A Phonetic Examination of Rhotics: Gestural Representation
Accounts for Phonological Behaviour

by

Phil Howson

A thesis submitted in conformity with the requirements
for the Degree of Doctor of Philosophy
Department of Linguistics
University of Toronto

© Copyright by Phil Howson, 2018
A phonetic examination of rhotics: gestural representation accounts for phonological behaviour

Phil Howson

Doctor of Philosophy

Department of Linguistics
University of Toronto
2018

Abstract

There is phonological evidence that rhotics form a natural class of sounds (Walsh-Dickey, 1997). However, the articulatory and acoustic properties that are common across all rhotics have been difficult to identify (Lindau, 1985; Ladefoged & Maddieson, 1996), creating a disconnect between phonetics and phonology. This thesis argues that the natural class of rhotics is united through articulatory and acoustic-perceptual characteristics.

Five studies were conducted: an acoustic examination of Upper and Lower Sorbian uvular (/r r/) and alveolar rhotics (/r r/), two ultrasound studies of the uvular rhotics in Upper Sorbian, an ultrasound study of variably realized rhotics ([ɾ ɻ ɹ ʁ χ]) in Brazilian Portuguese, and a perceptual study of rhotics by native English speakers. The rhotics in Sorbian and Brazilian Portuguese all differ in place/manner configurations but were found to be united by the occurrence of a tongue root gesture coordinated with a tongue tip or body gesture. These findings are consistent with previous studies of rhotics which show a secondary pharyngeal constriction (Delattre, 1971; Sproat & Fujimura, 1993). The articulatory findings suggest that the synchronic variation and diachronic changes of rhotics in Brazilian Portuguese result from the manipulation of the settings of the tongue gestures in the underlying representation. Palatalization in Upper Sorbian revealed antagonistic gestures for rhotic articulation and secondary palatalization,
suggesting that rhotics’ avoidance of secondary palatalization is related to tongue root constraints (Kavitskaya, Iskarous, Noiray, & Proctor, 2009).

The perceptual experiment revealed an underlying similarity in perception between the rhotics examined, /r ɻ r/. The perceptual similarities were interpreted as the result of a similarity in F2. The perceptual similarity is also argued to be the reason why large rhotic inventories are cross-linguistically avoided and for the reason the cross-linguistic distribution of rhotics and laterals is also different. Inventories with more than 2 rhotics and no laterals do not exist, while inventories of laterals up to and including 6 different segments and no rhotics, do exist (Maddieson, 1984). This disparity between the two types of liquids is argued to be related to the perceptual difficulty associated with identifying rhotic segments across place and manner differences.
Acknowledgements

First and foremost, I would like to thank my parents for their tremendous support throughout my life; it has really allowed me to pursue my goals and dreams. I truly would not be here today without their love and support. Thank you so much for your patience, guidance, and mentorship throughout my life. I also wish to express my sincere gratitude to Kate Komova for being such a tremendous friend since I first met her during my undergraduate at UBC. Her constant teasing about being unable to properly pronounce ř triggered my pursuit of phonetic research and helped launch an entire shift in my academic trajectory. I can also produce ř now, despite her continued teasing.

Throughout the years at UofT, I have been fortunate enough to work with many high-calibre researchers. My supervisor, Dr. Alexei Kochetov, has been helpful in not only suggesting possible avenues of inquiry, but pointing out key grants to apply for to support my research. I am grateful that due to the suggested inclusion of Sorbian, I was able to spend several months in Germany. It was a great experience working with the Sorbian community and gave me the opportunity to travel to many amazing European locations. I also had the good fortune to be supervised by Dr. Philip Monahan, who, despite my constantly getting side tracked in our meetings, was always a pleasure to work with. I appreciate being able to bounce my ideas off you and the many unrelated conversations we had through the years. It was always enjoyable. Dr. Peter Jurgec was also a super helpful and I appreciate all the times we discussed ideas unrelated to my thesis. It has been great to have you around and I appreciate all the great discussions we had despite our many disagreements.

I have also been lucky to have been given tremendous support and guidance for my research on Sorbian. Without the help that Dr. Edward Wornar and Dr. Lechosław Jocz provided me, I would not have been able to conduct my research. Dr. Wornar and Dr. Jocz gave me extensive support in creating word-lists and finding participants for my studies. I especially want to thank Dr. Jocz for his many helpful comments on my illustration of Upper Sorbian, and for guiding me through the key literature on the phonology of Upper and Lower Sorbian. I couldn’t have done it without his assistance. I also want to thank Dr. Sven Grawunder and Dr. Bernard Comrie for making me at home at MPI: Eva and allowing me use of the facilities. I was able to collect extensive acoustic and ultrasound data at MPI: Eva and I appreciate the tremendous
support. I also greatly miss the cafeteria there. I enjoy German cuisine quite a bit and it’s unfortunate that it’s nearly impossible to find Laugenbrötchen in Toronto.

During my time as a Ph. D. student, I also worked with Dr. Yoonjung Kang for both of my Generals Papers. She was always kind, and her comments and criticisms were always insightful and thought provoking. I appreciate the fact that the caliber of my research projects always improved as a result of working with her.

Instructorships are an important aspect of career development and I was able to teach several courses which helped develop my portfolio. During this time, Dr. Keren Rice was my supervisor and I appreciate the guidance she gave me for improving my teaching. Her helpful suggestions built in a systematic way of approaching course and assignment design. I feel that my way of thinking about teaching and the way I approach student interactions in the classroom has improved a great deal and I am thankful that she put up with me as her supervisee.

I would also like to take a moment to extend my appreciation to the external committee for my oral defense, Dr. Khalil Iskarous, Dr. Anabela Rato, and Dr. Nathan Sanders. I appreciate the short notice that the committee was given due to my surprise completion deadline. I have always respected Dr. Iskarous and his work. It was a pleasure to have him on my committee. Dr. Rato and Dr. Sanders were also a pleasure to have at the defense and I appreciate their many comments and questions. Dr. Sanders also provided some thought provoking questions about the phonological consequences of many of my theories. Dr. Rato made possible my research on Brazilian Portuguese by helping me develop a suitable word-list and helping me find participants. So, I would like to extend my gratitude to her for that. I would also like to thank Luke Zhou for help with some statistical programming issues for the Brazilian Portuguese data and to Youyee Chen for her help tracing tongue contours.

As a graduate student with a desire to pursue a career in academia, I have always felt that it was important to develop a research portfolio. I have been fortunate enough to have received tremendous academic freedom and support from the faculty at UofT. Dr. Jessamyn Schertz was kind enough to collaborate with me on an ongoing project unrelated to my Ph. D. and I have felt that my knowledge base and experimental designs have improved quite a bit as a result. I also want to thank her personally for inspiring me to be a better person.

I also had many questions throughout the years on administrative matters and I want to thank Jill Given-King, Mary Hsu, Deem Waham, and Jennifer McCallum for their help. I would
have been unable to navigate all the department related duties without their support. I especially want to thank Jennifer McCallum for helping organize my oral defense in a very tight time period.

I would also like to extend my gratitude to Dr. Bryan Gick for his constant support and guidance. During my undergraduate years, a disagreement between Kate and I over the nature of the articulation of ř, coupled with our collective distrust for its designated IPA transcription, led us to pursue a phonetic experiment using the ultrasound machine. It was at this juncture that Dr. Gick kindly agreed to supervise us and help us develop a project plan. This experiment would go on to result in two conference presentations and a publication in the Journal of the International Phonetic Association (JIPA). To this day, I am much obliged to Dr. Gick for encouraging my intellectual pursuits and for giving me the opportunity to prove myself. His help came at a time that was critical to my career development and which allowed me to demonstrate my research ability.

Finally, I would like to thank Dr. Melissa Redford for offering me a post-doc position with her at the University of Oregon. I am very excited to take the next step in my career and continue to pursue a position in academia. Allowing me to have a creative outlet through my work is truly a blessing.

All in all, I have had a tremendous experience here at UofT and I am grateful for all the amazing colleagues I have met and had the pleasure to work with, the friendships I have made, and the opportunities I have been given. Working through my Ph. D. gave me the opportunity to travel the world and during my time here I have presented at numerous conferences around the world. I have had the pleasure of having the world open up before my eyes. I lived in Germany, travelled Europe, have been to Asia, South America, and America as a result of the numerous presentations and side projects I have ventured on. This has filled me with a lifetime of beauty and has certainly fulfilled me in many ways. I have certainly learned through the course of my life to appreciate what one has, and I don’t want to understate my gratitude and appreciation for all the beauty that my pursuit has brought me.

This work was funded by the German/Europe Fund, the Michael Smith Foreign Studies Supplement, and SSHRC#771-2015-0048. The ultrasound equipment used for the Brazilian Portuguese data collection was made available through the Social Sciences and Humanities Research Council of Canada Standard Grant #435-2015-2013 to Alexei Kochetov.
Table of Contents

Acknowledgements ........................................................................................................ iv
Table of Contents ........................................................................................................... vii
List of Tables ................................................................................................................ x
List of Figures ............................................................................................................... xii
List of Appendices ...................................................................................................... xxi
Introduction .................................................................................................................. 1
  1.1 Aims ....................................................................................................................... 1
  1.2 Acoustics of Rhotics ............................................................................................ 3
  1.3 Acoustics of Palatalized Rhotics .......................................................................... 5
  1.4 Articulation of Rhotics ....................................................................................... 6
  1.5 Articulation of Palatalized Rhotics ..................................................................... 10
  1.6 Rhotic Perception ............................................................................................... 11
  1.7 Theoretical Framework ...................................................................................... 13
  1.8 Organization of the Dissertation ....................................................................... 17
Rhotics and Palatalization: An Acoustic Examination of Upper and Lower Sorbian .......... 20
  2.1 Introduction ......................................................................................................... 21
  2.2 Hypotheses ......................................................................................................... 23
  2.3 Methods ............................................................................................................. 25
    2.3.1 Speakers ...................................................................................................... 25
    2.3.2 Procedure .................................................................................................... 25
    2.3.3 Materials ..................................................................................................... 25
    2.3.4 Analysis ...................................................................................................... 26
  2.4 Results .................................................................................................................. 27
    2.4.1 Upper Sorbian ............................................................................................ 27
    2.5.2 Lower Sorbian ............................................................................................ 29
    2.5.3 Comparison of Upper and Lower Sorbian .................................................. 30
  2.5 Discussion ........................................................................................................... 31
Rhotic articulation involves tongue root retraction: evidence from Upper Sorbian ............. 33
  3.1 Introduction ......................................................................................................... 33
    3.1.1 Hypothesis .................................................................................................. 33
  3.2 Methods ............................................................................................................. 34
There is a tongue root conflict between secondary palatalization and rhotics: evidence from Upper Sorbian

<table>
<thead>
<tr>
<th>3.2.1 Participants</th>
<th>34</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.2 Materials</td>
<td>34</td>
</tr>
<tr>
<td>3.2.3 Procedure</td>
<td>35</td>
</tr>
<tr>
<td>3.2.4 Analysis</td>
<td>35</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>37</td>
</tr>
<tr>
<td>3.3.1 SSANOVA Results</td>
<td>37</td>
</tr>
<tr>
<td>3.3.2 Coarticulatory Analysis</td>
<td>42</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>47</td>
</tr>
</tbody>
</table>

Rhotic perception provides evidence for classhood and accounts for cross-linguistic distribution of rhotics in phonemic inventories

<table>
<thead>
<tr>
<th>4.1 Introduction</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1 Hypothesis</td>
<td>48</td>
</tr>
<tr>
<td>4.2 Methods</td>
<td>49</td>
</tr>
<tr>
<td>4.2.1 Materials</td>
<td>49</td>
</tr>
<tr>
<td>4.2.2 Analysis</td>
<td>50</td>
</tr>
<tr>
<td>4.3 Dynamic Results</td>
<td>51</td>
</tr>
<tr>
<td>4.4 Discussion</td>
<td>59</td>
</tr>
</tbody>
</table>

Lenition of gestural settings for rhotics accounts for synchronic and diachronic changes in Brazilian Portuguese

<table>
<thead>
<tr>
<th>5.1 Introduction</th>
<th>61</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.1 Brazilian Portuguese</td>
<td>61</td>
</tr>
<tr>
<td>5.1.2 Hypothesis</td>
<td>62</td>
</tr>
<tr>
<td>5.2 Methods</td>
<td>62</td>
</tr>
<tr>
<td>5.2.1 Participants</td>
<td>62</td>
</tr>
<tr>
<td>5.2.2 Instrumentation</td>
<td>63</td>
</tr>
<tr>
<td>5.2.3 Materials</td>
<td>63</td>
</tr>
<tr>
<td>5.2.4 Analysis</td>
<td>65</td>
</tr>
<tr>
<td>5.3 Results</td>
<td>66</td>
</tr>
<tr>
<td>5.3.1 SSANOVA Results</td>
<td>66</td>
</tr>
<tr>
<td>5.3.2 Dynamic Results</td>
<td>70</td>
</tr>
<tr>
<td>5.3.3 Coarticulatory Results</td>
<td>76</td>
</tr>
<tr>
<td>5.4 Conclusions</td>
<td>77</td>
</tr>
</tbody>
</table>

viii
6.1 Introduction .................................................................................................................................................. 79
5.1.1 Hypothesis ............................................................................................................................................. 79
6.2 Methods ....................................................................................................................................................... 80
  6.2.1 Participants ........................................................................................................................................... 80
  6.2.2 Stimuli ................................................................................................................................................... 80
  6.2.3 Procedure ............................................................................................................................................ 82
  6.2.4 Analysis ............................................................................................................................................. 83
6.3 Results ......................................................................................................................................................... 84
  6.3.1 d-prime Results ................................................................................................................................. 84
  6.3.2 MDS Results ..................................................................................................................................... 88
  6.3.3 kmeans Results ............................................................................................................................... 92
6.4 Discussion ................................................................................................................................................... 95
General Discussion and Conclusion .................................................................................................................. 96
  7.1 Discussion ................................................................................................................................................ 96
  7.2 Discussion of Acoustic and Articulatory Properties Uniting the Rhotics ................................................. 96
  7.3 Discussion of the relationship between the Gestural Composition of Rhotics and their Phonological Behaviour and Typological distribution ................................................................. 101
    7.3.1 Discussion of Cross-Linguistic Distribution of Rhotics ................................................................. 105
    7.3.2 Discussion of the Dispreference for Rhotics and Palatalization .................................................... 106
  7.4 Summary of Primary Findings ................................................................................................................ 109
  7.5 Summary of Theoretical Implications ..................................................................................................... 111
  7.6 Limitations and Future Research ........................................................................................................... 112
References ......................................................................................................................................................... 113
Appendix A Upper Sorbian SSANOVAs ........................................................................................................... 121
Appendix B Brazilian Portuguese SSANOVAs ............................................................................................... 126
List of Tables

Table 1. Summary of the vocal tract variables and articulators ................................................................. 15
Table 2. Summary of possible constriction degrees for each articulator .................................................... 15
Table 3. Consonant inventory for Upper Sorbian. Hard indicates unpalatalized and soft indicates palatalized consonants (Howson, 2017). .................................................................................. 22
Table 4. Consonant inventory for Lower Sorbian. Hard indicates unpalatalized consonants and soft indicates palatalized consonants. ................................................................................................. 23
Table 5. The Sorbian word sets used in this study ........................................................................................ 26
Table 6. Target stimuli for Upper Sorbian ................................................................................................... 35
Table 7. Summary of tongue shapes for /r/ for each speaker. + indicates that the tongue had a particular shape in the environment listed, while – indicates the lack of that tongue shape in the environment. .................................................................................................................. 42
Table 8. Means and standard deviations (in brackets) of the coarticulatory measures (mm$^2$) for each phoneme in each word position ................................................................................................................. 43
Table 9. Means and standard deviations (in brackets) for each phoneme in each environment. .......................... 44
Table 10. Means and standard deviations (in brackets) for each phoneme in each environment. .......................... 46
Table 11. Means and standard deviations (in brackets) for each phoneme in each environment. .......................... 47
Table 12. Target Stimuli for Upper Sorbian ................................................................................................... 50
Table 13. Consonant inventory of Brazilian Portuguese (from Barbosa & Albano, 2004, p. 228) ......................... 62
Table 14. Brazilian Portuguese Stimuli ........................................................................................................... 64
Table 15. Summary of the Realizations of /R/ .................................................................................................... 70
Table 16. Means and standard deviations of each phoneme in each environment. ........................................ 77
Table 17. Summary of the mixed effects model (** indicates significant p-values) ........................................... 77
Table 18. Mean formant frequencies (Hz) for the rhotics used in the experiment .............................................. 82
Table 19. Experimental Conditions ............................................................................................................... 83
Table 20. d-prime means and standard deviations for rhotic and stop comparisons ........................................ 85
Table 21. d-prime means and standard deviations for rhotic and nasals comparisons ..................................... 86
Table 22. d-prime means and standard deviations for rhotic and fricative comparisons ................................... 87
Table 23. d-prime means and standard deviations for rhotic and laterals comparisons ..................................... 88
Table 24. Results of the linear regression model (** indicates a significant result) .................. 91
Table 25. Distribution of the original groupings compared to their retribution in the kmeans
analysis......................................................................................................................................... 93
Table 26. Hypothesized tract variable specifications for the rhotics examined in this dissertation
........................................................................................................................................................ 101
List of Figures

Figure 1. IPA chart including the proto-typical members of the class of rhotics (red) and the members that typically involve synchronic or diachronic shifts to be considered rhotics (blue)... 1

Figure 2. Chart of the acoustic relationships shared by all members of the rhotic class (from Lindau, 1985, p. 165). a1 = pulse pattern (trill); a2 = closure duration; a3 = presence of formants (sonorant); a4 = presence of noise; a5 = distribution of spectral energy (place of articulation). ... 4

Figure 3. Intervocalic /ʁ/ in French (Delattre, 1971; p. 146). ................................................................. 7

Figure 4. Gestural score showing the activation and timing of different articulators involved in the production of the world ‘palm’ (from Browman & Goldstein, 1989, p. 212). .................. 16

Figure 5. Upper Sorbian is spoken south the solid black line, with the cultural capital of Budyšín. Image taken from Howson (2017). ............................................................................................................. 21

Figure 6. Map of the Lower Sorbian speaking region. The cultural capital is in Chóśebuz........... 22

Figure 7. Formant measures were taken from the onset of the rhotic to the midpoint of the vowel. Image shows [ɾa] from the Lower Sorbian word rjagotaš 'to stutter.' ................................. 27

Figure 8. Formant (F1 to F3, left; F2-F1 right) dynamics of the Upper Sorbian liquids. The left edge indicates the onset of the liquid and the right edge of the graph indicates the mid-point of the vowel, /a/. ............................................................................................................................... 28

Figure 9. Formant (F1 to F3, left; F2-F1, right) dynamics of the Lower Sorbian liquids. The left edge indicates the onset of the liquid and the right edge of the graph indicates the mid-point of the vowel, /a/. ............................................................................................................................... 29

Figure 10. Comparison of the F1, F2 and F3 for unpalatalized (left) and palatalized (right) trills in Upper and Lower Sorbian .................................................................................................................. 30

Figure 11. Ultrasound image loaded into Edgetrak and example contour tracing (left) and Tongue shape divided into the different areas referred to within this paper (right, modified from Catford 1988). ................................................................................................................................. 36

Figure 12. Example of the points taken for the Tongue Root and Tongue Body. ...................... 37

Figure 13. Tongue contours for /g/ in each environment /a/, /e/, and /o/ in each word position, word-initial (left), intervocalic (center), and word-final (right) for S1 .................................................. 39

Figure 14. /ʁ/ in each environment /a/, /e/, and /o/ in each word position, word-initial (left), intervocalic (center), and word-final (right) for S1. ................................................................. 40
Figure 15. SSANOVA plots of the point of maximum constriction for /ɡ/ vs. /r/ in Ca (left), aCa (middle), aC (right) for S1. Tongue tip is on the right.

