicopee, and especially Holyoke became chief destinations. In fact, in the mid-1960s, the character of Holyoke changed significantly for two main reasons. On the one hand, Springfield, MA started an urban renewal project in the North End to make way for Interstate 91 (Rivera, 2001), which displaced many Puerto Rican families. As a consequence, these families started to search for affordable housing in the area. Simultaneously, Holyoke was experiencing important changes in its industry. Most of the paper mills and other industries which flourished Holyoke in the previous century were closing and moving South in search of a cheaper unskilled labor force. The next generations of the Irish and French-Canadian immigrants no longer wanted those jobs and young, educated Anglos found opportunities outside of Holyoke. Those who stayed, however, left their original ethnic neighborhoods and began to acquire houses in other areas of the city, such as the Highlands (characterized by Victorians and Colonial houses) or Elmwood (see Figure 16). As such, there was soon a
vacancy in the elderly housing complexes built in the 19th and 20th centuries by mill owners and companies to house Irish or French-Canadians factory workers: The Flats, South Holyoke, Churchill and Prospects Heights.

Figure 16: Holyoke’s neighborhoods. (Source: Holyoke Collection, box 8, folder 11, p. 64, in Wistariahurst Museum, Holyoke, MA)

Therefore, as Puerto Ricans were being uprooted, not only from the island of Puerto Rico but also from the city of Springfield, Holyoke provided the perfect characteristics to receive new immigrant population: vacant buildings and a need for minimally-skilled workers who would work for low wages in factories and seasonal agriculture (Carlos Vega Collection). Hence, the decline of the industry in Holyoke helped receive the Puerto Rican population, who joined the dying manufacturing sectors of Holyoke, getting jobs characterized by low wages, instability, and poor working conditions, naturally resulting in high poverty rates. It is important to emphasize that they settled and have lived since then in specific neighborhoods of Holyoke (the southern part of the city, in the predominantly French-Canadian neighborhood known as South Holyoke and the previously-Irish neighborhoods called The Flats). White flight (the
process by which white Americans leave an area because people of color are moving in) and conflict have been key elements of these changes, demonstrating the political and social tensions that accompany the incorporation of immigrant groups in the host community (Borges-Mendez & Uriarte, 2003, in Holyoke Collection). Because of their different background, lack of training and poverty, Puerto Rican were often met with mistrust, fear, discrimination and racism (Borges-Mendez & Uriarte, 2003, in Holyoke Collection). Harper (1973, p. 220), refers to this problem as follows: “The new group is meeting with difficulties perhaps more complex than those of the other earlier migrants. Not the least of the problems encountered by this group is a lack of understanding of its customs and problems by the people among whom it is seeking to live”. Ultimately, in the first Puerto Rican waves, as Borges-Mendez and Uriarte (2003) state, the Latino integration in Holyoke did not imply assimilation through the generations, but rather social exclusion and residential segregation. A newspaper clipping from 1982 states:

“Latin Americans have been settled in Holyoke for over 30 years. At the same time that the number of Latins has increased, so has the mistrust and misunderstanding between the Latin and Anglo communities. The foundation for the misunderstandings is not just cultural, even though this does account for some of the irritation felt by both groups. A large part of the clash comes from the lack of understanding on the part of the Anglos as to who their new Latin neighbors are. The racism and hostility toward the Latin community can be felt and heard throughout our city. These attitudes are not restricted to certain areas or people. They can be heard on the streets and shops, in restaurants and bars, at the Mall or
in a housing project, in the city work yards as well as executive board rooms.
Racist attitudes are held by social workers, teachers, bankers, executives, city
officials, and the whole gamut of society” (Holyoke Collection).

Nevertheless, despite this documented hostility, Puerto Ricans continued to move to
Holyoke throughout the decades. In fact, by 1980, those already settled in the city began
to assist their families and friends in making a transition from Puerto Rico or other areas
in the mainland U.S. to a better life in Holyoke (Borges-Mendez & Uriarte, 2003, in
Holyoke Collection). At that point, the Holyoke Hispanic population was 13% of the
city’s total population. In the following years, another wave of Puerto Ricans moved
from the island with middle-class skills and education, trying to find, as the previous
waves, a better standard of living. By 1990, the Puerto Rican population had increased to
31.1%, by 2000, to 41.1% and, as previously mentioned, by 2016 the Puerto Rican
population makes up almost half of the overall population of the city. It is important to
mention that those numbers are currently increasing due to the high rates of Puerto
Ricans that moved to Holyoke after Hurricane Maria.

Furthermore, although the Puerto Rican population makes up roughly 50% of the
city’s residents, it comprises close to 80% of public school enrollment. Nonetheless, only
27% of Holyoke Public Schools staff are people of color and only 10% of the teachers are
Hispanic. Due to such high presence of Puerto Rican students in schools, some of the
recommendations to the commissioner and receiver from the Holyoke Public Schools are
hiring more teachers who speak Spanish or incorporating Puerto Rican Cultural
Education to the curriculum (Massachusetts Department of Elementary and Secondary Education, 2015).

One of the Puerto Rican cities from which Holyoke has received a high rate of immigrants throughout the decades is Salinas (Harper, 1973; Rivera, 2001).

In his book *The story of Holyoke*, Harper states that “those who are Spanish speaking and have come from Puerto Rico, come from two principal towns, Comerio and Salinas” (Harper, 1973, p. 220). While the methodology driving this statement is left unspecified, it is important to recognize that there is at least a sense in the 1970s that the Puerto Rican population in Holyoke comes from specific regions of the island. Furthermore, in a comparative study between Holyoke and Salinas, Rivera (2001) analyzes the influence of Puerto Rican culture in South Holyoke, investigating the use of public space (specifically, *plazas* and streets) in these two locations. The close ties between Holyoke and Salinas seem evident for the author: “signs such as Salinas mini Market and Salinas Club, including even Salinas car stickers, are indicators of a relation between Salinas and the South Holyoke population” (Rivera, 2001, p. 33). “By entering one of the Puerto Rican mainland neighborhoods, such as South Holyoke, Massachusetts, one feels that some parts of Puerto Rico have been transported. Certainly, this neighborhood keeps facts of Puerto Rican culture and identity as part of their adaptation process” (Rivera, 2001, p. 3). The author’s observations reveal a parallel in the use of public space in Holyoke and Salinas as a result of the adaptation process. Puerto Ricans in Holyoke recreate spaces that did not exist when they moved to Massachusetts. Hence, instead of using the *plaza* as a recreational space and center of the community
relationships, they use the streets for the same purpose. Rivera concludes that this space recreation process serves as a reaffirmation of Puerto Rican identity, thus introducing “resistance spaces against the dominant culture” (2001, p. 95).

Evidence of the link between Salinas and Holyoke has further been exposed through survey data as a part of this current project. For this study, an informal online survey was created in which Puerto Ricans from Holyoke answered questions related to their family origin as well as their perception as to where the majority of the Puerto Rican community come from. The most common responses about where participants’ families come from is Salinas and the general perception is that the Puerto Rican community comes from Salinas and San Juan.

Another piece of evidence to support the close ties between Salinas and Holyoke and the high immigration rates from this Puerto Rican town in Holyoke, MA is the fact that they became sister cities in 2012. Participant testimonies, collected as part of this current project, also reflect the close relationship between these two cities. Numerous subjects, both in Holyoke and in Salinas, mentioned knowing a family member, friend or other acquaintance in the other location.

### 3.2.2 Puerto Rico: Salinas & San Juan

Salinas and San Juan are the Puerto Rican locations relevant to the present study. Following results from an online survey administered in 2017 among the community, those are the main Puerto Rican provenances represented in Holyoke, MA. Including both geographical locations allows comparisons to be made between metropolitan and
non-metropolitan areas, a variable broadly examined in previous studies on PRS. The following sections present a brief demographic description of both settings.

### 3.2.2.1 Salinas

Salinas is a municipality (Sp. *municipio*) located on the southern coast of the island of Puerto Rico, approximately 23 miles east of Ponce, the second largest municipality on the island, and 50 miles southwest of San Juan, the capital of Puerto Rico. It covers an area of 69.7 sq mi (170 km$^2$) (Figure 17). The neighboring municipalities are Aibonito, Cayey, Coamo, Santa Isabel and Guayama. Salinas is a fishing center and one of the main agricultural producers on the southern region of the island. It is divided in 8 *barrios* ‘wards/districts’ (Aguirre, Coquí, Palmas, Quebrada Yeguas, Río Jueyes, and Pueblo) (see Figure 18). Due to a limited economy, job generation is stagnant, forcing many in the area to leave in search of better opportunities. In fact, from 1945 to 1965, Salinas lost 7% of its overall population (Santana, 1996, in Rivera, 2001).

According to the 2016 ACS 5-Year Population Estimate, Salinas has a population of 29,722 residents. The median per capita income in 2016 is $8,535 and the median household income is $16,540, which is less than the annual income in the United States (see Figure 19). Moreover, 53.2% of individuals are below the poverty level (U.S. Census, 2012-2016 American Community Survey 5-Year Estimates). In terms of education, 72.4% of the population has a high school degree or higher and 17.5% has a bachelors’ degree or higher. Considering that the 99.3% is Hispanic or Latino, it is expected that 95.4% of the population speak Spanish.
**Figure 19: Median Household Income in Salinas. Graph taken from Data USA**

(https://datausa.io/profile/geo/salinas-pr/)

Following the Department of Labor and Human Resources, in 2016, Salinas had a high unemployment rate, 22.4%, which has been a consistent problem in recent decades, forcing Puerto Ricans to leave Salinas (Rivera, 2001). As mentioned before, one of the places where the *Salinenses* (people from Salinas) have relocated to was Holyoke, Massachusetts, where “this group of Puerto Ricans had to confront complexities of urban life because they went directly from a small town with a simple lifestyle to a bustling city full of frenetic energy. They faced challenges as they began to adapt to a new language, society, and culture along with economic barriers” (Rivera, 2001, p. 36).

### 3.2.2.2 San Juan

San Juan, the capital and also the most populous municipality Puerto Rico, borders Carolina, Cataño, Guaynabo and Trujillo Alto. It belongs, along with its adjacent municipios and Bayamón, the metropolitan area of Puerto Rico, considered the most competitive and dynamic economic area in Puerto Rico. During the time period from
2012 to 2016, San Juan had a population of 33,148 people and the median household income is $22,553, higher than in Salinas ($16,540). Forty percent of its population live below the poverty line, a higher number than the national average of 14%, but lower than the Salinas average (53.2%) during the same time frame (U.S. Census, 2012-2016 American Community Survey 5-Year Estimates).

![Map of San Juan and Salinas](http://www.proyectosalonhogar.com/link%20p.r/www.linktopr.com/salinas.html)

**Figure 20:** Location of San Juan and Salinas in Puerto Rico. (Source http://www.proyectosalonhogar.com/link%20p.r/www.linktopr.com/salinas.html)

### 3.2.3 Motivation for choosing Holyoke, MA and Salinas, PR

Due to the diverse cultural origins of its population, studies on behavioral patterns consider the United States a “melting pot”, where new groups are constantly added and eventually assimilated, thus, losing their precedent lifestyle and ethnic identity throughout the Americanization process. Although this mechanism can be applied to some continental Puerto Ricans, in general, there is a continued affirmation of the Puerto Rican identity in the United States, preserving the cultural and linguistic Puerto Rican manifestations (Hernández, 1984, p. 13). Holyoke, MA is a perfect example of this Puerto Rican maintenance in the Western Massachusetts diaspora. With the largest per
capita population of Puerto Ricans outside of Puerto Rico, the Spanish spoken here remains almost entirely unexplored.

As Rivera (2001) states, Puerto Ricans shaped a new environment similar to that of Puerto Rico, and especially to that of Salinas. They maintain their self-identity and cultural traditions in Holyoke, MA, which is evidenced through language (Spanish), music (plena, aguinaldo, bomba, salsa), religion (Catholic and Pentecostal) and the Puerto Rican flag. All of them serve as a vehicle for cultural survival, affirming a Puerto Rican identity. Moreover, throughout the years, “bilingual education, Puerto Rican studies, […] the Puerto Rican Day Parade, Latin salsa and the Latin hustle” also demonstrated this affirmation (Rodríguez et al., 1984, p. 7). Ultimately, as Rivera declares:

“The “hanging-out” behavior of the Puerto Rican in South Holyoke was demonstrated to be a cultural behavior with roots in Puerto Rico. Such cultural behavior is transmitted from generation to generation and maintained. This behavior which often is considered in the United States as out of the ordinary is only a way of reinforcing their culture and identity in a place where they are a minority in the society” (2001, p. 95).

Language, as Rivera (2001) affirms, is the primary indication of the Puerto Rican identity, “it is the fundamental tool used for communication and also serves as a rallying point of Puerto Rican identity, in direct opposition to English” (Morris, 1995, in Rivera 2001, p. 60). Spanish, in fact, can be heard through the different Puerto Rican neighborhoods in Holyoke, being the most common language heard in the streets of
South Holyoke. Many signs of bars, churches, institutions, organizations are in Spanish. Some examples are Salinas Mini Market, Los Jibaritos Club, Medina Market, Nuestras Raíces, Enlace de Familia, Nueva Esperanza, Salsarengue, El Rincón Boricua, etc (see Figures 21-24 below).

Figure 21: “El Rincón Boricua” Restaurant in Holyoke, MA (on the left). Figure 22: Public Art Mural “Isla del Encanto” in Holyoke, MA. Source http://larespuestamedia.com/rotten-holyoke/

Figure 23: Nueva Esperanza Inc., a community based program in South Holyoke (on the left). Source http://www.nuevaofholyoke.org/about-us/ Figure 24: “Nuestras Raíces”, urban agriculture organization in South Holyoke (on the right). Source: https://nuestras-raices.org/en/our-mission/

“By entering one of the Puerto Rican mainland neighborhoods, such as South Holyoke, Massachusetts, one feels that some parts of Puerto Rico have been transported”
This statement is unsurprising considering the large Puerto Rican population in Holyoke (45.9%, n=18,478) which makes this city a high-density Puerto Rican area as compared with other Spanish-English bilingual communities, as the ones analyzed by Valentín-Márquez (2007) or Ramos-Pellicia (2004). In his comparative analysis, Valentín-Márquez (2007) shows that the demographic composition of Cabo Rojo and Grand Rapids is quite different. To be more specific, only 1.2% of Grand Rapids’ population is Puerto Rican, in comparison with the 97.4% of Cabo Rojo’s population. It explains that while the majority of Cabo Rojo residents are Puerto Ricans, embedded in Puerto Rican networks, Puerto Rican inhabitants in Grand Rapids are in contact with speakers of other national and ethnic groups, such as Mexicans (the largest Hispanic group, constituting 8.4% of the total population), Dominicans (0.4%) or Cubans (0.2%). Ramos-Pellicia (2004), shows that 21% inhabitants of Lorain are Hispanic, of whom 15.3% are Puerto Ricans, 3.5% are Mexicans, 0.1% is Cuban and 2% belong to other Hispanic groups. The Puerto Rican population in Lorain, thus, constitutes the biggest Hispanic group; however, Lorain’s Puerto Rican population is still not as dense as that of Holyoke, MA.

Although Spanish is visible in the public space, some of the comments provided by speakers in the sociolinguistic interviews conducted for the present dissertation reveal that younger speakers are being corrected by other Puerto Rican Speakers as well as in the educational setting: “Me corren porque no uso bien las erres, o las eles. Tengo problemas con eso” (“they correct me because I do not correctly use my erres (r’s) or eles (l’s)”) (Participant 2, Holyoke). Such corrections favor negative language attitudes towards specific features: “Yo sé que arrastrar las erres está mal” (“I am aware that
dragging the erres (r’s) is incorrect”) (Participant 8, Holyoke), “Arrastrar las erres es un problema” (“Dragging the erres (r’s) is a problem”) (Participant 12, Holyoke) or “No hablo con propiedad, tengo que pronunciarlo bien” (“I do not speak knowledgeably, I have to pronounce it correctly”) (Participant 14, Holyoke). Salient features, like the backed /r/ in PRS are the ones that mark a specific variety. As such, they can be stigmatized when associated with groups of less prestige, enabling teaching corrections. The comments show that such stigma exists in Holyoke, MA, indicating that the city is not an exception in this regard.

To conclude, Holyoke, MA was selected for the present study due to its large Puerto Rican community, almost half of the overall Holyoke population. Moreover, it is crucial to highlight that this location presents the highest Puerto Rican population per capita outside the island. As a consequence, Puerto Rican music, language, religion and studies (among others), are maintained in this diasporic community as a reaffirmation of their Puerto Rican identity (Rivera, 2001). Furthermore, the strong connection between Holyoke, MA and Salinas, PR prompted the decision to examine these two cities specifically as part of this study. Finally, San Juan participants were included since it was relevant for the purpose of this study to compare data from metropolitan and non-metropolitan areas in Puerto Rico. Moreover, following results from the online survey administrated in 2017, San Juan, along with Salinas, are the main Puerto Rican provenances in Holyoke, MA.
CHAPTER 4

METHODOLOGY

4.1 Introduction

This chapter presents the research questions concerned with the PRS /r/ and /h/ variation in production. It also provides a description of the methodology used to test the hypotheses and the analyses performed on the data. Section 4.2 conveys the overall goals, research questions and hypothesis of the present study. Section 4.3 presents an overview of the participants considered in both settings under study, Holyoke Puerto Rican speakers and Puerto Rican Islanders. It also provides a description of recruitment procedures. Section 4.4 details the materials: three experimental production tasks designed to answer the research questions. Section 4.5 focuses on procedures. Section 4.6 describes the acoustic analysis and coding, illustrating the linguistic and sociolinguistic variables for the dissertation project as well as data codification. Finally, section 4.7 provides the analysis undertaken.

4.2 Research questions and goals

Following the variationist framework (Labov, 1963, 1966; Wolfram, 1969) and incorporating the main theoretical claims from the sociolinguistics of globalization (Blommaert, 2003; Blommaert & Dong, 2007; Coupland, 2003), the present study contributes to the body of research on language use in language contact situations, documenting and analyzing the transmission of the trill realization in two different Puerto Rican communities: in the Western Massachusetts diaspora and on the island of Puerto Rico. Particular attention is paid to realizations of /r/ that are backed, that is to say,
realizations produced with the dorsum against the velar or uvular region. Linguistic and sociolinguistic factors such as AGE, SOCIOECONOMIC CLASS or POSITION are examined in order to shed light on the potential backed /r/ variation and to investigate whether or not there are differences among the realizations found in Western Massachusetts in comparison to Puerto Rico. Moreover, changes in apparent-time (Labov, 1972) are examined across two different generations of Puerto Rican speakers in the diaspora.

Phonemic /r/ and /h/ share allophones and a contextual neutralization of these phonemes in perception has been observed in Puerto Rico (Delgado-Díaz & Galarza, 2015). As a secondary goal, once the descriptive analysis is done, if /h/ allophones are also present among the data, it is worth exploring to what extent there is overlap, resulting in a /r/-/h/ merger, at least in production, in Puerto Rico and in the diaspora. Because categorical descriptions have obscured decisive facts about phonological variation and the factors that favor it, Erker (2010) uses continuous measurements (duration and center of gravity) to analyze coda /s/ in the speech of Dominican speakers. Likewise, acoustic continuous measurements will be used in this project: center of gravity, skewness, and kurtosis. These measurements are used with the reasoning that they will allow for the description of the variation present in speech and determine whether the allophones produced for /r/ and /h/ are classified under one of those phonological categories (/r/ vs. /h/), or on the contrary, whether there is a continuum among backed realizations which might be favoring the neutralization in PRS. Following Willis et al.’s (2015) and Delgado-Díaz and Galarza’s (2015) findings, voicing and phonological context should be analyzed in the interest of verifying whether these linguistic variables resolve the overlap between place and manner of articulation in
another PRS variety. It would thus be possible to shed more light onto how these variants affect the PRS phonological system more generally. Moreover, incorporating sociolinguistic factors is crucial in determining the distinction between backed /r/ and /h/.

Finally, comparing data from the island to data in the diaspora provides the empirical evidence necessary to determine whether this possible change in progress on the island occurs in another sociolinguistic context where Spanish is a minority language. Specifically, for this diasporic setting, this dissertation focuses on the city of Holyoke given its large Puerto Rican population. This area has the largest per capita population of Puerto Ricans outside of Puerto Rico and the Spanish spoken here remains almost entirely unexplored.

As stated in Chapter 1, since diasporas are changing communities (Hall, 1990) open to “lateral connections” (Clifford, 1994, p. 306) that are negotiated through language, the concepts of space and mobility receive enormous importance in the recent field of sociolinguistics of globalization. Migration entails the change in the spatial organization of someone’s life, “emigrating and immigrating”, leaving and settling (Blommaert & Dong, 2007). Thus, mobility implies different dimensions that can collapse in specific spaces where speech communities live and interact with one another (Blommaert & Dong, 2007). Consequently, this process allows the mechanisms of assimilation and acculturation as well as language change, hybridizations or even creolization (Gubitosi and De Oliveira, in press). In order to provide a more profound analysis of this diasporic variety, the current project connects the phonetic and phonological findings to the social component (gathered in the language background questionnaire –see Appendices D, E, F and G–). Based on Hulsen, De Bot, & Weltens
(2002) and Stoessel (2002), taking a more detailed look at the social context in the diaspora will allow us to see the extent to which factors such as the contact with the Puerto Rican community might be influential in language maintenance and shift. In this manner, since language is essential in order to construct and negotiate diaspora identities and relationships (Canagarajah & Silberstein, 2012), it is necessary to analyze the effect of participants’ linguistic practices as well as their relationship with a) other members of their diasporic community, b) members of other communities that coexist in the same diasporic setting and c) the island of Puerto Rico. Thus, participants’ responses related to their contact with the Puerto Rican community are crucial in order to address the possible role of bilingualism and language maintenance and change affecting the rhotic variation as well as the potential neutralization between /r/ and /h/.

By presenting and proving a systematic variation in terms of how Puerto Rican speakers in Western Massachusetts make use of the trill depending on different social variables such as gender or generation, and comparing those findings with data recorded from their families’ home towns in Puerto Rico, the present research project will help to address larger questions in the field about the social factors underlying language variation and change in the diaspora. Moreover, paying special attention to the concept of mobility and analyzing the contact that PRS speakers in the diaspora maintain with other members of the Puerto Rican community (in the diaspora and in Puerto Rico), this dissertation contributes to the recent field of sociolinguistics of globalization (Blommaert & Dong, 2007; Márquez Reiter & Martín Rojo, 2015; Bodomo, 2012). Finally, the results regarding the comparison between backed /r/ trill and /h/ fills a gap in the literature on
PRS in the United States, showing interesting implications for the body of research on changes in progress.

In addition to the theoretical goals, this dissertation has a practical objective: to help instructors, learners, and speakers in general to understand that stigmatized realizations are legitimate variation that exists, not only in PRS but in any other variety. Salient features, like the backed /r/ in PRS are the ones that mark a specific variety. As such, they can be stigmatized when associated with groups of less prestige. As detailed in Chapter 3, PRS is visible in the public space, however younger speakers are being corrected in the educational setting: “Me corrigen porque no uso bien las erres, o las eles. Tengo problemas con eso” (“they correct me because I do not correctly use my erres (R’s) or eles (L’s)”) (Participant 2, Holyoke). Therefore, to avoid teaching corrections, the linguistic situation in a U.S. community involving the Latino population needs to be studied and documented, promoting the incorporation of stigmatized realizations in its educational policy to the classroom and informing instructors working in the design of teaching materials.

Considering the literature reviewed in Chapter 2, this study aims to provide an answer to the following research questions:

**Research question 1:** Is there variation in the production of trills among PRS speakers and are there differences among those realizations found in the diaspora and the corresponding communities on the island (Salinas and San Juan)? Specifically, is the backed /r/ ([x]) among the allophones produced?
H1.1. Since previous studies have shown that there is variation in trill production in Puerto Rico as well as in other diaspora settings (Graml, 2009; Hammond, 2000; Navarro Tomás, 1948; Ramos-Pellicia, 2004; Valentín-Márquez, 2007; Vaquero & Quilis, 1989), it was hypothesized that there will also be trill variation in the PRS spoken in Western Massachusetts.

H1.2. Based on the sociolinguistics of globalization, the mobility of communities also implies “the mobility of sociolinguistic and linguistic resource” (Blommaert & Dong, 2007:4). Given that back and forth migration waves between Puerto Rico and Massachusetts have been increasing constantly since 1950 (Center for Puerto Rican Studies, 2016), many Puerto Rican cultural practices have been maintained in the diaspora (Rivera, 2001). Therefore, we predict that the same trill variants found in the aforementioned Western Massachusetts communities will occur in Puerto Rico. Such a result would be in line with Valentín-Márquez’s (2007) findings, who found a similar overall distribution of rhotic variants in Michigan and Puerto Rico.

H1.3. Since Graml (2009), Lamboy (2004) and Valentín-Márquez (2007) reported the presence of backed /r/ not only in Puerto Rico, but also in the diaspora (New York City area and Grand Rapids, Michigan), we predict that this variant will be also produced in Western Massachusetts.

**Research question 2:** If we find similar trill variation, are the predictors (sociolinguistic and linguistic) of the use of the backed /r/ similar for the two settings?
H2. Once again, given the back and forth migration waves previously mentioned (Center for Puerto Rican Studies, 2016), suggesting close ties between Puerto Rico and Holyoke, MA, the hypothesis is that the factors that might predict the use of /r/ variants will be similar in both settings. Specifically, considering previous studies on Puerto Rican Spanish, we predict that linguistic and sociolinguistic predictors affect trill variation. Linguistic factor groups include STRESS (Graml, 2009; Valentín-Márquez, 2007); WORD POSITION (Graml, 2009; López Morales, 1983), PREVIOUS and FOLLOWING SOUNDS (Flores, 2016; Graml, 2009). The sociolinguistic factor groups are ORIGIN (Graml, 2009; López Morales, 1983), GENERATION (Ramos-Pellicia, 2004, 2007); SEX (López Morales, 1983; Matta de Fiol, 1981); SOCIOECONOMIC STATUS (Dillard, 1962; Graml, 2009; Valentín-Márquez, 2007), AGE (Graml, 2009; Medina-Rivera, 1997), TYPE OF TASK (Medina-Rivera, 1997) and CONTACT WITHIN THE PUERTO RICAN COMMUNITY (Valentín-Márquez, 2007).

Linguistic variables:

- For STRESS, following Graml (2009) and Valentín-Márquez (2007) findings, we expect that the backed /r/ is more common in unstressed syllables.

- As for WORD POSITION, in consonance with Graml (2009), Hammond (1991), López Morales (1983) and Terrell (1980), we predict that the backed /r/ is more frequent in intervocalic position.

- With regard to PREVIOUS and FOLLOWING SOUNDS, in line with Flores (2016), we believe that coarticulation effects affect /r/ variation. Specifically, we
expect that previous or following backed vowels (/o,u/) favor the realization
of the backed /r/.

Sociolinguistic variables:

- Concerning ORIGIN, taking into account previous studies on PRS (Graml, 2009; López Morales, 1983) which reveal that the backed /r/ is more frequent in rural areas of the island of Puerto Rico, we expect to find more backed realizations among PRS speakers originally from Salinas (non-metropolitan area) than from San Juan (metropolitan area).

- In relation to GENERATION, Ramos-Pellicia (2004) reveals an increasing rate of non-normative trills in her younger generations in Lorain, OH. Similarly, we would expect that second-generation speakers in Holyoke, MA would produce more backed /r/ realizations than first generation speakers.

- With reference to SEX, a rich body of sociolinguistic research has showed that men are more likely to produce non-standard variants than women since men are more indifferent with regards to the stigma that sounds carry. However, women are more aware of linguistic prestige and stigmatization (Labov, 1994). Therefore, we expect men to produce more backed /r/ than women, in line with Alers-Valentín (1999), Hammond (1991), López Morales (1983), or Matta de Fiol (1981).

- As to SOCIOECONOMIC STATUS, following Dillard (1962), Graml (2009), Matta de Fiol (1981) and Valentín-Márquez (2007), we would expect more backed realizations among speakers in the lower socioeconomic class that in the upper socioeconomic class.
- With regard to AGE, we would expect that the older speakers produce more overall backed realizations, a result that is consistent with the general opinion about the use of the backed /r/ on the Island, in line with Valentín-Márquez (2007) research.

- For TYPE OF TASK, previous studies (Dillard, 1962; Graml, 2009; Guitart, 1981; Medina-Rivera, 1997) revealed that speakers produce vernacular realizations in more informal contexts. Therefore, we expect to find more backed realizations when the task involves spontaneous speech, that is to say, within the Picture Description task than in the Reading task.

- Finally, as to CONTACT WITHIN THE PUERTO RICAN COMMUNITY, given the back and forth migration waves mentioned above (Center for Puerto Rican Studies, 2016), we would expect that Puerto Ricans with closer ties with the PR community would produce more backed /r/ realizations, a salient feature in PRS.

**Research question 3**: Regarding the production of the backed /r/, is it possible to capture its variation on a continuum using continuous acoustic variables (center of gravity, skewness, kurtosis)? Do the potential linguistic and sociolinguistic factors that predict the appearance of the backed /r/ (in both settings under study) have the same predictive power on the acoustic characteristics (kurtosis, skewness and center of gravity values) of this fricative sound?

H3.1. Following Erker (2010), Flores (2016) and Jongman et al. (2000), who used continuous measurements (e.g. duration and center of gravity) to analyze fricative
sounds, we expect that acoustic continuous measurements will be also relevant to
describe the variation present in the current data and determine whether the
allophones produced for the backed /r/ are classified under different categories
(such as the velar [x], uvular [χ], or glottal [h] realizations) or on the contrary,
whether there is a continuum among backed realizations which might be favoring
the neutralization with /h/ in PRS. Analyzing fricative sounds is of particular
interest because not all Spanish varieties present the opportunity to use these
measures with rhotics.

H3.2. Previous studies (Flores, 2016; Jongman et al., 2000) reported that the
spectral properties (center of gravity, skewness, kurtosis) that differentiate fricative
realizations with respect to place of articulation can be affected by sex, voicing
pattern or vowel context. Similarly, it is predicted that the acoustic features that
characterize the backed /r/ are also affected by different linguistic and
sociolinguistic factors. In fact, it is expected that the aforementioned factors,
particularly the linguistic ones, would be the same ones that might predict the use
of the backed /r/. Specifically, in line with Flores (2016), with regard to Previous
and Following Sounds, it is possible that the backed /r/ continuum is affected by
c��ariculation: it is expected that realizations of backed /r/ are more posterior when
the sound appears before or after a back vowel.