Figure 16. SSANOVA plots of the point of maximum constriction for /ɡ/ vs. /r/ in Ce (left), eCe (middle), eC (right) for S1. Tongue tip is on the right.

Figure 17. SSANOVA plots of the point of maximum constriction for /ɡ/ vs. /r/ in Co (left), oCo (middle), oC (right) for S1. Tongue tip is on the right.

Figure 18. (left) comparison of /ɡ/ vs. /r/ in the /#_e/ environment for S1, showing a “double bunched” rhotic articulation; and (right) /ɡ/ vs. /r/ in the /#_e/ environment for S2, showing a “retracted” rhotic articulation.

Figure 19. Overall effects of coarticulation on the tongue root for each phoneme (in mm²).

Figure 20. Mean retraction (mm) for each phoneme.

Figure 21. Mean tongue root retraction (mm) in each environment.

Figure 22. The amount of tongue body retraction (mm) for each phoneme.

Figure 23. The amount of tongue body retraction (mm) for each environment.

Figure 24. The height of the tongue body (mm) for each phoneme.

Figure 25. The height of the tongue body (mm) for each environment.

Figure 26. Ultrasound image loaded into Edgetrak and example contour tracing (left) and Tongue shape divided into the different areas referred to within this paper (right, modified from Catford 1988).

Figure 27. Gestural dynamics for /r/ in the /#_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 28. Gestural dynamics for /r/ in the /a_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 29. Gestural dynamics for /r/ in the /a_#/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 30. Gestural dynamics for /r/ in the /#_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 31. Gestural dynamics for /r/ in the /e_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 32. Gestural dynamics for /r/ in the /e_#/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
Figure 33. Gestural dynamics for /ɾ/ in the /#_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 55
Figure 34. Gestural dynamics for /ɾ/ in the /o_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 55
Figure 35. Gestural dynamics for /ɾ/ in the /o_#/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 56
Figure 36. Gestural dynamics for /ɾʃ/ in the /#_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 57
Figure 37. Gestural dynamics for /ɾʃ/ in the /a_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 57
Figure 38. Gestural dynamics for /ɾʃ/ in the /#_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 58
Figure 39. Gestural dynamics for /ɾʃ/ in the /e_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 58
Figure 40. Gestural dynamics for /ɾʃ/ in the /#_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 59
Figure 41. Gestural dynamics for /ɾʃ/ in the /o_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 59
Figure 42. Example of the AAA interface and with tongue trace (left) and the division of the tongue into different regions as discussed in this paper (right, modified from Catford, 1988).... 65
Figure 43. SSANOVA results for BP5’s articulation of /k/ in all vocalic environments for the word-initial position (left), intervocalic position (center), and word-final position (right)........... 67
Figure 44. SSANOVA for BP5’s articulation of /t/ in all vocalic environments for the word-initial position (left), intervocalic position (center), and word-final position (right)........... 67
Figure 45. SSANOVA results for BP5’s articulation of /ɾ/ in word-initial position (left) and intervocalic position (right)........................................................................................................... 68
Figure 46. SSANOVA results for BP5’s articulation of /ɾʃ/ in intervocalic position (left) and /R/ in word-final position (right)........................................................................................................... 68
Figure 47. Gestural dynamics for /k/ for BP1 in the environment /#_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 71
Figure 48. Gestural dynamics for /k/ for BP1 in the environment /e_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 71
Figure 49. Gestural dynamics for /k/ for BP1 in the environment /e_#/ , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 72
Figure 50. Gestural dynamics for /t/ for BP1 in the environment /#/e , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 72
Figure 51. Gestural dynamics for /t/ for BP1 in the environment /e_e/ , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 73
Figure 52. Gestural dynamics for /t/ for BP1 in the environment /e_#/ , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 73
Figure 53. Gestural dynamics for /ɾ/ for BP1 in the environment /#/e , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 74
Figure 54. Gestural dynamics for /ɾ/ for BP1 in the environment /e_e/ , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 74
Figure 55. Gestural dynamics for /ɾ/ for BP1 in the environment /e_#/ , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 75
Figure 56. Gestural dynamics for coda /R/ for BP1 in the environment /e_#/ , including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.... 75
Figure 57. Bar plot of the mean root mean square for each phoneme in each environment. ....... 76
Figure 58. Spectrogram for the token /ɾ/ as produced by the female native Russian speaker. The left edge is aligned to the onset of the tongue-palate contact. The occlusion for trill contact can be clearly observed on the left edge................................................................. 81
Figure 59. F1-F3 of the rhotics used in the perceptual experiment. The left edge is the onset of the rhotic, the mid-point is the end of the rhotic, and the right edge is the mid-point of the vowel. ............................................................................................................. 82
Figure 60. Bar plot of the d-prime results for the comparison between rhotics and stops. ........ 85
Figure 61. Bar plot of the d-prime results for the comparison between rhotics and nasals........... 86
Figure 62. Bar plot of the d-prime results for the comparison between rhotics and fricatives..... 87
Figure 63. Bar plot of the d-prime results for the comparison between rhotics and laterals. ....... 88
Figure 64. a. Bar plot of the mean perceptual units for the rhotics vs. stops comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and stops. ................................................................. 89

Figure 65. a. Bar plot of the mean perceptual units for the rhotics vs. nasals comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and nasals................................................................. 89

Figure 66. a. Bar plot of the mean perceptual units for the rhotics vs. fricatives comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and fricatives................................................................. 90

Figure 67. a. Bar plot of the mean perceptual units for the rhotics vs. lateral comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and laterals................................................................. 91

Figure 68. Bar plot comparing the perceptual space for the rhotics in each of the experimental conditions (comparison with stops, nasals, fricatives, and laterals). ................................................................. 92

Figure 69. kmeans plot for the rhotics and stops (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right)................................................................. 93

Figure 70. kmeans plot for the rhotics and nasals (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right)................................................................. 94

Figure 71. kmeans plot for the rhotics and fricatives (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right)................................................................. 94

Figure 72. kmeans plot for the rhotics and laterals (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right)................................................................. 95

Figure 73. Gestural score for /oɡo/. The articulators involved include the tongue root (TR) and the tongue body (TB) with the target constriction location (uvular-pharyngeal and pharyngeal).97

Figure 74. Gestural score for /oɾo/. The articulators involved include the tongue root (TR) and the tongue body (TB) with the target constriction location (uvular-pharyngeal and pharyngeal).98

Figure 75. Gestural representation for /ʁ/. .......................................................................................... 99

Figure 76. Gestural representation for /ɾ/.......................................................................................... 100

Figure 77. Gestural score for /eɾ/. The articulators involved include the tongue body (TB) and the tongue root (TR) with the target constriction location (coronal, palatal, uvular-pharyngeal, and pharyngeal).......................................................................................... 102
Figure 78. Gestural score for /i/ for speakers who produce word-final [χ]. .................................................. 103
Figure 79. Gestural score for /oʊ̩/, demonstrating intervocalic lenition of /ʁ/ > [h]. ......................... 104
Figure 80. Gestural representation for /ʁ/. The articulators involved include the the tongue body (TB) and the tongue root (TR) with the target constriction location (coronal, palatal, uvular-pharyngeal, and pharyngeal)................................................................. 108
Figure 81. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S2. Tongue tip is on the right ................................................................. 121
Figure 82. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S2. Tongue tip is on the right ................................................................. 121
Figure 83. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S2. Tongue tip is on the right ................................................................. 121
Figure 84. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S3. Tongue tip is on the right ................................................................. 122
Figure 85. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S3. Tongue tip is on the right ................................................................. 122
Figure 86. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S3. Tongue tip is on the right ................................................................. 122
Figure 87. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S4. Tongue tip is on the right ................................................................. 123
Figure 88. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S4. Tongue tip is on the right ................................................................. 123
Figure 89. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S4. Tongue tip is on the right ................................................................. 123
Figure 90. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S5. Tongue tip is on the right ................................................................. 124
Figure 91. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S5. Tongue tip is on the right ................................................................. 124
Figure 92. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S5. Tongue tip is on the right ................................................................. 124
Figure 93. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S6. Tongue tip is on the right ................................................................. 125
Figure 94. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S6. Tongue tip is on the right. ................................. 125
Figure 95. SSANOVA plots of the point of maximum constriction for /g/ in aC (left), eC (middle), oC (right) for S6. Tongue tip is on the right. ........................................... 125
Figure 96. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP1. Tongue tip is on the right. ......................... 126
Figure 97. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP1. Tongue tip is on the right. ......................... 126
Figure 98. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP1. Tongue tip is on the right. ......................... 126
Figure 99. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP1. Tongue tip is on the right. ......................... 127
Figure 100. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP2. Tongue tip is on the right. ......................... 127
Figure 101. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP2. Tongue tip is on the right. ......................... 127
Figure 102. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP2. Tongue tip is on the right. ......................... 128
Figure 103. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP2. Tongue tip is on the right. ......................... 128
Figure 104. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP3. Tongue tip is on the right. ......................... 128
Figure 105. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP3. Tongue tip is on the right. ......................... 129
Figure 106. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP3. Tongue tip is on the right. ......................... 129
Figure 107. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP3. Tongue tip is on the right. ......................... 129
Figure 108. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP4. Tongue tip is on the right. ......................... 130

xviii
Figure 109. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP4. Tongue tip is on the right. .................................................. 130
Figure 110. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP4. Tongue tip is on the right. .................................................. 130
Figure 111. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP4. Tongue tip is on the right. .................................................. 131
Figure 112. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP5. Tongue tip is on the right. .................................................. 131
Figure 113. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP5. Tongue tip is on the right. .................................................. 131
Figure 114. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP5. Tongue tip is on the right. .................................................. 132
Figure 115. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP5. Tongue tip is on the right. .................................................. 132
Figure 116. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP6. Tongue tip is on the right. .................................................. 132
Figure 117. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP6. Tongue tip is on the right. .................................................. 133
Figure 118. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP6. Tongue tip is on the right. .................................................. 133
Figure 119. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP6. Tongue tip is on the right. .................................................. 133
List of Appendices

Appendix A
  Upper Sorbian SSANOVAs

Appendix B
  Brazilian Portuguese SSANOVAs
Chapter 1

Introduction

1.1 Aims

The rhotics (‘r-like’ sounds, typically represented with the Roman ‘r’; from the Greek letter ‘rho’) consist of a diverse set of sounds that span different manners and places of articulation. Figure 1 presents an IPA chart (International Phonetic Association, 2015) with the standard set of rhotics appearing in red boxes and segments which are typically only considered rhotics as a result of synchronic variation or diachronic sound change in blue boxes. These include trills, taps or flaps, coronal approximants, as well as certain fricatives and the bilabial approximant /ʋ/. Furthermore, rhotics are not strictly limited to consonants, but also encompass a group of vowels typically referred to as ‘rhotacized’ vowels (e.g. /ɜ/). Ladefoged and Maddieson (1996) suggest that the diversity in both acoustic and articulatory properties is one of the primary reasons that class characteristics have been so difficult to capture and that it may be the case rhotics are united purely by synchronic and diachronic phonological processes (see Walsh-Dickey [1996] and Wiese [2001] for extensive phonological evidence for classhood). However, this would be undesirable because it would set rhotics apart from other natural classes of sounds, which have clear acoustic and/or articulatory properties which capture the class.

**THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)**

<table>
<thead>
<tr>
<th>CONSONANTS (PULMONIC)</th>
<th>© 2015 IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plosive</strong></td>
<td><strong>B</strong></td>
</tr>
<tr>
<td><strong>Nasal</strong></td>
<td><strong>m</strong></td>
</tr>
<tr>
<td><strong>Trill</strong></td>
<td><strong>B</strong></td>
</tr>
<tr>
<td><strong>Tap or Flap</strong></td>
<td><strong>نبي</strong></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td><strong>نبي</strong></td>
</tr>
<tr>
<td><strong>Lateral fricative</strong></td>
<td><strong>نبي</strong></td>
</tr>
<tr>
<td><strong>Approximant</strong></td>
<td><strong>نبي</strong></td>
</tr>
<tr>
<td><strong>Lateral approximant</strong></td>
<td><strong>نبي</strong></td>
</tr>
</tbody>
</table>

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

*Figure 1. IPA chart including the proto-typical members of the class of rhotics (red) and the members that typically involve synchronic or diachronic shifts to be considered rhotics (blue).*

The purpose of this dissertation is to examine the acoustic, articulatory, and phonological properties of rhotics as a natural class and search for evidence of class membership from a low-
level phonetic perspective. From the perspective of Articulatory Phonology (Browman and Goldstein, 1986; 1988; 1989; 1992; 1993; 1995), a gestural representation for rhotics is used to account for the synchronic variation and diachronic changes that specifically target rhotics as a class. To facilitate the goal of understanding the phonetics and phonology of rhotics, one acoustic study, three articulatory studies, and one perceptual study, for a total of five studies, were conducted: First, an acoustic examination of unpalatalized and palatalized rhotics in Upper and Lower Sorbian; second, an ultrasound examination of the unpalatalized rhotic in Upper Sorbian; third, an ultrasound study of the unpalatalized and palatalized rhotics in Upper Sorbian; fourth, an ultrasound study of uvular and alveolar rhotics in Brazilian Portuguese; and fifth, the perception of non-native rhotics by native speakers of English.

Lower Sorbian was chosen because it has an unpalatalized and a palatalized alveolar rhotic and has had little phonetic work done on it. Upper Sorbian was included because it is also an endangered language with little phonetic work done on it and because of the typologically rare contrast between a unpalatalized and palatalized uvular rhotic /r/ and /ɾ/ (Howson, 2017). They are also closely related languages, allowing for strong comparisons between the rhotics in their inventories.

Brazilian Portuguese was chosen because it has two phonemic rhotics, /ɾ/ and /ɾ/, which cover both uvular and alveolar places of articulation and fricative and approximant manners of articulation. It also has a coda rhotic, /R/, which has a great deal of variation in its phonetic realization across both speakers and dialects, ranging from a tap to a retroflex approximant. The fact that Brazilian Portuguese has such diversity in its rhotic inventory allows for a direct comparison of the articulatory properties involved in different places and manners of articulation. Furthermore, Brazilian Portuguese is well known for a wide range of positional and prosodic variants of the realization of each of the rhotics (Silva and Albano, 1999; Mateus and d’Andrade, 2000; Silva, 2003; Barbosa and Albano, 2004). Examining the rhotics in Brazilian Portuguese allows for significant testing of whether low-level gestural representations are sufficient for accounting for synchronic alternations and diachronic changes. I achieve this by examining the phonetic properties of each segment across word and syllable positions. I hypothesize that manipulation of the tongue tip and the tongue body gestures should account for synchronic alternations as well as diachronic changes.
A perceptual study was also conducted to test native English speakers’ discrimination of three non-native rhotics, /r/, /ɻ/, and /ʀ/. The goal of the study was to identify acoustic/perceptual correlates to the underlying articulatory properties of rhotics as a class. The purpose of all the studies combined is to approach the question of what acoustic-articulatory characteristics unite the class of rhotics and how those properties affect their phonology. One of the overall aims of this dissertation is to include uvular rhotics in both the articulatory and perceptual study of rhotics, which have been largely excluded from studies examining the acoustic and articulatory correlates to rhotic articulation.