**Research question 4:** Is there evidence for /r/ and /h/ neutralization in production among
PRS speakers, not only in Puerto Rico but in the diaspora? Do the same linguistic or
sociolinguistic factors affect the distinction of /r/ and /h/ in the two settings under study?
H4.1. There are claims suggesting the neutralization of the /r/ and /h/ phonemes in PRS (Delgado-Díaz & Galarza, 2015; Dillard, 1962; Lipski, 1994) due to the similarities in the allophonic realizations of those phonemes. Along these lines, Delgado-Díaz and Galarza (2015) reveal in a perception experiment that the contrast is only maintained in some phonological contexts. Following those results, we expect to find some evidence for a /h/-/r/ neutralization in production on the island of Puerto Rico. If so, given the circular movement between Holyoke, MA and Puerto Rico, we predict that this process of language change on the island would also be occurring in Massachusetts.

H4.2. Considering Delgado-Díaz and Galarza’s (2015) and Willis et al.’s (2015) findings, the perceptual contrast between the backed /r/ and /h/ tends to be maintained in intervocalic position and lost in post-pausal position. Correspondingly, we expect that POSITION will have an effect on the overlap between place of articulation in both PRS varieties under study. Moreover, Delgado-Díaz and Galarza (2016) found that the backed /r/ is also influenced by speaker’s sex and age: older male listeners do not neutralize the backed /r/ with /h/, in line with the production patterns and attitudes related with the backed /r/ in Puerto Rico (Graml, 2009). Similarly, we expect that AGE and SEX will have an effect on the possible neutralization in production. Furthermore, considering the effect of speaker’s generation in Ramos-Pellicia’s research (2004) on PRS speakers in Lorain (Ohio), where the preference of the normative trill declines across the three generations, we expect speaker’s GENERATION to have a principal role on this neutralization in Western Massachusetts.
4.3 Participants

This study is comprised of recordings from 45 participants. These speakers represent two main groups: (1) Puerto Ricans in Holyoke and (2) Puerto Ricans in Puerto Rico. All participants were self-identified as members of one of those communities. Since previous studies (Graml, 2009; López Morales, 1983) show differences in rhotic variation depending on speaker’s origin (metropolitan vs. non-metropolitan), this variable was considered in this dissertation for each location under study (Puerto Rico and Holyoke, MA). In line with Ramos-Pellicia (2004) study, which considered five of the rural hometowns of the Puerto Rican community living in Lorain, the main interest of the present study is to compare spontaneous and formal PRS of those who live in Holyoke and that of speakers in Puerto Rico who live in the same municipios from which the Holyoke speakers originated. In addition to speaker’s origin, the sample is also balanced by generation and gender.

The sample for the production study was selected through community networks and research assistants. All speakers taking part in the study do so voluntarily and received incentives of $10 for their participation. Funds were provided by a National Science Foundation Doctoral Dissertation Research Improvement Grant. A detailed description of the recruitment process is included in the next subsections.

Results from five speakers had to be discarded from analysis. Four were discarded due to a history of either hearing or speech impairment. The fifth to be discarded is a second-generation participant who did not demonstrate sufficient proficiency of Spanish during the recording to allow for an adequate analysis.
The data presented in the following sections come from the responses from a sociolinguistic questionnaire distributed among the participants of this study. The questionnaire is described in detail in section 4.4.4.

4.3.1 Holyoke Puerto Rican speakers

A total of 21 participants were recorded in Holyoke, MA. Only participants that have been living in Holyoke for at least 5 years have been included in the study. As stated previously, the members of the Puerto Rican community in Holyoke, MA are divided into two different groups based on their origin: Salinas (a non-metropolitan municipio located in the south of Puerto Rico) or San Juan (the capital), which are the two more common provenances of the Holyoke community in an online survey carried out by the researcher. Moreover, subjects originally from Salinas (n=12) and San Juan (n=9) are divided by generation (1st and 2nd), following Silva-Corvalán’s (1994) criteria in her study on Spanish in the Mexican community in Los Angeles. The first generation encompasses individuals who were born in Puerto Rico and moved to the United States after age 12. This means that an individual who arrived in Holyoke in 1970 when s/he was 11 years old will belongs to the same generation group as a participant who arrived to Holyoke 11 years ago as an adult. The second generation consists of people who were born in the United States and whose parents came from the island of Puerto Rico (either from San Juan or Salinas). Age twelve was the cut-off point considered since, according to Lenneberg (1967), after this age it is more challenging to acquire a language without an accent as a result of neurological adjustments such as loss of brain plasticity or hemispheric specialization. Moreover, it is considered that at age 12 the period of
linguistic maturity in the individual’s native language has been mostly concluded.

Following Silva-Covalán (1994), the third generation consists of people who were born in Mainland United States, whose parents were also born in Mainland United States, but their grandparents were born on the island of Puerto Rico. However, the third generation was not included in the analysis. Since Puerto Rican immigration to Holyoke started in the 1960s, there are not many third-generation speakers over the age of 18 and thereby eligible for this study. The average age of first-generation speakers is 44.72 years old and for the second-generation speakers the average age is 31.88.

All 12 participants originally from Salinas were divided into evenly distributed groups, 3 females and 3 males for each generation. Participants originally from San Juan (n=9) were also divided by generation: 3 women and 3 men in the first generation; 3 women in the second generation.

In the present study, participants fall into one of three life stages, which were divided in three contrasting categories: young adults (18-25 years old), middle-aged adults (26-45 years old) and older adults (older than 45) (see Table 8 below).

Table 8: Holyoke sample based on life stage

<table>
<thead>
<tr>
<th>Life stage</th>
<th>Holyoke, Massachusetts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young adults</td>
<td>4 participants (average age=22.75)</td>
</tr>
<tr>
<td>Middle-aged adults</td>
<td>10 participants (average age=34.6)</td>
</tr>
<tr>
<td>Older adults</td>
<td>7 participants (average age= 48.8)</td>
</tr>
<tr>
<td>Average age</td>
<td><strong>39.47</strong></td>
</tr>
</tbody>
</table>

In terms of linguistic proficiency (Table 9), considering participants’ responses in the sociolinguistic questionnaire, where they had to rate how well they speak, read, write
and understand spoken Spanish as well as English, on a scale from 1 to 6 (see Appendices D, E, F and G), the majority of the participants classify themselves as Spanish native speakers (85.71%, n=18), followed by near-native speakers (10%, n=2 participants).

Table 9: Holyoke sample based on linguistic proficiency

<table>
<thead>
<tr>
<th>Language proficiency</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>(n=0/21)</td>
<td>(n=0/21)</td>
<td></td>
</tr>
<tr>
<td>Basic</td>
<td>0%</td>
<td>5%</td>
</tr>
<tr>
<td>(n=0/21)</td>
<td>(n=1/21)</td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td>4.16%</td>
<td>19.04%</td>
</tr>
<tr>
<td>(n=1/21)</td>
<td>(n=4/21)</td>
<td></td>
</tr>
<tr>
<td>Advanced</td>
<td>0%</td>
<td>9.52%</td>
</tr>
<tr>
<td>(n=0/21)</td>
<td>(n=2/21)</td>
<td></td>
</tr>
<tr>
<td>Near native</td>
<td>10%</td>
<td>23.80%</td>
</tr>
<tr>
<td>(n=2/21)</td>
<td>(n=5/21)</td>
<td></td>
</tr>
<tr>
<td>Native</td>
<td>85.71%</td>
<td>42.85%</td>
</tr>
<tr>
<td>(n=18/21)</td>
<td>(n=9/21)</td>
<td></td>
</tr>
</tbody>
</table>

With regards to participants’ language use (Table 10), most Puerto Ricans in our sample speak both English and Spanish on an everyday basis. The percentage, however is higher for Spanish (90%, n=19) than for English (71.42%, n=15). These numbers evidence the vitality of Spanish in Holyoke, MA, which speakers use every day, both inside and outside the home, and with greater frequency than English.
Table 10: Holyoke sample based on language use

<table>
<thead>
<tr>
<th>Language Use</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>0% (n=0/21)</td>
<td>0% (n=0/21)</td>
</tr>
<tr>
<td>1-2 a month</td>
<td>4.76% (n=1/21)</td>
<td>10% (n=2/21)</td>
</tr>
<tr>
<td>Once a week</td>
<td>4.76% (n=1/21)</td>
<td>14.28% (n=3/21)</td>
</tr>
<tr>
<td>2-3 times a week</td>
<td>0% (n=0/21)</td>
<td>4.76% (n=1/21)</td>
</tr>
<tr>
<td>Every day</td>
<td>90% (n=19/21)</td>
<td>71.42% (n=15/21)</td>
</tr>
</tbody>
</table>

Subjects in Holyoke were primarily recruited through community networks. Online social media, such as Facebook, provided another important recruitment resource to connect to people living in Holyoke. Additionally, more traditional recruitment measures were taken hanging flyers in the city’s restaurants, churches, libraries, schools, and streets.

### 4.3.2 Puerto Rican Islanders

Subjects on the island were recorded in the two locations of interest in this study: in San Juan (n=12) and Salinas (n=12). As previously state, the main interest is to investigate spontaneous and formal PRS by speakers in Holyoke and speakers who live in the Holyoke speakers’ municipios of origin. Each group consisted of 6 men and 6 women. All participants are natives of the cities in which they were recorded. The average age for Puerto Rico participants is 31.62, as can be seen in Table 11.
Table 11: Puerto Rico sample based on life stage

<table>
<thead>
<tr>
<th>Life stage</th>
<th>Puerto Rico</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young adults</td>
<td>12 participants (average age=19.8)</td>
</tr>
<tr>
<td>Middle-aged adults</td>
<td>6 participants (average age=32.5)</td>
</tr>
<tr>
<td>Older adults</td>
<td>6 participants (average age=54.3)</td>
</tr>
</tbody>
</table>

In terms of linguistic proficiency (Table 12), all participants in our sample identify themselves as native Spanish speakers (100%, n=24). More variability is found with regards to English, where only 16.66% (n=4) consider themselves native English speakers.

Table 12: Puerto Rico sample based on linguistic proficiency

<table>
<thead>
<tr>
<th>Language proficiency</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning</td>
<td>0% (n=0/24)</td>
<td>8.3% (n=2/24)</td>
</tr>
<tr>
<td>Basic</td>
<td>0% (n=0/24)</td>
<td>8.3% (n=2/24)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>0% (n=0/24)</td>
<td>29.16% (n=7/24)</td>
</tr>
<tr>
<td>Advanced</td>
<td>0% (n=0/24)</td>
<td>8.3% (n=2/24)</td>
</tr>
<tr>
<td>Near native</td>
<td>0% (n=0/24)</td>
<td>29.16% (n=7/24)</td>
</tr>
<tr>
<td>Native</td>
<td>100.00% (n=24/24)</td>
<td>16.66% (n=4/24)</td>
</tr>
</tbody>
</table>

Subjects in Puerto Rico were also mainly recruited through networks of friends or “friends of friends” (Milroy, 1987) as well as through networks of a research assistant.
who was a student at the University of Puerto Rico, Río Piedras with an interest in sociolinguistics and phonetics. As with the Mainland U.S. portion of the study, online social media, such as Facebook, provided an important recruitment tool and flyers advertising the study were hung in the University of Puerto Rico, Río Piedras and at the Salinas library. In Salinas, participants were also recruited at the plaza, talking directly to the people found in the street.

4.4 Materials

In order to answer the research questions stated above, three experimental production tasks were designed and employed: a picture description task, a map task and a reading task (see the Appendices A, B and C). Finally, participants completed a written language background questionnaire (see Appendices D, E, F and G).

4.4.1 Picture description task

The first task was the narration of the children’s picture book by Mercer Mayer, *Frog Where are you?* (1969). This task is common in the research on trill variation in Spanish (Henriksen & Willis, 2010; Willis, 2007; Willis & Bradley, 2008). The story is about a frog (rana /rá.na/), a dog (perro /pé.ro/), and a little boy, which motivated several productions of trill segments. The phonemic /h/ tokens used for comparison with the backed trills were also produced in this story. Target words that contain phonemic /h/ are *bee* (abeja /a.bé.ha/) or *leaf* (hoja /ó.ha/).

4.4.2 Map task
Participants also completed a modified version of the HCRC Map Task (Anderson et al., 1991) (see Appendix B). In the task, participants collaborated with the experimenter to reproduce a route shown on his/her map. Participants produced target landmark label phrases written on the map that contain the phonemic /t/ segment in intervocalic position (n=3) (i.e., parroquia del pueblo, town’s parish), word-initial position (n=3) (i.e., rosas de la bahía, bay’s roses) and after a heterosyllabic consonant (n=3) (i.e., avión israelí, Israeli’s plane), as well as the phonemic /h/ segment in word-initial position (n=3) (i.e., Juguetes de Pedro, Peter’s toys), in intervocalic position (i.e. pájaros de colores, colorful birds) (n=3) and after a heterosyllabic consonant (i.e., ángeles de la Gloria, Angels of Glory) (n=3). In order to avoid accommodation as much as possible and to encourage naturalistic Puerto Rican Spanish, the researcher limited her interventions to small questions such as ¿Cómo?/¿Qué? ¿Quién?, what?, sorry? The eighteen landmarks from the map task (see Appendix B) were verified with two Puerto Rican Spanish speakers (from San Juan and from Holyoke) to corroborate that they were regular used words in this Spanish variety.

4.4.3 Reading task

The third facet of the experiment is a formal reading task, a common task in phonetic research (Gordon & Maclagan, 1990; Horwarth & Horvath, 2000; Maclagan & Gordon, 1996). This task was slightly modified to present each word with a picture linked to it. This decision was made with the purpose of giving more guidance to participants and avoiding alienation, helping them to associate the word written with the picture before eliciting the item. A total of 36 target words and 12 fillers were produced (see Appendix
The choice to not consider the same number of fillers than target words was made to keep the task short and to avoid participants’ inattention during its completion. The reading task involves the same 6 conditions of the map task: /t/ in intervocalic position (n=6) (i.e., burro, donkey), in word-initial position (n=6) (i.e., rubí, ruby), and after heterosyllabic consonant (n=6) (i.e., honra, honor/honour). Target words also include /h/ in intervocalic position (n=6) (i.e., lujo, luxury), in word-initial position (n=6) (i.e., jabón, soap), and after heterosyllabic consonant (n=6) (i.e., álgebra, algebra).

Two different carrier sentences are used in order to avoid possible coarticulation in word-initial position. Thus, words in intervocalic position and after heterosyllabic consonant were produced in the carrier phrase Diga ____ otra vez (Say ____ again) while words in word-initial position were produced in the carrier sentence ____ es lo que digo (____ is what I say).

This task was controlled for stress and the preceding/subsequent vowels. In intervocalic position, target sounds were produced after /a/ (n=2), /i/ (n=2) and /u/ (n=2). Furthermore, those vowels were controlled for stress (i.e., berra/narrar, bar/narrate; caja/bajar, box/go down). In word-initial position, target sounds were produced before /a/ (n=2), /i/ (n=2) and /u/ (n=2). Those vowels were also controlled for stress (i.e., riñón/risa, kidney/laugh; gimo/girar, I moan/spin around). For intervocalic and word-initial positions, only disyllabic words were considered. Due to logistic limitations, for the heterosyllabic consonant context, target sounds were only controlled for stress (i.e., enredo/alrededor, tangle, around; monja/angina, nun/tonsils).

The task was distributed in a PowerPoint where fillers and target words were presented in random order. As in the map task, the 36 words were checked with two
Puerto Rican Spanish speakers (from San Juan and from Holyoke) to confirm that all tokens are familiar in the Puerto Rican variety.

4.4.4 Sociolinguistic questionnaire

Finally, participants completed a written questionnaire on language background to get all linguistic and extralinguistic information for each participant. The questionnaire use differs slightly depending on participant’s origin (Puerto Rico vs. Holyoke). Subjects were each given the version corresponding to their location. A copy of the questionnaires can be found in Appendices D, E, F and G.

The initial section of the questionnaire collected data related to general background information such as place of birth, length of residence in the cities under study, generation, education or current occupation. In addition, this section inquires as to the number of people living in participant’s household.

The second section of the questionnaire asked for speakers’ language proficiency and domains in which the participants speak English and Spanish. For example, the questions in this section regard the language spoken inside and outside the home with their partners, children, relatives, friends, classmates or workmates. In addition, participants were asked about where they went to school and knowledge of languages other than Spanish or English.

The final section explores participants’ contact with the Puerto Rican community, considering the communities in Massachusetts, mainland U.S. and on Puerto Rico with whom they talk most frequently. This section was slightly modified depending on
participant’s origin. Questions related to language proficiency follow Ramos-Pellicia’s (2004) questionnaire for the Puerto Rican community in Lorain, OH.

4.5 Procedures

Mainland U.S. data were collected between May 2017 and February 2018. Puerto Rican data were collected in June 2017. All 45 participants met with the researcher to complete the experimental task. Moreover, although the researcher is a native Spanish speaker, she is not a native speaker of Puerto Rican Spanish. Due to the Observer’s Paradox (Labov, 1972) showed in previous pilot studies, the researcher was accompanied by a native speaker of Puerto Rican Spanish in both settings under study: Puerto Rico and Massachusetts. The fieldwork assistant in Puerto Rico was a student at the University of Puerto Rico, Río Piedras with an interest in sociolinguistics. In Holyoke, MA, the fieldwork assistant was a member of the Puerto Rican community. In both settings, the assistants helped the researcher to find participants and accompanied her during the experimental tasks. As stated previously, all efforts to avoid the Observer’s Paradox (Labov, 1972) and the linguistic accommodation to the researcher’s variety (Asturian Spanish) were taken. In Holyoke, MA, due to scheduling limitations, there were some cases where the participants met with the researcher alone. In those cases, the researcher gave instructions in English.

In terms of procedure, the same steps were taken for all groups. Participants were first engaged in an informal or casual conversation with the assistant researcher about him or herself to solicit natural conversation prior to performing the linguistic tasks. Afterwards, the research assistants explained the experimental instructions. As mentioned
above, in the few cases in which the assistant was not present, the researcher disclosed
them in English. Following the variationist approach, “The aim of linguistic research in
the community must be to find out how people talk when they are not being
systematically observed; yet we can only obtain these data by systematic observation”
(Labov, 1972, p. 209). Therefore, in addition to the detail explanation of the tasks, it was
stressed to participants that they were supposed to talk how they would in a natural
situation. All efforts were made to collect data that is as close to participants’ natural way
of speaking as possible.

Finally, participants were presented with the experiment, which employed use of
the book by Mercer Mayer *Frog, where are you?* (1969) for the picture description task,
hard copies for the map task and a laptop for the reading task (see Appendices A, B and C).

As described above, the reading task was distributed in a PowerPoint where fillers
and target words were presented in random order. Each word (with its picture below it)
was presented in a different slide. Before starting the task, participants received
instructions written in Spanish and English and completed a training with two examples.
Once participants’ felt comfortable with the task, they started the experiment. Speakers
were able to move to the next slide/word at their own pace. They were only allowed to
read a single repetition. Then, participants were asked to complete the linguistic
background questionnaire, which was distributed in English or Spanish depending of
participants’ preference (Appendices D, E, F, G).

Participants were recorded in a quiet room or in a laboratory setting. In Puerto
Rico, researcher and fieldwork assistant used the facilities of The University of Puerto
Rico, Río Piedras and the Salinas Public Library. In occasional cases, recordings were made at participant’s place of work, such as the City Hall of Salinas. In Holyoke, MA the researcher met most participants at the Holyoke Public Library, where quiet study rooms were booked. Recordings were also carried out at participants’ homes when they were not available to meet at the library. All efforts were done to reduce and avoid environmental noise.

Subjects were recorded as they performed the experimental task using a Zoom H4n digital audio recorder and an AKG C520 condenser microphone placed around 10 cm away from the corner of the speaker’s mouth in order to avoid turbulence due to direct airflow from impinging on the microphone. The recordings were digitized at a 44,100 Hz sample rate and 16-bit amplitude resolution. A total of 45 .wav files were analyzed with an average duration of 45 minutes.

4.6 Acoustic analysis and coding

To control for consistency, the researcher only selected the recordings associated with the three experimental tasks. The first minutes of conversation before starting the experimental task were not transcribed.

Each onset /r/ and /h/ specified in the phonological representation of the words produced by the participants was analyzed. On average, a total of 97.6 tokens were elicited from each participant. As a result, a total of 4,393 phonemic /r/ and /h/ were analyzed. All data elicited with possible instrument malfunction or user error were not included in the analysis.
Each token was extracted from the recording and analyzed in PRAAT (Boersma & Weenink, 2014), a free scientific acoustic software program well known in phonetic and phonological research for the phonetic analysis of speech. Tokens were bounded off using interval tiers on a textgrid. Once all words in which the target sounds are embedded were transcribed orthographically, /r/ and /h/ were described segmentally on the basis of perceptual coding. That is to say, /r/ and /h/ tokens were listened to and then assigned to one of the categories considered (see Figure 25 below). Those tokens were selected to later run the script which extracted the acoustic measurements (center of gravity, skewness and kurtosis) as well as the duration in seconds of each segment marked off in the text grid. In order to be consistent, each segmentation starts right in the point the waveform goes upwards through the zero crossing.

![Waveform](image)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>Trill</td>
</tr>
<tr>
<td>e</td>
<td>r</td>
</tr>
<tr>
<td>perro</td>
<td></td>
</tr>
</tbody>
</table>

Figure 25: Normative trill in the utterance *perro* ‘dog’.
Figure 25 shows the visual representation of the acoustic-phonetic properties of the word *perro* ‘dog’ which was extracted from an interview with a female Puerto Rican participant in Holyoke, MA. From bottom to top the word in which the target sound is produced is orthographically transcribed in the first tier. In the second tear, preceding and subsequential sounds are transcribed. In the third tear, a description of the target’s sound phonetic realization (e.g. trill, approximant, tap, retroflex) is written. Finally, in the fourth tear, phonemic /r/ or /h/ were transcribed.

Formant structure, waveform amplitude, voicing, occlusions and spectral moments were used in PRAAT order to better identify the various phonetic realizations. Prescriptive trills were identified by the presence of one or more lingual contacts in the spectrogram, a reduction of amplitude of the waveform, changes in intensity, and a transition in F3 and F4 formant structure (Martínez Celdrán, 1998). Pre-aspirated or pre-breathy voiced (segmental portion preceding the first closure of the segment) realizations were identified based on the presence of a brief burst in the spectrogram, the appearance/absence of voicing and a reduction in amplitude as well as formant structure (Willis, 2006). Voicing was analyzed in terms of presence/absence based on pitch tracking, periodicity of the waveform, appearance of the voicing bar and pulse analysis (Willis et al., 2015).

Determining the boundaries of the fricative realizations (voiced and voiceless) included both the consultation of the waveform and the wideband spectrogram of each sound. Following Erker’s (2010) segmentation, the onset was determined by the start of high frequency frication and the cessation of the F2 in the previous vowel. The offset was regulated by the cessation of frication and by the start of the F2 of the following vowel.
Cases where the target sound is a phonemic /h/ and the previous sound is another fricative, like the coda /s/ aspiration in *las joyas*, ‘the jewelry’, were not included in the analysis due to co-articulation or transitional effects.

Spectral moments (center of gravity, skewness and kurtosis) have been shown to be useful in differentiating place of articulation in fricatives (Erker, 2010; Gordon, 2002; Gradoville, 2011; Haley et al., 2010). Therefore, in order to differentiate among all possible backed realizations of phonemic /r/ and phonemic /h/ and shed light onto the potential neutralization in production, a PRAAT script which produces the first spectral moments from fricative spectra was run (DiCanio, 2013). The script averages the Discrete Fourier Transform for the signal (DFTs) using time-averaging (Shadle, 2012). For each /r/ or /h/ realization, DFTs are averaged and the spectral moments are calculated. Due to co-articulation or transitional effects, the duration of the sound analyzed is equivalent to the center 80% of the total duration. Since the signal is sampled at 44.1 kHz, the Resampling rate is 44100. The window number was set to 6 and the window size was 15ms, so there was only up to 50% overlap between adjacent analysis windows. Low-pass filter at 300Hz was considered, as well as a high-pass filter to control for voicing.

The center of gravity of a spectrum is the frequency at which sonic energy is most concentrated (Erker, 2010). Following the center of gravity literature (Flores, 2016; Jongman et al., 2000; Silbert & de Jong, 2008), lower center of gravity values reflect a place of articulation further back along the vocal tract (see Chapter 2). Thus, we expect the lower the value is, the more backed the realization will be. The center of gravity “is not a standardized measure; therefore, the exact mean center of gravity values cannot be
mapped onto a specific place along the palate; it is the difference in mean center of gravity between the variants that reflects a change in place of articulation” (Flores, 2016, p. 12).

Skewness refers to the distribution’s asymmetry (spectral tilt or slant of the energy distribution). As Jongman et al. (2000, p. 1253) claim, “positive skewness suggests a negative tilt with a concentration of energy in the lower frequencies”. Therefore, we expect that the higher the value, the more backed the realization. “Negative skewness is associated with a positive tilt and a predominance of energy in the higher frequencies”. Finally, zero values are associated with symmetric distribution of lower and higher frequencies, that is to say, “a symmetrical distribution around the mean” (Jongman et al., 2000, p. 1253).

Kurtosis captures the distribution’s peakedness. “Positive values indicate a relatively high peakedness (the higher the value, the more peaked the distribution), while negative values indicate a relatively flat distribution” (Jongman et al., 2000, p. 1253). Therefore, well-resolved peaks are expected among more backed sounds.

Once all acoustic measurements were extracted and, therefore, once described both instrumentally and segmentally, tokens were transferred into an excel document and coded for the study’s independent variables, which were either linguistic or sociolinguistic. For the later ones, they were completed considering the data collected in the sociolinguistic questionnaire.

4.6.1 Linguistic variables
Taking into account the previous literature on rhotic variation (Beaton, 2015; Delgado-Díaz & Galarza, 2015; Graml, 2009; Valentín-Márquez, 2007) as well as on fricative variation, (Flores, 2016; Jongman et al., 2000), the linguistic variables considered to analyze the factors that might predict different realizations as well as the values of center of gravity, skewness and kurtosis are SEGMENT DURATION, VOICING, STRESS, WORD POSITION, PREVIOUS and FOLLOWING SOUND.

For SEGMENT DURATION, as stated before, a script was run to extract the duration in seconds of each segment (/h/ or /r/) marked off in the first tier of the text grid. VOICING was analyzed based on pitch tracking, periodicity of the waveform, appearance of the voicing bar and pulse analysis (Willis et al., 2015). Thus, this variable was coded as either yes (Y) or no (N). Syllable STRESS (Flores, 2016; Graml, 2009) is also coded as either yes (Y) or no (N).

Regarding the phonological context, each token was coded for WORD POSITION. (Flores, 2016; Graml, 2009; Valentín-Márquez, 2007). This variable distinguishes whether the /r/ or /h/ occurs in word initial position (I), as in rosa ‘rose’, in intervocalic position (V), as in perro ‘dog’ or after a heterosyllabic consonant (C), that is to say /n,l,s/, as in alrededor “surrounding/around”. Those three contexts were chosen since those are the positions where phonemic /r/ is expected to be produced. To be consistent, the same contexts were analyzed for the variable /h/.

Finally, PREVIOUS and FOLLOWING SOUND were examined (Flores, 2016; Graml, 2009). Those variables identify whether the segments surrounding /r/ or /h/ are consonants or vowels. For PREVIOUS SOUND, /a,e,i,o,u/ or consonant (c) were coded and, for FOLLOWING SOUND, as expected, only /a,e,i,o,u/ were studied. When /r/ or /h/ occurred
phrase initially, the PREVIOUS SOUND was coded as a pause (0). For the statistical analysis, vowels were grouped into central (/a/), fronted (/e/-/i/) or backed (/o/-/u/).

4.6.2 Sociolinguistic variables

Taking into account the information gathered in the sociolinguistic questionnaire, the sociolinguistic variables include GENERATION, AGE GROUP, SEX, PLACE, ORIGIN ON THE ISLAND, SOCIOECONOMIC CLASS, TASK, TIES WITH THE PUERTO RICAN COMMUNITY, TRAVELS TO PUERTO RICO, CONTACT WITH OTHER COMMUNITIES, LINGUISTIC PROFICIENCY IN ENGLISH, LINGUISTIC PROFICIENCY IN SPANISH, ENGLISH USE, SPANISH USE, CORRECTED PRONUNCIATION and YEARS IN THE U.S. (Hulsen et al., 2002; Ramos-Pellicia, 2004; Valentín-Márquez, 2007).

As explained in the participant’s section above, the division for GENERATION followed Silva-Corvalán’s criteria (1994) in her study on Mexican Spanish in Los Angeles. Either 1 or 2 were coded for first or second generation.

For AGE GROUP, three contrasting categories were considered: young adults (Y) (18-25 years old), middle-aged adults (M) (26-45 years old) and older adults (O) (older than 45). This division deviates slightly from standard age group divisions (<30, 30-50 and 50+) in other sociolinguistic studies (Silva-Corvalán, 2001) in order to better accommodate the sample population.

The variable SEX (Flores, 2016; Graml, 2009, López Morales, 1983; Matta de Fiol, 1981; Medina-Rivera, 1997; Valentín-Márquez, 2007), is coded as female (F) or male (M). Furthermore, the variable PLACE divides participants based on their place of residence: either Puerto Rico (PR) or Holyoke, MA (H). For ORIGIN (Graml, 2009; López
Moral, 1983), the categories considered are: metropolitan (M), being those from San Juan, or non-metropolitan (N), which are those from Salinas.

The classification of the speakers by AGE GROUP, SEX, PLACE and ORIGIN is straightforward. However, the division in terms of SOCIOECONOMIC STATUS categories is more complex. In fact, as Guy (1988) explains, there are a variety of factors that could be taken into account in order to determine the socioeconomic status such as occupation, education, place of residence, ethnicity, race, etc. Since, as seen before, most of Puerto Ricans live in the same areas of Holyoke, place of residence was not considered in this study. Following the recommendations provided by the National Committee on Vital and Health Statistics, The American Psychological Association explains that education, income, occupation and family size are the best practices for measuring socioeconomic status. Similarly, The Census Bureau considers income thresholds that change depending on composition and family size to classify population as impoverished. Following such organizations, the present study considers education, occupation and family size in calculating a SOCIOECONOMIC STATUS category.

For education, participants were divided into three groups considering whether participants have elementary-school education (E), high school education (S), or college education (C). One point was given to those participants who had elementary school, two points to those who have high school and three points to those who have college education.

Regarding occupation, jobs were divided in terms of manual laborers (L), middle professionals (M) and professionals (P), a classification that was also considered by
Rohena-Madrazo (2008). One point was given to manual laborers, two points to semi-professionals and three points to professionals.

As stated in Chapter 3, Puerto Rican residents are the ethnic group receiving the least income in the city of Holyoke. In fact, there is no upper socioeconomic class among this community. Considering that income can be a very sensitive topic, it was not considered in the current study. The American Psychological Association explains that examining household composition and family size are necessary in order to calculate poverty. Therefore, such information was gathered in the sociolinguistic questionnaire to establish the SOCIOECONOMIC STATUS variable. One point was given to participants who live alone or with another person in the same household, two points were given to participants who live with two or three more people in the same household and three points were given when four or more people live in the same household.