1.2 Acoustics of Rhotics

Acoustic correlates of rhotics as a class have been elusive, leading some to suggest that any connection between rhotic segments as a natural class is based purely on phonological grounds (c.f. Ladefoged & Maddieson, 1996). This claim seems justified given previous large-scale acoustic studies (e.g. Lindau, 1985). However, there could be some characteristic that has been missed.

Lindau (1985) examined eleven languages (4 Indo-European languages and 7 West African languages) and concluded that there was no single acoustic characteristic that unites all rhotic consonants. One interesting finding was that a cross-linguistic tendency for a low F3 (typically attributed to rhotics) was not present in the data. There were languages which utilized a characteristically low F3, such as English, but this was not a cross-linguistic trend. This finding suggests that there are ‘dark’ rhotics and ‘clear’ rhotics, similar to laterals. The ‘dark’ variants are likely produced with a constriction in the lower pharyngeal or post-alveolar/palatal region, which is predicted by acoustic theory (Fant, 1970). This finding is inconsistent with Ladefoged’s (1971) assertion that lowered F3 (close to F2) is an acoustic characteristic common across rhotics. Given that there was no single acoustic characteristic uniting all rhotics, Lindau (1985) proposes a theory of rhotics that incorporates overlapping acoustic characteristics which link all rhotics together. Figure 2 presents a chart outlining the acoustic connection between each of the family members, creating a web of familial characteristics. The approximants all have formant structures. The trills are further united by a pulsing pattern. While the fricatives are related by pulsing pattern (for the trilled ones) and the presence of spectral energy. Finally, there is also a relation in the distribution of energy for rhotics of a similar place. In this way, there is acoustical
overlap among all the rhotics through a web of similar characteristics that connect them to other members of the family, but no single acoustic characteristic that all rhotics share.

![Diagram of acoustic relationships]

Figure 2. Chart of the acoustic relationships shared by all members of the rhotic class (from Lindau, 1985, p. 165). 
- $a_1$ = pulse pattern (trill)
- $a_2$ = closure duration
- $a_3$ = presence of formants (sonorant)
- $a_4$ = presence of noise
- $a_5$ = distribution of spectral energy (place of articulation).

Fricative rhotics also present an interesting area of research for understanding rhotics as a class. Jesus and Shadle (2005) performed an acoustic examination of European Portuguese, which is known to have a wide range of phonetic realizations for the rhotics (Matues & Andrade, 2000). The authors’ focused on [ɻ] and [ʁ] (allophones of /ɾ/), and [ɾ] (an allophone of /ɾ/). They found that [ɾ̥] had a short duration, indicating a stop-like closure, but also had turbulent noise characteristics, which were different from transient burst noise that stops have, indicating that it was fricated. The frication typically had a spectral peak centred between 8.9 kHz and 14 kHz. They note that it was difficult to ascertain that [ɾ] had an alveolar place of articulation; however, due to the fact that it is an allophone of /ɾ/, it is reasonable to assume it is also alveolar in place. This contrasts with [ɻ], which had peaks in the 1.0-1.6 kHz, 2.1-2.8 kHz, and 3.2-4 kHz ranges, which is typical of uvular fricatives. The voiced variant, [ʁ], only appeared in the corpus two (out of 115) times, but the peaks and spectral characteristics appear to be comparable to [ɻ], indicating it is a voiced uvular fricative. These spectral properties of the fricative variants of rhotics are important to understand because they give us clues to their articulatory properties. This line of analysis can help to approach the question of what articulatory characteristics rhotics all have in common.
1.3 Acoustics of Palatalized Rhotics

Palatalized rhotics are an important research area due to an overwhelming cross-linguistic tendency to avoid them (Hall, 2000) and the general rarity of phonemic palatalized rhotics (Maddieson, 1984; Bateman, 2011). Despite there being a well-documented phonological avoidance and multiple observed repair strategies, the mechanisms that cause the dispreference are not well understood. In this dissertation, the avoidance of rhotics and secondary palatalization is interpreted as a class characteristic and so studying this phenomenon can reveal some clues as to what properties unites the class.

Toda, which has three unpalatalized and palatalized rhotic pairs /ɾ/, /ɽ/, /ɭ/, /ɭ/ and /ɽ/, is an excellent example language for palatalized rhotics. Spajić, Ladefoged, and Bhaskararao (1996) conducted an acoustic examination of the acoustic difference between the unpalatalized and palatalized rhotics in Toda. Spajić et al. (1996) measured time-course formant frequencies at the onset, middle, and offset of each rhotic and preceding vowel. The unpalatalized rhotics showed a clear difference in F1 and F2 for the alveolar as compared to the dental. The difference in formant frequencies led Spajić et al. (1996) to conclude that the primary difference is not primary place of articulation, but rather degree of retraction of the tongue root. The primary difference between the retroflex and the other two is that the retroflex exhibited a significantly lowered F3. The palatalized variants of each rhotic introduced two predictable changes in formant structure: (1) there was a lowering of F1; and (2) there was an increase in F2 that began during the proceeding vowel and continued until the palatal release during rhotic articulation. The delayed increase in F2 is of particular interest because it is not observed with other palatalized sounds. This suggests there is some incompatibility of rhotics with secondary palatalization, and this characteristic appears to be shared by rhotic consonants in general.

The delayed increase in F2 associated with palatalized rhotics is not unique to Toda. Iskarous and Kavitskaya (2010) found that the palatalized rhotic in Russian also demonstrated a significant increase in F2 during articulation. Not only was this increase present in all word positions, but it showed significant variability depending on word position. For example, contrary to expectations, in the word-final position, palatalized rhotics exhibited an increase in the degree of the F2 raising during articulation and into the off-glade. The authors note that this result is unexpected both from a gestural coordination hypothesis (Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007) and from a C-Center hypothesis (Sproat & Fujimura, 1993; Browman
It is unclear what articulatory or perceptual function the increase in palatalization intensity in word-final position provides. Furthermore, word-medial palatalization was not weakened, as might be expected due to gestural undershoot (Lindblom, 1963). Palatalization was always fully achieved, but it is not entirely clear why it did not undergo weakening. It could be that due to the vocalic nature of secondary palatalization, it is not susceptible to boundary effects. However, Iskarous and Kavitskaya (2010) also note that it could be that palatalization weakens the trilling gesture and that weakening allows for palatalization to always be fully realized. In short, the interaction between palatalization and rhotics represents a unique conflict that does not seem to apply to other segments. Therefore, this phenomenon is examined in this dissertation as a proxy for an articulatory class characteristic. The articulation of palatalized rhotics are discussed in Section 1.5 and the acoustics and articulation of palatalized rhotics are examined directly in Chapter 2 and Chapter 4 of this dissertation through the use of acoustic and ultrasound technology.

1.4 Articulation of Rhotics

One of the primary difficulties in identifying articulatory characteristics common among rhotics stems from the wide range of tongue shapes associated with the canonical realization in a given language. For example, the North American English approximant, /ɹ/, has “retroflex” and “bunched” varieties and even variations between the two extremes of articulatory possibilities (Delattre & Freeman, 1968). Alwan, Narayanan, and Haker (1997) found that whether “retroflex” or “bunched” was used corresponds strongly to the location of the constriction, with retroflex approximants being produced further forward in the mouth and bunched varieties being produced further back. In all cases, a degree of concavity posterior to the constriction location was accompanied by a sublingual cavity, although the amount of tongue concavity was quite variable. Interestingly, amongst the variation, there are similarities that can be observed such as the presence of a pharyngeal constriction in both “retroflex” and “bunched” varieties, suggesting there may be a consistent class characteristic. The location of pharyngeal constriction was variable across varieties: more anterior rhotics had a more superior pharyngeal constriction, while more posterior rhotics had a more inferior pharyngeal constriction.

Across different phonological inventories, rhotics also have a wide range of articulatory configurations, but some of these configurations appear to be more specific to rhotics. Narayanan, Byrd, and Kaun (1999) used MRI to conduct an experiment on Tamil, which has
three rhotics: two trills, /r/ and /ɻ/, and an approximant, /ɻ/. The authors found similarity between /r/ and /ɻ/ in overall tongue concavity. There was a slight or flat degree of concavity and lateral bracing in the palatal region which created a mid-sagittal channel. The tongue shape was partially a result of the tongue-palatal contact. The retroflex rhotic, /ɻ/, had a narrow constriction in the palatal region, with no medial contact. The tongue tip was raised upward and pulled inward, creating a “pitted” (p. 1996) cavity in the middle of the tongue body. This articulation was supported by lateral tongue bracing. The retroflex rhotic was distinguished from the other two rhotics by a much larger back-cavity volume and the “pitted” (p. 1996) tongue shape. Narayanan, Byrd, and Kaun (1999) suggest that there is a phonetic correlate to rhotic articulations due to the similarity in tongue shapes across the rhotics examined in their study.

Most of the previous studies have focused on apical rhotics (trills or taps) or the English approximant /ɹ/ (e.g. Delattre & Freeman, 1968; Tiede, Boyce, Holland, & Choe, 2004; Proctor, 2011; Van Lieshout, Merrick, & Goldstein, 2008) and as a result uvular rhotics have been underrepresented in the articulatory literature. However, Delattre (1971) performed a landmark x-ray study on French and German uvular rhotics and the English alveolar rhotic, making interesting observations from the X-ray tracings. Both the French and German rhotics were characterized by a circular motion to the retraction of the tongue root towards the point of maximum constriction (evidenced in Figure 3 for French). This caused a secondary pharyngeal constriction. The tongue root and then raised and fronted slightly to achieve the maximum point of displacement of the tongue body. Figure 3 presents the x-ray tracings for French /r/. It can be observed that as time proceeds (images from left to right) there is a measurable pharyngeal constriction. The presence of a pharyngeal constriction in all the rhotics examined by Delattre (1971) might suggest a consistent class feature.

Figure 3. Intervocalic /r/ in French (Delattre, 1971: p. 146).
Current articulatory work has also begun to find commonalities between the gestural coordination of rhotics, which may contribute to class membership. Boyce, Hamilton, and Rivera-Campos (2016) explored the possibility of the presence of a secondary pharyngeal constriction as a correlate to rhotic articulation by examining rhotics in English, Malayalam, Spanish, and French. They point out that previous clinical research has indicated one of the articulatory difficulties that child learners have in acquiring a rhotics is the inability to coordinate the tongue tip gesture with a secondary pharyngeal constriction (Hamilton, Rivera-Campos, Rivera Perez, Boyce, McNiell, & Schmidlin, 2012; Klein, McAllister Byun, Davidson, & Grigos, 2013). This was best observed in the lack of a pharyngeal constriction for children with speech acquisition disorders compared to typically developing children. Boyce et al.’s (2016) study also suggested that there is a consistent pharyngeal constriction in the languages studied. The location of the constriction varied by language and rhotic articulated, but the consistent presence of a constriction in the pharynx produced by a tongue root gesture is argued by the authors to be a class feature for rhotics. However, their study was based on a rather limited sample (single speakers of each language), and, therefore, their proposal requires further investigation.

In a related line of research, Recasens and Espinosa (2007) and Proctor (2011) suggest that rhotics have a strong tongue root component, which is why the root is highly constrained across environments. These findings are congruent with models of rhotics that postulate that these segments consist of a tongue tip gesture coordinated with a tongue root gesture (Delattre & Freeman, 1968). Proctor (2011) posits that the complex coordination of the tongue gestures accounts for the phonological behaviour of rhotics in Russian and Spanish. For instance, vocalization is the result of lenition of the tongue tip gesture, while liquid metathesis in Proto-Slavic (Carlton, 1990) is described as a reorganization of the gestural timing between the rhotic and the vowel. In both onset and coda positions, a uvular-pharyngeal tongue body gesture is involved for the entire duration, but metathesis occurs when the timing of coordination with the tongue tip gesture is changed from the 180° to 0°. This line of research suggests that there is a common articulatory setting for rhotics that effects their phonological behaviour.

Rhotics (and laterals) also share a commonality in intergestural timing across syllable positions. Gick, Campbell, Oh, and Tamburri-Watt (2006) conducted a cross-linguistic examination of gestural timing among liquids in six languages (Canadian English, Quebec French, Serbo-Croatian, Korean, Beijing Mandarin, and Squamish Salish) in three word
positions (syllable onset, intervocalic, coda), although the rhotics were limited to Mandarin and Korean. While there was great variation in syllable patterns and gestural timing, four cross-linguistic tendencies were identified: (1) there was a measurable dorsal constriction for postvocalic rhotics; (2) there was an asymmetrical relationship between the magnitude and gestural timing of rhotics in pre- and postvocalic positions; (3) there were multiple simultaneous gestures in intervocalic positions; (4) intergestural timing occurred in such a way that the less anterior, vowel-like dorsal gestures occurred closer to the syllable nucleus and more anterior, consonant-like tongue tip gestures occur more peripherally. Gick et al.’s (2006) research thus suggests that the rhotics tested here have shared articulatory settings and gestural coordination that effects their phonotactic distribution and intergestural timing.

Trills are a subgroup of rhotics, which have tightly constrained articulatory configurations. Trills rely on aerodynamic factors to cause vibration of the active articulator against a passive articulator (Catford, 1977; McGowan, 1992; Ladefoged & Maddieson, 1996) which sets them apart from taps/flaps which involve muscular activation to create the tongue-palate contact (Spajic et al., 1996; Ladefoged & Maddieson, 1996). In order to facilitate trilling (for tongue tip articulations), the tongue body must be stiffened, while maintaining a flexible tongue tip. Fine control over the tongue body and tongue tip stiffness is achieved via lateral tongue bracing (McGowan, 1992). The tongue-palate closure causes pressure to build up behind the tongue tip, which results in it opening. The opening of the tongue-palate closure is followed by a pressure drop which causes the flexible articulator to be sucked back into the contact position via a Bernoulli effect (Solé, 2002). The aerodynamic forces required to produce trills also result in wide variation in their phonetic realization. Solé (2002) showed that small variations in air pressure – which was measured by the displacement of H₂O using a hydrostat – resulted in vastly different realizations of apical trills. Air pressure between 2.5-3.5 cm H₂O resulted in loss of trilling and can result in approximant productions. Furthermore, if the tongue-palate configuration produces imperfect articulatory positioning, tongue tension, and mass, trills may alternate with fricatives when there is not enough oral-pharyngeal pressure to permit trilling. Solé’s (2002) analysis explains the cross-linguistic tendency for synchronic variation and diachronic changes from trills to fricatives. These phonological processes are precisely due to failure to achieve correct tongue-palate configuration and airflow. A slight deviation from the configuration for trilling will result in a fricative production. The tight constraints on both tongue
position and airflow make trills difficult to produce and easy to fail. Due to the ease of fricative productions are common cross-linguistic variants and are often the result of historical sound changes and allophonic variation.

The previous research provides evidence for articulatory similarities in tongue shape (Narayanan et al., 1999), presence of a secondary pharyngeal constriction (Delattre, 1971; Boyce et al., 2016), and intergestural coordination (Gick et al., 2006; Proctor, 2011). To account for this generalization, Proctor (2011) suggest that rhotics consist of the coordination of a tongue body and a tongue tip gesture. The posterior tongue gesture is primarily realized as a pharyngeal constriction accompanied by a tongue tip constriction. This holds true across places of articulation and – as far as has been examined – the different manners of articulation as well. Despite having different places of articulation, the rhotics all seem to be of a similar tongue shape configuration: there is concavity formed by the combination of a tongue tip and tongue body gesture. There are strong constraints imposed on rhotic production despite manner differences, although the constraints imposed also involve a strong requirement for tongue tip and tongue root articulations to be persevered despite environmental conditions. In trills, for example, the tight constraint on articulation manifests itself as extremely precise positioning of both the tongue root and the tongue tip to facilitate the type of aerodynamics necessary for trilling to occur. Finally, the complex coordination of the tongue tip and the tongue root has a timing pattern dependent on syllable position. In the initial position, the tongue tip gesture is typically realized first, while in the coda, the tongue root gesture is typically realized first. In the intervocalic position, however, the two gestures are typically realized at the same time (Gick et al., 2006).

1.5 Articulation of Palatalized Rhotics

As mentioned in Section 1.3, palatalized rhotics are typologically uncommon, being avoided both synchronically and diachronically (Maddieson, 1984; Hall, 2000; Bateman, 2011). However, it is not well understood why this dispreference should exist for rhotics and not other classes. Studying the articulation of palatalized rhotics is therefore important for understanding both the articulatory correlates to rhotic articulation and to uncovering the reason for phonological avoidance of palatalized rhotics.

Kochetov (2005) used electromagnetic articulography (EMA) to examine the difference in the realization of palatalization across the class of liquids in Russian, /l/, /l̥/ /l/, and /l̊/. The
data revealed a striking contrast in the articulation of the palatalization gesture for laterals vs. rhotics: laterals exhibited nearly simultaneous realization of palatalization and the tongue tip closure.

However, rhotic articulation caused a delay on the tongue body gesture associated with secondary palatalization. Furthermore, the tongue tip gesture was fronted for the palatalized rhotic, while it was retracted for the palatalized lateral (compared to the unpalatalized counterpart). The results suggest that there is a conflict in the posterior region of the tongue due to the delay in fronting for secondary palatalization. The delay may indicate that tongue retraction is required for rhotic articulation.