Considering the previous measurements, a scale from 0 to 9 was created. Participants who received 4-6 points were included in the lower middle class and participants with 7-9 points were included in the upper middle class. None of the participants’ scores fall into the category of working class (0-3). Therefore, considering Kahl’s social class divisions (1957), two socioeconomic groups were examined in the present study: lower middle class (L) (usually high school or some college education participants and semi-professionals or manual laborers) and upper middle class (U) (college graduates and professionals) (see Table 13 below).

<table>
<thead>
<tr>
<th></th>
<th>Holyoke, MA</th>
<th>Puerto Rico</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upper middle class</strong></td>
<td>12 participants</td>
<td>18 participants</td>
</tr>
<tr>
<td><strong>Lower middle class</strong></td>
<td>9 participants</td>
<td>6 participants</td>
</tr>
</tbody>
</table>
Previous sociolinguistic research has shown that participants pay more attention to their speech when they are involved in more formal contexts, while speakers produce innovative realizations in more informal contexts (Dillard, 1962; Guitart, 1981; Medina-Rivera, 1997). In order to take context into account, the type of task was also analyzed and coded as P (Picture description task), M (Map task) or R (Reading task). The Picture description task involves the most spontaneity of the three tasks, followed by the Map task, and finally the Reading task.

In line with Valentín-Márquez (2007), contact within the Puerto Rican community is also controlled in the present study. For participants living in Holyoke, Massachusetts, two different variables are considered: TIES WITH THE PUERTO RICAN COMMUNITY and TRAVELS TO PUERTO RICO. The first one was coded considering three different questions from the background questionnaire:

1. Do you have contact with other Puerto Ricans in Massachusetts?
2. Do you have contact with other Puerto Ricans in the rest of Mainland US?
3. Do you have contact with other Puerto Ricans in Puerto Rico?

Answers for each question were coded as follows: 1 (Once a year), 2 (Once or twice a month), 3 (Once a week), 4 (Every day). The numbers for each question were added up, creating a scale from 1 to 12 which was simplified into 1 to 4. TRAVELS TO PUERTO RICO was coded in a similar way: 1 (the participant travels to Puerto Rico every 10 years or more), 2 (every 5 years), 3 (every 2 years), 4 (every year). Finally, numbers for each variable were added up, creating a scale from 1 to 8 which was simplified to 1 to 4 (from lower to higher contact with the Puerto Rican community).
For participants living in Puerto Rico, the coding was slightly different. Since it is expected that Puerto Ricans would have a high level of contact with other Puerto Ricans in Puerto Rico, TIES WITH THE PUERTO RICAN COMMUNITY comprised CONTACT WITH PUERTO RICANS IN THE UNITED STATES and TRAVELS TO MAINLAND UNITED STATES. The first one was coded as 1 (Once a year), 2 (Once or twice a month), 3 (Once a week), 4 (Every day). TRAVELS TO THE UNITED STATES was coded in a similar way than TRAVELS TO PUERTO RICO: 1 (the participant travels to mainland U.S. every 10 years or more), 2 (every 5 years), 3 (every 2 years), 4 (every year). Numbers for each variable were added up, creating a scale from 1 to 8 which was simplified to 1 to 4 (from lower to higher contact with the Puerto Rican community in Mainland US).

Moreover, following Valentín-Márquez (2007) dissertation study, in both settings under study, CONTACT WITH OTHER COMMUNITIES was analyzed (Mexican, Dominican, Guatemalan, etc.). It was coded in a similar way than the previous variables: as 1 (Once a year), 2 (Once or twice a month), 3 (Once a week), 4 (Every day).

In terms of LINGUISTIC PROFICIENCY, participants were asked to rate how well they speak, read, write and understand spoken Spanish as well as English, on a scale from 1 to 6, where the highest number represented native fluency and the lower number beginning fluency. The numbers for each linguistic skill was added up creating a scale from 1 to 24, which was simplified into the following scale: 1 (beginning fluency), 2 (basic fluency), 3 (intermediate fluency), 4 (advanced fluency), 5 (near-native fluency) 6 (native fluency).

Participants also reported their ENGLISH AND SPANISH USE in and outside their home on a scale from 0 to 4. Once again, numbers for both questions were added up and
a simplified scale was considered: 0 (never); 1 (once or twice a month); 2 (two or three times a week); 4 (every day).

Finally, YEARS IN THE U.S. and CORRECTED PRONUNCIATION were considered. Although these are not crucial variables for the research questions in the present study, the researcher finds it prudent to control for them in the statistical analysis.

4.7 Data analysis

Qualitative and quantitative results are presented. The former shows figures of all /r/ and /h/ realizations found among the data. Moreover, it details frequency distribution and average durations for both settings under study.

The statistical analysis of variance in the data was performed with R statistical software (R Core Team, 2014) with the R packages lmerTest, lme4 and stats. Logistic and Linear Regression Models were run to distinguish between fricatives, and to determine any group differences for these measures. Models and interactions were compared using the anova function in R.

In the first place, different series of mixed effects logistic regression analysis with presence vs absence of backed /r/ as dependent variable and SPEAKER as random effect were performed. All realizations were collapsed to be backed /r/ vs. all other realizations. Data from Holyoke, MA and Puerto Rico were included in the first mixed effects logistic regression model. Afterwards, data were divided into two groups depending on the nature of the independent variables examined: linguistic or sociolinguistic. For each series, data were divided in terms of origin: Holyoke, MA and Puerto Rico.
To continue, only backed /r/ realizations were analyzed. Series of three mixed effect linear regression analysis with Center of gravity, Skewness and Kurtosis and SPEAKER as random effect were run. Those analyses had the objective of capturing backed /r/ variation in a continuum using acoustic measurements. Data were also divided in terms of linguistic and sociolinguistic variables as well as place.

Finally, phonemic /r/-/h/ neutralization was examined. In line with previous sections, series of three mixed effect linear regressions analysis with Center of gravity, Skewness and Kurtosis and SPEAKER as random effect were run. In the current series, both phonemic /h/ and /r/ realizations were included. Concretely, this section examines the interaction terms for PHONEMIC SOUND with each other linguistic and sociolinguistic variables. PHONEMIC SOUND includes either phonemic /r/ (coded as /r/) or phonemic /h/ (coded as /h/). Once again, in order to have parallel models, data were divided in terms of linguistic and sociolinguistic variables as well as place.

The linguistic variables considered in the statistical analysis are VOICING, STRESS, POSITION, PREVIOUS and FOLLOWING SOUND (Flores, 2016; Hulsen et al., 2002; Ramos-Pellicia, 2007; Valentín-Márquez, 2007). The sociolinguistic variables examined were: GENERATION, AGE GROUP, SEX, PLACE, ORIGIN ON THE ISLAND, SOCIOECONOMIC CLASS, TASK, TIES WITH THE PUERTO RICAN COMMUNITY, TRAVELS TO PUERTO RICO, CONTACT WITH OTHER COMMUNITIES, LINGUISTIC PROFICIENCY IN ENGLISH, LINGUISTIC PROFICIENCY IN SPANISH, ENGLISH USE, SPANISH USE, PRONUNCIATION CORRECTED and YEARS IN THE U.S. (Hulsen et al., 2002; Ramos-Pellicia, 2004; Valentín-Márquez, 2007).
CHAPTER 5
PRODUCTION EXPERIMENT

This chapter presents the analyses performed on the data in order to test the hypotheses presented in the previous chapter. Qualitative results are shown and discussed first, followed by the quantitative findings. To recap, this dissertation investigates the PRS onset trill variation found in Holyoke, MA and in Puerto Rico, analyzing the different linguistic and sociolinguistic factors that might affect such variation, and compares the backed realizations of /r/ with the allophonic inventory of /h/ in order to investigate whether or not a neutralization occurs in the communities under study.

5.1 Qualitative results
In the next section, figures of all /r/ realizations found among the data analyzed are presented. Examples of /h/ are also included. Afterwards, frequency distribution and average durations for both settings under study (Puerto Rico and Holyoke) are detailed.

5.1.2 Phonetic realizations
Figure 26 presents spectrographic output for a word-medial normative trill, containing two apical occlusions, that is to say, two consecutive instances in which the apex of the tongue (active articulator) approaches the alveolar area (passive articulator), blocking airflow in the vocal tract. Figure 26 also shows the corresponding amplitude reduction of the waveform.
As previously classified, and following Graml’s (2009) categorization, trills are separated from approximant trills. This latter realization involves the approximation of two articulators (the apex of the tongue and the alveolar area) but without enough precision to create the necessary turbulent airflow and, consequently, complete occlusions. The approximant trill falls between vowels and fricatives, therefore, its spectrographic formant structure, as seen in Figure 27, is not as precisely defined as that of the vowel that follows it. Graml (2009) refers to this realization as *approximante postalvéolaire*, ‘postalveolar approximant’. Figure 27 shows an example of an 0-occlusion approximant trill for *rotito* ‘little hole’. 

---

Figure 26: Word-medial trill production with two apical occlusions for *narrar* ‘narrate’
For the production of the approximant trill, it is possible that the friction between the apex of the tongue and the alveolar area is reinforced (Graml 2009). Graml (2009, p. 56) describes this realization as *fricative postalvéolaire*, ‘fricative post–alveolar’. Due to the few examples found of this realization and the difficulty to differentiate between the approximant and fricative alveolar/post-alveolar in the spectrogram, both realizations were unified for both, the quantitative and statistical analysis. It further exemplifies the importance of continuous measurements in descriptions of phonetic inventories. Figure 28 shows an example of the fricative postalveolar with friction in the higher frequencies.
Figure 28: Fricative trill production for *recostado* ‘leaned back’.

Figure 29 presents spectrographic output for an intervocalic tap, containing one apical occlusion. The apex of the tongue blocks airflow in the alveolar area only once. Valentín-Márquez (2007) and Graml (2009) are the first researchers that bring visibility to the importance of the tap as one of the main possible realizations of the phonemic rhotic trill (Graml, 2009).
Figure 30 presents a pre-aspirated /r/. It starts with pharyngeal friction and finishes with one [hr] or two or more apical occlusions [hr]. This pharyngeal friction can be voiced or voiceless. When it is voiced, it correlates with the pre-breathy voiced tap/trill [ɦɾ-ɦɾ] analyzed by Willis (2006) in his paper on Dominican Spanish. The aspiration was distinguished in terms of voicing; as such, 4 different aspirated variants are distinguished: [hr], [hr], [ɦɾ], or [ɦɾ]. However, since there are a few examples, for following analysis, they were grouped in either pre-aspirated trill and pre-aspirated tap (including both, voiceless and voiced aspirations). Figure 30 shows an example of a pre-breathy segmental portion (appearance of the voicing bar and reduction in amplitude as
well as formant structure) followed by multiple occlusions [fir] in the word *mirra* ‘myrrh’. Figure 31 shows an example of a pre-breathy segmental portion followed by a tap. Finally, Figure 32 illustrates a pre-aspirated segmental portion (voiceless) followed by one occlusion. In this last example, the spectrogram depicts the aspirated portion followed by a short vocalic segment which facilitates the transition towards the tap, which Graml (2009) also anticipated in her analysis.

Figure 30: Pre-breathy segmental portion (voiced) followed by three occlusions in *mirra* ‘myrrh’
Figure 31: Pre-breathy segmental portion (voiced) followed by one occlusion in *Israel*

‘Israel’
Figure 32: Pre-aspirated segmental portion (voiceless) followed by one occlusion in *riñón* ‘kidney’

Another frequently occurring variant in the corpus is a backed phonemic trill, presented in Figure 33, where the back of the tongue approaches the soft palate or the uvula. Even though several realizations of backed /ɾ/ have been reported (Graml, 2009), in this section, all single voiceless fricative realizations produced in the posterior oral cavity were grouped into the same category (without distinguishing between uvular vs. velar places of articulation). In order to shed light on this issue and be able to analyze the backness of each variant, the statistical analysis considers skewness, kurtosis and center of gravity. These acoustic parameters are detailed in Chapter 4. Figure 33 shows spectrographic output for a word-medial backed or velar realization, with a long period of
voiceless frication and the reduction of amplitude of the waveform that characterizes this realization.

Navarro Tomás (1948) finds other mixed backed variants of the phonemic trill. Instead of a pharyngeal friction, this segment starts with a backed or velar friction and ends with a trill [xr] or a tap [xɾ]. As Graml’s (2009) points out, the articulation of this segment is characterized by being produced in two completely different places: the soft palate and the alveolar area. Therefore, it involves the use the back of the tongue for the first phase of the articulation and the apex for the second phase of this complex articulation. In the case of the [xr] realization, Graml (2009) explains that sometimes there is a short vowel segment in the transition between the tap and the fricative. As she
states (2009, p. 63): *Cet élément vocalique est une conséquence nécessaire du changement du lieu d’articulation*, ‘this vocalic segment is a necessary consequence of the place of articulation change’. In our data, no cases of [xr] were found. However, a few instances of [xr] occurred. Figure 34 shows an example of [xr].

![Figure 34: Backed /r/ realization followed by a tap in corriendo ‘running’](image)

Finally, there were several cases where a rhotic segment (tap or trill) was followed by an alveolar approximant or fricative. This miscellaneous variant was coded and analyzed among the data. This is illustrated in Figure 35.
For the phonemic /h/, two different realizations were found and analyzed: a voiced [ɦ] and a voiceless [h] variant. Figures 36 and 37 provide two examples. Voicing was analyzed based on pitch tracking, periodicity of the waveform, appearance of the voicing bar and pulse analysis (Willis et al., 2015).
Figure 36: Voiced /h/ realization in ajo ‘garlic’
5.1.3 Frequency distribution

With regards to trill variation, the spectrographic analysis revealed eight different realizations: trill, approximant trill, backed trill, pre-aspirated trill, pre-aspirated tap, tap, backed trill followed by a tap and a post-tap and trill frication. As mentioned above, this division groups together fricative and approximant trills, as well as pre-breathy and pre-aspirated taps/trills under the same category. Figure 38 below shows the overall frequency distribution of /r/ variants. The most common realization is the normative trill, reaching almost half of the total /r/ variants (44%, n=1149), followed, as we could expect since it is a weaker version of the normative /r/, by the approximant trill (19%=500). As
mentioned above, this latter realization involves the approximation of two articulators but without enough precision to create the necessary turbulent airflow and complete the occlusions. The third most common variant is the backed /r/, the main sound under study in this dissertation, reaching 15% (n=392) of the overall /r/ realizations. It is followed by the tap (10%, n=272). The least common variants are the pre-aspirated trills and taps (4% each one, n=105, n=100), the post tap/trill frication (3%, n=80) and the backed /r/ followed by a tap (1%, n=15).

![/r/ Realizations Pie Chart]

Figure 38: Overall frequency distribution of /r/ variation

In both settings under study, the island of Puerto Rico and the Puerto Rican diasporic community in Holyoke, MA, the same eight realizations are found. Furthermore, a more detailed analysis of the results (Figure 39 and Table 14) reveals that there is little difference in frequency distribution in relation to participants’ origin. Hence,
in both settings, the most commonly occurring trill variant is the normative articulation: [r], followed by the approximant trill, the backed /ɾ/ and the tap. While the fifth most frequent variant in Holyoke is the pre-aspirated tap, the pre-aspirated trill occupies this place in Puerto Rico. For Holyoke, the next most common realizations are the pre-aspirated trill and the post tap/trill fricative. For Puerto Rico, the post tap/trill fricative is followed by the pre-aspirated tap. Finally, in both settings, the least produced variant is the backed /ɾ/ followed by a tap. Although backed realizations are slightly more common in Holyoke (15.65%, n=208) than in Puerto Rico (14.46%, n=184), this difference is minimal (1.19%, n=24), suggesting that the backed /ɾ/ behaves virtually identically in both settings under study.

Figure 39: Frequency distribution of /ɾ/ variation by place.
Table 14: Hierarchy of /r/ variation by place

<table>
<thead>
<tr>
<th>Puerto Rico</th>
<th>Holyoke</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.</strong> Trill (50.76%, n=730/1439)</td>
<td>Trill (35.65%, n=419/1199)</td>
</tr>
<tr>
<td><strong>2.</strong> Approximant trill (16.27%, n=243/1439)</td>
<td>Approximant trill (22.63%, n=266/1199)</td>
</tr>
<tr>
<td><strong>3.</strong> Backed /r/ (14.46%, n=184/1439)</td>
<td>Backed /r/ (15.65%, n=208/1199)</td>
</tr>
<tr>
<td><strong>4.</strong> Tap (10.08%, n=145/1439)</td>
<td>Tap (10.8%, n=127/1199)</td>
</tr>
<tr>
<td><strong>5.</strong> Pre-aspirated trill (3.4%, n=49/1439)</td>
<td>Pre-aspirated tap (5.7%, n=67/1199)</td>
</tr>
<tr>
<td><strong>6.</strong> Post tap/trill fricat. (2.5%, n=36/1439)</td>
<td>Pre-aspirated trill (4.76%, n=56/1199)</td>
</tr>
<tr>
<td><strong>7.</strong> Pre-aspirated tap (3.4%, n=49/1439)</td>
<td>Post tap/trill fricat. (3.74%, n=44/1199)</td>
</tr>
<tr>
<td><strong>8.</strong> Backed + tap (0.22%, n=3/1439)</td>
<td>Backed + tap (1.02%, n=12/1199)</td>
</tr>
</tbody>
</table>

With regards to phonemic /h/, considering voicing, two realizations were distinguished. As seen in Figure 40, the majority of the /h/ realizations analyzed were voiced (71%, n=1255), in comparison to the voiceless variant (27%, n=483). There were also a few cases of /h/ deletion (2%, n=42). Moreover, although the voiceless realization is more produced in Puerto Rico and the voiced variant is more common in Holyoke, this difference is minimal (see Figure 41). We find, thus, the same pattern in both locations.
With regards to /h/ variation, Willis et al. (2015) presented the distribution of voiceless and voiced realizations according to position (post-pausal or phrase medial context), showing that in absolute initial position there is a preference for a voiceless realization. The same pattern is found in our sample (see Table 15), /h/ is typically produced as voiceless when it follows a pause, while when it is produced after a consonant or vowel, it is generally voiced. This result, therefore, gives more evidence to the research indicating that phonemic /h/ is conditioned by context.

Table 15: Frequency of /h/ realizations considering POSITION.

<table>
<thead>
<tr>
<th>Position</th>
<th>Voiceless /h/</th>
<th>Voiced /h/</th>
</tr>
</thead>
<tbody>
<tr>
<td>After /n,l,s/</td>
<td>19.43% (n=41/211)</td>
<td>20.13% (n=120/596)</td>
</tr>
<tr>
<td>Initial position</td>
<td>56.39% (n=119/211)</td>
<td>2.01% (n=12/596)</td>
</tr>
<tr>
<td>Intervocalic position</td>
<td>24.17% (n=51/211)</td>
<td>77.85% (n=464/596)</td>
</tr>
</tbody>
</table>
5.1.4 Average durations

Regarding the overall average durations (Figure 42 and Table 16), in line with Graml’s results (2009), we find that pre-aspirated (or pre-breathy) trills (119 ms) and pre-aspirated taps (105 ms) are, along with the backed /r/ + tap (111 ms) the three /r/ variants with the longest duration. This result is unsurprising since those realizations involve different phases in their production. It is noteworthy to mention that, for the pre-aspirated realizations, the average of the aspiration is 56 ms while the rhotic portion presents an average of 57 ms of the total segment duration. This means that the aspirated section comprises approximately half the duration of the trill, which deviate slightly from Graml’s (2009) study on PRS and Willis’s (2006, 2007) work on Dominican Spanish. Both researchers report a larger duration for the aspirated segment (60% of the overall rhotic). The backed /ɾ/ and the normative trill occupy the next places in our hierarchy, with 78 ms and 68 ms, respectively, followed by the approximant variant (62 ms) and the post tap/trill frication realization (59 ms). In opposition to the trill, the approximate realization does not have different phases in its articulation (Graml, 2009); therefore, it has a shorter expected duration.

Finally, it is generally claimed that the trill is longer than the tap. In fact, Quilis (1993) reported average durations of 20 ms and 85 ms for the Peninsular Spanish tap and trill, measurements that are comparable to the averages addressed by Blecua (2001), 23 ms and 64 ms. Not surprisingly, trills (68 ms) are also longer than taps (11 ms) in our data, in line with previous studies (Blecua, 2001; Graml, 2009; Quilis, 1993).
Table 16: Overall average durations in milliseconds.

<table>
<thead>
<tr>
<th>Overall durations</th>
<th>Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-aspirated trill</td>
<td>0.11971</td>
</tr>
<tr>
<td>Backed /r/ + tap</td>
<td>0.1114</td>
</tr>
<tr>
<td>Pre-aspirated tap</td>
<td>0.10585</td>
</tr>
<tr>
<td>Backed /r/</td>
<td>0.07838</td>
</tr>
<tr>
<td>Trill</td>
<td>0.06857</td>
</tr>
<tr>
<td>Approximant trill</td>
<td>0.0626</td>
</tr>
<tr>
<td>Post tap/trip frication</td>
<td>0.0591</td>
</tr>
<tr>
<td>Tap</td>
<td>0.03009</td>
</tr>
</tbody>
</table>

Regarding the normative trill [r], Figure 43 shows the average durations depending on the number of occlusions. As expected, duration increases with the number of occlusions.
To conclude this section, with regards to phonemic /h/, voiceless realizations present a higher mean average duration (72ms) than the voiced variants (54ms), in line with Willis et al.’s findings (2015) which also report a longer mean duration for the voiceless allophone.

5.2 Quantitative results: Statistical analysis

In order to assess the role of different linguistic and sociolinguistic variables in relation to the production of the backed /r/, data are analyzed using R (R Core Team, 2014) with the R packages lmerTest, lme4 and stats. Section 5.2.1 presents three different series of mixed effects logistic regression analyses with presence vs. absence of backed /r/ as dependent variable and SPEAKER as a random effect. As explained in Chapter 4, all realizations are collapsed into two categories: backed /r/ vs. all other realizations. In the first mixed effects logistic regression model, data from both places under study (Holyoke, MA and Puerto Rico) are included in order to have a general understanding about the
main predictors that affect the dependent variable. However, the resulting model is too big and reliable results cannot be generated. To make the model more stable, data are divided into two groups: (i) linguistic and (ii) sociolinguistic variables. For each series, data are divided in terms of origin: Holyoke, MA and Puerto Rico. In this way, the different factors that might affect trill variation in each location are compared.

Secondly, in order to capture backed /r/ variation in a continuum using acoustic measurements, section 5.2.2 presents a series of three mixed effect linear regression analyses with center of gravity, skewness and kurtosis as dependent variables and \textit{SPEAKER} as random effect. In order to mirror the previous series of regressions, data are divided in terms of linguistic and sociolinguistic variables as well as place (Holyoke, MA data and Puerto Rico data).

Lastly, in order to test the possible /r/-/h/ neutralization in production among PRS speakers, not only in Puerto Rico, but in the diaspora, section 5.2.3 presents another series of three mixed effect linear regressions analyses with center of gravity, skewness and kurtosis as dependent variables and \textit{SPEAKER} as random. While in the previous analyses, only backed /r/ realizations are analyzed, both phonemic /h/ and /r/ realizations are included here. Once again, in order to have parallel models, data are divided in terms of linguistic and sociolinguistic variables in addition to geographic location.

5.2.1 Results for presence vs absence of backed /r/ as dependent variable

5.2.1.1 General results

A mixed modal linear regression analysis with presence vs. absence of backed /r/ as the dependent variable and \textit{SPEAKER} as a random effect is run for all data. Since including all
linguistic and sociolinguistic variables makes the model unstable, only \textit{PLACE} and \textit{ORIGIN} are considered as independent variables. The purpose of this first analysis is to have an overall idea of the data and examine whether \textit{PLACE} is a significant predictor, implying significant differences in the production of the backed /r/ realizations between the Puerto Rican community on the island of Puerto Rico and in Holyoke, MA. Table 17 shows the Estimate, Standard Error, z values and p values for the model with absence of backed /r/ as dependent variable and with metropolitan area and Puerto Rico as reference baseline levels. For the baseline participant, it is significantly less likely to produce a backed /r/ than a non-backed /r/.

Table 17: Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and \textit{SPEAKER} as a random effect, for the all the data combined.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backed /r/</td>
<td>(Intercept) Baseline=M\textsuperscript{9}, PR</td>
<td>-5.94300</td>
<td>.65068</td>
<td>-9.134</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>4.47674</td>
<td>.65383</td>
<td>6.847</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Place (H)</td>
<td>-.09776</td>
<td>.54195</td>
<td>-.180</td>
<td>.857</td>
</tr>
</tbody>
</table>

Findings show that \textit{PLACE} is not significant, suggesting that there are not main differences in the realization of the backed /r/ between Puerto Rico and the Holyoke diaspora. However, \textit{ORIGIN} is significant: speakers in Holyoke and Puerto Rico who are originally from non-metropolitan areas (p<.001) produce more backed /r/ realizations than speakers from the metropolitan area, as can be seen in Figure 44.

\footnote{Abbreviations are in correspondence with the description of variables described and explained in the methodology section.}
Although the linear regression presented above does not show significant differences for \textit{PLACE}, the next step was to study in more detail the variation for each location. Therefore, two different analyses are run. As mentioned above, the first model analyzes the role of the linguistic variables and the second model analyzes the role of the sociolinguistic factors. For each model, Holyoke data is presented first, followed by the Puerto Rico data.

\textbf{5.2.2 Linguistic variables}

In the following section, the role of the linguistic factors in relation to the backed /r/ are analyzed. The linguistic variables considered are \textit{STRESS}, \textit{POSITION}, \textit{PREVIOUS} and \textit{FOLLOWING SOUND}. Models and interactions are compared using ANOVA.

\textbf{5.2.2.1 Results for Holyoke}
The best-fit model for Holyoke include STRESS and POSITION as fixed effects (AIC: 631.04). Only POSITION was a significant predictor. Table 18 shows the Estimate, Standard Error, z values and p values for the model with absence of backed /r/ as dependent variable and with unstressed and initial position as baseline predictors. For the baseline participant, it is significantly less likely to produce a backed /r/ than a non-backed /r/.

Table 18: Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and SPEAKER as a random effect for the linguistic variables among the Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backed /r/</td>
<td>(Intercept)</td>
<td>4.6900</td>
<td>.9739</td>
<td>6.183</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Baseline=N, I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stress (Y)</td>
<td>0.4206</td>
<td>.2220</td>
<td>1.895</td>
<td>.0581</td>
</tr>
<tr>
<td></td>
<td>Position (C)</td>
<td>-1.3321</td>
<td>.6007</td>
<td>-2.218</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Position (V)</td>
<td>-2.8395</td>
<td>.5511</td>
<td>-5.152</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

As can be seen in Figure 45, POSITION is indeed a significant predictor of backed /r/: backed /r/ is more likely in intervocalic position (p<.001) or after an heterosyllabic consonant (p<.01) when compared to initial position. The model was afterwards releveled in order to better understand the data, showing that the backed /r/ is more likely in intervocalic position (p<.001) when compared to after heterosyllabic position.
5.2.2.2 Results for Puerto Rico

The best-fit model for Puerto Rico only includes POSITION as fixed effects (AIC: 829.05).

Table 19 shows the Estimate, Standard Error, z values and p-values for the model with absence of backed /r/ as dependent variable and with initial position as reference baseline level.

Table 19. Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backed /r/</td>
<td>(Intercept)</td>
<td>5.0127</td>
<td>0.8738</td>
<td>5.737</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Baseline=I</td>
<td>Position (C)</td>
<td>-.9947</td>
<td>0.4022</td>
<td>-2.473</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Position (V)</td>
<td>-1.2828</td>
<td>0.3545</td>
<td>-3.618</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Figure 45: Presence of Backed /r/ as a function of POSITION in Holyoke
Similarly to the model for Holyoke, **STRESS** is not a significant predictor, while **POSITION** is (see Figure 46): it is more likely to produce a backed /r/ in intervocalic position ($p<.001$) or after a consonant ($p<.01$) than in initial position. The model was also releveled in order to compare after heterosyllabic consonant position to intervocalic position, revealing that there are not statistically significant differences in the production of the backed /r/ between those two positions.

![Presence of Backed /r/ based on Position](image)

*Figure 46: Presence of Backed /r/ as a function of POSITION in Puerto Rico*

**5.2.1.2.3 Summary – the effect of linguistic variables on the production of backed /r/**

In Holyoke, only **POSITION** is a significant predictor. The model shows that it is more likely to produce the backed /r/ in intervocalic position than after heterosyllabic position or in initial position, in line with Graml (2009), Hammond (1991), Lopéz Morales (1983)
and Terrel (1980). Among the Puerto Rico data, POSITION is also significant, revealing the same result as in Holyoke: intervocalic position favors the appearance of the backed /r/.

5.2.1.3 Sociolinguistic variables

The following series of regressions analyze the role of the sociolinguistic factors in relation to the backed /r/. Like the models considering linguistic variables, two models are run for each location. The sociolinguistic variables considered are ORIGIN, AGE GROUP, SEX, GENERATION, SOCIOECONOMIC STATUS, TASK, TIES WITH THE PR COMMUNITY, TRAVELS TO PR, YEARS IN THE US, PRONUNCIATION CORRECTED, and CONTACT WITH OTHER COMMUNITIES. Models and interactions are compared using ANOVA.

5.2.1.3.1 Results for Holyoke

When backed /r/ is selected as the dependent variable, the best-fit model for Holyoke, MA includes ORIGIN, AGE GROUP, SEX, GENERATION, SOCIOECONOMIC STATUS, TASK, TIES WITH THE PR COMMUNITY, TRAVELS TO PR, YEARS IN THE U.S. AND PRONUNCIATION CORRECTED as fixed effects (AIC: 567.5). Table 20 shows the Estimate, Standard Error, z-values and p-values for the model with presence of backed /r/ as dependent variable and with metropolitan area, older, female, first-generation, lower middle class, contact with the Puerto Rican community once or twice a month, travels to PR every year, no pronunciation corrected and Picture Description task as reference baseline levels. For the baseline participant, it is significantly less likely to produce a backed /r/ than a non-backed /r/. 
Table 20. Summary of the linear mixed-effect regression model with presence of backed /r/ as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backed /r/</td>
<td>(Intercept) Baseline=M, O, M, L, 1, P, 1, 1, N</td>
<td>15.17927</td>
<td>3.30225</td>
<td>4.597</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>-4.60215</td>
<td>1.39309</td>
<td>-3.304</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Age group (M)</td>
<td>-4.18993</td>
<td>1.22891</td>
<td>-3.409</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Age group (Y)</td>
<td>-3.17818</td>
<td>1.60953</td>
<td>-1.975</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Sex (M)</td>
<td>-1.08882</td>
<td>1.01211</td>
<td>-1.076</td>
<td>&lt;.282</td>
</tr>
<tr>
<td></td>
<td>Socioeconomic status (U)</td>
<td>1.67522</td>
<td>.89188</td>
<td>1.878</td>
<td>&lt;.1</td>
</tr>
<tr>
<td></td>
<td>Ties PR community (once a week)</td>
<td>4.39277</td>
<td>1.85720</td>
<td>-2.365</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Ties PR community (every day)</td>
<td>-10.73726</td>
<td>4.13384</td>
<td>-2.597</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Task (R)</td>
<td>3.43186</td>
<td>.42278</td>
<td>8.117</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Task (M)</td>
<td>.96837</td>
<td>.25602</td>
<td>3.782</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Years US</td>
<td>-.23307</td>
<td>.06845</td>
<td>-3.405</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Travels to PR (2)</td>
<td>-15.26603</td>
<td>105.62974</td>
<td>-.145</td>
<td>.885</td>
</tr>
<tr>
<td></td>
<td>Travels to PR (3)</td>
<td>5.33114</td>
<td>3.20752</td>
<td>1.662</td>
<td>&lt;.1</td>
</tr>
<tr>
<td></td>
<td>Travels to PR (4)</td>
<td>5.06203</td>
<td>3.05985</td>
<td>1.654</td>
<td>&lt;.1</td>
</tr>
<tr>
<td></td>
<td>Generation (2)</td>
<td>4.87026</td>
<td>1.65304</td>
<td>2.946</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Pronunciation corrected (Y)</td>
<td>15.97178</td>
<td>105.61803</td>
<td>.151</td>
<td>.879801</td>
</tr>
</tbody>
</table>

The model shows that ORIGIN, AGE GROUP, GENERATION, SOCIOECONOMIC STATUS, TASK, TIES WITH THE PR COMMUNITY, TRAVELS TO PR, and YEARS IN THE U.S. are significant. SEX and PRONUNCIATION CORRECTED are not significant. Results for the significant predictors are discussed below.
ORIGIN is a significant predictor for the Holyoke data: backed /r/ is more likely to be produced in Holyoke, MA when the speaker is originally from a non-metropolitan area in Puerto Rico (Salinas), than from a metropolitan area (San Juan) \((p<.001)\) (Figure 47). Hence, those findings correspond with the overall data analysis previously reported.

![Presence of Backed /r/ based on Origin](image)

Figure 47: Presence of Backed /r/ as a function of ORIGIN in Holyoke

In terms of AGE GROUP, a backed /r/ is more likely to be produced in Holyoke, MA when the speaker is a middle-aged adult \((p<.001)\) or a young adult \((p<.05)\), than an older adult. GENERATION is also a significant predictor: it is more likely that a backed /r/ be produced in Holyoke, MA when the speaker is from the second generation than the first generation \((p<.01)\) (see Figure 48 below).
SOCIOECONOMIC STATUS is another significant predictor (Figure 49). Although the difference is marginal, the model shows that it is more likely to produce a backed /r/ when the speaker is from the lower middle class than from the upper middle class (p<.1).
As for TASK, speakers produce more backed realizations when they are completing the Picture Description Task than when they are performing the Reading Task (p<.001), or the Map Task (p<.001) (see Figure 50. The model was relevered in order to compare the Reading Task to the Map Task, showing that the backed /r/ is less likely to be produced when participants are performing the Reading Task (p<.001). This is an expected result since the Picture Description Task and the Map Task involve more spontaneous speech in comparison to the Reading Task.
With regards to TIES WITH THE PR COMMUNITY, the baseline case refers to a participant that speaks with the PR community once a month. As illustrated in Figure 51, as their ties with the PR community increase, they are more likely to produce a backed /r/. That is to say, when the ties with the PR community are closer, such as when speakers have contact with other members of the community every day (p<.01) or once a week (p<.05), use of a backed /r/ is more likely than when the contact occurs only once a month.
As stated in the methodology, this variable is related to TRAVELS TO PR, in which the baseline refers to a speaker that has not traveled to Puerto Rico since moving to Holyoke, MA. As exhibited in Figure 52, when the speaker travels to Puerto Rico every year (p<.1) or every other year (p<.1), the backed /r/ is produced with more frequency than when the speaker rarely travels to Puerto Rico. Hence, when the participant travels back and forth to Puerto Rico, the probability of producing a backed /r/ increases. When the speaker’s travels to Puerto Rico are limited to every five years, there is no effect on the realization of the backed /r/ despite having a large coefficient. This result is due to the fact that the standard error is very high (SE: 105.62). Is it possible, thus, that this factor is correlated with other variables, looking like an increase in the figure but a decrease in the model.
Finally, YEARS IN THE U.S. is significant. For every additional year that the speaker is in the US, the probability of producing a backed /r/ increases. As stated in the methodology, although the model controls for it, this variable does not have the same theoretical importance as the rest of factors in this study.

### 5.2.1.3.2 Results for Puerto Rico

GENERATION, TRAVELS TO PR and YEARS IN THE U.S. are not included in the model since they are not relevant for Puerto Rico’s data. When Backed /r/ is selected as the dependent variable, the best-fit model for Puerto Rico included ORIGIN, AGE GROUP, SEX, SOCIOECONOMIC STATUS, TASK and CONTACT WITH OTHERS COMMUNITIES as fixed effects (AIC: 757.69). Table 21 shows the Estimate, Standard Error, z-values and p-values for
the model with absence of backed /r/ as dependent variable and with metropolitan area, older, female, contact with other communities once or twice a month and Picture Description task as reference baseline levels. For the baseline participant, the production of a backed /r/ is significantly less likely than the production of a non-backed /r/.

Table 21. Summary of the linear mixed-effect regression model with absence of backed /r/ as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backed /r/ (Intercept)</td>
<td>-5.4835</td>
<td>.9784</td>
<td>-5.605</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Baseline=M, O, F, L, P, 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Origin (N)</td>
<td>5.1618</td>
<td>.8297</td>
<td>6.222</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age group (M)</td>
<td>-2.3811</td>
<td>.6291</td>
<td>-3.785</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age group (Y)</td>
<td>-1.7298</td>
<td>.5260</td>
<td>-3.288</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Sex (M)</td>
<td>2.1719</td>
<td>.6600</td>
<td>3.291</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Socioeconomic status (U)</td>
<td>-2.0666</td>
<td>.6463</td>
<td>-3.198</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Task (R)</td>
<td>1.7682</td>
<td>.2666</td>
<td>-6.633</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Task (M)</td>
<td>.4685</td>
<td>.2116</td>
<td>-2.214</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Contact with other communities (2)</td>
<td>.6845</td>
<td>.4932</td>
<td>1.388</td>
<td>.165</td>
</tr>
<tr>
<td>Contact with other communities (3)</td>
<td>1.8291</td>
<td>.6847</td>
<td>2.671</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Contact with other communities (4)</td>
<td>-.4862</td>
<td>.5275</td>
<td>-.922</td>
<td>.356</td>
</tr>
</tbody>
</table>

The model shows that ORIGIN, AGE GROUP, SEX, SOCIOECONOMIC STATUS, TASK, and CONTACT WITH OTHER COMMUNITIES are significant. In this section, figures of the most important theoretical variables are included.

ORIGIN is selected as a significant predictor: in Puerto Rico, a backed /r/ is more likely to be produced when the speaker is originally from Salinas, the non-metropolitan area, than from San Juan, the metropolitan area (p<.001), as can be seen in Figure 53.
Those findings correspond with the overall data analysis as well as with the Holyoke analysis.

With regards to AGE, the model shows a slightly different pattern as opposed to Holyoke. In fact, in Puerto Rico, a backed /r/ is less likely to be produced when the speaker is a middle-aged adult (p<.001) or a young adult (p<.01) than an older speaker (Figure 54). Therefore, in contrast to the diasporic data (middle-aged adults and younger speakers produce more backed /r/), in Puerto Rico older speakers as well as young adults are the two age groups that most use this stigmatized sound. The model was afterwards relevered in order to compare middle-aged adults to young adults, revealing no statistically differences between those two groups in the production of the backed /r/.

**Figure 53: Presence of Backed /r/ as a function of ORIGIN in Puerto Rico**
Figure 5: Presence of Backed /r/ as a function of AGE in Puerto Rico

SEX is another significant predictor, unlike in the Holyoke data: in Puerto Rico, female speakers are more likely to produce a backed /r/ than male speakers (<.001) (Figure 55).
In relation to socioeconmic status, in consonance with the Holyoke data, the model that a backed /r/ is more likely to be produced when the speaker is from a lower middle class than from an upper middle class (p<.01), as can be seen in Figure 56.

**Figure 55: Presence of Backed /r/ as a function of sex in Puerto Rico**
In relation to TASK, the model shows the same pattern as in the Holyoke analysis (Figure 57). Puerto Ricans produce more backed realizations when they are completing the Picture Description task than when they are performing the Reading task \((p<.001)\), or the Map Task \((p<.05)\). The model was afterwards re-leveled in order to compare the Reading Task to the Map Task, showing that the backed /r/ is less likely to be produced when participants are performing the Reading Task \((p<.001)\) in comparison to the Map Task, that is to say, when the task involves more spontaneous speech.

Figure 56: Presence of Backed /r/ as a function of SOCIODEMOMIC STATUS in Puerto Rico
Presence of Backed /r/ based on Task

Figure 57: Presence of Backed /r/ as a function of TASK in Puerto Rico

5.2.1.3.3 Summary – the effect of sociolinguistic variables on the production of backed /r/

Results for Holyoke show that ORIGIN, AGE GROUP, SEX, GENERATION, SOCIOECONOMIC STATUS, TASK, TIES WITH THE PR COMMUNITY, TRAVELS TO PR, YEARS IN THE U.S. are significant. Findings for ORIGIN reveal the same pattern as in the overall data analysis, an unsurprising result considering the previous body of literature on PRS (Graml, 2009; López Morales, 1983). The model presents that middle age-adults and young adults produce the backed /r/ more than older speakers, not in line with Valentin-Márquez (2007) findings in Grand Rapids, MI, which show no differences among generational
groups in the diaspora. In the current data, it is possible that the older age group is more aware of the socioindexical meaning that this sound carries on the island, avoiding its realization. Furthermore, second generation speakers are more likely to use the backed /r/ than first generation. This result could be explained considering the relationship between age and generation: second generation speakers are mostly middle-aged adults as can be seen in the cross tabulation below (Table 22). Moreover, those findings suggest that first-generation speakers (who are mostly older adults) may be aware of the socioindexical meaning/stigma that the backed /r/ carries on the island, avoiding its use, in line with Ramos-Pellicia (2004) results. In her study, first generation PRS speakers in Lorain, OH are aware of the negative prestige that lateralizing their /r/ carries in PRS and use other alternatives, such as deletion or a retroflex instead of a [l].

Table 22. Frequency of backed /r/ realizations by age and generation.

<table>
<thead>
<tr>
<th></th>
<th>Older adults</th>
<th>Middle-aged adults</th>
<th>Young adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>First generation</td>
<td>472</td>
<td>224</td>
<td>62</td>
</tr>
<tr>
<td>Second generation</td>
<td>0</td>
<td>290</td>
<td>126</td>
</tr>
</tbody>
</table>

In regard to SOCIODEMOCRATIC STATUS, although the difference is minimal, the lower middle class produces more backed /r/ than the upper middle class, a finding in line with the rich body of sociolinguistic research that began with Labov’s study (1972) on New York city, where he illustrates the social stratification of /r/ in different department stores. In his study, those speakers with higher socioeconomic status produce more /r/ than those
with lower status, suggesting that more prestigious variants are more frequently used among speakers with higher social class. Similarly, focusing on the backed /r/, Dillard (1962, p. 423) mentions: [...] el fenómeno es tanto más perceptible cuanto más bajamos en la escala sociocultural [...] (‘the phenomenon is more perceptible the more we move down the sociocultural class’). Although it was not statistically analyzed, data show a pattern of higher frequencies of [x] when the speaker has lower level of social class in Graml’s (2009) and Valentin-Márquez’s (2007) dissertation projects. The current results, hence, mirror this previous research on PRS.

As for TASK, the model shows that participants produced more backed realizations when the task involves more spontaneous speech, which is a common result in sociolinguistic research. Likewise, Dillard (1962) reveals that in formal contexts, speakers produce the standard realization more frequently while a familiar/informal style favors the velar or backed realization. Along the same lines, Guitart (1981) shows that although the normative or standard variant can be found in all contexts, the backed /r/ is only produced in spontaneous speech. Furthermore, Graml’s (2009) findings report that when participants read a list of words or a poem the backed /r/ (15.8%) is less produced than when the task involves more spontaneous speech. Similarly, the current results show that the backed /r/ is produced more when the speaker is involved in more spontaneous tasks (Picture Description Task and Map Task) than in controlled tasks (Reading task).

Finally, two more variables were considered in order to test the effect of contact with the Puerto Rican community in the realization of the backed /r/: TIES WITH THE PR COMMUNITY and TRAVELS TO PUERTO RICO. The model demonstrates that those Puerto Ricans with increased ties to the PR community are more likely to produce a backed /r/. 
Moreover, when the participant travels back and forth to Puerto Rico, the probability of producing a backed /r/ increases. This result does not mirror those of Valentín-Márquez’s analysis, which does not reveal differences related to speakers’ networks and community integration in the use of the backed /r/.

As reported by Graml (2009) and Valentín-Márquez (2007), SEX is not significant (see Figure 58). However, it is possible to see a tendency: male speakers produce a higher percentage of backed /r/ realizations than female speakers. This result does not corroborate Valentín-Márquez’s (2007) findings: although his analysis did not yield statistical significance, in Cabo Rojo [x] was more frequent among male speakers while in Grand Rapids it was used as frequently by women as by men.

![Figure 58: Presence of Backed /r/ as a function of SEX in Holyoke](image)
Finally, data for Puerto Rico shows that ORIGIN, AGE GROUP, SEX, SOCIOECONOMIC STATUS, and TASK are significant.

Results for ORIGIN, SOCIOECONOMIC STATUS, and TASK mirror those from Holyoke, Massachusetts: participants from Salinas who have a lower socioeconomic status and are involved in more spontaneous tasks produce more backed /r/ than participants from San Juan who have a higher socioeconomic status and are involved in controlled tasks.

AGE GROUP is also significant; however, this pattern differs slightly from that of Holyoke. While in Holyoke middle-aged adults and younger speakers produce more backed /r/, in Puerto Rico older speakers as well as young adults are the two age groups who use this stigmatized sound more frequently. This finding does not completely correlate with Valentín-Márquez’s results: the backed /r/ is gradually diminishing in Cabo Rojo since adolescents produce 7% of [x] occurrences, middle-age adults articulate 19% and older speakers produce 45% of the overall backed realizations.

With regard to SEX, and in contrast to Holyoke data, this variable significantly affects the production of backed /r/. Contrary to the hypothesis, female speakers use more backed /r/ than male speakers in Puerto Rico, in line with Medina-Rivera’s (1997) work in Caguas, Puerto Rico.

5.2.2 Results for the acoustic analysis: center of gravity, skewness and kurtosis

As previously explained, after assessing the role of all linguistic and sociolinguistic variables in relation to the presence versus absence of backed /r/, the role of the same predictors is analyzed in relation to three known measurements that acoustically
characterize fricatives and, therefore, the backed /r/: center of gravity, skewness and kurtosis. The first objective is to investigate whether the backed /r/ realizations cluster into distinct groups or whether they present a continuum which may be favoring the neutralization of /r/ and /h/ in PRS. Consequently, the distribution of center of gravity, skewness and kurtosis are examined. Figure 59 shows the distribution of center of gravity, revealing that values are overall dispersed in a continuum from, approximately, 1000 Hz to 3500 Hz. Interestingly, although in both areas under study the majority of the backed /r/ realizations are produced with a center of gravity of 1500 Hz, values are more widely distributed in Holyoke. The statistical analyses will reveal if this difference is, in fact, significant.

Figure 59: Distribution of center of gravity
As with center of gravity, Figure 6 shows that skewness values are also dispersed in a continuum from, approximately, 0.3 to 10. However, values do not seem as spread out as in the case of center of gravity. As for differences depending on PLACE, the histogram suggests that the majority of the backed /r/ realizations are produced with an approximate value of 0.4 for skewness in both places. Potential statistical differences will be verified with statistical analyses.

Interestingly, the range for kurtosis values is more reduced than for skewness or center of gravity, as can be seen in Figure 61. In both settings under study, values disperse from 0 to, approximately 200, although the graph shows more backed /r/ realizations with a kurtosis of 0.2.
Therefore, this first approximation to center of gravity, skewness and kurtosis reveal that values are dispersed on a continuum which, depending on the acoustic measurement can be more (center of gravity & kurtosis) or less dispersed (kurtosis). For this reason, analyzing the backed /r/ on a continuum without having to prove conclusively that backed /r/ realizations can be either velar, uvular or both, gains importance.

In order to capture backed /r/ variation, three series of mixed effect linear regression analyses with center of gravity, skewness and kurtosis as dependent variables and SPEAKER as random effect are performed.

The first series include the overall data, that is to say, backed /r/ realizations in both areas, Holyoke and Puerto Rico. Afterwards, in order to mirror the previous
analyses, data are divided in terms of linguistic and sociolinguistic variables as well as place (Holyoke, MA vs Puerto Rico).

5.2.2.1 General results

Three mixed modal linear regression analysis with center of gravity, skewness and kurtosis as dependent variables and SPEAKER as a random effect are run among all data. To make the model more stable, only PLACE and ORIGIN are considered as independent variables. The objective of this first analysis is to analyze whether PLACE is a significant predictor, implying significant differences in the acoustic characteristics of the backed /r/ realizations between Puerto Rico and the diasporic setting of Western Massachusetts.

Center of gravity

Table 23 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable, and with metropolitan area and Puerto Rico as reference baseline levels.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>(Intercept)</td>
<td>2558.71</td>
<td>300.48</td>
<td>8.515</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Baseline=M, PR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>-599.96</td>
<td>294.16</td>
<td>-2.040</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Place (H)</td>
<td>104.52</td>
<td>166.81</td>
<td>.627</td>
<td>.539</td>
</tr>
</tbody>
</table>

In line with the analyses taking presence vs absence of backed /r/ as the dependent variable, PLACE is not significant, suggesting that there are not differences in the center of gravity mean of backed /r/ realizations between Puerto Rico and Holyoke. However,
ORIGIN is significant: speakers who are originally from non-metropolitan areas (p<.05) produce lower values for center of gravity than speakers from the metropolitan area, as can be seen in Figure 62. Therefore, the frequency at which sonic energy is most concentrated in the production of backed /r/ is lower among participants from non-metropolitan areas. Since the lower the value, the more backed the realization, these findings suggest that the backed /r/ is produced more posterior among speakers originally from non-metropolitan areas than among speakers from metropolitan areas.

Figure 62: Boxplots for center of gravity values for ORIGIN. (I refers to the island of Puerto Rico and M refers to Massachusetts).

**Skewness**

Table 24 shows the Estimate, Standard Error, t-values and p-values for the model with skewness as the dependent variable, and with metropolitan area and Puerto Rico as reference baseline levels.
Table 24: Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect, for the all the data combined.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept)</td>
<td>2.9012</td>
<td>1.0907</td>
<td>2.660</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Baseline=M, PR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>1.7951</td>
<td>1.0727</td>
<td>1.674</td>
<td>&lt;.1</td>
</tr>
<tr>
<td></td>
<td>Place (H)</td>
<td>.1344</td>
<td>.4944</td>
<td>.272</td>
<td>.788</td>
</tr>
</tbody>
</table>

Similar to the results from previous analyses (presence vs absence of backed /r/) as well as with the results for center of gravity, PLACE is not selected as a significant factor, suggesting that there are not differences in the skewness values of the backed /r/ realizations between the two locations under study. On the contrary, ORIGIN is marginally significant: speakers who are originally from non-metropolitan areas (p<0.1) produce higher values for skewness than speakers from the metropolitan area, as can be seen in Figure 63. Since the higher the value, the more backed the realization, findings suggest that the backed /r/ is produced more posterior among speakers originally from non-metropolitan areas than among speakers from San Juan. This result correlates with the center of gravity findings presented above (speakers from non-metropolitan areas present a concentration of energy in the lower frequencies, in comparison to speakers from San Juan).
Figure 63: Boxplots for skewness values for ORIGIN

Kurtosis

Table 25 shows the Estimate, Standard Error, t-values and p-values for the model with kurtosis as the dependent variable, and with metropolitan area and Puerto Rico as reference baseline levels.

Table 25: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect, for the all the data combined.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept)</td>
<td>20.727</td>
<td>24.314</td>
<td>.852</td>
<td>.395</td>
</tr>
<tr>
<td></td>
<td>Baseline=M, PR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>24.631</td>
<td>24.012</td>
<td>1.026</td>
<td>.306</td>
</tr>
<tr>
<td></td>
<td>Place (H)</td>
<td>2.876</td>
<td>9.152</td>
<td>.314</td>
<td>.757</td>
</tr>
</tbody>
</table>

Neither PLACE nor ORIGIN are significant, suggesting that there are not differences in the kurtosis values of the backed /r/ between speakers from Holyoke and Puerto Rico nor are there differences between speakers from the metropolitan and speakers from the non-metropolitan area.
5.2.2.2 Linguistic variables

In the following section, the role of the linguistic factors in relation to center of gravity, skewness and kurtosis are analyzed. The linguistic variables considered are STRESS, POSITION, PREVIOUS and FOLLOWING SOUND.

5.2.1.2.1 Results for Holyoke

Center of gravity

Table 26 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable and with unstressed, previous /e/-/i/ and following /a/ as reference baseline levels. PREVIOUS and FOLLOWING SOUND are selected as significant fixed effects.

Table 26: Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>(Intercept)</td>
<td>2824.77</td>
<td>229.80</td>
<td>12.292</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Baseline=N,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>/e/-/i/, /a/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress (Y)</td>
<td></td>
<td>122.14</td>
<td>94.15</td>
<td>1.297</td>
<td>.196</td>
</tr>
<tr>
<td>Previous (0)</td>
<td></td>
<td>192.88</td>
<td>249.58</td>
<td>.773</td>
<td>.441</td>
</tr>
<tr>
<td>Previous (c)</td>
<td></td>
<td>-125.63</td>
<td>128.10</td>
<td>-.981</td>
<td>.328</td>
</tr>
<tr>
<td>Previous (/a/)</td>
<td></td>
<td>-37.30</td>
<td>110.45</td>
<td>-.338</td>
<td>.736</td>
</tr>
<tr>
<td>Previous (/o/-/u/)</td>
<td></td>
<td>-715.23</td>
<td>134.98</td>
<td>-5.299</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Following (/e/-/i/)</td>
<td></td>
<td>-164.06</td>
<td>190.60</td>
<td>-.861</td>
<td>.390</td>
</tr>
<tr>
<td>Following (/o/-/u/)</td>
<td></td>
<td>-1070.27</td>
<td>196.83</td>
<td>-5.437</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>
PREVIOUS and FOLLOWING SOUND (see Figure 6.4 below) are indeed significant predictors of the values of center of gravity among Holyoke data. The model reveals that when the backed /r/ is produced after /o/-/u/ (p<.001) the values for center of gravity are significantly lower than when it is produced after /e/-/i/. Similarly, those values are significantly lower when the backed /r/ is produced before either /o/ or /u/ in comparison to the baseline case (/a/) (p<.001). Therefore, findings with regard to PREVIOUS and FOLLOWING SOUND suggest that when the backed /r/ is produced after or before a back vowel (/o,u/), the values for center of gravity are significantly lower, implying that the sound of interest is produced more posterior.

Figure 6.4: Boxplots for center of gravity values for FOLLOWING SOUND (lower part of the Figure) and PREVIOUS SOUND (upper area) in Holyoke. (c) refers to consonant and (0) refers to absence of sound produced before the backed /r/.

Skewness
The best fit model when skewness is selected as the dependent variable include POSITION and FOLLOWING SOUND as fixed factors. Table 27 shows the Estimate, Standard Error, t-values and p-values for the model with skewness as the dependent variable, and with after consonant position and following /a/ as reference baseline levels. FOLLOWING SOUND is selected as significant fixed effects.

Table 27: Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept) Baseline=N, C, /a/</td>
<td>3.259</td>
<td>1.008</td>
<td>3.232</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Position (I)</td>
<td>-.616</td>
<td>1.274</td>
<td>-.484</td>
<td>.629</td>
</tr>
<tr>
<td></td>
<td>Position (V)</td>
<td>.035</td>
<td>.587</td>
<td>.060</td>
<td>.952</td>
</tr>
<tr>
<td></td>
<td>Following (/e/-/i/)</td>
<td>-.362</td>
<td>.841</td>
<td>-.430</td>
<td>.667</td>
</tr>
<tr>
<td></td>
<td>Following (/o/-/u/)</td>
<td>2.626</td>
<td>.811</td>
<td>3.235</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

As can be seen in Figure 65, the model shows that FOLLOWING SOUND is significant, reporting similar results in correlation to center of gravity findings. The regression reveals that skewness values are significantly higher when the backed /r/ is produced before /o,u/ (p<0.1) in comparison to the baseline case (/a/). Recall from the methodology section that the higher the value, the more backed the realization; therefore, results suggest that the backed /r/ is produced more posterior when it appears after a back vowel.
Figure 65: Boxplots for skewness values for FOLLOWING SOUND in Holyoke.

Kurtosis

For the third acoustic measure, kurtosis, none of the independent variables are significant, as can be seen in Table 28.

Table 28: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the linguistic variables among Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept)</td>
<td>32.603</td>
<td>23.800</td>
<td>1.370</td>
<td>.172</td>
</tr>
<tr>
<td></td>
<td>Baseline=N, C, /a/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress (Y)</td>
<td>-4.624</td>
<td>10.258</td>
<td>-.451</td>
<td>.652</td>
<td></td>
</tr>
<tr>
<td>Position (I)</td>
<td>-27.204</td>
<td>30.115</td>
<td>-.903</td>
<td>.367</td>
<td></td>
</tr>
<tr>
<td>Position (V)</td>
<td>-2.800</td>
<td>13.789</td>
<td>-.203</td>
<td>.839</td>
<td></td>
</tr>
<tr>
<td>Following (/e/-/i/)</td>
<td>-9.651</td>
<td>19.681</td>
<td>-.490</td>
<td>.624</td>
<td></td>
</tr>
<tr>
<td>Following (/o/-/u/)</td>
<td>38.349</td>
<td>19.969</td>
<td>1.920</td>
<td>.056</td>
<td></td>
</tr>
</tbody>
</table>
5.2.1.2.1 Results for Puerto Rico

Center of gravity

Table 29 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as dependent variable, and unstressed, previous sound /e/-/i/ and following sound /a/ as reference baseline levels. PREVIOUS and FOLLOWING SOUND are selected as significant fixed effects.

Table 29: Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response of gravity</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>Baseline=N, /e/-/i/, /a/</td>
<td>2252.6</td>
<td>139.39</td>
<td>16.161</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Stress (Y)</td>
<td></td>
<td>71.22</td>
<td>79.41</td>
<td>.897</td>
<td>.370</td>
</tr>
<tr>
<td>Previous (0)</td>
<td></td>
<td>86.9</td>
<td>161.61</td>
<td>.538</td>
<td>.591</td>
</tr>
<tr>
<td>Previous (c)</td>
<td></td>
<td>-167.57</td>
<td>110.05</td>
<td>-1.523</td>
<td>.129</td>
</tr>
<tr>
<td>Previous (/a/)</td>
<td></td>
<td>-174.67</td>
<td>93.53</td>
<td>-1.867</td>
<td>.063</td>
</tr>
<tr>
<td>Previous (/o/-/u/)</td>
<td></td>
<td>-363.41</td>
<td>134.35</td>
<td>-2.705</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Following (/e/-/i/)</td>
<td></td>
<td>97.92</td>
<td>106.81</td>
<td>.917</td>
<td>.360</td>
</tr>
<tr>
<td>Following (/o/-/u/)</td>
<td></td>
<td>-619.73</td>
<td>119.96</td>
<td>-5.166</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

With regard to PREVIOUS and FOLLOWING SOUND (see Figure 66 below), the model reveals similar results to the findings for Holyoke: when the backed /r/ is produced after /o/-/u/ (p<.01), the values for center of gravity are significantly lower than when it is produced after /e/-/i/. Furthermore, when the backed /r/ is produced before a backed vowel (/o,u/), the values for center of gravity are significantly lower (p<.001) in
comparison to the baseline case (/a/). Those results imply that the sound of interest is produced more posterior when it is produced between backed vowels.

Figure 6.6: Boxplots for center of gravity values for FOLLOWING SOUND (lower part) and PREVIOUS SOUND (upper area) in Puerto Rico.

Skewness

Table 30 shows the Estimate, Standard Error, t-values and p-values for the model with skewness as dependent variable, and with previous sound /e/-/i/ and following sound /a/ as reference baseline levels. PREVIOUS and FOLLOWING SOUND are selected as significant mixed effects.
Table 30: Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept)</td>
<td>3.538</td>
<td>.566</td>
<td>6.242</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Baseline= N, (/e/-/i/), (/a/)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Previous (0)</td>
<td>.703</td>
<td>.758</td>
<td>.928</td>
<td>.354</td>
</tr>
<tr>
<td></td>
<td>Previous (c)</td>
<td>.272</td>
<td>.510</td>
<td>.534</td>
<td>.593</td>
</tr>
<tr>
<td></td>
<td>Previous (/a/)</td>
<td>.374</td>
<td>.434</td>
<td>.861</td>
<td>.390</td>
</tr>
<tr>
<td></td>
<td>Previous (/o/-/u/)</td>
<td>1.577</td>
<td>.633</td>
<td>2.489</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Following (/e/-/i/)</td>
<td>-.153</td>
<td>.499</td>
<td>-.308</td>
<td>.758</td>
</tr>
<tr>
<td></td>
<td>Following (/o/-/u/)</td>
<td>2.176</td>
<td>.537</td>
<td>4.047</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

As can be seen in Figure 67, and in line with previous results, the model shows that PREVIOUS and FOLLOWING SOUND are significant, reporting similar results in correlation to center of gravity findings: when the backed /r/ is produced after or before a backed vowel (/o,u/), the values for skewness are significantly higher, suggesting that the backed /r/ is produced more posterior when it appears before or after a velar vowel.
Figure 67: Boxplots for skewness values for FOLLOWING SOUND (lower part of the Figure) and PREVIOUS SOUND (upper area) in Puerto Rico.

**Kurtosis**

Table 31 shows the Estimate, Standard Error, t-values and p-values for the model with kurtosis as dependent variable, and with unstressed, after consonant, and following sound /a/ as reference baseline levels.