Stoll, Harrington, and Hoole (2015) also performed an EMA examination of Russian unpalatalized and palatalized liquids and found a similar delay in the tongue body gesture for the rhotics when compared to the laterals. However, this distinction was variable depending on both speech rate and word position for the rhotics (but not the laterals). The lag between the tongue tip and tongue body gestures was the greatest in word-initial position at a slow speech rate. However, in intervocalic position the gestural timing was nearly simultaneous. Stoll et al. (2015) argue the differences in realization of secondary palatalization was due to perceptual recoverability. In word-initial position, palatalization can only be perceived on C-to-V transitions, but intervocally, palatalization can be perceived on V-to-C and C-to-V transitions. This also raises the question of whether or not trilling and palatalization are actually incompatible as previous research has suggested (e.g. Kavitskaya, Iskarous, Noiray, and Proctor, 2009). This leaves open the possibility that the conflict between trilling and palatalization is not actually related to articulatory constraints but is rather rooted in perceptual constraints.

In summary, rhotics, unlike other segments, display a dispreference for secondary palatalization. The dispreference manifests itself as a delay between the onset of rhotic articulation and the realization of secondary palatalization. This quality of rhotics suggests an articulatory similarity across different rhotics, and, given this, the postulation of a natural class of rhotics is justified.

1.6 Rhotic Perception
In studies of rhotics, relatively little has been done to identify precise acoustic and perceptual properties of these sounds, linking these two together (c.f. Lindau, 1985). It has therefore been difficult to understand what exactly listeners are perceiving when they are hearing r-sounds.
However, if the articulatory evidence presented in Section 1.4 is correct, we would anticipate a strong acoustic-perceptual correlate.

Engstrand, Frid, and Lindblom (2007) performed an identification and discrimination task with 22 native speakers of Swedish attempting to find a characteristic that links alveolar and uvular rhotics (which are dialectal variants of a single rhotic phoneme) together in the perceptual space. Five of the speakers produced an [ɾ] in their own speech and the remaining 17 produced an [r] in their own speech. The perceptual experiment was carried out using synthetically manipulated [r] and [ɾ] productions to create a continuum of 8 steps. The difference between F2 and F3 was manipulated in 8 equal steps ranging from the typically F2 and F3 for a prototypical [r] to the final step, the F2 and F3 for a prototypical [ɾ]. Although there was a clear “categoric” tendency, even the most “front” of productions were only identified correctly approximately 60% of the time. The same was true for the most “back” productions. For the discrimination task, none of the pairs were discriminated reliably as almost all the comparisons failed to breach chance levels of identification. Furthermore, the discrimination experiment revealed a complete lack of categorical perceptual boundaries between uvular and alveolar rhotics. Engstrand et al. (2007) also observed that listeners’ perception was influenced by their own production. Speakers with alveolar productions were more likely to identify or discriminate sounds with a front bias, while the opposite was true for speakers with a back production. The results overwhelmingly suggest that rhotics at different places of articulation are linked by a distinct acoustic-perceptual characteristic.

Engstrand et al.’s (2007) finding may be somewhat surprising because lowered F3 is often described as the main distinguishing feature for rhotic segments (Fujimura & Erickson, 1997; Espy-Wilson, Boyce, Jackson, Narayanan, & Alwan, 2000; Ladefoged, 2003). We might expect differences in F3 to help distinguish rhotics from other classes of sounds; however, this prediction was not supported in perceptual studies. Heselwood (2009) conducted a perception experiment with native English listeners where F3 had been completely filtered out. The results indicated that speakers preferred tokens with no F3 and categorized them as more rhotic. Heselwood (2009) suggested that the low F3 found in English rhotics actually acts as a suppressant and enhances F2. F3 is lowered to such an extend that it almost merges with F2 acting to intensify its perception. Based on Heselwood’s (2009) findings, F2 is possibly the strongest perceptual correlate to rhotic classhood.
Rhotics also exhibit strong coarticulatory effects on the surrounding environments and are generally resistant to these effects themselves (Recasens & Pallarès, 1997; 1999). The tendency for rhotics to exert coarticulatory effects on the surrounding environment also manifests itself in the listeners’ perception. West (1999) conducted a perceptual study comparing English minimal pairs containing /l/ and /ɹ/. The experimental design involved replacing both the /l/ and /ɹ/ and surrounding vowels and consonants with noise. In conditions where /ɹ/ and the adjacent vowels were replaced with noise, /ɹ/ was still found to be identifiable above chance. The high perceptual recovery for the rhotic suggested that /ɹ/ has long-distance coarticulatory effects that are clearly perceptible to listeners. This differed from /l/, which was found not to show the same types of long-distance perception when significant portions of the surrounding vocalic environment were replaced with noise. This suggests a key difference between /l/ and /ɹ/, in that /ɹ/ exhibits long-distance coarticulatory affects, while /l/ is relatively local in its effects. West (1999) describes the long-distance effect as a global F3 lowering; however, rhotics have also been shown to have an “r-lowering” effect on F2 (Hawkins & Slater, 1994), suggesting that the long-distance acoustic/perceptual cues for rhotics may be deeply rooted in F2. It is also interesting to note that this probably correlates to some articulatory factor because long-distance effects such as these have been shown to be largely planned rather than the result of automatic processes (Whalen, 1990). It is likely that a specific articulatory property of rhotics may in fact be present and causing these coarticulatory effects and the long-distance lowering of F2 and F3. The F2 lowering is most likely connected to the half-wavelength resonance of the posterior cavity or a Helmholtz resonance for a pharyngeal constriction involved in the production of rhotics (Alwan et al., 1997). Taken together, the data suggests an acoustic/perceptual cue to rhotics as a class, which is probably linked to an articulatory property of rhotics due to a relationship between articulatory gestures and acoustic outputs (Al-Tamimi, 2017).

1.7 Theoretical Framework

Increasing attention has been paid to speech dynamics and phonetic details in theories of phonology, due to recent work that has shown the need to incorporate dynamics and phonetics into phonological theories to account for observable phenomena. In one of the earlier works addressing this question, Lisker (1974) demonstrated that the temporal structure of speech plays a crucial role in phonological processes. For example, it is only possible to contrast voiced and voiceless stops in the word final position because of temporal differences in their production.
Lisker (1974) also showed that the temporal organization of the articulators varies from language to language, which inhibits the use of predictable physiological laws. Therefore, there has been an increasing interest in including speech dynamics into phonological theory (although not all phonologists agree with the necessity to do this). While there have been numerous revisions to phonological theory (Anderson, 1974; Ewen, 1982; Prince & Smolensky, 2004), few theoretical approaches have incorporated a strong link between phonology and phonetics (but see Hayes, Kirchner, & Steriade, 2004), and even fewer have incorporated temporal information into phonological representations or constraints (see Gafos, 2002).

Browman and Goldstein (1986; 1989; 1992; 1993; 1995; 2000) developed a theory of Articulatory Phonology, which accounts for temporal and dynamic changes in speech, while maintaining a strong link between phonology and the physical aspects of speech. The authors discard the notion of featural representation, instead implementing a system which provides a direct, explicit description of temporal and dynamic articulatory movements.

Articulatory Phonology treats gestures as the primitive phonological units. Gestures are modeled as the actions of the vocal tract articulators and are identified by a formation of a constriction in an area of the vocal tract. Gestures are produced with sets of independent articulators: the jaw, the lips, the teeth, the tongue tip/blade, the tongue body, the tongue root, the velum, and the glottis. Gestures encode in them both the movement and through space and time. Gestures are combined – as the basic building block of phonology – to form phonological structures, or “constellations” (Browman & Goldstein, 1989, p. 201). The articulators are combined with vocal tract variables to produce different constriction locations and degrees (CL and CD, respectively). Constriction locations can be anywhere in the vocal tract, from the lips (e.g. lip protrusion) to the glottal aperture. The degree of constriction is also encoded in the system to distinguish between different manners of articulation. [clo(sed)] refers to a constriction that is characterize by a complete closure, while [crit(ical)] refers to a constriction degree tight enough to produce frication. Constriction degrees used for sonorants can be described under the broader constriction degree [open], which can be broken into three different degrees, [nar(row)], [mid], [wide], in order of how tight the constriction is. Table 1 summarizes the tract variables and the active articulators. Table 2 summarizes the active articulators and the possible constriction degrees.
Table 1. Summary of the vocal tract variables and articulators

<table>
<thead>
<tr>
<th>Vocal Tract Variables</th>
<th>Model Articulators</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP</td>
<td>lip protrusion</td>
</tr>
<tr>
<td></td>
<td>upper and lower lips</td>
</tr>
<tr>
<td>LA</td>
<td>lip aperture</td>
</tr>
<tr>
<td></td>
<td>upper and lower lips, jaw</td>
</tr>
<tr>
<td>LTH</td>
<td>lower tooth height</td>
</tr>
<tr>
<td></td>
<td>jaw</td>
</tr>
<tr>
<td>TTCL</td>
<td>tongue tip constriction location</td>
</tr>
<tr>
<td></td>
<td>tongue tip, tongue body, jaw</td>
</tr>
<tr>
<td>TTCD</td>
<td>tongue tip constriction degree</td>
</tr>
<tr>
<td></td>
<td>tongue tip, tongue body, jaw</td>
</tr>
<tr>
<td>TBCL</td>
<td>tongue body closure location</td>
</tr>
<tr>
<td></td>
<td>tongue body, jaw</td>
</tr>
<tr>
<td>TBCD</td>
<td>tongue body closure degree</td>
</tr>
<tr>
<td></td>
<td>tongue body, jaw</td>
</tr>
<tr>
<td>TRCL</td>
<td>tongue root closure location</td>
</tr>
<tr>
<td></td>
<td>tongue root, jaw</td>
</tr>
<tr>
<td>TRCD</td>
<td>tongue root closure degree</td>
</tr>
<tr>
<td></td>
<td>tongue root, jaw</td>
</tr>
<tr>
<td>VEL</td>
<td>velic aperture</td>
</tr>
<tr>
<td></td>
<td>velum</td>
</tr>
<tr>
<td>GLO</td>
<td>glottal aperture</td>
</tr>
<tr>
<td></td>
<td>glottal width</td>
</tr>
</tbody>
</table>

Table 2. Summary of possible constriction degrees for each articulator

<table>
<thead>
<tr>
<th>Articulator</th>
<th>Constriction Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lips</td>
<td>[clo]</td>
</tr>
<tr>
<td></td>
<td>[crit]</td>
</tr>
<tr>
<td></td>
<td>[nar]</td>
</tr>
<tr>
<td>Lower Teeth</td>
<td>[clo]</td>
</tr>
<tr>
<td></td>
<td>[crit]</td>
</tr>
<tr>
<td></td>
<td>[nar]</td>
</tr>
<tr>
<td>Tongue Tip</td>
<td>[clo]</td>
</tr>
<tr>
<td></td>
<td>[crit]</td>
</tr>
<tr>
<td></td>
<td>[nar]</td>
</tr>
<tr>
<td></td>
<td>[mid]</td>
</tr>
<tr>
<td></td>
<td>[wide]</td>
</tr>
<tr>
<td>Tongue Body</td>
<td>[clo]</td>
</tr>
<tr>
<td></td>
<td>[crit]</td>
</tr>
<tr>
<td></td>
<td>[nar]</td>
</tr>
<tr>
<td></td>
<td>[mid]</td>
</tr>
<tr>
<td></td>
<td>[wide]</td>
</tr>
<tr>
<td>Tongue Root</td>
<td>[clo]</td>
</tr>
<tr>
<td></td>
<td>[crit]</td>
</tr>
<tr>
<td></td>
<td>[nar]</td>
</tr>
<tr>
<td></td>
<td>[mid]</td>
</tr>
<tr>
<td></td>
<td>[wide]</td>
</tr>
<tr>
<td>Velum</td>
<td>[clo]</td>
</tr>
<tr>
<td></td>
<td>[nar]</td>
</tr>
<tr>
<td>Glottis</td>
<td>[clo]</td>
</tr>
<tr>
<td></td>
<td>[crit]</td>
</tr>
<tr>
<td></td>
<td>[nar]</td>
</tr>
<tr>
<td></td>
<td>[mid]</td>
</tr>
<tr>
<td></td>
<td>[wide]</td>
</tr>
</tbody>
</table>

The vocal tract variables and model articulators are combined with the constriction degree to yield a gestural representation of different segments. Gestures are coordinated together in phase relationships (Browman and Goldstein, 1987; Byrd, 1996). Each gesture has an abstract
360° cycle, which represents the dynamic unfolding of the gesture over time. 0° represents the absolute onset of the gesture, while 360° represents the completion of a gesture. Browman and Goldstein (1987) also state that the point at which a gesture achieves its target is at 240° with respect to the abstract 360° in the gestural representation. The phase relationship between different gestures thus describes how different articulatory motions are coordinated together to produce speech. In an abstract relationship, segments can be represented in a 0° relationship to each other (see Proctor, 2011 for examples); in other words the onset, achievement of target, and offset are in synch with each other. However, in online speech production, articulatory cycles interact and overlap. A gestural score can be used to model the dynamic phase relationship between segments, as well as coarticulatory effects they exert on each other and various connected speech processes, such as assimilation and deletion (Browman & Goldstein, 1987). A gestural score, in an abstract sense, acts like a speech plan that the articulators follow. An example gestural score for the word ‘palm’ is presented in Figure 4 (from Browman & Goldstein, 1989, p. 212).

![Gestural score showing the activation and timing of different articulators involved in the production of the word 'palm' (from Browman & Goldstein, 1989, p. 212).](image)

In Figure 4, the gestural score can be read in such a way that the vertical axis displays the articulatory sets and the horizontal axis displays time. Each gesture is marked in row which corresponds to the active articulatory set that is forming the constriction and the constriction location is marked inside the box. The x-axis corresponds to the duration of the gesture. It is also important to note that gestures can and do overlap, and that the entire box represents the onset
(the movement towards the gestural target), the point of maximum constriction (the maximum
displacement of the gestural target) and the offset (the movement away from the gestural target).
In this way, dynamic movement over time are captured and their interactions can more
accurately be represented, at the expense of economy.

To summarize, Articulatory Phonology is a departure from abstract phonological features
and purely theoretical approaches to phonology; it focuses primarily on how the articulatory
characteristics of different segments explain their phonological behaviour. The smallest unit of
Articulatory Phonology is a gesture and differs from feature-based theories because each gesture
is considered a dynamic movement towards a target. However, not all movements of the tongues
are considered gestures, only movements of the tongue that play a phonological role
(functionally meaningful) in production are considered a gesture under Articulatory Phonology.
Thus, the primary goal is to explain phonology in terms of different articulatory interactions.

Articulatory Phonology has been successful in accounting for a range of phonological
processes (see Proctor, 2011 for a discussion of liquid metathesis, for example) and made
substantial progress in capturing the phonological behaviour of rhotics. Despite this work, the
question of gestural representations of rhotics as a class has not been addressed in a systematic
way (although see Proctor’s 2011 work on representations for liquids, including apical rhotics).
It seems, however, that it is possible that rhotics in general (both apical and uvular) can be linked
together as a class through the presence of a tongue root gesture.

In this dissertation, the experimental results will be examined through the lens of
Articulatory Phonology with the express purpose of formulating a gestural representation for
rhotics that accounts for the phonological behaviour observed in the data.

1.8 Organization of the Dissertation
To investigate the common articulatory, acoustic, and perceptual correlates of rhotic consonants,
the dissertation is organized as follows. Chapter 2 presents an acoustic examination of Upper and
Lower Sorbian. Upper and Lower Sorbian have phonemic unpalatalized and palatalized rhotic
pairs, /r, ř/ and /r, ř/, respectively. The comparison of unpalatalized and palatalized uvular and
alveolar rhotics provides an important case study because it allows for a systematic approach to
both the question of rhotics as a class and the dispreference of secondary palatalization and
rhotics. The questions of rhotic classhood and the dispreference for rhotics and palatalization are
approached through a dynamic acoustic analysis of the formants for rhotics and laterals to
examine: (1) do rhotics of different places of articulation share acoustic commonalities; and (2) do rhotics of different places show the same dispreference for palatalization. Implications for articulatory causes of observed acoustics are discussed.

In Chapter 3, underlying articulatory causes for the acoustic patterns observed for the unpalatalized uvular rhotic in Upper Sorbian in Chapter 2 are examined. Ultrasound technology is utilized to examine the coarticulatory patterns of the uvular rhotic, /ᵝᵝ/, with the direct purpose of uncovering the gestures relevant to its articulation. This chapter is aimed at addressing the rarity of articulatory studies in the literature that explore uvular rhotics and determine if there is a link between the articulation of uvular and alveolar rhotics.

The articulatory examination of the unpalatalized and palatalized rhotics in Upper Sorbian, is presented in Chapter 4. The aim of this chapter is to address the articulatory asymmetries that are hypothesized to explain the dispreference for rhotics and secondary palatalization. A dynamic comparison of /ᵝᵝ/ and /ᵝ()/ using ultrasound technology was employed for this aim. A direct articulatory comparison of how the two segments unfold over time can provide significant insight into how the gestural representation proposed in the previous chapters conflicts with the articulatory requirements for secondary palatalization.