Table 31: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept) Baseline=N, C, /a/</td>
<td>25.689</td>
<td>15.914</td>
<td>1.614</td>
<td>.108</td>
</tr>
<tr>
<td></td>
<td>Stress (Y)</td>
<td>1.213</td>
<td>10.098</td>
<td>-.120</td>
<td>.904</td>
</tr>
<tr>
<td></td>
<td>Position (I)</td>
<td>5.660</td>
<td>21.257</td>
<td>-.266</td>
<td>.790</td>
</tr>
<tr>
<td></td>
<td>Position (V)</td>
<td>4.772</td>
<td>12.414</td>
<td>-.384</td>
<td>.701</td>
</tr>
<tr>
<td></td>
<td>Following (/e/-/i/)</td>
<td>1.102</td>
<td>13.084</td>
<td>-.084</td>
<td>.933</td>
</tr>
<tr>
<td></td>
<td>Following (/o/-/u/)</td>
<td>34.282</td>
<td>14.300</td>
<td>2.397</td>
<td>&lt;.05</td>
</tr>
</tbody>
</table>
FOLLOWING SOUND is selected as significant fixed effects. When the backed /r/ is produced before /o/-/u/ (p<.05), the values for kurtosis are significantly higher than when it is produced before /a/, as can be seen in Figure 68. Kurtosis captures the distribution’s peakedness. Since the higher the value, the more peaked the distribution, results suggest that the backed /r/ is produced more posterior when it is followed by a backed vowel.

Figure 68: Boxplots for kurtosis values for FOLLOWING SOUND in Puerto Rico.

5.2.2.3 Summary – the effect of linguistic variables on acoustic aspects of backed /r/

Previous literature on PRS has analyzed backed /r/ variation considering it a phonetic category, looking at the different linguistic and sociolinguistic factors affecting its frequency. However, in spite of finding evidence for the production of different
allophones (velar, uvular, glottal), none of the previous research has examined such
variation on a continuum of backedness. Since categorical descriptions have obscured
decisive facts about phonological variation and the factors that favor it (Erker, 2010), this
section assesses the role of different linguistic and sociolinguistic variables in relation to
center of gravity, skewness and kurtosis.

For the linguistic analysis, PREVIOUS and FOLLOWING SOUND are selected as
significant factors when center of gravity and skewness are the dependent variables in
each location under study (Holyoke and Puerto Rico) and when kurtosis is selected as
dependent variable in Puerto Rico. Results show the same tendency for each model,
demonstrating that when the backed /r/ is produced after or before a back vowel (/o,u/),
the acoustic characteristics of this fricative sound are different than when is produced
between central or front vowels. Concretely, the backed /r/ is produced more posterior
when it appears before or after a backed vowel, findings attributed to coarticulation
effects: the presence of a backed sounds contributes to the implementation of the backed
/r/. Similarly, in her study on velar palatalization in Chilean Spanish, Flores (2016) shows
that palatalization is favored when /x/ follows a front vowel. Therefore, in both cases,
findings can be explained considering coarticulatory processes.

5.2.2.3 Sociolinguistic variables

After assessing the role of the linguistic factors in relation to center of gravity and
skewness, the next section shows the role of the sociolinguistic factors. Like previous
sections, two models are run for each location. The sociolinguistic variables considered
are ORIGIN, AGE GROUP, SEX, GENERATION, SOCIOECONOMIC STATUS, TASK, TIES WITH THE
PR COMMUNITY, TRAVELS TO PR, YEARS IN THE US, PRONUNCIATION CORRECTED, and CONTACT WITH OTHER COMMUNITIES. Models and interactions were compared using ANOVA.

5.2.2.3.1 Results for Holyoke

Center of gravity

When Center of gravity is selected as the dependent variable, the best-fit model for Holyoke, MA includes ORIGIN, AGE GROUP, GENERATION, SOCIOECONOMIC STATUS, and TASK as fixed effects. Table 32 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable, and with metropolitan area, older, first-generation, lower middle socioeconomic status and Picture Description task as reference baseline levels.

Table 32. Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of Gravity</td>
<td>(Intercept) Baseline= M, O, 1, L, P</td>
<td>2825.01</td>
<td>402.54</td>
<td>7.018</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Origin (N)</td>
<td>-335.03</td>
<td>426.52</td>
<td>-0.786</td>
<td>0.4527</td>
<td></td>
</tr>
<tr>
<td>Age group (M)</td>
<td>1.86</td>
<td>448.76</td>
<td>0.004</td>
<td>0.9968</td>
<td></td>
</tr>
<tr>
<td>Age group (Y)</td>
<td>633.63</td>
<td>549.76</td>
<td>1.153</td>
<td>0.2726</td>
<td></td>
</tr>
<tr>
<td>Generation (2)</td>
<td>-866.22</td>
<td>435.81</td>
<td>-1.988</td>
<td>&lt;.1</td>
<td></td>
</tr>
<tr>
<td>Socioeconomic status (U)</td>
<td>-534.51</td>
<td>439.83</td>
<td>-1.215</td>
<td>0.2563</td>
<td></td>
</tr>
<tr>
<td>Task (R)</td>
<td>210.21</td>
<td>264.45</td>
<td>0.795</td>
<td>0.4278</td>
<td></td>
</tr>
<tr>
<td>Task (M)</td>
<td>73.96</td>
<td>118.63</td>
<td>0.623</td>
<td>0.5338</td>
<td></td>
</tr>
</tbody>
</table>
The model reveals that only GENERATION is marginally significant (Figure 69): the frequency at which sonic energy is most concentrated for a backed /r/ is lower when the participant is a second-generation speaker in comparison to a first-generation speaker (p<0.1). This finding suggests that the backed /r/ is produced more posterior among second-generation PRS speakers in Holyoke, MA.

![Boxplots for center of gravity values for GENERATION in Holyoke.](image)

Figure 69: Boxplots for center of gravity values for GENERATION in Holyoke.

**Skewness**

When skewness is selected as the dependent variable, the best-fit model for Holyoke, MA includes ORIGIN, AGE GROUP, GENERATION, SOCIOECONOMIC STATUS, TASK and TIES WITH THE PUERTO RICAN COMMUNITY as fixed effects. Table 33 shows the Estimate, Standard Error, t-values and p-values for the model with skewness as the dependent variable, and
with metropolitan area, older, first-generation, lower middle socioeconomic status, Picture Description task and Ties with the PR community once or twice a month as reference baseline levels.

Table 33: Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept)</td>
<td>-0.10625</td>
<td>1.76541</td>
<td>-0.060</td>
<td>0.9525</td>
</tr>
<tr>
<td></td>
<td>Baseline= M, O, 1, L, P, 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>1.06017</td>
<td>1.41922</td>
<td>0.747</td>
<td>0.4596</td>
</tr>
<tr>
<td></td>
<td>Age group (M)</td>
<td>0.05606</td>
<td>1.45660</td>
<td>0.038</td>
<td>0.9695</td>
</tr>
<tr>
<td></td>
<td>Age group (Y)</td>
<td>-2.31230</td>
<td>1.86624</td>
<td>-1.239</td>
<td>0.2224</td>
</tr>
<tr>
<td></td>
<td>Generation (2)</td>
<td>2.92788</td>
<td>1.44306</td>
<td>2.029</td>
<td>&lt;.1</td>
</tr>
<tr>
<td></td>
<td>Socioeconomic status (U)</td>
<td>2.25974</td>
<td>1.48657</td>
<td>1.520</td>
<td>0.1399</td>
</tr>
<tr>
<td></td>
<td>Task (R)</td>
<td>-1.98188</td>
<td>1.00801</td>
<td>-1.966</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Task (M)</td>
<td>-0.69693</td>
<td>0.46875</td>
<td>-1.487</td>
<td>0.1389</td>
</tr>
<tr>
<td></td>
<td>Ties PR community (3)</td>
<td>2.86502</td>
<td>1.20284</td>
<td>2.382</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Ties PR community (4)</td>
<td>2.45629</td>
<td>1.15006</td>
<td>2.136</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

The linear regression conveys that GENERATION, TASK and TIES WITH THE PUERTO RICAN COMMUNITY are significant. Skewness values are marginally higher when speakers are second-generation (p<0.1) than when they are first-generation, suggesting, in line with center of gravity findings, that the backed /r/ is produced more posterior among second-generation PRS speakers in Holyoke, MA.

Moreover, skewness values are significantly different with regards to TASK: speakers produce lower skewness values when they are performing the Reading task.
(p<.01) than when they are involved in the Picture Description task. These findings suggest that the backed /r/ is produced more posterior when speakers are accomplishing more spontaneous tasks in comparison to more controlled tasks. Afterwards, the model was releveled in order to compare the Reading Task with the Map Task, but no statistical differences were found.

Finally, as for TIES WITH THE PUERTO RICAN COMMUNITY, skewness values are significantly higher when speakers have contact with other members of the community every day (p<.01) or once a week (p<.05), than when the contact occurs only once a month. These results imply that the backed /r/ is produced more posterior when the ties with the PR community increase.

Kurtosis

For kurtosis, none of the independent variables are significant, as can be seen in Table 34.

Table 34: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among Holyoke data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept)</td>
<td>21.618</td>
<td>29.476</td>
<td>.733</td>
<td>.466</td>
</tr>
<tr>
<td></td>
<td>Baseline= M, O, L, P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>6.565</td>
<td>30.166</td>
<td>.218</td>
<td>.829</td>
</tr>
<tr>
<td></td>
<td>Age group (M)</td>
<td>36.96</td>
<td>12.843</td>
<td>2.878</td>
<td>.124</td>
</tr>
<tr>
<td></td>
<td>Age group (Y)</td>
<td>1.157</td>
<td>29.80</td>
<td>.039</td>
<td>.969</td>
</tr>
<tr>
<td></td>
<td>Socioeconomic status (U)</td>
<td>.084</td>
<td>12.174</td>
<td>.007</td>
<td>.995</td>
</tr>
<tr>
<td></td>
<td>Task (R)</td>
<td>-32.036</td>
<td>22.102</td>
<td>-1.449</td>
<td>.150</td>
</tr>
<tr>
<td></td>
<td>Task (M)</td>
<td>-12.363</td>
<td>10.121</td>
<td>-1.222</td>
<td>.224</td>
</tr>
</tbody>
</table>
5.2.2.3.2 Results for Puerto Rico

Center of gravity

When Center of gravity is selected as the dependent variable, the best-fit model for Puerto Rico includes ORIGIN, SEX, SOCIOECONOMIC STATUS, and TASK as fixed effects. Table 35 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable, and with metropolitan area, female, lower middle socioeconomic class, and Picture Description task as baseline reference levels.

Table 35: Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>(Intercept)</td>
<td>2195.93</td>
<td>465.64</td>
<td>4.716</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Baseline= M, F, L, P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>-368.21</td>
<td>456.97</td>
<td>-0.806</td>
<td>0.42552</td>
</tr>
<tr>
<td></td>
<td>Sex (M)</td>
<td>64.80</td>
<td>200.86</td>
<td>0.323</td>
<td>0.76157</td>
</tr>
<tr>
<td></td>
<td>Socioeconomic status (U)</td>
<td>-229.86</td>
<td>269.52</td>
<td>-0.853</td>
<td>0.42783</td>
</tr>
<tr>
<td></td>
<td>Task (R)</td>
<td>383.50</td>
<td>135.07</td>
<td>2.839</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Task (M)</td>
<td>280.42</td>
<td>98.62</td>
<td>2.844</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

The model only selects TASK as a significant factor: center of gravity values are higher when the speaker is involved in the Reading task (p<.01) or the Map task (p<.01) than in the Picture Description task. In line with Holyoke data, those results imply that the backed /r/ is produced more posterior when speakers are accomplishing more spontaneous tasks (Picture description task) in comparison with more controlled tasks.

Skewness
The best-fit model for Puerto Rico include ORIGIN, AGE GROUP, SEX, SOCIOECONOMIC STATUS and TASK as fixed effects. Table 36 shows the Estimate, Standard Error, t-values and p-values for the model with skewness as the dependent variable, and with metropolitan area, older, female, lower middle socioeconomic status and Picture description task as reference baseline levels. The model only selects AGE GROUP and TASK as significant factors.

Table 36: Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept) Baseline= M, O, F, L, P</td>
<td>4.47650</td>
<td>1.76128</td>
<td>2.542</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Origin (N)</td>
<td>-0.01869</td>
<td>1.73539</td>
<td>-0.011</td>
<td>0.9914</td>
</tr>
<tr>
<td></td>
<td>Age group (M)</td>
<td>-1.01297</td>
<td>0.74629</td>
<td>-1.357</td>
<td>0.1762</td>
</tr>
<tr>
<td></td>
<td>Age group (Y)</td>
<td>0.99293</td>
<td>0.45617</td>
<td>2.177</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Sex (M)</td>
<td>0.17254</td>
<td>0.45214</td>
<td>0.382</td>
<td>0.7032</td>
</tr>
<tr>
<td></td>
<td>Socioeconomic status (U)</td>
<td>0.47618</td>
<td>0.47618</td>
<td>0.721</td>
<td>0.4719</td>
</tr>
<tr>
<td></td>
<td>Task (R)</td>
<td>-0.69820</td>
<td>0.57517</td>
<td>-1.214</td>
<td>0.2262</td>
</tr>
<tr>
<td></td>
<td>Task (M)</td>
<td>-0.70556</td>
<td>0.39795</td>
<td>-1.773</td>
<td>&lt;.1</td>
</tr>
</tbody>
</table>

As for AGE GROUP, the regression reveals that younger speakers (p<.05) produce higher skewness values in comparison to older speakers, suggesting that the backed /r/ is produced slightly more posterior among younger speakers on the island of Puerto Rico (Figure 70). Then, the model was releveled in order to compare the younger speakers with the middle-aged speakers, revealing that the backed /r/ is produced even more posterior among middle-aged speakers than the younger adults (p<.01).
Lastly, results for TASK report the same tendencies as the findings for Holyoke: skewness values are marginally lower when the speaker is involved in the Reading task (p<0.1) than when they are performing the Picture Description task. These data suggest, in correlation with center of gravity results, that the backed /r/ is produced more posterior when speakers are accomplishing more spontaneous tasks (Picture description task) in comparison with more controlled tasks. After revealing the data, no statistically differences were found between the Map Task and the Reading task.

**Kurtosis**

Finally, the best fit model when kurtosis is selected as the dependent variable includes ORIGIN, not being selected as significant (Table 37).
Table 37: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect for the sociolinguistic variables among the Puerto Rico data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis (Intercept)</td>
<td>Baseline= M</td>
<td>30.28</td>
<td>44.06</td>
<td>.687</td>
<td>.493</td>
</tr>
<tr>
<td>Origin (N)</td>
<td>Origin (N)</td>
<td>15.08</td>
<td>44.38</td>
<td>.340</td>
<td>.734</td>
</tr>
</tbody>
</table>

5.2.2.3.3 Summary – the effect of sociolinguistic variables on acoustic aspects of backed /r/

Taking into account data from Holyoke, MA, GENERATION is a marginally significant factor affecting both, center of gravity and skewness values: backed /r/ realizations are produced marginally more posterior among second-generation than first-generation speakers in Holyoke, MA. Furthermore, when skewness is selected as a dependent variable, the model demonstrates that TASK and TIES WITH THE PUERTO RICAN COMMUNITY are significant factors. Speakers produce lower skewness values when they are performing the Reading task than when they are involved in the Picture description task, implying that the backed /r/ might be produced more posterior when speakers are accomplishing spontaneous tasks in comparison to more controlled tasks. With regard to TIES WITH THE PUERTO RICAN COMMUNITY, findings suggest that the backed /r/ is produced more posterior when the ties with the PR community increase.

Among Puerto Rico data, TASK is a significant predictor for skewness and for center of gravity. Findings suggest the same pattern as in Holyoke, MA: the backed /r/ is produced more posterior when speakers are accomplishing spontaneous tasks in
comparison to more controlled tasks. In addition, AGE GROUP is selected as another significant factor regarding skewness, revealing that the backed /r/ is produced slightly more posterior among middle-aged and younger speakers than older speakers.

In sum, the sociolinguistic approach reveals that when Holyoke speakers are second-generation, involved in spontaneous speech, and when they have regular contact with other members of the PR community, they are more likely to produce the backed /r/ more posterior. The same occurs in Puerto Rico among younger and middle-aged speakers. Consequently, the next step is to better understand whether those groups of speakers might be favoring the neutralization /h/ -/r/ in PRS, an issue assessed in the next section.

These previous series of mixed modal linear regression analyses confirm that center of gravity, skewness and kurtosis are helpful acoustic measurements capturing backed /r/ variation. Those continuous values are affected not only by different linguistic factors but also by sociolinguistic factors, in line with previous literature on fricatives (Flores, 2016; Jongman et al., 2000). This is not to say that there is a clear categorical distinction in terms of place of articulation (maybe [x] or [χ]) among the current data, but that these possible allophones are produced on a continuum of backed /r/ realizations that is affected by different linguistic and sociolinguistic variables, such as a PREVIOUS and FOLLOWING SOUND, GENERATION or TASK.

Thus, this section offers an acoustic alternative for a more accurate explanatory and descriptive analysis of backed /r/ variation, adding more information to the findings shown in the previous series of regressions, which considered the backed /r/ a discrete unit or categorical allophone.
5.2.3 Results for the possible /h/-/r/ neutralization

In the previous section, it was reported that it is possible to capture backed /r/ variation on a continuum using center of gravity and skewness as acoustic measurements. In line with the hypothesis, findings show that there are linguistic and sociolinguistic factors that affect the degree of backedness in the sound under study. Moving further, as mentioned in Chapter 2, there are claims suggesting the neutralization of the /r/ and /h/ phonemes in PRS (Delgado-Díaz & Galarza, 2015; Dillard, 1962; Lipski, 1994) due to the similarities in their allophonic realizations:

/r/: [r], [x], [xr], [xr], [hr], [hr], [h] (Graml, 2009; Navarro Tomás, 1948)

/h/: [h], [ɦ], or [x] (Hualde, 2005; Willis et al., 2015)

Therefore, both, the /r/ and /h/ phonemes, share [h] and [x] as allophones, favoring the neutralization, which could imply a sound change in progress in Puerto Rico. The fourth goal of this analysis is to shed light on this issue and determine whether or not there is a continuum among backed realizations which might be favoring the neutralization in PRS. Among all data containing backed /r/ and phonemic /h/, a series of two mixed effect linear regression were performed with center of gravity, skewness and kurtosis as dependent variables and SPEAKER as random effect. The first series included both areas, Holyoke and Puerto Rico. Afterwards, in order to mirror the previous sections, data were divided in terms of linguistic and sociolinguistic variables as well as place (Holyoke, MA data and Puerto Rico data).

5.2.3.1 General results
Firstly, in order to shed light onto the neutralization issue, a series of three mixed modal linear regression models were run. In the first regression, center of gravity was the dependent variable; in the second, skewness was the dependent variable; and in the third, kurtosis was the dependent variable. PHONEMIC SOUND and PLACE were included as fixed effects for each model, along with the interaction between the two. The means for each of these acoustic measurements are shown for phonemic /r/\textsuperscript{10} and phonemic /h/ in Table 38.

Table 38: Means for each acoustic measurement for phonemic /r/ and phonemic /h/

<table>
<thead>
<tr>
<th>Mean</th>
<th>Phonemic /r/</th>
<th>Phonemic /h/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>1941.907</td>
<td>1412.7319</td>
</tr>
<tr>
<td>Skewness</td>
<td>4.904472</td>
<td>7.8392</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>47.86429</td>
<td>144.2557</td>
</tr>
</tbody>
</table>

Center of gravity

Table 39 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable, and with phonemic /h/ and Puerto Rico as reference baseline levels.

Table 39: Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect, for the overall /h/-/r/ data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>(Intercept) Baseline=</td>
<td>1350.46</td>
<td>76.42</td>
<td>17.671</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>/h/, PR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/)</td>
<td>559.33</td>
<td>54.15</td>
<td>10.330</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Place (H)</td>
<td>121.19</td>
<td>110.49</td>
<td>1.097</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Place</td>
<td>12.20</td>
<td>79.35</td>
<td>0.154</td>
<td>0.878</td>
</tr>
<tr>
<td></td>
<td>(H)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{10} In this section, the phonemic /r/ category only includes backed /r/ tokens. It does not include the non-backed realizations.
Results show a significant difference in center of gravity values for phonemic /h/ vs. phonemic /r/ (p<0.001), indicating that speakers keep the two sounds distinct in terms of center of gravity, as can be seen in Figure 71. Specifically, values for /r/ are significantly higher than they are for /h/ meaning that phonemic /h/ is produced more backed than phonemic /r/. PLACE was not found to be significant, nor was the interaction between PLACE and PHONEMIC SOUND, which indicates that both Holyoke and Puerto Rico speakers produce center of gravity values for each sound in similar ways.

Figure 71: Boxplots for center of gravity of phonemic /r/ and /h/ in Holyoke and Puerto Rico

Skewness
Table 40 shows the Estimate, Standard Error, t-values and p-values for the model with skewness as the dependent variable, and with phonemic /h/ and Puerto Rico as reference baseline levels.

Table 40: Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect, for the overall /h/-/r/ data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept) Baseline=/h/, PR</td>
<td>8.2856</td>
<td>0.4345</td>
<td>19.068</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/)</td>
<td>-3.523</td>
<td>0.3836</td>
<td>-9.186</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Place (H)</td>
<td>-0.822</td>
<td>0.6287</td>
<td>-1.309</td>
<td>0.202</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Place (H)</td>
<td>0.668</td>
<td>0.5618</td>
<td>1.190</td>
<td>0.234</td>
</tr>
</tbody>
</table>

Results show a significant difference in skewness values for phonemic /h/ vs. phonemic /r/ (p<0.001), indicating that speakers keep the two sounds distinct in terms of skewness, as can be seen in Figure 72. Specifically, values for /r/ are significantly lower than they are for /h/ meaning that phonemic /h/ is produced more backed than phonemic /r/. PLACE was not found to be significant, nor was the interaction between PLACE and PHONEMIC SOUND, which indicates that both Holyoke and Puerto Rico speakers produce skewness values for each sound in similar ways. This finding correlates with the center of gravity results presented above (phonemic /h/ is produced with a concentration of energy in lower frequencies than the backed /r/). Thus, considering center of gravity and skewness means, /h/ is produced more backed than phonemic /r/, implying the inexistence of a merge in production.
Figure 72: Boxplots for center of gravity of phonemic /r/ and /h/ in Holyoke and Puerto Rico

*Kurtosis*

Table 41 shows the Estimate, Standard Error, t-values and p-values for the model with kurtosis as the dependent variable, and with phonemic /h/ and Puerto Rico as reference baseline levels.

Table 41: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and SPEAKER as a random effect, for the overall /h/-/r/ data.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kurtosis</strong></td>
<td>(Intercept) Baseline=/h/, PR</td>
<td>158.75</td>
<td>17.74</td>
<td>8.946</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/)</td>
<td>-111.13</td>
<td>17.28</td>
<td>-6.431</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Place (H)</td>
<td>-28.67</td>
<td>25.68</td>
<td>-1.116</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Place (H)</td>
<td>14.41</td>
<td>25.30</td>
<td>0.569</td>
<td>0.569</td>
</tr>
</tbody>
</table>
For Kurtosis, all phonemic /r/ and phonemic /h/ analyzed presented positive values, demonstrating high distribution’s peakedness. This is an expected result since backed sounds are characterized for having “a clearly defined spectrum with well-resolved peaks” (Jongman et al., 2000, p. 1253). Results reveal that PHONEMIC SOUND is significant (p<.001): moving from phonemic /h/ to phonemic /r/, the values for kurtosis decrease significantly, indicating more peaked distribution for phonemic /h/, as can be seen in Figure 73. Therefore, findings suggest that /h/ is produced more posterior than the backed /r/. Moreover, in line with center of gravity and skewness findings, PLACE was not found to be significant, nor was the interaction between PLACE and PHONEMIC SOUND, which indicates that both Holyoke and Puerto Rico speakers produce kurtosis values for each sound in similar ways.

Figure 73: Boxplots for kurtosis of phonemic /r/ and /h/ in Holyoke and Puerto Rico
These first series of regressions showed that there are statistically significant differences in the production of /h/ and backed /r/ with regards to all three acoustic characteristics examined here. That is to say, the means of center of gravity, skewness and kurtosis are significantly different between /h/ and /r/, suggesting an absence of neutralization in the data analyzed here. However, examining the difference in means does not entirely reveal a possible overlap with regards to the values of such acoustic continuous measurements. To better inspect the data analyzed, the distribution of center of gravity (Figure 74), skewness (Figure 75) and kurtosis (Figure 76) are presented next.

Figure 74: Center of gravity distribution of phonemic /r/ and /h/ in Holyoke and Puerto Rico
Figure 75: Skewness distribution of phonemic /r/ and /h/ in Holyoke and Puerto Rico
As seen in Figures 74, 75 and 76, there is an overlap in the values of center of gravity, skewness and kurtosis based on PHONEMIC SOUND. In fact, the distribution for each measurement reveals that any value that the backed /r/ takes in this sample can also be taken by /h/. Therefore, although the difference in means for each acoustic characteristic is different with regards to /h/ and /r/, there is still an overlap of these values. These findings lead us to take a further step and examine the interaction terms for PHONEMIC SOUND with each other linguistic and sociolinguistic variables.

5.2.3.2 Linguistic variables
In the following section, the role of the linguistic factors in relation to center of gravity and skewness are analyzed. In keeping consistent with previous statistical models, STRESS, POSITION, PREVIOUS and FOLLOWING SOUND were included in the model. PHONEMIC SOUND and VOICING are also considered. As mentioned above, in order to examine the differences between /h/ and /r/, interaction terms for PHONEMIC SOUND with each of the other variables included. Models and interactions are compared using ANOVA.

5.2.3.2.1 Results for Holyoke

Center of gravity

Table 42 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable, and with phonemic /h/, unstressed and after consonant as reference baseline levels. Only PHONEMIC SOUND:STRESS, was selected as significant interaction.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>(Intercept) Baseline=/h/, N, C</td>
<td>1726.28</td>
<td>101.19</td>
<td>17.060</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Stress (Y)</td>
<td>358.74</td>
<td>117.31</td>
<td>3.058</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Position (I)</td>
<td>-124.57</td>
<td>362.84</td>
<td>-0.343</td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Position (V)</td>
<td>118.90</td>
<td>171.19</td>
<td>0.695</td>
<td>0.487</td>
</tr>
</tbody>
</table>
The production of /h/ and /r/ in terms of center of gravity is significantly different depending on stress (p<.01), as can be seen in Figure 77. The distance between /h/ and /r/ is greater when these sounds are produced in stressed syllables compared to unstressed syllables.

![Center of gravity](image)

*Figure 77: Boxplots for center of gravity values for PHONEMIC SOUND and STRESS in Holyoke*

*Skewness and Kurtosis*

For the other two acoustic measures, skewness and kurtosis, none of the interactions are significant, as can be seen in Tables 43 and 44.
Table 43: Summary of the linear mixed-effect regression model with skewness as dependent variable and speaker as a random effect for the linguistic variables among the Massachusetts data. Only interactions are included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept) Baseline=/h/, N, C</td>
<td>6.4396</td>
<td>.6045</td>
<td>10.652</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Stress (Y)</td>
<td>-.6208</td>
<td>.8210</td>
<td>-.756</td>
<td>.449</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Position (I)</td>
<td>.5385</td>
<td>2.5398</td>
<td>-.212</td>
<td>.832</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Position (V)</td>
<td>-.4423</td>
<td>1.1989</td>
<td>-.369</td>
<td>.712</td>
</tr>
</tbody>
</table>

Table 44: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and speaker as a random effect for the linguistic variables among the Massachusetts data. Only interactions are included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept) Baseline=/h/, N, C</td>
<td>113.835</td>
<td>25.610</td>
<td>4.445</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Stress (Y)</td>
<td>2.443</td>
<td>37.696</td>
<td>.065</td>
<td>.948</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Position (I)</td>
<td>67.218</td>
<td>116.636</td>
<td>.576</td>
<td>.564</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Position (V)</td>
<td>-12.904</td>
<td>55.095</td>
<td>-.234</td>
<td>.814</td>
</tr>
</tbody>
</table>

5.2.3.2.2 Results for Puerto Rico

Center of gravity
The best fit model when center of gravity is selected as the dependent variable include STRESS, PREVIOUS and FOLLOWING SOUND as fixed factors. Table 45 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable, and with phonemic /h/, unstressed, following sound /a/ and previous sound /e/-/i/ as reference baseline levels. Only PHONEMIC SOUND:FOLLOWING SOUND is selected as significant interaction.

Table 45: Summary of the linear mixed-effect regression model with center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Interactions are only included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>(Intercept) Baseline=/h/, N, /a/, /e/-/i/</td>
<td>1747.90</td>
<td>99.68</td>
<td>13.52</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Stress (Y)</td>
<td>-107.74</td>
<td>98.07</td>
<td>-1.09</td>
<td>.272</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Following (/e/-/i/)</td>
<td>-15.35</td>
<td>129.89</td>
<td>-.118</td>
<td>.905</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Following (/o/-/u/)</td>
<td>-320.06</td>
<td>138.35</td>
<td>-2.313</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (/a/)</td>
<td>8.008</td>
<td>121.58</td>
<td>.066</td>
<td>.947</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (0)</td>
<td>-124.67</td>
<td>187.75</td>
<td>.664</td>
<td>0.506</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (c)</td>
<td>-30.802</td>
<td>414.303</td>
<td>-.218</td>
<td>.827</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (/o/-/u/)</td>
<td>-181.621</td>
<td>164.630</td>
<td>-1.103</td>
<td>.270</td>
</tr>
</tbody>
</table>

The interaction between PHONEMIC SOUND and FOLLOWING SOUND is significant (p<.05): the production of phonemic /h/ and /r/ in terms of center of gravity is
significantly different depending on FOLLOWING SOUND. Figure 78 shows how the center of gravity values for /h/ and /r/ are lower when the following sound is a backed vowel in comparison to central or fronted vowels. This result can be explained in terms of coarticulation effects—back vowels lead to more backed realizations of /h/ and /r/. The model reveals that the effect of the following vowel is larger in /r/ than in /h/ when these sounds are followed by either /o/ or /u/. Moreover, Figure 78 demonstrates that the distance between /h/ and /r/ is smaller when the following sound is /o,u/ than when it is a central vowel, suggesting that /h/ and /r/ are produced in more similar ways when they are followed by backed vowels, and are less similar when they are followed by non-back vowels.

Figure 78: Boxplots for center of gravity values for PHONEMIC SOUND and FOLLOWING SOUND in Puerto Rico

Skewness and Kurtosis
For the other two acoustic measures, skewness and kurtosis, none of the interactions are significant, as can be seen in Tables 46 and 47.