Chapter 5 contains an ultrasound examination of Brazilian Portuguese. This language was selected because of the large variation in the production of coda rhotics and the presence of a phonemic uvular fricative, /ᵝᵝ/, and an alveolar tap, /ɾ/. The aim for this chapter is two-fold: (1) to see if the gestural characterization of the uvular rhotics in Chapter 3 extends to Brazilian Portuguese rhotics; and (2) to see if the gestural characterization accounts for the variation observed in the language.

Chapter 6 examines the perception of non-native rhotics by native speakers of Canadian English. The aim of this chapter is to determine if the underlying articulatory similarities discovered across rhotics in the previous chapters have an acoustic-perceptual correlate. To this end, a discrimination task was employed to examine whether English speakers can reliably perceive rhotics of different place and manner configurations. Three rhotics, /ᵝ/, /ɻ/, and /ᵝ()/, were compared against other non-native sounds from four other natural classes (stops, nasals, fricatives, and laterals) to determine if the rhotics group together in the perceptual space. Given the hypothesized articulatory link between rhotics of different place and manner configurations, a systematic acoustic-perceptual correlate is anticipated.
Finally, Chapter 7 provides a broad theoretical discussion of how the facts presented in previous chapters bear on the hypotheses tested. Specifically, this chapter explores gestural representations of rhotics in Sorbian and Brazilian Portuguese, accounting for the phonetic findings and phonological representations of these segments. The chapter also addresses broader questions about rhotics as a class, while also addressing limitations of the current studies and issues for further research.
Chapter 2
Rhotics and Palatalization: An Acoustic Examination of Upper and Lower Sorbian

Contents of this chapter have been published in *Phonetica*, Karger Publishers:


A link to the published paper can be found at:
[https://www.karger.com/Article/Abstract/481783](https://www.karger.com/Article/Abstract/481783)
2.1 Introduction

Sorbian consists of two west Slavic languages, Upper and Lower Sorbian, and is spoken in Eastern Germany. Upper Sorbian is spoken in the south of the area known as Łužica, within the state of Saxony. Figure 5 shows a map of the Upper Sorbian speaking area within Germany.

Figure 5. Upper Sorbian is spoken south of the solid black line, with the cultural capital of Budyšin. Image taken from Howson (2017).

Upper Sorbian has a total of 29 consonants and is special among the Slavic languages because it has adopted the uvular trill and, as a result, has a typologically rare palatalized uvular trill, /ʁj/ (Schaarschmidt, 1998; Howson, 2017). However, both are often produced as an approximant [ʁ] (e.g. rat [ʁat] ‘good’) and [ʁj] (e.g. rjadka [ʁjatka] ‘line, verse’). /ʁj/ is also frequently realized as either a voiced or a voiceless uvular fricative, [ʁ] (e.g. rat [ʁat] ‘gladly’) or [χ] (e.g. rat [χat] ‘gladly’), respectively (Howson, 2017). /ʁ/ and /ʁj/ are more likely to be realized as a trill in clear and careful pronunciation (Jocz, 2013). The palatalized rhotic is also occasionally dropped from speech completely, leaving only a palatal glide (e.g. rjadka [jatka] ‘line, verse’). The rhotic, /ʁ/ undergoes vocalization in coda positions (e.g. por [por] ‘couple, pair’), while /ʁj/ does not appear in coda positions. Consonants are palatalized before /i/ (e.g. ličak [liʧakʰ] ‘calculator’), /e/ (e.g. wón so lehnje [wuʰn so lehniə] ‘(he) lies’), and /ɪ/ (e.g. bělić so [bʲlʲiʧ so] ‘to peel’), while unpalatalized consonants appear before /i/ /e/ (Howson, 2017). However, Jocz (2013) notes that palatalization for all consonants before /i/, /e/, and /ɪ/ is becoming optional. Table 3 shows a consonant chart for Upper Sorbian.
**Table 3. Consonant inventory for Upper Sorbian.** Hard indicates unpalatalized and soft indicates palatalized consonants (Howson, 2017).

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hard</td>
<td>Soft</td>
<td>Hard</td>
<td>Soft</td>
<td>Hard</td>
<td>Soft</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plosive</td>
<td>p</td>
<td>b</td>
<td>pʲ</td>
<td>bʲ</td>
<td>t</td>
<td>d</td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td>Affricate</td>
<td>ts</td>
<td>dfish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>mʲ</td>
<td>n</td>
<td>nʲ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>r</td>
<td>rʲ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f v s z</td>
<td>j 3</td>
<td>x</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>wʲ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
</tr>
</tbody>
</table>

Lower Sorbian is spoken in the north of the area known as Łužica, within the state of Brandenburg and is primarily only spoken by the grandparent generation (Elle, 2003). Figure 6 shows a map of the Lower Sorbian speaking area.

*Figure 6. Map of the Lower Sorbian speaking region. The cultural capital is in Chóšebsz.*

Lower Sorbian has 39 consonants and has retained the alveolar rhotic and has an unpalatalized and palatalized pair, /r/ and /rʲ/ (Stone, 1993). Unlike Upper Sorbian, Lower Sorbian has a three-way sibilant contrast (dental, retroflex, and alveolopalatal) and palatalization of consonants before /i/, /e/, and /ɪ/ is not optional. Table 4 shows the consonant inventory of Lower Sorbian.
Table 4. Consonant inventory for Lower Sorbian. Hard indicates unpalatalized consonants and soft indicates palatalized consonants.

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Alveolar</th>
<th>Retroflex</th>
<th>Alveolo-palatal</th>
<th>Palatal</th>
<th>Velar</th>
<th>Laryngeal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p b</td>
<td>t d</td>
<td>d\textsuperscript{3}</td>
<td>t s dz</td>
<td>t s dz\textsuperscript{3}</td>
<td>t c</td>
<td>k\textsuperscript{3}</td>
<td>k\textsuperscript{3}</td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m m\textsuperscript{3}</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td></td>
<td>r</td>
<td>r\textsuperscript{3}</td>
<td>f v s z</td>
<td>s z</td>
<td>c z</td>
<td>x</td>
<td>h</td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximate</td>
<td>w w\textsuperscript{3}</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Upper and Lower Sorbian were chosen because they have unpalatalized and palatalized rhotics at the uvular and alveolar place, respectively. This allows an examination of the formant frequencies in order to approach two of the larger issues surrounding rhotics that were presented in the chapter 1: (1) what acoustic-articulatory characteristic, if any, unites the rhotics as a class? and (2) why is there a dispreference for rhotics and palatalization?

2.2 Hypotheses

There are two hypotheses that this paper will test based on the previous literature. (1) There will be a discernible delay in the peak F2 achievement for both apical and uvular palatalized rhotics; and (2) Despite being different places of articulation, both the alveolar and uvular rhotics have similar formant frequencies and trajectories for F1, F2 and F2-F1. F3 is anticipated to be different for uvular and alveolar rhotics.

For hypothesis (1), the expectation for a delay in an increase in F2 as a result of secondary palatalization for rhotics follows other studies which have correlated the two together (e.g. Spajić et al., 1996; Kochetov, 2005). Thus, a delay in the peak of F2 should obtain for both alveolar and uvular rhotics. Kochetov (2005) and Stoll et al. (2015) have also correlated the delayed increase in F2 observed in rhotics with a delay in the palatalization gesture for rhotic segments. Therefore, a delay in F2 for Sorbian rhotics is likely to correlate with a delay in the palatalization gesture.

Hypothesis (1) is rooted in previous acoustic and articulatory work which suggests that secondary palatalization and rhotics involve inherently antagonistic gestures. Spajić et al. (1996) and Iskarous and Kavitskaya (2010) both found a delayed increase in F2 for palatalized rhotics (in Toda and Russian, respectively). This delay has been connected specifically to rhotics as
opposed to other palatalized segments (Kochetov, 2005). This is likely the acoustic consequence of an inherent instability in the production of palatalized rhotics. Kavitskaya et al. (2009) suggest that this instability originates with an inherent conflict between the tongue root during rhotic production and secondary palatalization. Rhotic production requires tongue root retraction, while secondary palatalization requires tongue root fronting. If this is the case, we expect to see a delay in the increase of F2 and F2-F1 when secondary palatalization occurs with rhotics. The hypothesis relates to the vocalic tongue root gesture, which Proctor (2009, 2011) posits as one of the primary components of rhotic production. This results in an inherent conflict in the target and dynamics between rhotic production and secondary palatalization.

Hypothesis (2) also follows from previous acoustic and articulatory work. Heselwood (2009) performed a perceptual identification experiment which consisted of the English rhotic with F3 removed in a variety of words. The results showed that participants were better at identifying the rhotic without the presence of F3 than with it. The authors suggest that F2 may be the primary acoustic correlate to rhotic classhood (in English). Engstrand, Frid, and Lindblom (2007) also performed a perceptual study of rhotics, focusing instead on alveolar and uvular variants. The experiment manipulated the F2 and F3 formant structure on a continuum from prototypically alveolar to prototypically uvular for Swedish speech. Swedish listeners participated in the study and the results for the discrimination task indicated that correct identification as alveolar or uvular was rarely above chance at any point of the continuum. This suggests an overall similarity in formant structure between the alveolar and uvular rhotics.

Previous articulatory data also suggests an articulatory similarity across places of articulation. Narayanan at al. (1999) suggested that the coronal rhotics in Tamil are all united by a similar concave tongue shape. While the uvular rhotic has not been examined specifically in this context, the expectation is that there are strong similarities in overall vocal tract shape between the alveolar and uvular rhotics. This should be observed in the acoustics as an overall similarity between the formant dynamics and frequencies, most specifically F1, F2, and F2-F1. If a similarity in the formants and trajectories is found, this would suggest a corresponding similarity in the vocal tract shape created by the tongue root, specifically in the area of the soft palate and the pharynx.

Proctor (2009, 2011) suggests that rhotics are linked by the coordination of the tongue root and tongue tip gestures. If Proctor’s (2009, 2011) analysis is on the correct path, it is
expected that there should be striking similarities between the formants and their dynamics for both the uvular and alveolar rhotic. Furthermore, if Narayanan et al.’s (1999) analysis of Tamil rhotics extends to other languages and places of articulation, we would expect to see similar acoustics between the uvular and alveolar rhotic. The assumption here is essentially that a uvular rhotic is composed of a strong dorsal component, like the alveolars, but lacks the coronal gesture. The dorsal component accompanied by a lack of a coronal gesture is still expected to create a similar vocal tract shape which would be reflected in the acoustics. Specifically, the expectations are that F2 and F2-F1 should share similar formant frequencies and trajectories between the two places of articulation as they correlate most strongly to the activity of the tongue root (especially F2-F1, see Sproat and Fujimura, 1993).

2.3 Methods
2.3.1 Speakers
Six native speakers of Upper Sorbian (three male and three female; ages 20-26) were recorded in a sound-attenuated booth at the Max Planck Institute for Evolutionary Anthropology, in Leipzig, Germany. Three other native speakers of Upper Sorbian (three female) were also recorded in a quiet room in Bautzen. Five native speakers of Lower Sorbian (two male and three female; all aged over 70) were recorded in a quiet room in Cottbus. All locations were within Lužica, the Sorbian speaking region of eastern Germany. There was a total of eight Upper Sorbian participants and five Lower Sorbian participants (thirteen total participants).

2.3.2 Procedure
The recordings of the Upper Sorbian participants at the Max Planck Institute for Evolutionary Anthropology were all done using an Audix HT5P microphone at 44100 Hz and 32-bits. All the data recorded in the field was taken with a Fostex FR-2 LE: Field Recorder and a Lavalier AT831b microphone at 44100 Hz and 32-bits.

2.3.3 Materials
Participants repeated a corpus of real words which contained each rhotic and lateral in a word-initial position followed by the low vowel /a/. A table of these words is presented in Table 5. The Upper Sorbian data set contained words with the liquids /r, r\j, l/. Each word was produced three times by each speaker for a total of 72 tokens (3 repetitions of each word x 3 words x 8 speakers). The Lower Sorbian data set contained words with the liquids /r, r\j, l/. Each word was produced three times by each speaker for a total of 45 tokens (3 repetitions of each word x 3
words x 5 speakers). Low F3 has been described as characteristic of rhotics (Delattre & Freeman, 1968), but not laterals; therefore, /l/ was included in both data sets in order to compare the F3 of both alveolar and uvular rhotics against a lateral to determine if it does distinguish rhotics from laterals.

Table 5. The Sorbian word sets used in this study

<table>
<thead>
<tr>
<th></th>
<th>Upper Sorbian</th>
<th>Lower Sorbian</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>rad [rad]</td>
<td>-</td>
</tr>
<tr>
<td>r̊</td>
<td>rjadka [r̊atka]</td>
<td>-</td>
</tr>
<tr>
<td>r</td>
<td>-</td>
<td>rada [rada]</td>
</tr>
<tr>
<td>r̊</td>
<td>-</td>
<td>rjagotaś [r̊jagotac]</td>
</tr>
<tr>
<td>l</td>
<td>lac [lats]</td>
<td>lac [lats]</td>
</tr>
</tbody>
</table>

The total number of tokens recorded may be somewhat small, but the data presented here provides unique insight into an understudied language with little phonetic work done on it. The acoustic analysis of Upper and Lower Sorbian allows a comparison of alveolar and uvular rhotics, both of which have unpalatalized and palatalized phonemes. This allows for both documentation an endangered language and significant generalizations about rhotics and palatalized rhotics.

2.3.4 Analysis

The analysis is composed of observations from smoothing spline (SS)ANOVA measures of the formants F1 to F3 and F2-F1 for the rhotics (and the lateral /l/). SSANOVA creates mean curves based on input splines. Confidence intervals are then calculated to locate areas of statistically significant difference along the two (or more) curves (Davidson, 2006).

Dynamic measures were taken for F1 to F3 at 10 equally spaced intervals for the entire liquid up to the temporal mid-point for the following vowel (Figure 7). This created an interpolated time for the dynamic measures analysis. The formant measures were extracted using Praat (Boersma & Weenink, 2014) and were normalized using a Lobanov correction (Adank, Smits, & van Hout, 2004) with the Vowels package (Thomas & Kendall, 2007) in R (R Development Core Team, 2016). The formant measures were then compared using a SSANOVA (Gu, 2002; Davidson, 2006) in R (2016). Liquids were grouped according to dialect (Upper
Sorbian, 3 levels: /ɾ, ɾ̠, ɻ/; Lower Sorbian, 3 levels: /ɾ, ɾ̠, ɻ/). The factor for comparison was liquid identity (Upper Sorbian: /ɾ, ɾ̠, ɻ/; Lower Sorbian: /ɾ, ɾ̠, ɻ/).

Figure 7. Formant measures were taken from the onset of the liquid to the midpoint of the vowel. Image shows [ɾ̠a] from the Lower Sorbian word r̠agotaś 'to stutter.'

2.4 Results

The Upper Sorbian data is first presented, followed by the Lower Sorbian data. The results presented here show the dynamics of the formants for the liquid and the transition into the following vowel, up to the mid-point of the following vowel. The solid lines indicate the mean curve for each of the formants and F2-F1. The dotted lines surrounding the mean indicate the 95% confidence intervals. If the dotted lines for two (or more) means do not cross at a particular point, the point is considered to be a location of statistically significant difference.

2.4.1 Upper Sorbian

Figure 8 presents the results of the dynamic analysis of the Upper Sorbian liquids. The results for the dynamics of F1 to F3 are presented on the left and the results for the dynamics of F2-F1 are presented on the right.
The formant trajectories for /ɾ/ indicated a relatively steady decrease in F2 over the duration of its articulation and during the following vowel, while F1 showed a steady rise for following vowel, /a/. Both F1 and F2 for /ɾ/ were significantly different from /ɾ̄/ for the entire duration. For /l/, F1 and F2 were also significantly different from /ɾ̄/ except the difference in F1 for the vowel following both /ɾ, l/ became non-significant after the transitions for /l/. /ɾ̄/ had a lower F1 than /ɾ/ and was briefly higher than /l/ during articulation of the liquid. F2 was higher for /ɾ̄/ than /ɾ, l/, but started at the same frequency as /l/. The off-glide following /ɾ̄/ resulted in a sharp rise in F2, which was then followed by a sharp drop towards the target for /a/. F1 was lower for /l/ than both trills but showed a sharp increase for the transition to the following vowel. F2 for /l/ showed a steady decrease over the duration.

The F3 trajectory for /ɾ/ showed a brief period of difference from /l/ during articulation of the liquid, but this difference is lost during the transition into the following vowel. F3 for /ɾ̄/ was not significantly different from /ɾ/ or /l/ during the majority of the duration of the liquid; however, F3 rose sharply before dropping again after F2 began to fall, separating it from /ɾ, l/. The F3 rise was also not as dramatic as F2. F3 for /l/ was significantly different from /ɾ/ during the articulation of the liquid, but the formant transitions into the vowel caused a drop in F3, resulting in it meeting with /ɾ/, but diverging from /ɾ̄/.

The dynamics of the F2-F1 measure indicated that the trajectory of /ɾ/ was lower than /l/ and /ɾ̄/ for the entire duration until it meets with /l/ at the mid-point of the vowel. The F2-F1 trajectory for /ɾ̄/ began as high as /l/ and continued upwards for the duration of the off-glide and
then fell sharply. /ɾ/) had a large difference between F1 and F2 causing it to be higher than /ɾ/ and /l/ even by the midpoint of the vowel. /l/ began with the same difference in F2-F1 as /ɾ/ but had a different dynamic trajectory. F2-F1 for /l/ quickly began to drop in order to reach the target for /a/, causing it to meet with /ɾ/ at the mid-point of the vowel.

The formant trajectory for each of the liquids was different, which likely reflects differences in place and tongue posture. These differences exerted different effects on the vowels following each liquid. Palatalization had a particularly strong effect on the following vowel, causing differences in F1 and F2, even at the mid-point of the vowel.

2.5.2 Lower Sorbian

Figure 9 presents the results of the dynamic analysis of the Lower Sorbian liquids. The results for the dynamics of F1 to F3 are presented on the left and the results for the dynamics of F2-F1 are presented on the right.

![Figure 9. Formant (F1 to F3, left; F2-F1, right) dynamics of the Lower Sorbian liquids. The left edge indicates the onset of the liquid and the right edge of the graph indicates the mid-point of the vowel, /a/.](image)

F1 for /ɾ/ steadily rose throughout articulation and during the following vowel. It remained statistically lower than the other two liquids until much closer to the mid-point of the vowel where the F1 for all three liquids started to converge. F2 remained stable over the duration of the liquid and vowel, lowering only slightly. F1 for /ɾ/) began lower than /ɾ/ and increased over the course. It remained lower than /ɾ l/ even at the mid-point of the following vowel. F2 for /ɾ/) began higher than /ɾ/ and rose during the off-glize following the trill, and then dropped towards the mid-point of the vowel, remaining higher than /ɾ/. The F2 was higher for /l/ than /ɾ/ and remained that way over the duration. F2 is relatively stable the entire time for both /ɾ l/.
F3 for /r/ and /rɪ/ began at the same relative formant frequency, lower than for the lateral. However, /r/ and /rɪ/ quickly diverged as the off-glide for /rɪ/ causes an upwards formant trajectory. F3 for /r/ lowered, causing it to be lower than the other two liquids. The formant frequency and trajectories for F3 of the lateral was higher than the rhotics, although it takes downward trajectory, which eventually meets with /rɪ/, but remains higher than the rhotics at the midpoint of the vowel.

F2-F1 for /r/ remains significantly lower than /rɪ/ and /l/ during the entire duration and has a downward trajectory. F2-F1 for /rɪ/ began higher than /r/ and then had an upward trajectory during the off-glide. It then fell again towards the mid-point of the vowel. However, it was still higher than for the other liquids by the mid-point, suggesting palatalization causes a substantial fronting effect of the following vowel. F2-F1 for /l/ began higher than /rɪ/ but had a downward trajectory, causing it lower as far as /r/ by the mid-point of the vowel. The difference between /l/ and the rhotics remained significant for the duration of production.

2.5.3 Comparison of Upper and Lower Sorbian

Figure 10 presents direct comparisons of the unpalatalized and palatalized rhotics in Upper and Lower Sorbian. These figures are presented here in order to facilitate ease of discussion about the similarities between uvular and alveolar rhotics. Figure 10 shows mean formant frequencies for Upper Sorbian, /r/ and /rɪ/, and Lower Sorbian, /r/ and /rɪ/.

![Figure 10. Comparison of the F1, F2 and F3 for unpalatalized (left) and palatalized (right) trills in Upper and Lower Sorbian.](image)

F1 for /r/ was slightly lower than that for /r/ for the duration of the rhotic production and most of the following vowel. Both of /r ɾ/ showed an increasing trajectory for F1, until the mid-
point of the following vowel. It should be noted that while these differences are significant, the absolute value of the difference in Hz is relatively small, within approximately 100 Hz at its largest point of difference. However, F2 showed no significant difference between /r r/. The trajectories and formant frequencies overlap for the entire duration of production. F3 is a marked point of difference between /r r/. F3 for /r/ is significantly higher than /r/ throughout the duration of articulation and remains significantly higher until the midpoint of the vowel. Both /r r/ have a dropping F3 over the duration of articulation, although the actual trajectories do differ to some degree. /r/ remains high slightly longer during the rhotic, before dropping, while /r/ drops much quicker and slightly increases into the midpoint of the vowel.

Both of the palatalized rhotics, /rj rj/, show identical frequencies and trajectories for F1. However, F2 and F3 differ in absolute Hz values. Although the formant frequencies are different, the trajectories share a striking number of similarities. Both /r j/ have a delayed increase in F2 which peaks approximately the same time before dropping again into the midpoint of the vowel. F3 also has a steady increase and eventual peak close to the same point as the peak in F2. However, /r j/ has a short area of a steady F3 before it peaks and falls again.

2.5 Discussion

This chapter examined the formant frequencies and trajectories of Upper, /r r/, and Lower, /r l/, Sorbian rhotics. The goal was to see if there are observable similarity between uvular and alveolar rhotics and if there is an observable conflict between rhotics and secondary palatalization. The SSANOVA analysis revealed that there was a minor (~50 Hz) difference in F1 between the two unpalatalized rhotics, /r/ and /l/, but there was no significant difference between F2 formants or trajectories for both unpalatalized rhotics, /r/ and /l/. A major point of difference between the two rhotics was discovered to be F3, where /r/ had a significantly higher F3 than /l/. The examination of the palatalized rhotics, /rj/and /l j/, overlapping formant frequencies and trajectories for F1. The frequencies for F2 were significantly different from each other, but the trajectories were the same: there was an increase in F2 over the duration until the transition into the following vowel, where F2 fell again. F3 followed a similar pattern as F2.

The overlapping F2 frequency and trajectory suggests there is a similarity in the vocal tract shapes, specifically with respect to tongue root retraction into the pharyngeal cavity for both uvular and alveolar unpalatalized rhotics. A pharyngeal constriction would cause a systematic lowering of F2 (Fant, 1960), which has also been observed by West (1999) for English rhotics.
This may also reveal why there is a dispreference for rhotics and palatalization. The F2 for both the uvular and alveolar palatalized rhotics were characterized by a steady increase in F2 up to a peak before the off-glide. This would suggest that there is an articulatory conflict between the tongue root requirements for palatalization and rhotic articulation (Kavitskaya et al., 2009). The tongue root must retract for rhotic articulation, but it also needs to advance to produce the characteristic tongue body raising associated with secondary palatalization.

Further investigation into the articulation of both uvular and alveolar rhotics is done in Chapter 3 and 4 to address these findings more directly. Discussion of the theoretical implications of the acoustic results are discussed in Chapter 7, Section 7.1.
Chapter 3
Rhotic articulation involves tongue root retraction: evidence from Upper Sorbian

3.1 Introduction
In the previous chapter, I examined the formant frequencies and trajectories of the uvular and alveolar rhotics in Upper and Lower Sorbian. Taken together, there was a striking similarity in the formant frequencies and trajectories of F2 for the unpalatalized rhotics. While the palatalized rhotics had different frequency ranges for the uvular and alveolar places of articulation, the trajectories were similar.

In this chapter, I examine the articulatory properties that may lead to the similarity in F2 for the uvular rhotics through ultrasound measures. The aim is to uncover an underlying gestural representation that unites rhotics as a class.

3.1.1 Hypothesis
Based on the acoustic research, this study aims to test the following hypothesis:

Rhotic segments are unified by the coordination of a consonantal tongue gesture and vocalic tongue root gesture (see Chapter 2).

The hypothesis above addresses one of the primary issues within the current research body on rhotics: what is the underlying gestural representation of rhotics? The prediction for this representation of rhotics is that there are two gestures associated with their articulation. Observation of these gestures should be evident in coarticulatory effects; specifically, high degrees of resistance to coarticulatory effects from the surrounding phonetic environment. Rhotics should show effects that other, non-complex segments such as stops or fricatives, do not exhibit. To this end, ultrasound data is being employed to examine both the major gestures involved in rhotics production.

The hypothesis that uvular rhotics are composed of a tongue body and tongue root gesture is tested using ultrasound with measurements at the point of maximum constriction. The expectation is that at the point of maximum constriction consistent gestural targets should be observed. This expectation is rooted in the hypothesis that the composition of a retracted tongue root and tongue body gesture. Therefore, there should be significant overlap in the tongue body
and root regardless of phonetic environment. This can be directly observed through confidence intervals overlapping in different environments, indicated that the tongue root and body gestures are present and resistant to differing articulatory demands. This is predicted to contrast with the control condition, /g/. The velar stop, /g/, contrary to the rhotics, is not expected to show resistance to phonetic environments for both the tongue body and tongue root gesture because it is hypothesized to be a simple articulation composed of only a tongue body gesture. This result is presumed to indicate that there is an underlying composition of a tongue body and tongue root gesture for rhotics.

3.2 Methods

3.2.1 Participants
Six native speakers of Upper Sorbian were recruited through personal contacts in the Sorbian community. All participants were undergraduates students (aged 18-24) at the University of Leipzig. Four participants were female and 2 were male. All participants provided informed consent and had no self-reported speaking or hearing problems.

3.2.2 Materials
The stimuli were composed of a uvular rhotic, /ʁ/, and a velar stop, /g/. The latter was included as the closest control for place of articulation (Upper Sorbian has no other uvular consonants). Distractor tokens including a lateral, /l/, fricatives, /z/ and /ʒ/, and bilabial stops, /p/ and /pʰ/, were also included. Target stimuli were produced in the word-initial, (#CV), word-medial, (VCV), and word-final, (VC#), positions using the vowels /a/, /e/, and /o/. Each of the words were embedded in a carrier phrase “Sym X měnil(a)” [(I) have thought X] to facilitate more natural pronunciations. Stimuli were randomized and participants produced each phrase a total of 8 times (9 conditions x 2 phonemes x 8 repetitions x 6 speakers = a total of 864 token, with 144 tokens per speaker). Table 6 presents the target stimuli; words with no glosses were nonsense words. /k/ was used word finally as opposed to /g/ due to the scarcity of word-final /g/ and because word-final devoicing would render the token as [k] regardless. The vowels were chosen on the basis of providing contrasting high front and high back phonetic environments with a more neutral low mid environment. /e/ was used instead of /i/ because there is a contrast neutralization before /i/, where only palatalized segments appear (Howson, 2017). In order to maintain similarity in the height of the two high vowels used, /o/ was used instead of /u/.
### Table 6. Target stimuli for Upper Sorbian

<table>
<thead>
<tr>
<th>#Ca</th>
<th>rad</th>
<th>[ɾat]</th>
<th>gladly</th>
<th>gagot</th>
<th>[ɡaɡot]</th>
<th>cackle</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Ce</td>
<td>retl</td>
<td>[ɾɛtl]</td>
<td>club,</td>
<td>gen</td>
<td>[ɡen]</td>
<td>gene</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cudgel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#Co</td>
<td>row</td>
<td>[ɾoʊu]</td>
<td>grave</td>
<td>goł</td>
<td>[ɡoʊu]</td>
<td>-</td>
</tr>
<tr>
<td>aCa</td>
<td>parać</td>
<td>[parASCII]aʧ</td>
<td>to trifle</td>
<td>bagaża</td>
<td>[baASCIIa]aʧ</td>
<td>baggage</td>
</tr>
<tr>
<td>eCe</td>
<td>ceremonija</td>
<td>[tsɛremɒniASCIIa]</td>
<td>ceremony</td>
<td>degeneracija</td>
<td>[deASCIIɛneraASCIItASCIIaASCII]</td>
<td>degeneration</td>
</tr>
<tr>
<td>oCo</td>
<td>dorost</td>
<td>[dɔrst]</td>
<td>new blood</td>
<td>hogosć</td>
<td>[hoASCIIgɒʃ]</td>
<td>-</td>
</tr>
<tr>
<td>aC#</td>
<td>dar</td>
<td>[daASCIIɾ]</td>
<td>gift,</td>
<td>dźak</td>
<td>[dʒak]</td>
<td>thanks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>present</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eC#</td>
<td>žer</td>
<td>[ʒɛɾ]</td>
<td>caries</td>
<td>pek</td>
<td>[pek]</td>
<td>-</td>
</tr>
<tr>
<td>oC#</td>
<td>por</td>
<td>[poASCIIɾ]</td>
<td>couple,</td>
<td>bok</td>
<td>[bok]</td>
<td>side</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pair</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.2.3 Procedure

Ultrasound recordings of the Upper Sorbian speakers were all preformed at the Max Planck Institute for Evolutionary Anthropology with a Terason 300T ultrasound and a veterinary transducer. Data was collected at a frame rate of 55 fps (frames per second). Audio data was recorded using a Fostex FR-2 LE: Field Recorder and a lavalier AT831b microphone at 44100 Hz and 32-bits. The participants wore a probe stabilization headset (Scobbie, Wrech, & van der Linden, 2008) which was designed to hold the ultrasound transducer in the same position while participants read a word list.

### 3.2.4 Analysis

The analysis consisted of two different statistical methods. (1) smoothing spline (SS)ANOVA (Gu, 2002; Davidson, 2006) comparisons of the point of maximum constriction; and (2) a coarticulatory analysis of the tongue root and tongue body. Both comparisons were performed using images taken from the point of maximum constriction – the peak movement of the articulatory towards a constriction location – for each of the target phonemes in each of the environmental conditions. This point was calculated based on both acoustic and visual inspection. The acoustic data indicated at what time points /ɾ/ and /ɡ/ were being articulated and...
a visual inspection of the images was used to determine which frame the peak movement of the tongue body and root occurred in.

There was a total of 27 statistical comparisons for each participant (27 x 6 = 162 total comparisons) in the SSANOVA analysis. Comparisons were performed for each of the target stimuli in each of the three vocalic environments (e.g. a comparison of a_a, e_e, and o_o, 3 comparisons for each phoneme, 6 total), across the three word positions (e.g. a comparison of #_a, a_a, and a_#, 3 comparisons for each phoneme, 6 total), and across target stimuli in the same vocalic environment (e.g. a comparison of [ara] and [aga], 6 total environmental conditions). The point of maximum constriction was determined by a combination of acoustic and visual cues. Acoustic measures were used to accurately determine where articulation of target phonemes occurred in the speech stream and visual inspection was performed to determine the maximum displacement of the tongue. Edgetrak (Li, Kambhamettu, and Stone, 2005) was used on the images exported from the ultrasound to calculate tongue contour coordinates, shown in Figure 11. SSANOVAs were performed with a script written by Mielke (2013) in R (R Development Core Team, 2017). This script was also used to convert the extracted coordinates into polar coordinates. The average tongue curve and the 95% confidence intervals were then calculated and plotted.

Figure 11. Ultrasound image loaded into Edgetrak and example contour tracing (left) and Tongue shape divided into the different areas referred to within this paper (right, modified from Catford 1988).

The maximum constriction results were also used to statistically compare the degree of coarticulatory resistance of /ɾ/ compared to /ɡ/. This was done by taking a coordinate at the tongue root and tongue body in each environment for the point of maximum constriction. The coordinates for each phoneme in each environment and each word position were measured and the area of the triangle produced for each phoneme in each environment was calculated using the
Surveyor’s Area Formula (Braden, 1986). Triangles were used because there were three phonetic environments and it allows horizontal and vertical displacement to be captured:

\[
area = \frac{1}{2} |x_1y_2 + x_2y_3 + x_3y_1 - x_2y_1 - x_3y_2 - x_1y_3|
\]

Figure 12. Example of the points taken for the Tongue Root and Tongue Body.

Figure 12 shows an example of three tongue contours for /ɡ/ and where the tongue body and tongue dorsum measurements were taken. Each area was then compared using two different repeated measures ANOVAs: the coarticulatory variation in (1) the tongue root and (2) the tongue body. Both ANOVAs had the factors Phoneme (2 levels: /ɡ/ and /r/) and Environment (3 levels: CV, VCV, VC). Post-hoc \( t \)-tests were preformed with Bonferroni correction. To further examine the coarticulatory differences, the x- and y-coordinates for the tongue root and body were compared using additional ANOVAs. Two ANOVAs (1 for the x-coordinates and 1 for the y-coordinates) were done on the tongue root and two on the tongue body. The factors were Phoneme (2 levels: /ɡ/ and /r/) and Environment (3 levels: a, e, o). Post-hoc \( t \)-tests were preformed with Bonferroni correction.

3.3 Results

The results are presented in the following order: (1) the results for the SSANOVA comparisons; and (2) the results of the coarticulatory analysis.

3.3.1 SSANOVA Results

The SSANOVA plots are derived from the statistical means of the tongue contours and plotted with 95% confidence intervals. The plots can be read by identifying where the 95% confidence intervals for two or more comparisons overlap. If they overlap, those areas are not considered to
be statistically different from each other. If the confidence intervals do not overlap the tongue curves are considered to be statistically different in those locations.

The first part of this section describes in detail the results for a single representative speaker (S1) separately across consonants and vowel contexts. Following that, cross-speaker generalizations are made. The results for the velar stop, /ɡ/, serve as a baseline because they are assumed to represent a simple articulation composed of one gesture – a tongue body gesture – and as a result should be more susceptible to phonetic environments.

Figure 13 presents the SSANOVA results for /ɡ/ produced by S1 in 3 intervocalic environments and 3 word positions. It can be seen that the velar stop was articulated with a raised and retracted tongue body gesture, presumably making a closure in the velar region. The constriction location was highly susceptible to the surrounding vowel environment. Specifically, the tongue root was most advanced next to /e/ and most retracted next to /o/; the environment /a/ was intermediate, although much closer to /o/. The tongue root frontness also corresponded strongly to the vocalic context: it was most fronted next to /e/ and most retracted next to /o/. This is most clearly observed by the total lack of overlap between for the confidence intervals for the tongue root of each trace. This difference was consistent across word-positions. That is, the advancement for /e/ environments occurred in word-initial, intervocalic, and word-final positions (#_e, e_e, e_#). The tongue retraction in the /o/ environments was also observed in all word positions. The tongue next to /a/ remained somewhere in between the other two environments for all word positions. The overall degree of advancement/retraction did not vary considerably by word position. This shows that the velar stop, /ɡ/, was produced a tongue body closure, but the tongue root was highly susceptible to phonetic environment, causing shifts in place of articulation. Therefore, /ɡ/ can be thought of as consisting of a single lingual (tongue body) gesture, rather than a complex articulation of multiple gestures.
Figure 13. Tongue contours for /ɡ/ in each environment /a/, /e/, and /o/ in each word position, word-initial (left), intervocalic (center), and word-final (right) for S1.