Table 46: Summary of the linear mixed-effect regression model with skewness as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Interactions are only included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept) Baseline=/h/, /a/, /e/-/i/</td>
<td>7.080</td>
<td>.713</td>
<td>9.926</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Following (/e/-/i/)</td>
<td>1.4727</td>
<td>1.065</td>
<td>1.383</td>
<td>.167</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Following (/o/-/u/)</td>
<td>-0.840</td>
<td>1.095</td>
<td>-0.767</td>
<td>.443</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (/a/)</td>
<td>-1.063</td>
<td>.997</td>
<td>-1.066</td>
<td>.286</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (0)</td>
<td>.730</td>
<td>1.551</td>
<td>.471</td>
<td>.637</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (c)</td>
<td>-1.139</td>
<td>1.168</td>
<td>-0.975</td>
<td>.329</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (/o/-/u/)</td>
<td>-.244</td>
<td>1.356</td>
<td>-0.180</td>
<td>.857</td>
</tr>
</tbody>
</table>
Table 47: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and Speaker as a random effect for the linguistic variables among the Puerto Rico data. Interactions are only included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept) Baseline=/h/, /a/, /e/-/i/</td>
<td>131.632</td>
<td>31.953</td>
<td>4.120</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Following (/e/-/i/)</td>
<td>29.469</td>
<td>50.354</td>
<td>.585</td>
<td>.558</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Following (/o/-/u/)</td>
<td>-71.842</td>
<td>51.805</td>
<td>-1.387</td>
<td>.165</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (/a/)</td>
<td>-17.355</td>
<td>47.152</td>
<td>-.368</td>
<td>.712</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (/o/-/u/)</td>
<td>20.615</td>
<td>64.151</td>
<td>.321</td>
<td>.748</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (0)</td>
<td>39.96</td>
<td>73.358</td>
<td>.545</td>
<td>.586</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Previous (c)</td>
<td>-16.109</td>
<td>55.256</td>
<td>-.292</td>
<td>.770</td>
</tr>
</tbody>
</table>

5.2.3.2.3 Summary – the effect of linguistic variables on acoustic aspects of /h/ and /r/

This section studied the possible neutralization of phonemic /r/ and /h/ in production. The main goal was to verify whether or not their allophones are classified under one phonological category (/r/ vs. /h/), or whether there is a continuum among backed
realizations which could favor this merge in PRS. Results for the overall data reveal that the means of center of gravity, skewness and kurtosis are significantly different for /h/ and the backed /r/, suggesting a lack of merge in production. However, the difference in means does not necessarily counter a possible overlap in the values of such acoustic measurements. In fact, the plots showing the distribution of center of gravity (Figure 74), skewness (Figure 75) and kurtosis (Figure 76) reveal an overlap in which any value that the backed /r/ takes can also be taken by /h/ but not the other way around.

The next linear regression analyses examine the interaction terms for PHONEMIC SOUND with each other linguistic and sociolinguistic variables in order to better understand the factors affecting the /h/ and /r/ distinction.

With regards to the linguistic variables, results for Holyoke reveal that taking center of gravity as a dependent variable, only PHONEMIC SOUND:STRESS was selected as significant interaction: the distance between /h/ and /r/ is bigger among stressed syllables than among unstressed syllables. Among Puerto Rico data, however, a different interaction was selected: PHONEMIC SOUND:FOLLOWING SOUND, suggesting that the distance between /h/ and /r/ is significantly different when the following sound is /o,u/ than when it is a fronted or central vowel. Therefore, /h/ and /r/ are produced in more analogous ways when they are followed by backed vowels and are less similar when they are followed by central or front vowels.

For the other two acoustic measures, skewness and kurtosis, none of the interactions are significant in any location under study, Puerto Rico or Holyoke. Consequently, they fail to show any statistical difference between phonemic /h/ and phonemic /r/ in production among the current data.
Contrary to our expectations, the effect of POSITION is the same for both phonemic sounds, indicating that the distance between /h/ and /r/ cannot be explained depending on whether the sound under study is produced in initial or intervocalic position. As explained in Chapter 2, Willis et al. (2015) present the distribution of voiceless and voiced /h/ according to position (post-pausal or phrase medial context), showing that in initial position there is a preference for a voiceless realization. In line with Willis, similar findings are found in the present sample (See qualitative section). Moreover, Delgado-Díaz and Galarza (2015) show that there is an effect of phonological context on the possible neutralization between /r/ and /h/ in perception: Puerto Ricans are more accurate when identifying the backed /r/ in intervocalic position than in word initial position. Considering Willis et al.’s results, the authors suggest that this tendency can be explained in terms of allophonic distribution and its relation with voicing: it might be more difficult for Puerto Rican listeners to distinguish /r/ from /h/ in word initial position only, where this context does not provide listeners with cues to the phonemic distinction (i.e. voicing). The current analysis, thus, cannot provide evidence to such claim.

5.2.3.3 Sociolinguistic variables

After assessing the role of the linguistic factors in relation to center of gravity and skewness, sociolinguistic factors were analyzed. The variables considered are ORIGIN, AGE GROUP, SEX, GENERATION, SOCIOECONOMIC STATUS, TASK, TIES WITH THE PR COMMUNITY, TRAVELS TO PR, YEARS IN THE US, PRONUNCIATION CORRECTED, and CONTACT WITH OTHER COMMUNITIES. As mentioned above, in order to examine the
differences between /h/ and /r/, interaction terms for PHONEMIC SOUND with each of the other variables are included. Models and interactions are compared using ANOVA.

5.2.3.3.1 Results for Holyoke

Center of gravity

The best fit model when center of gravity is selected as the dependent variable includes ORIGIN, AGE GROUP, GENERATION, SOCIOECONOMIC STATUS, TASK and TIES WITH THE PR COMMUNITY as fixed effects. Table 48 shows the Estimate, Standard Error, t-values and p-values for the model with center of gravity as the dependent variable, and with phonemic /h/, metropolitan area, older speaker, first-generation, lower middle socioeconomic status, Map task and ties with the PR community once or twice a month as reference baseline levels. Three interactions are significant: PHONEMIC SOUND:ORIGIN, PHONEMIC SOUND:GENERATION, and PHONEMIC SOUND:SOCIOECONOMIC STATUS.
Table 48: Summary of the linear mixed-effect regression model with center of gravity as dependent variable and *Speaker* as a random effect for the sociolinguistic variables among the Holyoke data. Interactions are only included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity</td>
<td>(Intercept) Baseline= M, O, 1, L, M, 2</td>
<td>1219.650</td>
<td>357.92</td>
<td>3.408</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Origin (N)</td>
<td>-723.427</td>
<td>361.759</td>
<td>-2.000</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (M)</td>
<td>118.487</td>
<td>369.993</td>
<td>.320</td>
<td>.7489</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (Y)</td>
<td>689.890</td>
<td>476.164</td>
<td>1.449</td>
<td>.1479</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Generation (2)</td>
<td>-604.115</td>
<td>366.515</td>
<td>-1.648</td>
<td>&lt;.1</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Socioeconomic status (U)</td>
<td>-760.005</td>
<td>378.608</td>
<td>-2.007</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (P)</td>
<td>105.901</td>
<td>146.071</td>
<td>.725</td>
<td>.4687</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (R)</td>
<td>-270.485</td>
<td>278.024</td>
<td>-.973</td>
<td>.3310</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Ties PR community (Once a week)</td>
<td>-422.395</td>
<td>328.826</td>
<td>-1.285</td>
<td>.1994</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Ties PR community (every day)</td>
<td>-311.532</td>
<td>312.955</td>
<td>-0.995</td>
<td>.3199</td>
</tr>
</tbody>
</table>

Results show a significant difference in the interaction PHONEMIC SOUND:ORIGIN (p<.05). The distance between /h/ and /r/ is bigger among speakers from the metropolitan area than from the non-metropolitan area, as can be seen in Figure 79. Hence, findings suggest that speakers from the metropolitan area (San Juan) keep the values for center of gravity between phonemic /r/ and phonemic /h/ more separated than speakers from the
non-metropolitan area, who are inclined to produce those sounds more similarly. That is to say, the potential for neutralization of /r/ and /h/ is greater among non-metropolitan speakers.

![Center of gravity](image)

Figure 79: Boxplots for center of gravity values for PHONEMIC SOUND and ORIGIN in Holyoke

With regards to GENERATION, the model reveals that the difference in production of phonemic /h/ and /r/ for the second-generation speakers is marginally different than the difference between /h/ and /r/ for first-generation speakers (p<.1). Concretely, such difference is bigger among first generation speakers, suggesting that they keep the difference in production between phonemic /h/ and phonemic /r/ more separated in comparison with second-generation speakers (Figure 80).
As for PHONEMIC SOUND:SOCIOECONOMIC STATUS, the difference in the production of /h/ and /r/ for the upper middle class is significantly different than the difference between /h/ and /r/ for lower middle class speakers (p<.05). Specifically, as shown in Figure 81, moving from /h/ to /r/, there are smaller effects for the lower middle class than the upper middle class, suggesting that lower middle-class speakers maintain greater contrast in the production of phonemic /h/ and phonemic /r/. 

Figure 80: Boxplots for center of gravity values for PHONEMIC SOUND and GENERATION in Holyoke
For the other two acoustic measures, skewness and kurtosis, none of the interactions are significant, as can be seen in Tables 49 and 50.
Table 49: Summary of the linear mixed-effect regression model with skewness as dependent variable and Speaker as a random effect for the sociolinguistic variables among the Holyoke data. Interactions are only included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>(Intercept) Baseline= M, O, 1, L, M, 2</td>
<td>8.59</td>
<td>2.158</td>
<td>3.98</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Origin (N)</td>
<td>2.895</td>
<td>2.423</td>
<td>1.195</td>
<td>.232</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (M)</td>
<td>0.558</td>
<td>2.478</td>
<td>.225</td>
<td>.821</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (Y)</td>
<td>-0.571</td>
<td>3.187</td>
<td>-.179</td>
<td>.857</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Generation (2)</td>
<td>0.050</td>
<td>2.454</td>
<td>.021</td>
<td>.983</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Socioeconomic status (U)</td>
<td>1.939</td>
<td>2.534</td>
<td>.765</td>
<td>.444</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (P)</td>
<td>-0.014</td>
<td>0.978</td>
<td>-.015</td>
<td>.988</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (R)</td>
<td>1.100</td>
<td>1.859</td>
<td>.592</td>
<td>.554</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Ties PR community (3)</td>
<td>-2.200</td>
<td>2.201</td>
<td>.999</td>
<td>.317</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Ties PR community (4)</td>
<td>-1.778</td>
<td>2.096</td>
<td>.849</td>
<td>.396</td>
</tr>
</tbody>
</table>
Table 50: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and Speaker as a random effect for the sociolinguistic variables among the Holyoke data. Interactions are only included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept) Baseline= M, O, 1, L, M, 2</td>
<td>134.473</td>
<td>81.361</td>
<td>1.653</td>
<td>.137</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Origin (N)</td>
<td>94.738</td>
<td>109.196</td>
<td>.868</td>
<td>.385</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (M)</td>
<td>-9.745</td>
<td>111.708</td>
<td>-0.087</td>
<td>.930</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (Y)</td>
<td>-26.276</td>
<td>143.483</td>
<td>-0.183</td>
<td>.854</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Generation (2)</td>
<td>-59.445</td>
<td>110.636</td>
<td>-0.537</td>
<td>.591</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Socioeconomic status (U)</td>
<td>68.558</td>
<td>114.007</td>
<td>.601</td>
<td>.547</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (P)</td>
<td>-15.776</td>
<td>44.072</td>
<td>-.358</td>
<td>.720</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (R)</td>
<td>71.843</td>
<td>83.500</td>
<td>.860</td>
<td>.389</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Ties PR community (3)</td>
<td>-8.377</td>
<td>99.179</td>
<td>-.084</td>
<td>.932</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Ties PR community (4)</td>
<td>-52.490</td>
<td>94.469</td>
<td>-.556</td>
<td>.578</td>
</tr>
</tbody>
</table>

5.2.3.3.2 Results for Puerto Rico

Center of gravity

When Center of gravity was selected as the dependent variable, the best-fit model for Puerto Rico included ORIGIN, AGE GROUP, SOCIOECONOMIC STATUS, TASK and SEX as fixed effects. Table 51 shows the Estimate, Standard Error, t-values and p-values for the model.
with center of gravity as the dependent variable, and with phonemic /h/, metropolitan area, older speaker, lower middle socioeconomic status, Reading task and female as reference baseline levels. Two interactions are significant: PHONEMIC SOUND:AGE GROUP and PHONEMIC SOUND:TASK.

Table 51: Summary of the linear mixed-effect regression model with Center of gravity as dependent variable and SPEAKER as a random effect for the linguistic variables among the Puerto Rico data. Only interactions are included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of gravity ( Intercept)</td>
<td>1221.336</td>
<td>171.836</td>
<td>7.108</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Baseline= /h/, M, O, L, R, F</td>
<td>-570.146</td>
<td>431.369</td>
<td>-1.322</td>
<td>.18652</td>
<td></td>
</tr>
<tr>
<td>Phonemic (/r/): Origin (N)</td>
<td>471.637</td>
<td>206.667</td>
<td>2.282</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Phonemic (/r/): Age group (M)</td>
<td>-39.127</td>
<td>133.707</td>
<td>-.293</td>
<td>.76985</td>
<td></td>
</tr>
<tr>
<td>Phonemic (/r/): Age group (Y)</td>
<td>-247.319</td>
<td>185.923</td>
<td>-1.330</td>
<td>.18370</td>
<td></td>
</tr>
<tr>
<td>Phonemic (/r/): Socioeconomic status (U)</td>
<td>-86.153</td>
<td>151.480</td>
<td>-.569</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Phonemic (/r/): Task (P)</td>
<td>129.723</td>
<td>154.797</td>
<td>.838</td>
<td>.402</td>
<td></td>
</tr>
<tr>
<td>Phonemic (/r/): Task (M)</td>
<td>-4.414</td>
<td>131.882</td>
<td>-.033</td>
<td>.97331</td>
<td></td>
</tr>
</tbody>
</table>

With respect to AGE GROUP, the difference in the production of /h/ and /r/ in terms of center of gravity for the middle-age adults is significantly different than for older speakers (<.05). Specifically, the model reveals that the difference is bigger among middle-aged adults than older adults, suggesting that middle-aged adults maintain greater
separation in the pronunciation of phonemic /h/ and phonemic /r/ than older speakers (Figure 82).

![Center of gravity](image)

**Figure 82:** Boxplots for center of gravity values for PHONEMIC SOUND and GROUP AGE in Puerto Rico

Results also show a significant difference in the interaction PHONEMIC SOUND:TASK ($p<.05$), indicating that in moving from /h/ to /r/, there are bigger effects among speakers performing the Picture Description task than the Reading task as can be seen in Figure 83. Specifically, values for /r/ and /h/ are produced more distinctly when participants are performing the Reading task than when they are involved in the Picture Description task.
Skewness and Kurtosis

For the other two acoustic measures, skewness and kurtosis, none of the interactions are significant, as can be seen in Tables 52 and 53.

84: Boxplots for center of gravity values for PHONEMIC SOUND and TASK in Puerto Rico
Table 52: Summary of the linear mixed-effect regression model with skewness as dependent variable and Speaker as a random effect for the linguistic variables among the Puerto Rico data. Only interactions are included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phonemic (/r/): Origin (N)</td>
<td>2.581</td>
<td>3.329</td>
<td>.775</td>
<td>.438</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (M)</td>
<td>-2.407</td>
<td>1.595</td>
<td>-1.509</td>
<td>.131</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (Y)</td>
<td>-.680</td>
<td>1.034</td>
<td>-.657</td>
<td>.511</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Socioeconomic (U)</td>
<td>1.737</td>
<td>1.431</td>
<td>1.214</td>
<td>.225</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (P)</td>
<td>1.271</td>
<td>.955</td>
<td>1.331</td>
<td>.183</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (R)</td>
<td>1.304</td>
<td>1.230</td>
<td>1.060</td>
<td>.289</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Sex (M)</td>
<td>.216</td>
<td>1.019</td>
<td>708.70</td>
<td>.832</td>
</tr>
</tbody>
</table>
Table 53: Summary of the linear mixed-effect regression model with kurtosis as dependent variable and Speaker as a random effect for the linguistic variables among the Puerto Rico data. Only interactions are included.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>(Intercept) Baseline= /h/, M, O, L, R, F</td>
<td>230.45</td>
<td>70.21</td>
<td>3.282</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Origin (N)</td>
<td>96.80</td>
<td>150.97</td>
<td>.641</td>
<td>.521</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (M)</td>
<td>-78.91</td>
<td>72.34</td>
<td>-1.091</td>
<td>.275</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Age group (Y)</td>
<td>-30.52</td>
<td>46.90</td>
<td>-.651</td>
<td>.515</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Socioeconomic (U)</td>
<td>51.94</td>
<td>64.91</td>
<td>.800</td>
<td>.423</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (P)</td>
<td>44.95</td>
<td>43.30</td>
<td>1.038</td>
<td>.299</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Task (R)</td>
<td>55.97</td>
<td>55.79</td>
<td>1.003</td>
<td>.316</td>
</tr>
<tr>
<td></td>
<td>Phonemic (/r/): Sex (M)</td>
<td>28.09</td>
<td>46.22</td>
<td>.608</td>
<td>.543</td>
</tr>
</tbody>
</table>

5.2.3.3.3 Summary – the effect of sociolinguistic variables on acoustic aspects of /h/ and /r/

Considering center of gravity as the dependent variable, results for Holyoke showed that PHONEMIC SOUND:ORIGIN, PHONEMIC SOUND:GENERATION, and PHONEMIC SOUND:SOCIOECONOMIC STATUS are significant predictors in the distinction between /h/ and /r/. For Puerto Rico, two other interactions were selected: PHONEMIC SOUND:AGE GROUP and PHONEMIC SOUND:TASK.

As for PHONEMIC SOUND:ORIGIN, speakers originally from the non-metropolitan area (Salinas) produce /h/ and /r/ more alike than speakers from the metropolitan area, a
group who might be more aware of the standard used in Puerto Rico, and, therefore, substantially maintain the /h/-/r/ distinction. With regards to PHONEMIC SOUND:GENERATION, second-generation speakers produce more similar center of gravity values for phonemic /h/ and /r/ than first generation speakers. Hence, those PRS speakers who were born on the island and emigrated to Holyoke maintain /h/ and /r/ differences in their production while those speakers who were born or raised in Mainland U.S. do not. Those results suggest that the second-generation group could be less aware of the standard or norm found on the island of Puerto Rico, in line with Ramos-Pellicia (2004) findings, which also show that first generation PRS speakers in Ohio know the negative prestige that lateralizing their /r/ has on the island of Puerto Rico and use other alternatives, such as deletion or a retroflex. With respect to PHONEMIC SOUND:SOCIOECONOMIC STATUS, lower middle-class speakers keep more contrast in the production of phonemic /h/ and phonemic /r/. This is a surprising result considering the large body of sociolinguistic research (Labov, 1966; Trudgill, 1974) demonstrating a relationship between social class and the use of a speech variety. Standard forms of speech, that is to say, those socially accepted or prestigious variants, are more commonly used among the higher socio-economic class while non-standard forms are more frequent among lower socio-economic classes. In line with this claim, we previously mentioned how the frequency of use of the backed /r/ is higher among lower socioeconomic status speakers (Graml, 2009; López Morales, 1983; Matta de Fiol, 1981). However, the current result regarding the /h/-/r/ neutralization does not corroborate the core claims in the previous literature on social class and language change (Labov, 1966; Trudgill, 1974).
On the other hand, findings for Puerto Rico showed that \textsc{Phonemic Sound:Age Group} and \textsc{Phonemic Sound:Task} are significant in regard to center of gravity. Specifically, the model reveals that the difference in production between /h/ and /r/ is bigger among middle-aged adults than older adults. Those findings are not in consonance with Delgado-Díaz and Galarza (2016), who found in perception that elderly male listeners do not neutralize the backed /r/ with /h/ (Graml, 2009). It is possible interpreting the current results, thus, as a consequence of age-grading and marketplace pressure among the middle-aged adults in Puerto Rico (Wagner, 2012). Conversely, \textsc{Sex} does not have an effect on the possible neutralization in the current study. Finally, with regards to \textsc{Phonemic Sound:Task}, the absolute difference between /r/ and /h/ is greater in the Reading Task than in the Picture Description task, suggesting that when speakers are involved in more spontaneous tasks, the differences in production between /h/ and /r/ become more alike. These findings provide more evidence to the sociolinguistic research regarding this phenomenon, implying that in formal contexts, speakers produce standard realizations more frequently; whereas familiar/informal situations favor the less prestigious realizations (Dillard, 1962; Graml, 2009).

In consonance with the linguistic analyses, for the other two acoustic measures, skewness and kurtosis, none of the interactions are significant in any location under study, Puerto Rico or Holyoke. Consequently, they fail to show any statistical difference between phonemic /h/ and phonemic /r/ in production among the current data.
CHAPTER 6
CONCLUSIONS

6.1 Summary of the results

Rhotic variation has long been of interest to linguists both in Puerto Rico (Graml, 2009; Medina-Rivera, 1997; Navarro Tomás, 1948) and in the diaspora (Ramos-Pellicia, 2004, 2007; Valentín-Márquez, 2007) but non-impressionistic research on the phonetic and phonological variation of Puerto Rican Spanish has not focused on PRS of Western Massachusetts. Combining auditory and acoustic analysis, this research project is unique in that it studies acoustic data on onset /r/ in an area of the United States which is experiencing a rapid growth in its Puerto Rican population. With the largest per capita population of Puerto Ricans living outside the island (U.S. Census, 2010), Holyoke, MA is ideal for this research.

The first research question asks whether there is trill variation in the Puerto Rican community in Western Massachusetts and, if so, whether it mirrors the variation found on the island of Puerto Rico. It also asks whether the backed /r/ is among the allophones produced. In line with previous studies on PRS (Graml, 2009; Hammond, 2000; Lamboy, 2004; Navarro Tomás, 1948; Valentín-Márquez, 2007; Vaquero & Quilis, 1989) and confirming hypothesis H1.1., results corroborate that there is such variation among the diasporic community of Holyoke, MA. Up to eight different variants are reported: trill, approximant trill, backed trill, pre-aspirated trill, pre-aspirated tap, tap, backed trill followed by a tap and a post-tap and trill frication. To our knowledge, these realizations have all been previously reported in PRS (Graml, 2009; Navarro Tomás, 1948; Valentín-
Márquez, 2007) with the exception of the miscellaneous category, which is found in Mexican Spanish of the Chicago diaspora (Henriksen, 2014). Confirming hypothesis H1.2, the same eight phonetic realizations are found on the island of Puerto Rico. Equivalently, Valentín-Márquez (2007) finds a similar overall distribution of rhotic variants in Michigan and Puerto Rico. In fact, not only does he find the same variants, but also the frequency distribution of these variants show almost no differences related to participant’s location. As reported by Valentín-Márquez (2007) and Graml (2009), the most common realization is the normative trill, which is followed by the approximant trill. This result is not surprising since, as mentioned in Chapter 5, the latter realization is considered a weaker version of the normative trill. More importantly, the variable of interest in this study, the backed /r/, occupies the third place in terms of frequency in Massachusetts and in Puerto Rico, confirming hypothesis H1.3 and in line with previous studies on PRS, not only in Puerto Rico but in the Puerto Rican diaspora (Graml, 2009; Lamboy, 2004; Valentín-Márquez, 2007). As mentioned previously, the difference in its frequency distribution between Puerto Rico and Massachusetts is minimal, behaving at this early stage of the study virtually identically in both settings under study.

The second research question asks whether the predictors (sociolinguistic and linguistic) for the use of the backed /r/ are similar for the two settings under study. When the overall data are included in the model, PLACE is not selected as a significant predictor, implying that there are not major differences in the linguistic and sociolinguistic factors affecting the production of the backed /r/ between Massachusetts and Puerto Rico. This result corresponds with the findings from the qualitative analysis summarized above as well as with previous research on comparative analyses between Puerto Rico and
mainland U.S. As presented in Chapter 2, Valentín-Márquez (2007) finds the same linguistic patterns in Grand Rapids, MI and Cabo Rojo, PR. Importantly, for the overall data, the model shows that ORIGIN is significant: speakers originally from the non-metropolitan area (Salinas) produce more backed /r/ realizations than speakers from the metropolitan area (San Juan), suggesting that there is regional variation that is maintained in the diaspora. It has been broadly demonstrated that speech indexes a speaker’s geographical identity (Foulkes et al., 2002). Along these lines, as described in Chapter 2, the backed /r/ is associated with rural areas on the island of Puerto Rico (Graml, 2009; López Morales, 1983; Medina-Rivera, 1997; Navarro Tomás, 1948). As Álvarez Nazario (1990, p. 124) states: “Tiene orígenes campesinos la erre velar” (‘the velar /r/ has rural origins’). This claim is further demonstrated among studies on PRS backed /r/, which demonstrate that this stigmatized sound is more produced in non-metropolitan areas than in the metropolitan area (Graml, 2009; López Morales, 1983). San Juan is the capital of Puerto Rico, center of the economic activity on the island, cultural, financial, tourism center and, most importantly, home to the majority of Puerto Rico’s institutions of higher education. Therefore, it is not surprising that in Puerto Rico the standard or prestigious form of the language, in this case, the normative trill is produced more frequently among speakers originally from San Juan while the backed /r/ is more frequent among speakers from the non-metropolitan area, Salinas. The circular movement between Puerto Rico and Massachusetts and recent technological advances in telecommunications along with the use of social media have facilitated the interaction and interference between those phonetic systems (Foulkes et al., 2010). In this fashion, the Puerto Rican community in Holyoke, MA maintain the same ORIGIN patterns as on the island.
With regard to the linguistic variables affecting the use of the backed /r/, the mixed effects logistic regression analyses reveal that POSITION is a significant predictor in Western Massachusetts and in Puerto Rico. Confirming our hypothesis, the model reveals similar patterns in both locations: the backed /r/ is more likely to be produced in intervocalic position than in initial position, in consonance with Graml (2009), Hammond (1991), Lopéz Morales (1983), and Terrel (1980). Against our hypothesis, neither PREVIOUS nor FOLLOWING SOUNDS are significant, suggesting that coarticulation effects do not affect the production of the backed /r/. Moreover, counter to our expectations, STRESS was not significant, which does not confirm previous research (Graml, 2009; Valentín-Márquez, 2007).

As for the sociolinguistic variables, findings reveal that ORIGIN, AGE GROUP, GENERATION, SOCIOECONOMIC STATUS, TASK, TIES WITH THE PUERTO RICAN COMMUNITY, TRAVELS TO PUERTO RICO, YEARS IN THE U.S. are significant in Holyoke while only ORIGIN, AGE GROUP, SEX, SOCIOECONOMIC STATUS, and TASK are significant in Puerto Rico. ORIGIN is significant in both analyses, revealing the same pattern in both locations. This is an unsurprising result since it was also identified as significant in the analysis for the overall data reported above.

AGE GROUP is a significant predictor in both locations. However, the model shows contradictory results depending on whether the speaker is from Holyoke or Puerto Rico. In Holyoke, middle-aged adults and young adults produce more backed /r/ than older speakers while in Puerto Rico the older adults are the age group who more frequently use this stigmatized sound. Regarding the older and middle-aged speakers, the former is the age group using more backed /r/ realizations in Puerto Rico (partially confirming our
hypothesis as well as previous research done on the island of Puerto Rico) while the middle-aged group is selected for Holyoke. Consequently, it is possible that those middle-aged speakers in the diaspora are the descendants of the older speakers on the island, who are reproducing the same pattern in Holyoke. An alternative explanation, which could be also applied to the younger speakers in Holyoke is that, since most of them have college degrees or some college instruction, they might use this stigmatized sound as an identity marker in order to sound more Puerto Rican (Lamboy, 2004). As previous research on diasporic communities states, the diaspora “reinforces the symbolic ties with their motherland, to which diasporic groups pledge their loyalty” (Gubitosi and De Oliveira, in press). Therefore, the appearance of this salient feature among younger and middle-age speakers in Holyoke may be related to identity factors: a community such as Holyoke is precisely the environment where a quintessentially Puerto Rican vernacular feature might flourish.

In relation to SOCIOECONOMIC STATUS, the lower middle class produce more backed /r/ realizations than the upper middle class in both locations. These findings corroborate the hypothesis as well as previous sociolinguistic research (Graml, 2009; Labov, 1972; Matta de Fiol, 1981; Valentín-Márquez, 2007). In addition to indexing rural origin, the backed /r/ seem to also be associated with lower socioeconomic status, a socioindexical meaning that the backed /r/ carries on the island of Puerto Rico and that could be maintained in Holyoke. In order to verify such socioindexical meanings further perception studies should be conducted in Holyoke. This could be tested with a matched guise task, a sociolinguistic tool that helps to better understand perceived social meaning of linguistic variables.
Supporting the hypothesis with regard to TASK, the same pattern is found in both settings under study: the production of backed /r/ realizations is more frequent when speakers are involved in more spontaneous tasks (Picture Description task and Map task) than in more controlled tasks (Reading task). The Picture Description task and the Map Task demand less attention to language use on the part of the participant, allowing more vernacular variants, hence, the production of the backed /r/. Those results support findings from previous sociolinguistic research (Dillard, 1962; Graml, 2009; Guitart, 1981; Medina-Rivera, 1997).

The model also reveals other significant variables specific to one area or the other. Among PRS speakers in Holyoke, GENERATION, TIES WITH THE PUERTO RICAN COMMUNITY and TRAVELS TO THE U.S. are significant. For the Puerto Rico data, SEX is significant.

With regards to GENERATION, second-generation speakers are more likely to use the backed /r/ than first generation speakers. These results confirm the hypothesis concerning GENERATION, as well as Ramos-Pellicia’s (2004) findings, who found an increasing rate of non-normative trills in her younger generations in Lorain, OH. As mentioned before, those findings might suggest that first-generation speakers are aware of the socioindexical meaning that the backed /r/ carries on the island, avoiding its use in the diaspora (Ramos-Pellicia, 2004). An alternative and more reliable explanation since this study does not consider perceptual analyses is that the covert prestige that the backed /r/ provides is much greater in a Puerto Rican ethnic enclave within the continental United States. Therefore, second-generation speakers might use the backed /r/, a quintessentially PRS feature, as an index to this Puerto Rican identity, that is to say, to
sound more Puerto Rican in this specific diasporic setting. This identity-based interpretation falls into the theoretical framework of the sociolinguistics of globalization, as will be explained in more detail at the end of the section.

As explained in the methodology (Chapter 4), CONTACT WITHIN THE PUERTO RICAN COMMUNITY in Holyoke involves two different variables: TIES WITH THE PUERTO RICAN COMMUNITY and TRAVELS TO THE U.S. Interestingly, both predictors were significant in the statistical model, suggesting that Holyoke speakers’ networks and community integration affect the use of the backed /r/: when the ties with the Puerto Rican community are closer the use of a backed /r/ is more likely. Similarly, when the participant frequently travels back to Puerto Rico, the probability of producing a backed /r/ increases. This result could be explained considering the va y ven (back and forth) or circular migratory movement between Puerto Rico and Massachusetts. The close ties with the island are affecting the frequency of the backed /r/ use, giving more evidence in favor of an identity-based usage of backed /r/ in a demographically special location such as Holyoke. Those findings corroborate in some way the importance of analyzing participants’ responses related to their contact with the Puerto Rican community in order to address the possible role of language maintenance and change affecting the rhotic variation in diasporic communities.