Figure 14 presents the SSANOVA results for the production of /ʀ/ for S1 in each vowel environment and each word position. The results indicate that the degree to which /ʀ/ was retracted was variable by environment; however, most striking was the fact that in word-initial position, the tongue root for the /o/ and /e/ environment did not show any difference in retraction (/a/ is slightly advanced). In the intervocalic position, the /a/ and /e/ environments showed no significant difference in advancement, but [o] showed more retraction, presumably due to the increased pressure of the intervocalic environment (Figure 14, left and middle panels). In both the word-initial and intervocalic positions, there was variability the height and location of the tongue body gesture. This suggests that the vocalic environment affected the degree of raising of the tongue body, although the consistent raising suggested there was an underlying tongue body gesture. The tongue tip/blade was highly variable by environment. It was significantly more raised in the /e/ environment and than contrasting /o/ and /a/ environments. This suggests that there is no tongue tip/blade involvement in the articulation of /ʀ/ for this speaker. The word-final position was drastically different due to the extreme lenition of the tongue body gesture. This resulted in a tongue shape involving retraction of the tongue root, but a much more vowel-like articulation, resembling vowel lengthening (Figure 14, right panel). Figures 15 – 17 shows a comparison of /ɡ/ and /ʀ/ in each environment.
Figure 14. /ɛ/ in each environment /a/, /ε/, and /o/ in each word position, word-initial (left), intervocalic (center), and word-final (right) for S1.

Figure 15. SSANOVA plots of the point of maximum constriction for /ɡ/ vs. /ɛ/ in Ca (left), aCa (middle), aC (right) for S1. Tongue tip is on the right.

Figure 16. SSANOVA plots of the point of maximum constriction for /ɡ/ vs. /ɛ/ in Ce (left), eCe (middle), eC (right) for S1. Tongue tip is on the right.
The results for the other 5 speakers were similar to S1, but there was some differences among speakers in the realization of /r/ suggesting the mid-sagittal tongue shapes can be categorized into two distinct types: (1) “double bunched” and (2) “retracted.” Figure 18 shows examples of each articulation with “double bunched” on the left and “retracted” on the right. In the “double bunched” tongue shape, the root and body were bunched to form a uvular constriction, while the tongue blade was slightly raised, forming a concave tongue shape. The “retracted” tongue shape involved a single bunching by the tongue root, but for these speakers, the tongue body/blade were typically retracted and together with the root formed a convex shape. These results were unexpected. Individual SSANOVA results are presented in the Appendix. A summary of tongue shapes used by each participant is presented in Table 7. It should be noted that none of the speakers – aside from S1 – produced the double bunched configuration in every phonetic environment. However, speakers labeled as double bunched had predominantly articulations of that shape. In word-final position, /r/ tended to be somewhat similar in shape to the preceding vowel. There was an absence of any uvular constriction, characteristic of the unvocalized variant. Following /e/, the tongue body was a more raised and fronted, but was somewhat reduced compared to /e/ itself. Following /o/, the tongue had a much more “double bunched” shaped for most speakers and resembled most closely /r/, with a drastic reduction in the amount the tongue body raised to create a uvular constriction. See Figures 81 – 95 in the Appendix for a plot of /g/ and /r/ in each environment and word position for each speaker.
Figure 18. (left) comparison of /ɡ/ vs. /ʀ/ in the /#_e/ environment for S1, showing a “double bunched” rhotic articulation; and (right) /ɡ/ vs. /ʀ/ in the /#_e/ environment for S2, showing a “retracted” rhotic articulation.

Table 7. Summary of tongue shapes for /ʀ/ for each speaker. + indicates that the tongue had a particular shape in the environment listed, while – indicates the lack of that tongue shape in the environment.

<table>
<thead>
<tr>
<th></th>
<th>#_V</th>
<th>V_V</th>
<th>overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>retracted:</td>
<td>a e o</td>
<td>a e o</td>
</tr>
<tr>
<td></td>
<td>double bunched:</td>
<td>+ + +</td>
<td>+ + +</td>
</tr>
<tr>
<td>S2</td>
<td>retracted:</td>
<td>+ + +</td>
<td>+ + +</td>
</tr>
<tr>
<td></td>
<td>double bunched:</td>
<td>- - -</td>
<td>- - -</td>
</tr>
<tr>
<td>S3</td>
<td>retracted:</td>
<td>+ + -</td>
<td>+ + -</td>
</tr>
<tr>
<td></td>
<td>double bunched:</td>
<td>- - +</td>
<td>- - +</td>
</tr>
<tr>
<td>S4</td>
<td>retracted:</td>
<td>+ + +</td>
<td>+ + -</td>
</tr>
<tr>
<td></td>
<td>double bunched:</td>
<td>- - -</td>
<td>- - +</td>
</tr>
<tr>
<td>S5</td>
<td>retracted:</td>
<td>- + +</td>
<td>+ + -</td>
</tr>
<tr>
<td></td>
<td>double bunched:</td>
<td>+ - -</td>
<td>- - +</td>
</tr>
<tr>
<td>S6</td>
<td>retracted:</td>
<td>- + -</td>
<td>- + -</td>
</tr>
<tr>
<td></td>
<td>double bunched:</td>
<td>+ - +</td>
<td>+ - +</td>
</tr>
</tbody>
</table>

In summary, the overall SSANOVA results suggest a few primary differences between /ʀ/ and /ɡ/: (1) the tongue root for the rhotic was highly resistant to phonetic environment, while /ɡ/ was quite susceptible to differences and (2) the uvular-pharyngeal constriction location was relatively invariable across environments for the rhotic, while the location of the tongue body constriction was highly susceptible to environments for /ɡ/.

3.3.2 Coarticulatory Analysis

A coarticulatory analysis was performed to measure the amount of resistance each phoneme had at the tongue root and tongue body had to the effects of the phonemic environment.

The results for the coarticulatory resistance of the root revealed the factor phoneme was found to be significant [F(1,5) = 7.48, p = 0.0410]. However, there was no effect of environment
[F(2,10) = 0.88, p = 0.4414] and no interaction between phoneme and environment [F(2,10) = 1.19, p = 0.3480]. /g/ was found to be more susceptible to coarticulation (50.14 mm²) compared to /r/ (2.86 mm²). Figure 19 presents a bar graph of the results for Phoneme. Table 8 shows the means and standard deviations for each phoneme in each environment.

![Bar graph showing coarticulation on tongue root for each phoneme.](image)

**Figure 19. Overall effects of coarticulation on the tongue root for each phoneme (in mm²).**

**Table 8. Means and standard deviations (in brackets) of the coarticulatory measures (mm²) for each phoneme in each word position.**

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>VCV</th>
<th>VC</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>/g/</td>
<td>51.04</td>
<td>(53.44)</td>
<td>55.25</td>
<td>(46.08)</td>
</tr>
<tr>
<td>/r/</td>
<td>2.67</td>
<td>(1.76)</td>
<td>2.54</td>
<td>(2.07)</td>
</tr>
<tr>
<td>mean</td>
<td>26.86</td>
<td>(27.60)</td>
<td>28.90</td>
<td>(24.08)</td>
</tr>
</tbody>
</table>

To further examine the coarticulatory resistance, the x- and y-coordinates were tested for significance using a repeated measures ANOVA. The results for the retraction of the tongue root (x-axis) indicated the factor phoneme was significant [F(1,5) = 15.94, p = 0.0104]. The factor environment was also significant [F(2,10) = 8.35, p = 0.0074]. The data for the interaction between phoneme and environment was not significant [F(2,10) = 3.68, p = 0.0634]. /g/ was found to be more advanced (114.36 mm) compared to /r/ (122.70 mm). Figure 20 shows a bar graph of the results for phoneme.
Post-hoc tests for the factor environment revealed that there was a significant difference between all environments (a-e: $p = 0.0031$; a-o: $p < 0.0001$; e-o: $p < 0.0001$). The primary reason for this was likely due to the variability in the tongue root for /ɡ/, especially as a result of advancement in the /e/ environments (109.23 mm) and retraction in the /o/ environments (120.03 mm), having a total range of 11 mm between the two. However, /ʁ/ had little variability with a range of 2.98 mm between environments. Figure 21 shows a bar graph of the results for Environment and Table 9 shows the means and standard deviations for each phoneme in each environment.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>e</th>
<th>o</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>ɡ</td>
<td>113.81</td>
<td>(8.83)</td>
<td>109.23</td>
<td>(9.79)</td>
</tr>
<tr>
<td>ʁ</td>
<td>122.09</td>
<td>(7.15)</td>
<td>121.51</td>
<td>(8.06)</td>
</tr>
<tr>
<td>Mean</td>
<td>117.95</td>
<td>(7.99)</td>
<td>115.37</td>
<td>(8.93)</td>
</tr>
</tbody>
</table>
The results for tongue root height (y-axis) revealed that there was no main effect of phoneme \( [F(1,5) = 0.12, p = 0.7474] \), or environment \( [F(2,10) = 2.64, p = 0.1202] \), and no interaction between the two factors phoneme and environment \( [F(2,10) = 0.38, p = 0.6954] \).

The results for the coarticulation of the tongue body indicated no main effect of phoneme \( [F(1,5) = 1.39, p = 0.2912] \), no main effect of environment \( [F(2,10) = 2.11, p = 0.1719] \), and no interaction between the two factors \( [F(2,10) = 0.20, p = 0.8200] \). To more closely examine differences in tongue body position, the x- and y-coordinates were also compared for the tongue body. The results for the tongue body x-axis, indicated that there was a main effect of phoneme \( [F(1,5) = 12.52, p = 0.0166] \) and environment \( [F(2,10) = 13.85, p = 0.0013] \), but no interaction between the two factors \( [F(2,10) = 1.99, p = 0.1875] \). There was a significant difference between /ğ/ (90.13 mm) and /r/ (99.96 mm). Post-hoc t-tests indicated that there was a significant difference between all environments (a-e: \( p = 0.0046 \); a-o: \( p < 0.0001 \); e-o: \( p < 0.0001 \)). Figure 22 presents a bar graph of the results for Phoneme and Figure 23 presents a bar graph of the results for Environment. Table 10 presents the means and standard deviations for each environment.

![Figure 22](image1.png)

*Figure 22. The amount of tongue body retraction (mm) for each phoneme.*

![Figure 23](image2.png)

*Figure 23. The amount of tongue body retraction (mm) for each environment.*
Table 10. Means and standard deviations (in brackets) for each phoneme in each environment.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>e</th>
<th>o</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>90.43 (10.28)</td>
<td>83.98 (11.32)</td>
<td>95.69 (7.54)</td>
<td>90.13 (9.71)</td>
</tr>
<tr>
<td>r</td>
<td>97.43 (10.40)</td>
<td>98.12 (9.24)</td>
<td>104.33 (8.14)</td>
<td>99.96 (9.26)</td>
</tr>
<tr>
<td>Mean</td>
<td>93.93 (10.34)</td>
<td>91.05 (10.28)</td>
<td>100.01 (7.83)</td>
<td>-</td>
</tr>
</tbody>
</table>

The results for the tongue body y-axis, indicated that there was a main effect of phoneme \( F(1,5) = 33.29, p = 0.0022 \), results presented in Figure 24, and environment \( F(2,10) = 7.88, p = 0.0087 \), results presented in Figure 25, but no interaction between phoneme and environment \( F(2,10) = 0.98, p = 0.4100 \). There was a difference in the height of /g/ (28.50 mm) and /r/ (40.96 mm). The post-hoc comparisons for environment revealed a significant difference between a-e (\( p = 0.0160 \)) and e-o (\( p = 0.0240 \)), not a-o (\( p = 1.0000 \)). Table 11 presents means and standard deviations for each phoneme in each environment and the grand means and stand deviations.

Figure 24. The height of the tongue body (mm) for each phoneme.

Figure 25. The height of the tongue body (mm) for each environment.
Table 11. Means and standard deviations (in brackets) for each phoneme in each environment.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>e</th>
<th>o</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>ɡ</td>
<td>28.89 (9.92)</td>
<td>26.80 (8.85)</td>
<td>29.82 (8.64)</td>
<td>28.50 (9.14)</td>
</tr>
<tr>
<td>ʀ</td>
<td>42.26 (6.69)</td>
<td>39.52 (7.91)</td>
<td>41.11 (7.60)</td>
<td>40.96 (7.40)</td>
</tr>
<tr>
<td>Mean</td>
<td>35.58 (8.31)</td>
<td>33.16 (8.38)</td>
<td>35.47 (8.12)</td>
<td>-</td>
</tr>
</tbody>
</table>

3.4 Discussion

In this chapter, I examined the velar stop and uvular rhotic in Upper Sorbian using ultrasound technology. The SSANOVA revealed high variability in the place of articulation for the velar stop. This was best observed through vowel environment effects causing more advancement of the tongue body constriction when in front vowel environments, and more retracted in back vowel environments. The tongue root was also highly variable in articulation. The same vowel place effects were observed for the tongue root as for the tongue body. This was confirmed with the coarticulatory analysis, which revealed large areas created by the tongue root and tongue body.

The ultrasound results revealed a highly resistant tongue root for the articulation of the uvular rhotic. This was observed in overlapping or nearly overlapping confidence intervals for the tongue root on SSANOVA plots across all environments and confirmed with the coarticulatory analysis: the tongue root across environments produced a small area. The tongue body was not as resistant to coarticulation, visible in variable posterior tongue body constriction location in the SSANOVA and measured by the higher area produced by the coarticulatory analysis.

The results suggest that there is a consistent tongue root gesture involved in the articulation of /r/, coordinated with a uvular-pharyngeal tongue body constriction. This contrasted with /ɡ/, which was composed solely of a tongue body constriction. The complexity of articulation highlights a difference between uvular rhotics and velar stops and forms the basis for notion that rhotics are complex articulations involving a vocalic tongue gesture coordinated with another gesture. In the case of uvular rhotics, this is the coordination between a tongue root and tongue body gesture. I discuss theoretical implications for these findings in Chapter 7, Section 7.2.1.
Chapter 4

There is a tongue root conflict between secondary palatalization and rhotics: evidence from Upper Sorbian

4.1 Introduction

The previous chapters presented acoustic and articulatory evidence that there is a consistent tongue root gesture associated with rhotics articulation in Sorbian. Importantly, this gesture appears to conflict with secondary palatalization, which typically involves fronting of the tongue root to accommodate the secondary constriction in the palatal region. This may present an explanation for the fact that rhotics have a high rate of avoiding secondary palatalization both through phonological processes and as an actual phonemic representation. The study presented in this chapter examines the gestural dynamics of rhotics in Upper Sorbian to ascertain if there is a conflict between the tongue root for rhotic articulation and secondary palatalization.

4.1.1 Hypothesis

Based on articulatory research, this study aims to test the following hypothesis:

Rhotics and palatalization are avoided due to conflicting constraints on the tongue root (Kavitskaya et al., 2009).

The hypothesis originates from Kavitskaya et al., (2009) who suggest that palatalization requires tongue root fronting to permit the raising and fronting of the tongue body, while rhotic articulation demands retraction of the tongue root. Hall and Hamann (2010) also suggest that tongue shape plays a role in the dispreference of rhotics and palatalization/high-front vocoids. Hall and Hamann (2010) assert that rhotics are produced with a concave tongue shape, while high-front vowels are produced with a convex tongue shape. This difference in tongue shape is argued to cause an inherent avoidance of these sequences because the need to rapidly change the entire tongue shape is difficult to accommodate. However, rhotics typically only have a convex shape when talking about tongue-tip rhotics and so it is unclear how Hall and Hamann’s (2010) approach would apply to uvular rhotics. Rhotics indeed have a great deal of tongue root retraction (see Delattre [1971] for discussion of German, French and English rhotics), but not necessarily a concave tongue shape. Therefore, a single, unifying theory such as the one posited by Kavitskaya et al. (2009) is desirable and will be the theoretical approach of this paper.
A dynamic analysis is being used to approach this question. The two main contrastive pairs will be /ʁ/ and /ʁʲ/. The dynamic analysis for /ʁ/ is expected to show a consistent tongue root retraction and tongue body raising towards the point of maximum constriction and then transition into the following vowel. /ʁʲ/, in contrast, is expected to show conflicting demands on the tongue root, observed by a tightly sequenced retraction and advancement of the tongue root gestures. This is to be interpreted as the result of a gestural conflict between the requirements for articulating a rhotic and the tongue root advancement associated with secondary palatalization.

4.2 Methods
The participants and procedure were the same as in chapter 3. This data was recorded in the same session as the data recorded in chapter 3.