In summation, results from the auditory analyses in which categorical variables are considered reveal: (1) there is trill variation among PRS speakers in Puerto Rico and Holyoke, MA, (2) the same realizations are found in both settings under study, (3) the same frequency distribution is reported for both locations, (4) the backed /r/ is the third most common realization, (5) there are similar linguistic and sociolinguistic predictors
that affect the presence of the backed /r/ in Holyoke and Puerto Rico, (6) there are also specific results for each location. Subsequently, these findings partially confirm the main hypothesis: given that back and forth migration waves between Puerto Rico and Massachusetts have been in constant increase since 1950 (Center for Puerto Rican Studies, 2016), many Puerto Rican cultural aspects have been maintained in the Massachusetts diaspora (Rivera, 2001), included linguistic practices. A more detailed discussion of these differences is presented at the end of the current section.

The third research question asks whether it is possible to capture backed variation on a continuum taking into account acoustic measurements that are known to differentiate fricative realizations with respect to place of articulation (Flores, 2016; Jongman et al., 2000). Those measurements are center of gravity, skewness and kurtosis. It also asks whether the linguistic and sociolinguistic factors that predict the appearance of the backed /r/ have the same predictive power on the acoustic characteristics of this fricative sound. A qualitative analysis of the distribution of such measurements reveals that the production of the backed /r/ realizations can fall on a continuum, being more dispersed for center of gravity and skewness than for kurtosis. Moreover, quantitative analyses confirm that such continuum is affected by different linguistic and sociolinguistic factors, in line with previous research on fricatives (Flores, 2016; Jongman et al., 2000).

When data from Holyoke and Puerto Rico are included in the model, findings show, unsurprisingly, that PLACE is not a significant predictor, implying that there are no differences in the center of gravity, skewness and kurtosis values in the production of the backed /r/ between Puerto Rico and Holyoke. Therefore, in consonance with the
segmental analyses, it is safe to claim that the production of backed /r/ realizations with regards to the acoustic characteristics that define them in Puerto Rico is overall maintained in the diasporic setting of Holyoke. Moreover, in line with the segmental results, ORIGIN is a significant factor with respect to center of gravity and skewness, suggesting that the backed /r/ is produced slightly more posterior among speakers originally from the non-metropolitan area (Salinas) than from the metropolitan area (San Juan).

When data is divided in terms of place (Holyoke, Puerto Rico), findings show similarities and differences between those locations with regards to the linguistic and sociolinguistic factors affecting the aforementioned continuum of backedness. With regard to the linguistic variables, contrary to the segmental analyses, PREVIOUS and FOLLOWING SOUND are selected as significant factors in both settings, affecting the values of center of gravity, skewness and kurtosis. The backed /r/ is produced more posterior when it appears before or after a back vowel, partially confirming hypothesis 3.2 and in line with previous research on fricative variation (Flores, 2016).

Regarding sociolinguistic variation, in line with the articulatory analyses, TASK affects center of gravity values in Holyoke and center of gravity and skewness values in Puerto Rico. Not only do speakers produce more backed /r/ in spontaneous speech, but the degree of backedness is affected. When speakers are accomplishing spontaneous tasks, results suggest that the backed /r/ is produced more posterior than when speakers are involved in more controlled tasks. No other common sociolinguistic variables are selected as significant in both settings. Among Holyoke data, in line with the articulatory analyses, GENERATION and TIES WITH THE PUERTO RICAN COMMUNITY affect center of
gravity and skewness values: the backed /r/ is produced more back among second-generation speakers and among those participants who maintain a closer contact with the Puerto Rican community. Similarly, those two groups are the ones who produce more backed /r/ realizations, as reported above. On the other hand, among the Puerto Rico data, AGE GROUP is significant. Backed /r/ realizations are produced slightly more posterior among middle-aged adults and younger speakers than the older adults. The same variable is significant in the analyses examining the presence of the backed /r/, however both, younger and older speakers are the age groups producing this sound more frequently in comparison to middle aged speakers.

In sum, Kurtosis is only affected by FOLLOWING SOUND in Puerto Rico. However, center of gravity and skewness values present significant variance in relation to different linguistic and sociolinguistic variables in both Holyoke and Puerto Rico. Therefore, it is safe to say that the analyses of continuous measurements (especially center of gravity and skewness) allow a description of the variation present in speech, determining that there is a continuum among backed realization and offering a multidimensional acoustic alternative for a more accurate explanatory mechanism in the research of rhotic variation. Findings do not support clear categorical distinctions in terms of place of articulation (maybe [x] or [χ]) among the current data, but rather that these possible allophones are produced on a continuum of backed /r/ realizations that is affected by different linguistic and sociolinguistic variables. Hence, hypothesis 3.1 is confirmed. Corroborating Erker’s (2010) statement, instrumental analyses are necessary to capture differences in speech that otherwise are obscured in segmental analysis. Hypothesis 3.2 is partially confirmed. Although center of gravity and skewness are affected by linguistic and sociolinguistics
factors, they are not exactly the same ones that predict the use of the backed /r/. In fact, AGE GROUP only affects skewness values in Puerto Rico.

Since the analyses presented above suggest that there is a continuum of backed /r/ realizations, such continuum could possibly result in a neutralization between the backed /r/ and the glottal fricative /h/ due to their acoustic or phonological similarity. Along these lines, the fourth research question asked whether there is evidence for /r/ and /h/ neutralization in production among PRS, not only in Puerto Rico but in the diaspora. It also asks whether there are similar linguistic and sociolinguistic factors affecting the distinction of phonemic /r/ and /h/ in the two settings under study. When all phonemic /r/ and /h/ produced in Puerto Rico and Holyoke are considered, findings show that the means of center of gravity, skewness and kurtosis are significantly different between /h/ and backed /r/, suggesting an absence of neutralization in the production data analyzed. However, the examination of the distribution of those continuous measurements reveal an overlap within their values: any value that the backed /r/ takes in the current data can also be taken by /h/. Subsequently, interaction terms for PHONEMIC /r/ OR /h/ with each other linguistic and sociolinguistic variables were examined in order to better understand the variables affecting the /h/ and /r/ distinction.

In relation to the linguistic variables, different interactions are selected as significant depending on place. When center of gravity is the dependent variable, only PHONEMIC SOUND:STRESS is significant for the Holyoke data, suggesting that the distance between /h/ and /r/ is greater among stressed syllables than among unstressed syllables. Previous research show that STRESS is a significant factor in the realization of fricatives
(Henriksen & Harper, 2016). In these lines, focusing on the realization of /s/, Henriksen and Harper (2016) show that unstressed syllables trend toward deletion and aspiration, both signs of lenition. Similarly, it is unsurprising that more variation in the acoustic characteristics of backed /r/ and /h/ that could lead to a neutralization would happen among unstressed syllables, whereas the standard /r/ and /h/ realizations would be more common in stressed syllables.

On the other hand, for Puerto Rico data, a different interaction was found to be significant: PHONEMIC SOUND: FOLLOWING SOUND, revealing that /h/ and backed /r/ are produced in more analogous ways when they are followed by back vowels, and are less similar when they are followed by central or front vowels. As opposed to /r/, /h/ has no specified lingual gesture; therefore, this indicates a more direct coarticulatory effect with the surrounding vocalic environment. Back vowels create a longer anterior chamber in the vocal tract, perhaps contributing to the lower center of gravity for the /h/ in this context.

Regarding the sociolinguistic variables, as with the linguistic ones, different interactions are selected as significant depending on place. Among Holyoke data, PHONEMIC SOUND: ORIGIN, PHONEMIC SOUND: GENERATION, and PHONEMIC SOUND: SOCIOECONOMIC STATUS are significant predictors in the distinction between /h/ and /r/. Speakers originally from the non-metropolitan area not only produce more backed /r/ but the acoustic characteristics of such fricative realizations are more similar to those of /h/ in comparison to speakers from the metropolitan area. Those results indicate that since the overlap in the non-metropolitan people is greater, the potential for neutralization of /r/ and /h/ is also greater. Therefore, if a merger is beginning to take
effect, it is among the non-metropolitan PRS speakers. Results can be explained in the same way as in the analyses evaluating presence vs absence of backed /r/. There is regional variation in which participants from San Juan (metropolitan area) might be more aware of the standard, which in this case is the /h/-/r/ distinction. Moreover, second-generation speakers produce more similar center of gravity values for phonemic /h/ and backed /r/ than first-generation speakers. This might suggest that, like non-metropolitan speakers, second-generation speakers are less aware of the norm found in Puerto Rico, in line with Ramos-Pelliccia (2004) findings. As explain in the results section, against our expectations and previous literature on change in progress (Labov, 1966; Trudgill, 1974), lower middle-class speakers maintain greater contrast the production of phonemic /h/ and the backed /r/. Among Puerto Rico data, results for PHONEMIC SOUND:AGE GROUP and PHONEMIC SOUND:TASK reveal that middle-aged adults produce /h/ and /r/ with a greater/wider distinction compared with the older adults group. Such findings seem to contradict those from the section 5.2.2, where backed /r/ realizations are produced slightly more posterior among younger speakers and middle-aged adults than the older speakers. Consequently, we would expect that middle-aged adults comprise the age group presenting more similarities in the production of /h/ and /r/. This conflict could be resolved considering that each analysis is taking a different dependent variable. In the former analysis, center of gravity is being affected by PHONEMIC SOUND:AGE GROUP, while in the latter, AGE GROUP is affecting skewness. Furthermore, the fact that middle-aged adults produce /h/ and /r/ with a greater distinction could be explained considering age-grading (Wagner, 2012). It is possible that the speakers of the community under study change their linguistic behavior throughout their lifetimes due to marketplace
pressures. That is to say, middle-age speakers use more conservative speech patterns that they had previously used as young adults or that they will use as older speakers due to the pressures of the socially prestigious variety of Spanish required in their lifetime as active and successful workers. Finally, when speakers are involved in more spontaneous tasks, the differences in production between /h/ and /r/ become more alike. This is an expected finding, given that variable mergers are less likely in the most formal reading tasks. Results from the three analyses run can be explained in a similar way: the production of backed /r/ realizations is more frequent (5.2.1), the backed /r/ is produced most posterior (5.2.2) and the differences between /h/ and backed /r/ become more alike in production (5.2.3) when speakers are involved in more spontaneous tasks (Picture Description task) than in more controlled tasks (Reading task). Those results corroborate findings from previous sociolinguistic research (Dillard, 1962; Graml, 2009; Guitart, 1981; Medina-Rivera, 1997).

In sum, the means of center of gravity, skewness and kurtosis are significantly different between /h/ and /r/ in both locations under study, suggesting an absence of neutralization in Holyoke and Puerto Rico. However, the distribution for each measurement reveals that there is an overlap between /h/ and /r/. Despite not being able to draw a solid conclusion about this matter, we can say that a similar pattern is found in Puerto Rico and in the diasporic setting of Holyoke, likely resulting from the circular movement between these two locations. Hence, hypothesis 4.1 is partially confirmed. The following analyses examining the interaction terms for PHONEMIC /r/ and /h/ reveal that there are linguistic and sociolinguistic variables affecting /h/-/r/ distinction. Against our expectations, and previous research (Delgado-Díaz & Galarza, 2015; Willis et al., 2015),
POSITION has not effect on the overlap between place of articulation in both PRS varieties under study: the distance between /h/ and /r/ cannot be explained by whether the sound under study is produced in initial or intervocalic position. Delgado-Díaz and Galarza (2016) find that the perception of backed /r/ and /h/ is influenced by speaker’s sex and age: older male listeners do not neutralize the backed /r/ with /h/. Similarly, we expected that AGE and SEX would have an effect on the possible neutralization in production. The current analysis cannot provide evidence for SEX. However, AGE GROUP, does have an effect on skewness: middle-aged adults produce /h/ and /r/ with greater distinction than the older adults group. Furthermore, considering the effect of speaker’s generation in Ramos-Pellicia’s research (2004) on PRS speakers in Lorain (Ohio), where the preference of the normative trill declines across the three generations, speaker’s GENERATION was expected to have a principal role in this neutralization in Western Massachusetts. Although GENERATION does not reach a significance of 0.5, the model shows some tendencies, revealing the same pattern as in Pellicia’s research is found among the current data. Therefore, considering results for POSITION, GENERATION, AGE and SEX, hypothesis 4.2 is partially confirmed.

To conclude, this dissertation reveals that, overall, the Puerto Rican community in Holyoke maintains the same linguistic practices as those on the island of Puerto Rico with regard to the production of phonemic /r/: (1) the same realizations are found, (2) the same frequency distribution is reported, (3) there are similar linguistic and sociolinguistic predictors that affect the presence of the backed /r/ in Holyoke and Puerto Rico, (4) there are not differences in the center of gravity, skewness and kurtosis mean of backed /r/
realizations, (5) the backed /r/ is produced more posterior among speakers originally from non-metropolitan areas than among speakers from metropolitan areas, (6) /h/ is produced more backed than phonemic /r/ with regard to center of gravity, skewness and kurtosis means but (7) there is an overlap in those acoustic measurement values looking at the distribution of the data.

Considering the sociolinguistics of globalization, the mobility of communities also implies the mobility of sociolinguistic and linguistic resource (Blommaert & Dong, 2007, p. 4). As explained in Chapter 1, such mobility can imply different dimensions that collapse in specific spaces where speech communities live and interact with one another (Blommaert & Dong, 2007). As a consequence, this process of mobility allows different outputs: the mechanisms of assimilation and acculturation as well as language change, hybridizations or even creolization (Gubitosi and De Oliveira, in press). Previous diaspora research explains that such minority diasporic communities are threatened by the majority population of their new country where they settle (Ladilova, 2015), allowing the assimilation and acculturation (Gubitosi and De Oliveira, in press). However, what is encountered in Holyoke is that the Puerto Rican community tries to maintain their language, one of the most noticeable signs of immigrants’ origin, to strengthen authenticity (Coupland, 2003) in the same way that they keep other Puerto Rican cultural experiences, such as mofongo, salsa, the Puerto Rican parade or other Puerto Rican manifestations (as shown in Chapter 3). This situation is different with respect to other diasporic communities. For instance, Nigerians are a minority community in Berchem, Belgium whose language is invisible in the public space. In fact, Nigerian families choose English-medium channels in television, such as BBC World or MTV because
they are trying to assimilate to the majority language (Blommaert & Dong, 2007). On the contrary, although English is necessary for success in the mainstream community, the Puerto Rican community in Holyoke has access to Latino radio stations and cable television, such as Telemundo and Univision, which broadcast programs in Spanish. Spanish is prevalent throughout the community, found on many signs in bars, churches, stores, markets, institutions, and organizations. PRS can be heard throughout the different Puerto Rican neighborhoods in Holyoke, being the most common language heard in the streets. In doing so, they avoid assimilation to the American dominant culture, in consonance with the Portuguese community analyzed by Gubitosi and De Oliveira in Eastern Massachusetts (in press). The importance of language as a marker of identity has been previously established (Coupland, 2003; Gubitosi and De Oliveira, in press); Gubitosi and De Oliveira stress this importance arguing that “language attaches a value of authenticity when it indexes authentic cultural membership” (in press, p. 26).

Therefore, by using the “authentic” language, the diasporic community reinforces their Puerto Rican identity. Puerto Rican Spanish speakers in Holyoke, hence, have rebuilt their community in the diaspora in a cohesive way, where their identity has been negotiated and established since 1950 by maintaining their traditions, as well as their language. Similarly, the Portuguese diaspora in Massachusetts mapped their linguistic and social landscape, developing a new Azores in the United States and becoming “a replica of the island of Saint Michael in the Azorean archipelago” (Gubitosi and De Oliveira, in press, p. 29). In this regard, the concept of mobility takes crucial importance at reterritorializing and deterritorializing such diasporic experiences (Gubitosi and De Oliveira, in press; Rosa, 2015). Those Portuguese communities do not show nostalgia for
being far away from their homeland. Contrarily, as the researchers explain, they have created an idea or sense of community that imitates what their ancestors (among younger generations) or they (first generation) have lived in their home country. As a consequence, following these processes of deterritorialisation and reterritorialisation (Rosa, 2015), such “diasporic communities re-create the new spaces reinforcing migrants’ identities and remapping the linguistic and social landscape of their new home” (in press, p. 29). Like the Portuguese community in Massachusetts or the Puerto Rican and Mexican communities in Chicago (Rosa, 2015), the Puerto Rican diaspora in Holyoke has also rebuilt a new sense of what being Puerto Rican conveys. Maintaining close ties with the Puerto Rican community, traveling back and forth to the island of Puerto Rico and recreating what they or their parents have learnt or experimented on the island, the Puerto Rican diasporic community of Holyoke have restructured and re-territorialized their new homeland (Holyoke) to identify it with the motherland (Puerto Rico). It is the new Puerto Rican imagined community (Anderson, 1991) built by those people who perceive themselves as part of the Puerto Rican group. Throughout this process of Puerto Rican replication and reinterpretation of the Puerto Rican identity, thus, the ties with the island of Puerto Rico get reinforced. Taking into account these reterritorialization practices (Rosa 2015) in which Puerto Rico is remapped as part of Holyoke, it is possible to better understand the Puerto Rican linguistic and cultural characteristics that have been reinforced in the diaspora. Furthermore, as Rosa (2015) claims among the Puerto Rican and Mexican diaspora in Chicago, such reterritorialization processes challenge even more the geographical borders between Puerto Rico and Mainland U.S. All Puerto Rican symbols that can be seen in Holyoke,
such as the Puerto Rican flag, in conjunction with the neighborhood segregation contributes to the remapping of national borders. Public art murals like the one presented in Figure 22 (page 79): “Holyoke, Isla del Encanto”, it is a Holyoke based Puerto Rican emblem that conveys this Puerto Rican identity “across local, national, and international scales” (Rosa, 2015, p. 34). Such design is not simply a Puerto Rican symbol, but a Holyoke-based representation of Puerto Ricanness. Therefore, the territorial displacement that usually characterizes diasporic communities are challenged when members of the community reconstruct their new homes as part of their motherland. Rosa (2015) does a further step, claiming that “by reframing spatial segregation, these forms of deterritorialization and reterritorialization counteract forces of internal colonialism” (2015, p. 39). That is to say, since diasporic areas have been struggling in order to resist some kind of exclusion (in terms of race, space or class), such communities create diasporic identities through the reterritorialization of the new land as part of their motherlands. That way, they are responding to those forms of exclusion that they face as minorities in the diaspora.

“By laying claim to parts of the city in which they dwell in masse, generations of Puerto Ricans and Mexicans valorize their national identities and the Chicago-based territories to which they are understood to correspond. Such diasporic imaginaries demonstrate students’ engagement with competing ideas about their identities and unsettle straightforward narratives of assimilation and transnationalism” (Rosa, 2015: 39).
Similarly, it might be possible to apply this idea to the diasporic community of Holyoke, where Puerto Ricans have located in specific geographical sections of Holyoke, being the South Side the area with the highest concentration of Puerto Ricans. It might be the case that by building this new and strong sense of Puerto Rican community, Puerto Ricans in Holyoke are facing some sort of race, space or class exclusion. Nonetheless, further studies are required in order to verify this hypothesis.

Aside from the similarities in rhotic variation between Puerto Rico and Holyoke, the three analyses also reveal some discrepancies between such geographical locations: although not many, different linguistic and sociolinguistic variables affect (1) the use of the backed /r/, (2) center of gravity and skewness values in the continuum of backedness, and (3) phonemic /h/-/r/ distinction. Although Holyoke maintains a close bond with Puerto Rico due to the back and forth migration waves, diasporas are still complex and heterogeneous (Feld & Basso, 1996), they are changing communities (Canagarajah & Silberstein, 2012) which create sites of super-diversity, with different patterns as a result of these new dialect contact situations (Vertovec, 2007). In this regard, the concept of mobility takes again crucial importance at reterritorializing and deterritorializing such diasporic experiences (Gubitosi and De Oliveira, in press; Rosa, 2015). Puerto Ricans have developed a replicated image of the island of Puerto Rico in Holyoke, resisting assimilation and acculturation and maintaining their language and traditions in the diaspora. However, that image or interpretation of what being Puerto Rican conveys does not entirely have to correspond to the patterns found on the island of Puerto Rico, being able to find different linguistic and sociolinguistic patterns in such settings. As an
example, De Oliveira (2016) studies the diasporic Azorean community in Rhode Island and Massachusetts. Her findings show that Portuguese migrants maintain their language. However, there are also innovative behaviors: they use features they perceive as more prestigious, such as replacing the form of gerund periphrasis in the Azores with the gerund from Lisbon.

Interestingly, among the current data, GENERATION has a significant effect on the three analyses run: first and second generation do not follow the same pattern in Holyoke. Not only second generation speakers produce significantly more backed /r/ realizations, but their degree of backedness in the production of such sound is more posterior than first generation speakers. Moreover, they are the generational group that less maintain backed /r/ and /h/ differences in their production, suggesting that a potential neutralization would start among second generation speakers. As explained before, it might be possible to interpret those results suggesting that second generation speakers are less aware of the standard or norm found on the island of Puerto Rico. However, it is also feasible to explain such finding taking into account the processes of reterritorializing and deterritorializing (Gubitosi and De Oliveira, in press; Rosa, 2015). As mentioned before, those people who perceive themselves as part of the Puerto Rican group build their own sense of what a Puerto Rican community conveys. Although second generation speakers were born and raised primarily within the Mainland U.S., they still identify strongly as Puerto Rican. Since the backed /r/ is a salient feature of PRS, and therefore, it indexes Puerto Ricanness, second generation speakers might use this identity marker to sound more Puerto Rican (Lamboy, 2004). It is what they have heard on Puerto Rico if they had the opportunity of traveling there or their sense of what their parents or other Puerto
Ricans speak. Similarly, students that were born and raised in the Chicago diaspora, identify themselves with being Mexican, Puerto Rican, or simply Hispanic/Latino, having or not traveled to those original countries throughout their lives (Rosa, 2015). There is a “reworking of geographical borders, in which parts of Chicago become linked to Puerto Rico and Mexico” in the same way that parts of Holyoke become linked to Puerto Rico (2015, p. 39).

To conclude this section, the present results indicate that there are substantial similarities in rhotic variation as well as some variation between Puerto Rico and Holyoke; similar patterns are present in other diasporic areas (De Oliveira, 2016; Ramos-Pellicia, 2004). Ramos-Pellicia (2004) concludes that the PRS community in Lorain maintains many phonological features from the island of Puerto Rico although some variation is found. She states the PRS in Lorain is a continuation of PRS in Puerto Rico “that exhibits some variation in its phonology using more alternatives that are available for (r) and (b) due to the influence of A[merican]E[nglish]. There is no change in LPRS” (2004, p. 173). In her study, the little variation found is explained considering the pressures from Mexican American Spanish and American English. In contrast, no retroflex realizations are found in the current data that could motivate such pressure from English. Similarly, the Latino community in Holyoke is mainly Puerto Rican and, as shown in the statistical analysis, the contact with other communities does not have a significant impact on the use of the backed /r/. Along these lines, the following citation by Blommaert and Dong (2007) is particularly significant, reminding us that language “belongs to a particular environment” but it is also “something translocal”: 238
“Language is traditionally seen as something that anchors people in a local context: it is described as something that belongs to a particular environment, is locked into local meanings and interactional dynamics. This insight is too important to be dismissed, and research on it has yielded important results. But it is a partial view, for language is also something translocal, it moves along with people across space and time, and it is being deployed locally in ways that reveal the translocal histories of the speaker’s resources. Language is not just a tool for the construction of locality, it is also a tool for mobility” (2007, p. 19).

Finally, the continuous back and forth movement between Holyoke and Puerto Rico as well as the continuous contact with the Puerto Rican community through social media or other type of communication allows us to predict that Puerto Rican Spanish will be kept alive in this diasporic setting. The role that the first generation has in the processes of linguistic loss or maintenance in crucial. In consonance with the PR community in Lorain, OH, first-generation speakers in Holyoke use Spanish at home, favoring that the second generation speaks it with them and, ultimately, supporting the maintenance of Spanish in the Latino community (Gonzales and Wherrit, 1990; Ramos-Pellicia, 2004). The opposite pattern, the lack of Spanish use by first-generation speakers leads to a Spanish loss among second and third generations, enduring a shift to English, as happens in Fortuna, California (Rivera-Mills, 2001).
6.2 Intellectual merit

Following the variationist framework (Labov, 1963, 1966; Wolfram, 1969) and incorporating the main theoretical claims from the sociolinguistics of globalization (Blommaert, 2003; Coupland, 2003; Blommaert & Dong, 2007), the present study contributes to the body of research on language use in language contact situations, documenting and analyzing the transmission of the trill realization in two different Puerto Rican communities: in the Western Massachusetts diaspora and on the island of Puerto Rico. Additionally, in light of this variation, it provides greater understanding of the processes that underlie phonemic /r/ and /h/ distinction in production by Puerto Rican Spanish speakers. Previous studies on Spanish sociophonetics have examined rhotic variation on the island (Graml, 2009; Medina-Rivera, 1997; Navarro Tomás, 1948) and in the diaspora (Valentín-Márquez, 2007) but have not shown a detailed acoustic description and comparison of the backed /r/ ([xa.món]) and /h/ ([ha.món]). Hence, this study is the first analysis to extract center of gravity, skewness and kurtosis measurements (Haley et al., 2010) to better understand the points of articulation involved in fricative realizations of /r/ and /h/. Continuous measurements are considered since, following Erker’s (2010) claim, they can reveal more adequately the relationship between the variation found in the acoustic signal and its conditioning factors and to better identify patterns of variation. Consequently, this study adds to the body of research on PRS phonology in the United States. Specifically, it provides substantial contributions on the mapping between phonetic categories and acoustic properties (Jongman et al., 2000) by offering a detail analysis of this mapping for phonemic /r/ and /h/ in Puerto Rican Spanish. Among the current results, kurtosis fails to explain backed /r/ variation as well as the phonemic /r/-/h/
distinction, in line with Jongman et al. (2000), whose results also show that kurtosis does not differentiate /f,v/ from /s,z/. However, center of gravity and skewness successfully offers a multidimensional acoustic alternative for a more accurate explanatory mechanism and richer descriptive adequacy in the analyses of phonemic /r/ and /h/. Therefore, the analyses of those measurements reveal that an acoustic phenomenon could be better understood along a continuum rather than in discrete units typical of segmental analysis.

This dissertation demonstrates that the incorporation of a new fricative realization (backed /r/), affects the PRS phonological system. Although the difference in center of gravity, skewness and kurtosis means between the backed /r/ and /h/ are significantly different, there are still cases of overlap. That is to say, phonemic /h/ and phonemic /r/, which are generally found in phonemic opposition, might fail to completely contrast in some specific environment: in onset position when phonemic /r/ is produced as a backed /r/. However, since the backed /r/ is produced in a 15% among the overall phonemic /r/ realizations among the current data, it is not accurate to talk about a complete /r/-/h/ neutralization. Moreover, this possible merge is affected by different linguistic (STRESS and FOLLOWING SOUND) and sociolinguistic factors (ORIGIN, GENERATION SOCIOECONOMIC STATUS, AGE and TASK). Such potential neutralization that affects the realization of phonemic /h/ and /r/ in specific contexts in a significant way is not an accidental phenomenon that occurs in PRS. Its phonological system undergoes different phonological processes such as the deletion of final /s/ and /n/; alternation of /l/ and /r/ in syllable/word final position; the nasalization of vowels; the simplification of consonant clusters; or the reduction of the front vowel [e] before aspiration in initial position (está >
ehtá, htá or tá) (Example taken from Santoro, 2007, p. 50). Previous literature has shown how there are other examples of phonemic contrast neutralization that affect the Spanish phonological system, as well as assimilation, dissimilation or weakening and deletion processes (Hualde, 2005). In these lines, as described above, in Chilean Spanish the velar fricative /x/ becomes a palatal fricative [ç] before a glide or front vowel (Flores, 2016).

Interestingly, many of the neutralized contrasts regarding the Spanish phonology happen in coda position (Hualde, 2005): neutralization of liquids, plosives, nasals or rhotics. However, the merger under study in the current research happens in the onset of the syllable, position where the other phenomena mentioned maintain a phonological contrast. Phonological processes can also be a consequence of language attitudes and identity, such as the velarized /s/ in Madrileño Spanish, a highly-stigmatized feature that conveys Madrileño identity (Wright, 2017). Similarly, the backed /ɾ/ realization is an identity marker used in Holyoke to express Puerto Ricanness, whose more posterior variants might be neutralized with the glottal phoneme /h/ under specific circumstances implying a linguistic change that affects not only the consonantal system of the PRS phonemic inventory but also the Spanish system in general.

Finally, the study of spoken variation also assists in pedagogical issues. By documenting a variety, in this case PRS, teachers are better equipped to determine educational needs in U.S. communities involving the Latino population, promoting the incorporation of stigmatized realizations in its educational policy to the classroom. Salient features, like the backed /ɾ/ in PRS are the ones that mark a specific variety. As such, they can be stigmatized when associated with groups of less prestige. Many of the
speakers of the current sample mentioned being corrected in school with regards to their /r/ pronunciation. They also apologize for their “incorrect” Spanish. As argued before, Puerto Ricans in Holyoke maintain PRS variety for identity reasons. Since the backed /r/ is a salient Puerto Rican feature, it is used to convey authentic Puerto Rican identity in the diasporic setting. As such, correcting its use in the classroom would result in the negation of the Puerto Rican identity which is transmitted through language. Therefore, this study aims to help teachers and users in general to understand the differences between local pronunciation norms and standard written forms, helping them to recognize the backed /r/ as a legitimate realization in PRS used to transmit their Puerto Ricanness. For this reason, it is crucial to avoid teaching corrections or the stigmatization of its use.

6.3 Limitations of the study and future research

One of the limitations of the present study is the uneven number of participants per place. Twenty-four participants were recorded in Puerto Rico while only 21 were recorded in Holyoke, MA. Such difference was due to the inability to recruit second-generation male speakers in the diaspora. In spite of receiving substantial support from the Puerto Rican community, who helped and introduced the researcher with more potential subjects, not being a member of such community affected the process of data collection. Moreover, future studies should incorporate third-generation speakers. The incorporation of this generational group in the study would bring meaningful responses with regard to Spanish maintenance in the diaspora. As Ramos-Pellicia claims (2004, p. 180) “it is really the third generation who will decide whether Spanish will survive the
AE influence, or if it will succumb to its pressures as many other languages in the United States and around the world have”.

Another substantial limitation is the inability to draw solid conclusions on the sociolinguistic variables affecting phonemic /h/-/r/ distinction. The distance between phonemic /h/ and /r/ is larger among speakers form the metropolitan area in both geographical locations, however GENERATION has a marginal effect on such distinction in Holyoke. Second-generation speakers digress from the pattern used by first generation speakers maintaining phonemic /h/-/r/ distinction less than first-generation speakers. Since second-generation speakers are a very dynamic group, strong claims cannot be generated in explaining the direction of the change. However, as Gimeno Menéndez insists (Gimeno Menéndez, 1995, p. 21), it is important to remember that:

“no toda variabilidad y heterogeneidad en la estructura lingüística envuelven cambio. Es más, no toda variación sincrónica implica un cambio “en curso” (...) Sin embargo, todo cambio lingüístico implica variabilidad y heterogeneidad sincrónica en la comunidad de habla”. (‘not all variability and heterogeneity in the linguistic structure involve change. Moreover, not all synchronic variation implies an "ongoing" change (...) However, any linguistic change implies variability and synchronic heterogeneity in the speech community’).

To improve the present research and obtain stronger implications with regard to the phonemic /h/-/r/ maintenance or merge and the direction of the change, subsequent studies should control for center of gravity, skewness and kurtosis in the vowels surrounding the variable of interest. Furthermore, it might be illuminating to measure the
center of gravity, skewness and kurtosis of the whole dataset (backed /r/ and non-backed 
/r/ realizations) in order to make a more solid case for this type of gradient phonetic 
analysis. Having duration as a dependent variable could also be decisive since it is a 
continuous measurement that reliably establishes phonemic distinctions cross-
linguistically (Erker, 2010; Jongman et al. 2000). Contrary to what has been proposed by 
Delgado-Diaz & Galarza (2015) and Willis et al., (2015), evidence of the interaction 
between PHONEME:POSITION was not found. Perhaps looking at the duration results might 
reveal significant differences between the two phonemes, which might account for the 
perceptual neutralization reported in previous studies. The examination of whether 
duration interacts with the dependent variables analyzed in this dissertation (center of 
gravity, skewness and kurtosis) would provide more light into the degree of 
neutralization that might occur. Moreover, the examination of other diasporic areas with 
different geographic areas of contact with the island of Puerto Rico may enhance the 
current results. Finally, as mentioned previously, analyzing answers on language attitudes 
and language identities with regard to backed /r/ variation could be helpful to better 
interpret the findings of the current study. As Auer, Hinskens, & Kerswill (2005) claim, 
examining the identity factor is crucial for explaining why and for whom varieties 
diverge or converge.

The majority of previous work analyzing the effect of social factors on phonetic 
variation has focused on variation in production. However, as Drager (2010) argues, 
speech perception research is necessary to better understand variation found in 
production, especially in areas such as language change. Not only do speakers use 
phonetic features in socially significant ways, but “they perceive speech differently
depending on trends in their own production, their previous experience with other
dialects, and the social characteristics that they attribute to the speaker” (Drager, 2010, p. 473).

As mentioned in the Methodology section, only data pre-Hurricane Maria are
considered in the present study. Nevertheless, the disruption caused by the Hurricane
could potentially leave effects on language use on and off the island; as such, it would be
interesting to obtain data from post-Hurricane migrants to compare with the current
results. It is likely that those findings would be substantially different. Furthermore,
future studies could also challenge the apparent time hypothesis by studying the speech
of the current sample in some years, in line with Harrington, Palethorpe, & Watson
(2000, 2005) work, which analyze Queen Elizabeth’s vowels during her Christmas
message through the years. That way, if the direction of participants’ language-change-
in-progress varies at different stages of their lives, it may provide evidence that such a
hypothesis “underestimate[s] the speed of change” (Hay & Drager, 2007, p. 91) rather
than negate its legitimacy all together. Therefore, returning to study the present
community a few years down the line would help to better understand the direction of the
change and whether we can talk about a change in progress. Finally, different domains of
the language need to be explored. This dissertation is limited to phonology, but other
areas such as syntax or morphology could be included in order to better examine the PRS
spoken in Holyoke, MA. As Ramos-Pellicia states: “There may not be any changes
taking place in one area of a dialect (e.g. phonology), while there are changes occurring
in others (e.g. morphology, syntax)” (2004, p. 172-173). These studies will allow us to
better understand language variation and change through space and time, providing us with better insight to diasporic communities through language.
APPENDIX A

PICTURE DESCRIPTION TASK

Carretera del ciego

Joyerías de la abuela

Refugio de animales

Avión israelí

Ángeles de la gloria

Pájaros de colores

Juguetes de Pedro

Rosas de la bahía
List of target sentences written on the map (n=18)

Conditions

- /t/
  - Intervocalic position (n=3)
    - Cementerio de perros
    - Carretera del ciego
    - Dos carros
  - Word-initial position (n=3)
    - Rosas de la bahía
    - Rebaño de ovejas
    - Refugio de animales
  - After heterosyllabic consonant (n=3)
    - El rincón de María
    - Avión israelí
    - Casa de Enrique

- /h/-/x/
  - Intervocalic position (n=3)
    - Un ajo
    - Pájaros de colores
    - Baraja de cartas
  - Word-initial position (n=3)
    - Joyas de la abuela
    - Jíbaros
    - Juegos de Pedro
  - After heterosyllabic consonant (n=3)
    - Ángeles de la gloria
    - Bandera de Bélgica
    - Dos jabones
APPENDIX C

READING TASK

Instrucciones/Instructions
Part 1:
• Por favor, lea las siguientes palabras insertadas en la siguiente oración:
• Please, read the following words embedded in the frame sentence:
  o Diga ___________ otra vez.
  
  (example)

• ¡EMPEZAMOS!
• WE ARE GOING TO START!

(list of words and pictures)

Instrucciones/Instructions
Part 2:
• Por favor, lea las siguientes palabras insertadas en la siguiente oración:
• Please, read the following words embedded in the frame sentence:
  o ___________ es lo que digo.
  
  (example)

• ¡EMPEZAMOS!
• WE ARE GOING TO START!

(list of words and pictures)

List of target words (n=36)

Conditions
• /ɾ/
  • Intervocalic position (n=6)
    • Barra
    • Narrar
    • Mirra
    • Virrey
    • Burro
• Turrón

• Word-initial position (n=6)
  • Ratón
  • Ramo
  • Riñón
  • Risa
  • Rubí
  • Ruso

• After heterosyllabic consonant (n=6)
  • Enredo
  • Alrededor
  • Honra
  • Desrizar
  • Enrosar
  • Israel

• /h/-/x/

• Intervocalic position (n=6)
  • Caja
  • Bajar
  • Mijo
  • Fijó
  • Lujo
  • Pujó

• Word-initial position (n=6)
  • Jabón
  • Jamón
  • Gimo
  • Girar
  • Jugar
  • Jugo

• After heterosyllabic consonant (n=6)
  Monja
  Angina
  Fingir
  Cisjordania
  Álgebra
  Nostalgia

• Fillers (n=12)
  • Pisar
• Casar
• Misa
• Musa
• Casa
• Usar
• Saco
• Silbar
• Sumar
• Silla
• Salón
• Subo

Examples for /t/:
Examples for /h/-/x/:
MONJA
Por favor, complete el cuestionario. Si hay preguntas que le resulten incómodas, puede dejarlas en blanco.

PARTICIPANTE #__________

• Información General

Edad: ___________ Sexo: ___________

Lugar de nacimiento: ___________
Lugar en el que creció: ___________
Lugar en el que vivió la mayor parte de su vida: ___________
Número de años viviendo en Holyoke: ___________

Escoja una:
- Nací en Puerto Rico
- Al menos uno de mis padres nació en Puerto Rico
- Al menos uno de mis abuelos nació en Puerto Rico

¿De qué parte de Puerto Rico procede su familia? ___________
¿De dónde son sus padres? ___________

Estudios (rodee una): Escuela elemental Escuela secundaria/instituto Universidad

Trabajo:

Viajes a países de habla hispana:
¿Dónde y por cuánto tiempo? ___________

Incluyéndose a usted, ¿cuántas personas viven en su casa?
- a. una/dos
d. cinco
- b. tres/cuatro
e. seis o más

• Competencia lingüística y contextos de uso

Evalúe su nivel en español e inglés de acuerdo a la siguiente escala (escriba un número al lado de cada competencia):

<table>
<thead>
<tr>
<th>ESPAÑOL</th>
<th>INGLÉS</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 = NATIVO</td>
<td>3 = INTERMEDIO</td>
</tr>
<tr>
<td>5 = CASI NATIVO</td>
<td>2 = BÁSICO</td>
</tr>
<tr>
<td>4 = AVANZADO</td>
<td>1 = INICIAL</td>
</tr>
<tr>
<td>Expresión oral</td>
<td>Expresión oral</td>
</tr>
<tr>
<td>Comprensión lectora</td>
<td>Comprensión lectora</td>
</tr>
<tr>
<td>Expresión escrita</td>
<td>Expresión escrita</td>
</tr>
</tbody>
</table>
¿A qué edad empezó a aprender español?

¿A qué edad empezó a aprender inglés?

Cuando era un NIÑO, ¿qué lengua hablaba en casa?:
- ESPAÑOL
- INGLÉS
- AMBOS

Cuando era un NIÑO, ¿qué lengua hablaba con sus amigos?:
- ESPAÑOL
- INGLÉS
- AMBOS

Conocimiento de otras lenguas además del español o inglés:
- a) Sí. Lenguas: ____________
- b) No

¿Fue o va a la escuela en los Estados Unidos?
- SÍ   NO

Si la respuesta es SÍ… desde los _____ años hasta los ______ años

¿Fue o va a la escuela en un país de habla hispana? (ej. Puerto Rico, México)
- SÍ   NO

Si la respuesta es SÍ… ¿Qué país? _________________ Desde los _____ años hasta los ______ años

¿Alguna vez sus profesores le han corregido su pronunciación? ¿Qué sonidos?

¿Cada cuánto…

<table>
<thead>
<tr>
<th>Cada</th>
<th>2 ó 3 veces</th>
<th>1 vez</th>
<th>1 ó 2 veces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nunca</td>
<td>día</td>
<td>a la semana</td>
<td>a la semana</td>
</tr>
</tbody>
</table>

- a. habla inglés en su casa?
  - ___
  - ___
  - ___
  - ___

- b. habla inglés fuera de casa?
  - ___
  - ___
  - ___
  - ___

- c. habla inglés con:
1. su marido/mujer? ___ ___ ___ ___ ___ ___
2. sus hijos? ___ ___ ___ ___ ___ ___
3. sus padres? ___ ___ ___ ___ ___ ___
4. resto de su familia? ___ ___ ___ ___ ___ ___
5. sus amigos? ___ ___ ___ ___ ___ ___
6. compañeros de clase? ___ ___ ___ ___ ___ ___
7. profesores? ___ ___ ___ ___ ___ ___
8. compañeros de trabajo? ___ ___ ___ ___ ___ ___
9. en su comunidad (mall, c. de comunidad, iglesia)? ___ ___ ___ ___ ___ ___

d. habla español en su casa? ___ ___ ___ ___ ___ ___

e. habla español fuera de casa? ___ ___ ___ ___ ___ ___

f. habla español con:
   1. su marido/mujer? ___ ___ ___ ___ ___ ___
   2. sus hijos? ___ ___ ___ ___ ___ ___
3. sus padres? ___ ___ ___ ___ ___

4. resto de su familia? ___ ___ ___ ___ ___

5. sus amigos? ___ ___ ___ ___ ___

6. compañeros de clase? ___ ___ ___ ___ ___

7. profesores? ___ ___ ___ ___ ___

8. compañeros de trabajo? ___ ___ ___ ___ ___

9. en su comunidad? ___ ___ ___ ___ ___

g. cambia de código entre español e inglés con:

1. su marido/mujer? ___ ___ ___ ___ ___

2. sus hijos? ___ ___ ___ ___ ___

3. sus padres? ___ ___ ___ ___ ___

4. resto de su familia? ___ ___ ___ ___ ___

5. sus amigos? ___ ___ ___ ___ ___

6. en su comunidad? ___ ___ ___ ___ ___

¿Qué lengua prefiere para las redes sociales?
Inglés Español Ambas

- Actitudes lingüísticas/Red de contactos
¿Tiene contacto con otros puertorriqueños Massachusetts?  
Rodee una: SÍ   NO  
Si la respuesta es SÍ... ¿Cada cuánto? Rodee una  
Cada día 1 vez a la semana 1 ó 2 veces al mes 1 vez al año

¿Tiene contacto con otros puertorriqueños en el resto de Estados Unidos continental?  
Rodee una: SÍ   NO  
Si la respuesta es SÍ... ¿de qué ciudades?  
¿Cada cuánto? Rodee una  
Cada día 1 vez a la semana 1 ó 2 veces al mes 1 vez al año

¿Y con otras comunidades hispanas? Ej. dominicanos, ecuatorianos, etc.  
Rodee una: SÍ   NO  
Si la respuesta es SÍ... ¿Cada cuánto? Rodee una  
Cada día 1 vez a la semana 1 ó 2 veces al mes 1 vez al año

¿De dónde son? ¿De qué países/ciudades?

¿Ha estado alguna vez en Puerto Rico?  
Rodee una: SÍ   NO  
Si la respuesta es SÍ... ¿Cada cuánto va?  
¿Cuánto duran las estancias?

¿Le gusta viajar allá?

¿Le gustaría regresar/ vivir allá? ¿Por qué/?Por qué no?  
SÍ   NO

¿Tiene contacto con otros puertorriqueños en Puerto Rico?  
Rodee una: SÍ   NO  
¿Cada cuánto habla con ellos (llamadas/redes sociales, etc.)? Rodee una  
Cada día 1 vez a la semana 1 ó 2 veces al mes 1 vez al año

¿Ve o sigue las noticias, periódicos, redes sociales puertorriqueños?  
Rodee una: SÍ   NO

¿Los puertorriqueños en Holyoke “arrastran la r”? ¿Y usted?  
SÍ   NO   SÍ   NO  
a) ¿Quién la arrastra?
En Holyoke (ej. gente joven/mayor; mujeres/hombres):
En Puerto Rico:
b) ¿Qué le parece esta manera de pronunciar la “rr”?

¿Le gustaría participar en una segunda parte?
Información de contacto:
Please, complete the questionnaire. You may skip any question(s) you are uncomfortable answering.

PARTICIPANT #__________

- Background information
  Age: _________ Sex: _________
  Place of birth: _________
    Place where you grew up:
    Place where you have lived most of your life:
    Number of years living in Holyoke:
  Pick one:
    I was born in Puerto Rico
    At least one of my parents was born in Puerto Rico
    At least one of my grand-parents was born in Puerto Rico

Which part of Puerto Rico is your family from? _________

Where are your parents from? _________

Studies (circle one):
  Elementary School
  High School
  College/University

Job:

Trips to Spanish speaking countries:
  When and for how long _________

Including yourself, how many persons live in your household?
  a. one/two
d. five
  b. three/four
e. six or more

- Language proficiency and domains

Rate your proficiency in Spanish and English according to the scale (write the number next to each skill):

<table>
<thead>
<tr>
<th>6 = NATIVE FLUENCY</th>
<th>3 = INTERMEDIATE FLUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 = NEAR (ALMOST) NATIVE FLUENCY</td>
<td>2 = BASIC FLUENCY</td>
</tr>
<tr>
<td>4 = ADVANCED FLUENCY</td>
<td>1 = BEGINNING</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SPANISH</th>
<th>ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>Speaking</td>
</tr>
</tbody>
</table>
At what age did you begin learning Spanish?
At what age did you begin learning English?

When you were a CHILD, language spoken at home:
SPANISH    ENGLISH    BOTH
When you were a CHILD, language spoken among friends:
SPANISH    ENGLISH    BOTH

Knowledge of languages other than Spanish or English:
a) Yes. Language ________________
b) No

Did/Do you go to school in the United States?
YES   NO
If YES, from age ______ to age ______

Did/Do you go to school in a Spanish-speaking country? (e.g., Puerto Rico, Mexico)
YES   NO
If YES…. What country? ________________ From age ______ to age ______

Have your teachers ever corrected your pronunciation? What sounds?

How often do you…

<table>
<thead>
<tr>
<th></th>
<th>Every</th>
<th>2 or 3</th>
<th>Once</th>
<th>1 or 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>day</td>
<td></td>
<td>a week</td>
<td>a week</td>
<td>a month</td>
</tr>
</tbody>
</table>

a. speak English in your
home?

b. speak English outside
your home?
c. speak English to:
   1. your husband/wife? 
      
      
      
      
   2. your children? 
      
      
      
      
   3. your parents? 
      
      
      
      
   4. your relatives? 
      
      
      
      
   5. your friends? 
      
      
      
      
   6. classmates? 
      
      
      
      
      
      
   7. teachers? 
      
      
      
      
      
      
   8. workmates? 
      
      
      
      
      
      
   9. at your community 
      (mall, community ctr, church)? 
      
      
      
      
      
      
      

   d. speak Spanish in your home? 
      
      
      
      
      
      

   e. speak Spanish outside your home? 
      
      
      
      
      
      
      

   f. speak Spanish to:
      1. your husband/wife? 
         
         
         
         
         
      2. your children? 
         
         
         
         
         
         
         

3. your parents?  ___  ___  ___  ___  ___  ___

4. your relatives?  ___  ___  ___  ___  ___  ___

5. your friends?  ___  ___  ___  ___  ___  ___

6. classmates?  ___  ___  ___  ___  ___  ___

7. teachers?  ___  ___  ___  ___  ___  ___

8. workmates?  ___  ___  ___  ___  ___  ___

9. at your community?  ___  ___  ___  ___  ___  ___

---

g. switch between Spanish and English with:

1. your husband/wife?  ___  ___  ___  ___  ___  ___

2. your children?  ___  ___  ___  ___  ___  ___

3. your parents?  ___  ___  ___  ___  ___  ___

4. your relatives?  ___  ___  ___  ___  ___  ___

5. your friends?  ___  ___  ___  ___  ___  ___

6. at your community?  ___  ___  ___  ___  ___  ___

---

Which language do you prefer for social media?
English  Spanish  Both

• Language attitudes/networking

Do you have contact with other Puerto Ricans in Massachusetts?
Circle one:  YES  NO
If you answer **YES**…. **How often?** Circle one
  - Every day
  - Once a week
  - Once or twice a month
  - Once a year

Do you have contact with other **Puerto Ricans in the rest of Mainland US?**
Circle one: **YES**   **NO**
*If you answer **YES**…. **From what cities?***
  **How often?** Circle one
  - Every day
  - Once a week
  - Once or twice a month
  - Once a year

And with other **Spanish-speaking communities?** E.g. Dominicans, Ecuadorians, etc.
Circle one: **YES**   **NO**
*If you answer **YES**…. **How often?** Circle one*
  - Every day
  - Once a week
  - Once or twice a month
  - Once a year
  Where are they from? What countries/cities?

**Have you ever been to Puerto Rico?**
Circle one: **YES**   **NO**
*If you answer **YES**…. **How often do you go?***
  **How long are the stays?**

  **Do you like traveling there?**

    **Would you like to come back/ live there? Why/Why not?**
      **YES**   **NO**

**Do you have contact with other Puerto Ricans in Puerto Rico?**
Circle one: **YES**   **NO**
*How often do you talk to them (calls/social media, etc.)? Circle one*
  - Every day
  - Once a week
  - Once or twice a month
  - Once a year

**Do you watch/follow Puerto Rican TV, newspapers, Puerto Rican social media?**
Circle one: **YES**   **NO**

Do Puerto Ricans **in Holyoke** “arrastran la r”? And you?
  **YES**   **NO**
  c) Who does it?
  **In Holyoke** (e.g. old/young people; women/men):
  **In Puerto Rico:**
  d) What do you think about this way of pronouncing the “rr”?
Would you like to participate in a second part of this research?
Contact info:
APPENDIX F

PUERTO RICO QUESTIONNAIRE
IN SPANISH

Por favor, complete el cuestionario. Si hay preguntas que le resulten incómodas, puede dejarlas en blanco.

PARTICIPANTE #__________

- Información General
  Edad: ___________ Sexo: ___________
  Lugar de nacimiento: ___________
  Lugar en el que creció:
  Lugar en el que vivió la mayor parte de su vida:
  Número de años viviendo en Salinas/San Juan:

Escoja una:
  Nací en Puerto Rico
  Al menos uno de mis padres nació en Puerto Rico
  Al menos uno de mis abuelos nació en Puerto Rico

¿De qué parte de Puerto Rico procede su familia? ___________
¿De dónde son sus padres? ___________

Estudios (rodee una): Escuela elemental Escuela secundaria/instituto
  Universidad

Trabajo:

Viajes a países de habla hispana:
  ¿Dónde y por cuánto tiempo? ___________

Incluyéndose a usted, ¿cuántas personas viven en su casa?
  a. una/dos
  b. tres/cuatro
  c. cinco
  d. seis o más

- Competencia lingüística y contextos de uso

Evalúe su nivel en español e inglés de acuerdo a la siguiente escala (escriba un número al lado de cada competencia):

6 = NATIVO  3 = INTERMEDIO
5 = CASI NATIVO  2 = BÁSICO
4 = AVANZADO  1 = INICIAL

<table>
<thead>
<tr>
<th>ESPAÑOL</th>
<th>INGLÉS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expresión oral</td>
<td>Expresión oral</td>
</tr>
<tr>
<td>Comprensión lectora</td>
<td>Comprensión lectora</td>
</tr>
<tr>
<td>Expresión escrita</td>
<td>Expresión escrita</td>
</tr>
</tbody>
</table>
¿A qué edad empezó a **aprender español**?
¿A qué edad empezó a **aprender inglés**?

Cuando era un **NIÑO**, ¿qué lengua hablaba en casa?:

<table>
<thead>
<tr>
<th>ESPAÑOL</th>
<th>INGLÉS</th>
<th>AMBOS</th>
</tr>
</thead>
</table>

Cuando era un **NIÑO**, ¿qué lengua hablaba con sus amigos?:

<table>
<thead>
<tr>
<th>ESPAÑOL</th>
<th>INGLÉS</th>
<th>AMBOS</th>
</tr>
</thead>
</table>

Conocimiento de otras **lenguas** además del español o inglés:

a) Sí. Lenguas: __________

b) No

¿Fue o va a la **escuela en Puerto Rico**?

| SÍ | NO |

**Si la respuesta es SÍ**… desde los ______ años hasta los ______ años

¿Fue o va a la **escuela en Estados Unidos continental**?

| SÍ | NO |

**Si la respuesta es SÍ**… Desde los ______ años hasta los ______ años

¿Alguna vez sus profesores le han corregido su **pronunciación**? ¿Qué sonidos?

**¿Cada cuánto…**

<table>
<thead>
<tr>
<th>Nunca</th>
<th>Cada día</th>
<th>2 ó 3 veces</th>
<th>1 vez</th>
<th>1 ó 2 veces</th>
<th>a la semana</th>
<th>al mes</th>
</tr>
</thead>
</table>

a. **habla inglés** en su casa?

____

b. **habla inglés** fuera de casa?

____

c. **habla inglés** con:

1. su marido/mujer?

____
2. sus hijos?   
___   ___   ___   ___   ___

3. sus padres?   
___   ___   ___   ___   ___

4. resto de su familia?   
___   ___   ___   ___   ___

5. sus amigos?   
___   ___   ___   ___   ___

6. compañeros de clase?   
___   ___   ___   ___   ___

7. profesores?   
___   ___   ___   ___   ___

8. compañeros de trabajo?   
___   ___   ___   ___   ___

9. en su comunidad
   (mall, c. de comunidad, iglesia)?   
___   ___   ___   ___   ___

d. **habla español** en su
   casa?   
___   ___   ___   ___   ___

e. **habla español** fuera
   de casa?   
___   ___   ___   ___   ___

f. **habla español** con:
   1. su marido/mujer?   
___   ___   ___   ___   ___

   2. sus hijos?   
___   ___   ___   ___   ___

   3. sus padres?   
___   ___   ___   ___   ___
4. resto de su familia? ___ ___ ___ ___ ___

5. sus amigos? ___ ___ ___ ___ ___

6. compañeros de clase? ___ ___ ___ ___ ___

7. profesores? ___ ___ ___ ___ ___

8. compañeros de trabajo? ___ ___ ___ ___ ___

9. en su comunidad? ___ ___ ___ ___ ___

g. cambios de código entre español e inglés con:

1. su marido/mujer? ___ ___ ___ ___ ___

2. sus hijos? ___ ___ ___ ___ ___

3. sus padres? ___ ___ ___ ___ ___

4. resto de su familia? ___ ___ ___ ___ ___

5. sus amigos? ___ ___ ___ ___ ___

6. en su comunidad? ___ ___ ___ ___ ___

¿Qué lengua prefiere para las redes sociales?
Inglés Españool Ambas

• Actitudes lingüísticas/Red de contactos

¿Ha estado alguna vez en Estados Unidos continental?
Rodee una: SÍ NO
Si la respuesta es SÍ... ¿Cada cuánto va?
¿Cuánto duran las estancias?

¿Le gusta viajar allá?

¿Tiene contacto con otros puertorriqueños en Estados Unidos continental?
Rodee una: SÍ   NO

Si la respuesta es SÍ... ¿de qué ciudades?
¿Cada cuánto? Rodee una
   Cada día  1 vez a la semana  1 ó 2 veces al mes  1 vez al año

¿Tiene contacto con otras comunidades hispanas? Ej. dominicanos, ecuatorianos, etc.
Rodee una: SÍ   NO

Si la respuesta es SÍ... ¿Cada cuánto? Rodee una
   Cada día  1 vez a la semana  1 ó 2 veces al mes  1 vez al año

¿De dónde son? ¿De qué países/ciudades?

¿Ve o sigue las noticias, periódicos, redes sociales puertorriqueños?
Rodee una: SÍ   NO

¿Los puertorriqueños en Salinas/San Juan “arrastran la r”? ¿Y usted?
SÍ   NO
   Sí   NO

a) ¿Quién la arrastra?
   (ej. gente joven/mayor; mujeres/hombres):

b) ¿Qué le parece esta manera de pronunciar la “rr”?

¿Le gustaría participar en una segunda parte?
Información de contacto:
APPENDIX G

PUERTO RICO QUESTIONNAIRE
IN ENGLISH

Please, complete the questionnaire. You may skip any question(s) you are uncomfortable answering.

PARTICIPANT #__________

• Background information
Age: __________ Sex: __________
Place of birth: _________
Place where you grew up:
Place where you have lived most of your life:
Number of years living in Salinas/San Juan:

Pick one:
I was born in Puerto Rico
At least one of my parents was born in Puerto Rico
At least one of my grand-parents was born in Puerto Rico

Which part of Puerto Rico is your family from? _________
Where are your parents from?_________

Studies (circle one): Elementary School High School
College/University

Job:

Trips to Spanish speaking countries:
When and for how long _________

Including yourself, how many persons live in your household?
a. one/two b. three/four
d. five e. six or more

• Language proficiency and domains
Knowledge of languages other than Spanish:
a) Yes. Language______________
b) No

Rate your proficiency in Spanish and English according to the scale (write the number next to each skill):

6 = NATIVE FLUENCY
3 = INTERMEDIATE
FLUENCY
5 = NEAR (ALMOST) NATIVE FLUENCY        2 = BASIC
4 = ADVANCED FLUENCY                    1 = BEGINNING
FLUENCY

<table>
<thead>
<tr>
<th></th>
<th>SPANISH</th>
<th>ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>Speaking</td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>Reading</td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>Writing</td>
<td></td>
</tr>
<tr>
<td>Listening</td>
<td>Listening</td>
<td></td>
</tr>
</tbody>
</table>

At what age did you begin learning Spanish?  
If you speak English, at what age did you begin learning English?

When you were a CHILD, language spoken at home:
SPANISH  ENGLISH  BOTH
When you were a CHILD, language spoken among friends:
SPANISH  ENGLISH  BOTH

Did/Do you go to school in Puerto Rico?  
YES   NO
If YES.... From age ______ to age ______

Did/Do you go to school in Mainland US?  
YES   NO
If YES, from age ______ to age ______

Have your teachers ever corrected your pronunciation? What sounds?

How often do you...

<table>
<thead>
<tr>
<th></th>
<th>Every</th>
<th>2 or 3</th>
<th>Once</th>
<th>1 or 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>day</td>
<td>a week</td>
<td>a week</td>
<td>a month</td>
</tr>
</tbody>
</table>

a. speak English in your home?  
   ___  ___  ___  ___  ___

b. speak English outside
c. **speak English to:**

   1. your husband/wife? __
   
   2. your children? __
   
   3. your parents? __
   
   4. your relatives? __
   
   5. your friends? __
   
   6. classmates? __
   
   7. teachers? __
   
   8. workmates? __
   
   9. at your community
      (mall, community ctr, church)? __


d. **speak Spanish** in your home?


e. **speak Spanish** outside your home?


f. **speak Spanish** to:

   1. your husband/wife? __
2. your children? ___ ___ ___ ___ ___ ___

3. your parents? ___ ___ ___ ___ ___ ___

4. your relatives? ___ ___ ___ ___ ___ ___

5. your friends? ___ ___ ___ ___ ___ ___

6. classmates? ___ ___ ___ ___ ___ ___

7. teachers? ___ ___ ___ ___ ___ ___

8. workmates? ___ ___ ___ ___ ___ ___

9. at your community? ___ ___ ___ ___ ___

g. switch between Spanish and English with:
   1. your husband/wife? ___ ___ ___ ___ ___ ___

   2. your children? ___ ___ ___ ___ ___ ___

   3. your parents? ___ ___ ___ ___ ___ ___

   4. your relatives? ___ ___ ___ ___ ___ ___

   5. your friends? ___ ___ ___ ___ ___ ___

   6. at your community? ___ ___ ___ ___ ___

Which language do you prefer for social media?
   English    Spanish    Both

• Language attitudes/networking
Have you ever been to Mainland US?
Circle one: YES NO
If you answer YES…. How often do you go?
How long are the stays?
Do you like traveling there?

Do you have contact with other Puerto Ricans in Mainland US?
Circle one: YES NO
If you answer YES…. From what cities?
How often do you talk to them (calls/social media, etc.)? Circle one?
Every day Once a week Once or twice a month Once a year

Do you have contact with other Spanish-speaking communities? E.g. Dominicans, Ecuadorians, etc.
Circle one: YES NO
If you answer YES…. How often? Circle one
Every day Once a week Once or twice a month Once a year
Where are they from? What countries/cities?

Do you watch/follow Puerto Rican TV, newspapers, Puerto Rican social media?
Circle one: YES NO

Do Puerto Ricans in Salinas/San Juan “arrastran la r”? And you?
YES NO YES NO

c) Who does it?
(e.g. old/young people; women/men; metropolitan/non-metropolitan area):

d) What do you think about this way of pronouncing the “rr”?

Would you like to participate in a second part of this research?
Contact info:
BIBLIOGRAPHY


Carlos Vega Collection of Latino History in Holyoke, Collection 2011.1.2, Wistariahurst Museum, Holyoke, MA.


Data USA. Holyoke, MA. Retrieved from https://datausa.io/profile/geo/holyoke-ma/

Data USA. Salinas, PR. Retrieved from https://datausa.io/profile/geo/salinas-pr/


Horvath, B., & Horvath, R. J. (2000). Place and apparent time: interpreting the effects of place and space. Article presented in the *New Waves of Analyzing Variation* (N WAV 29), Michigan State University.


Rohana-Madrazo, Marcos. (2008). Buenos Aires Spanish in real time: the advancement of palatal fricative devoicing. Article presented in New Ways of Analyzing Variation (NWAV37), Rice University Houston, TX.