4.2.1 Materials
The stimuli, shown in Table 12, were composed of 2 uvular rhotics, /ʁ/ and /ʁʲ/. Distractor tokens including a lateral, /l/, fricatives, /z/ and /ʒ/, and bilabial stops, /p/ and /pʲ/, were incorporated to mask the target stimuli. Target stimuli were produced in the word-initial, /#CV/, and word-medial, /VCV/, positions using the vowels /a/, /e/, and /o/. Word-final position were omitted for /ʁʲ/ because it is prohibited in that position. Each of the words were embedded in a carrier phrase “Sym X měnil(a)” [(I) have thought X] to facilitate more natural pronunciations. Stimuli were randomized and presented to the participants in a list format. The list was randomized and was read a total of 8 times by each participant (6 conditions x 2 phonemes x 8 repetitions x 6 speakers = a total of 576 token, with 96 tokens per speaker). In Table 12, words with no glosses were nonsense words. The vowels were chosen on the basis of providing contrasting high front and high back phonetic environments with a more neutral low mid environment. /e/ was used instead of /i/ because there is a contrast neutralization before /i/, where only palatalized segments appear. In order to maintain similarity in the height of the two high vowels used, /o/ was used instead of /u/.
Table 12. Target Stimuli for Upper Sorbian

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>R^j</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#Ca</td>
<td>rad</td>
<td>[rəd]</td>
<td>gladly</td>
<td>rjad</td>
</tr>
<tr>
<td>#Ce</td>
<td>retl</td>
<td>[rɛtl]</td>
<td>club, cudgel</td>
<td>rjec</td>
</tr>
<tr>
<td>#Co</td>
<td>row</td>
<td>[rəʊ]</td>
<td>grave</td>
<td>rjot</td>
</tr>
<tr>
<td>aCa</td>
<td>parać</td>
<td>[pəraʧ]</td>
<td>to trifle</td>
<td>parjak</td>
</tr>
<tr>
<td>eCe</td>
<td>ceremonija</td>
<td>[tseremoniːja]</td>
<td>ceremony</td>
<td>šerjenje</td>
</tr>
<tr>
<td>oCo</td>
<td>dorost</td>
<td>[dorost]</td>
<td>new blood</td>
<td>horjo</td>
</tr>
<tr>
<td>aC#</td>
<td>dar</td>
<td>[dar]</td>
<td>gift, present</td>
<td>-</td>
</tr>
<tr>
<td>eC#</td>
<td>žer</td>
<td>[ʒɛɾ]</td>
<td>caries</td>
<td>-</td>
</tr>
<tr>
<td>oC#</td>
<td>por</td>
<td>[pɔɾ]</td>
<td>couple, pair</td>
<td>-</td>
</tr>
</tbody>
</table>

4.2.2 Analysis

The dynamic analysis of gestural components was performed using a total of 9 frames for each articulation. Four frames were taken from directly before the point of maximum constriction and 4 frames were taken after (a total of 9 frames) using Edgetrak (Li et al., 2005). Average tongue contours were calculated for each frame and plotted in R (R Core Development Team, 2017). Dynamic plots were broken into two separate plots for each phoneme in each environment (9 environments x 2 plots x 6 speakers = 108 total plots). One plot for each environment is for the temporal dynamics leading up to the point of maximum constriction and 1 plot is for the dynamics leading away. This means each plot has 5 frames total (4 frames + the maximum constriction frame = 5 frames). Figure 26 shows images loaded into Edgetrak with tongue tracing and the schema for the divisions of the tongue as they are talked about in this chapter.

Figure 26. Ultrasound image loaded into Edgetrak and example contour tracing (left) and Tongue shape divided into the different areas referred to within this paper (right, modified from Catford 1988).
4.3 Dynamic Results

The analysis consisted of a dynamic examination of the tongue gestures during articulation of both rhotics, /r/ and /ɾ/.

The results for S1 are shown in Figures 27 – 41 and discussed, but the same general tendencies held for each speaker. Figures 27 – 35 present the gestural results for /r/ and revealed that it was most similar in gestural composition and dynamics to the articulation of /o/. This can be observed in Figure 33, where the retraction and raising of the tongue body does not appear to change significantly through the transition from /ɾ/ to /o/. The articulatory dynamics of /ɾ/ unfolded with a characteristic retraction/raising of the tongue body towards the uvular place of articulation, accompanied by retraction of the tongue root. This was observed in two specific ways: (1) retraction of the tongue body from preceding vowels (aside from /o/ which was already retracted). Figure 31 presents the gestural dynamics for the /eɾe/ environment, and shows that retraction occurred, but raising was not generally observed because the body was already raised. For /aɾa/, there was both a raising and retraction of the tongue body. This can be observed in Figure 28. (2) There was a consistent tongue root retraction. This was best observed in the word-final environments /aɾ/ and /eɾ/, which are presented in Figure 29 and 32. There was a clear lenition of the tongue body gesture, observable by a lowering of the tongue body. However, the lowering of the tongue body was accompanied by a retraction of the tongue root into the pharyngeal cavity, suggesting the presence of an unlenited tongue root gesture. In the case of /oɾ/, the tongue was already retracted a great deal and so there was simply lenition of the tongue root gesture.
Figure 27. Gestural dynamics for /n/ in the /#_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 28. Gestural dynamics for /n/ in the /a_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
Figure 29. Gestural dynamics for /n/ in the /a#/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 30. Gestural dynamics for /n/ in the #_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
Figure 31. Gestural dynamics for /a/ in the /e_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 32. Gestural dynamics for /a/ in the /e_/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
Figure 33. Gestural dynamics for /n/ in the /#_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 34. Gestural dynamics for /n/ in the /o_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
Figure 35. Gestural dynamics for /r/ in the /o_/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

The results for /r/ are presented in Figures 36 – 41 and revealed a complex set of gestures and coordination between them. Similar to /r/, there was a retraction/raising of the tongue body accompanied by a strong retraction of the tongue root. The raising and retraction was visible in every intervocalic environment and in many cases the tongue blade was also raised with the tongue body (e.g. /ɾe/ and /ɾo/) which can be observed in Figures 38 and 40. There was also a significant delay between the gestures associated most strongly secondary palatalization – raising and fronting of the tongue body/blade – and the components clearly comparable to /r/, retraction and raising of the tongue body accompanied by retraction of the tongue root. The tongue body and root gestures were coordinated more sequentially, such that tongue root and the posterior tongue body retracted, typically accompanied with a raising effect across the entire tongue body. After the tongue body and root reached the point of maximum constriction, there was a simultaneous fronting of the tongue root and raising of the tongue body towards a target more consistent with secondary palatalization. The tongue root/body gestures were coordinated concurrently and occur nearly simultaneously.
Figure 36. Gestural dynamics for /ɨ/ in the /#_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 37. Gestural dynamics for /ɨ/ in the /a_a/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
Figure 38. Gestural dynamics for /ɾ/ in the /#_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 39. Gestural dynamics for /ɾ/ in the /e_e/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
Figure 40. Gestural dynamics for /r/ in the /#_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 41. Gestural dynamics for /r/ in the /o_o/ environment for S1, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

4.4 Discussion

The goal of this chapter was to test the hypothesis that rhotics have a dispreference for secondary palatalization because of conflicting demands on the tongue root (Kavitskaya et al., 2009). The dynamic results revealed that for /r/, there was retraction and raising of the tongue body towards a uvular-pharyngeal constriction location. A tongue root gesture was also observable for /r/ in all cases. There was retraction into the pharyngeal cavity, which was even present in word-final position where lenition of the tongue body gesture occurred. The observable tongue root retraction was still visible by a movement of the tongue root towards the pharyngeal cavity.
/ᵣʲ/ also exhibited tongue body raising and retraction towards a uvular-pharyngeal place of articulation and a retraction of the tongue body into the pharyngeal cavity. However, this was often accompanied by raising of tongue blade/tip during articulation, marking a distinct difference between /ᵣ/ and /ᵣʲ/. The offset, after the point of maximum constriction at the uvular place of articulation, was characterized by tongue root fronting and raising of the anterior tongue body for secondary palatalization.

The results suggest that tongue root retraction into the pharyngeal cavity is a component for rhotics and that there is a coordination between the tongue root and tongue body gestures for uvular rhotics. The data presented here also supports the notion that there is a conflict between the tongue root and secondary (Kavitskaya et al., 2009): rhotics require tongue root retraction and secondary palatalization requires tongue root advancement. The result of the conflict is that the coordination between the tongue root retraction and tongue root advancement is delayed. Pivot and arch transitions (Iskarous, 2005) were both observed for the articulation of /ᵣʲ/. Retraction into the pharyngeal cavity target was first achieved via a pivot transition followed by an arch transition towards the palatal region. This contrasted /ᵣ/, which was uniformly achieved by a pivot transition. The difference between the articulation of /ᵣ/ and /ᵣʲ/ most likely stems from the necessity for the tongue to perform two antagonistic movements, retraction and advancement, is likely the source of the global dispreference of rhotics and secondary palatalization. Theoretical implications are discussed in Chapter 7.2.1.
Chapter 5
Lenition of gestural settings for rhotics accounts for synchronic and diachronic changes in Brazilian Portuguese

5.1 Introduction
The previous chapters set the stage for an articulatory link between uvular and alveolar rhotics. The underlying gestural composition of rhotics is theorized to be the core of all the synchronic and diachronic sound changes observed in natural languages. In order to test the ability of the gestural representation to account for the phonological behaviour of rhotics, Brazilian Portuguese is employed because of the large number of both synchronic variation and diachronic changes observed across dialects and idiolects. The goal is to determine if the rhotics in Brazilian Portuguese can also be described as having a tongue root gesture coordinated with a tongue body/tip gesture. The secondary goal is to determine if the gestural representation can account for the synchronic variation and diachronic changes that occur in Brazilian Portuguese.

5.1.1 Brazilian Portuguese
Brazilian Portuguese has 19 consonants, including two phonemic rhotics, /ɾ/ and /ʁ/ (ca[ɾ]o ‘expensive;’ ca[ʁ]o ‘car’). Barbosa and Albano (2004) refer to /ʁ/ as /ɣ/ (see Table 13 for a complete list of consonants), but due to the overwhelming tendency for the speakers examined in this study to produce the voiced variant /ʁ/, I will be referring to it as /ʁ/. /ɾ/ and /ʁ/ only contrast intervocally and the contrast is neutralized in all other word-positions. /ɾ/ occurs in syllable onset clusters (at[ɾ]ibuto ‘attribute’), while /ʁ/ appears word-initially (/ʁ]ato ‘mouse’) or syllable initially when preceded by /l s n/ (/Is[ʁ]ael ‘Israel’). /ʁ/ can be realized as [ɾ] under strong prosodic conditions. However, this effect is gradient, correlating to how strong the syllable stress is. In strong conditions, [ɾ] emerges, but in slightly weaker stress conditions a gradient effect can emerge, causing [ɾ] to surface. Intervocally, /ʁ/ is often realized as a glottal fricative, [h] (Albano, 2001).

Barbosa and Albano (2004) also describe the coda rhotic as having a variable realization (e.g. carta /kaRta/ ‘letter’) that is highly variable along regional dialect and idiolects. It may be realized as a tap, [ɾ] (e.g. ca[r]ta ‘letter’), but it may also be realized as a uvular fricative, [χ], a velar fricative, [ɣ], a glottal fricative, [h], an alveolar approximant, [ɹ], a trill, [r], or it may be deleted completely (Mascaró, 2003). Barbosa and Albano (2004) note that deletion is especially
common word-finally. The realization is partially related to region, with areas in São Paulo and southwards having higher frequency of the tap and areas in Rio de Janeiro and northwards having higher frequency of the velar and glottal fricative. As a result, the contrast between /ɾ/ and /ʁ/ is neutralized in the coda positions and it is not possible to discern which phoneme is the underlying form.

Table 13. Consonant inventory of Brazilian Portuguese (from Barbosa & Albano, 2004, p. 228)

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Postalveolar</th>
<th>Palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td></td>
<td></td>
<td>k</td>
</tr>
<tr>
<td>Affricates</td>
<td>m</td>
<td>n</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td>j</td>
</tr>
<tr>
<td>Tap</td>
<td>f</td>
<td>v</td>
<td>s</td>
<td>z</td>
<td>f</td>
<td>3</td>
<td>y</td>
</tr>
<tr>
<td>Fricative</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.1.2 Hypothesis

The goal of this chapter is to examine the articulatory gestures and the dynamics associated with rhotics in Brazilian Portuguese and to explain their synchronic and diachronic alternations. The core hypothesis is that rhotics composed of a tongue root gesture coordinated with tongue tip/body gesture (c.f. Proctor, 2011; chapter 2). The alternations between rhotics in Brazilian Portuguese are hypothesized to result from the manipulation of the coordination, timing, and intensity of these two gestures. The coda rhotic, /R/, (Barbosa & Albano, 2004) is hypothesized to be underlying /ɾ/ since most of the realizations of the word-final rhotic involved the tongue tip. As a result, I anticipate that the allophony between coda /R/ and [ɾ χ] in the coda position is the result of lenition of the tongue tip gesture. Intervocalic allophonic changes between /ʁ/ and [h] are expected to be a result of complete lenition of tongue root and body gestures, resulting in a placeless articulation.

5.2 Methods

5.2.1 Participants

Six native speakers of Brazilian Portuguese were recruited from the Portuguese community in Toronto. All participants were undergraduate students aged 18–24 years of age and were originally born and raised in São Paulo, Brazil, until the age of at least 14. All participants
provided informed consent and had no self-reported speaking or hearing problems. All participants were L2 speakers of English.

5.2.2 Instrumentation

Recordings were done at the Phonetics Lab at the University of Toronto using a Telemed EchoB ultrasound (60 fps) with audio and video synchronization through AAA software (Articulate Instruments Ltd., 2012). The participants wore specially designed headgear (Articulate Instruments Ltd., 2008) to hold the ultrasound transducer in a stable position while participants read a word list. The list was randomized and was read a total of 8 times by each participant (12 conditions x 2 phonemes x 8 repetitions x 6 speakers + 6 conditions x 2 phonemes x 8 repetitions x 6 speakers = a total of 1,728 tokens, 288 tokens per speaker).

5.2.3 Materials

The target stimuli consisted of two rhotics, /ɾ/ and /ʁ/, and two stops, /t/ and /k/. /t/ and /k/ were included as control tokens. Distractor tokens included two phonemic palatals, /ʃ/ and /ɲ/ in the same environments as the other tokens. Stimuli were produced in the word-initial, /#CV/, word medial, /VCV/, and word-final, /VC#/ positions using the vowels /ɪ/, /ɛ/, /a/, and /o/. Syllable stress was designed to be on the first syllable for consistency of experimental design. Barbosa and Albano (2004) note that post-tonic vowels are often reduced: /ɪ/ remains [ɪ], while /ɛ/ and /a/ are more centralized [ɛ] and [a]; finally, /u/ is more often reduced to [o]. /t/ and /k/ contrast intervocally, but only /ʃ/ appears word-initially. Word-finally, the coda rhotic /R/ appears (Barbosa & Albano, 2004). As a result, the dataset only includes /ʃ/ in word-initial, /ɪ/ and /u/ in intervocalic position, and /R/ in word-final position. /t/ is palatalized before /i/, surfacing as [ʃ] (Barbosa & Albano, 2004). Stimuli was produced in the carrier phrase “diga target para si mesmo” [Say target for yourself] to facilitate more natural productions. For the control phoneme /kl/, /loco, picnic, parsec, kayak, and wok are all loan words. Table 14 summarizes the target stimuli.
<table>
<thead>
<tr>
<th></th>
<th>/ɾ R/</th>
<th>/ʁ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Ci</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>#Ce</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>#Ca</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>#Co</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>iCi</td>
<td>irisa</td>
<td>[iɾisɐ]</td>
</tr>
<tr>
<td>eCe</td>
<td>ereto</td>
<td>[eɾeɾo]</td>
</tr>
<tr>
<td>aCa</td>
<td>arado</td>
<td>[aɾəɾo]</td>
</tr>
<tr>
<td>oCo</td>
<td>ouro</td>
<td>[oɾo]</td>
</tr>
<tr>
<td>iC#</td>
<td>vir</td>
<td>[viɾ]</td>
</tr>
<tr>
<td>eC#</td>
<td>der</td>
<td>[deɾ]</td>
</tr>
<tr>
<td>aC#</td>
<td>dar</td>
<td>[dɐɾ]</td>
</tr>
<tr>
<td>oC#</td>
<td>dor</td>
<td>[dɔɾ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>/t/</th>
<th>/k/</th>
</tr>
</thead>
<tbody>
<tr>
<td>#Ci</td>
<td>tico</td>
<td>[tikʊ]</td>
</tr>
<tr>
<td>#Ce</td>
<td>teco</td>
<td>[tekʊ]</td>
</tr>
<tr>
<td>#Ca</td>
<td>taco</td>
<td>[takʊ]</td>
</tr>
<tr>
<td>#Co</td>
<td>toco</td>
<td>[tɔkʊ]</td>
</tr>
<tr>
<td>iCi</td>
<td>aditivo</td>
<td>[adيثɪvʊ]</td>
</tr>
<tr>
<td>eCe</td>
<td>sete</td>
<td>[seɾe]</td>
</tr>
<tr>
<td>aCa</td>
<td>data</td>
<td>[daɾa]</td>
</tr>
<tr>
<td>oCo</td>
<td>coto</td>
<td>[kɔto]</td>
</tr>
<tr>
<td>iC#</td>
<td>bit</td>
<td>[bik]</td>
</tr>
<tr>
<td>eC#</td>
<td>jet</td>
<td>[dɛɾɛɾ]</td>
</tr>
<tr>
<td>aC#</td>
<td>chat</td>
<td>[ʃaɾ]</td>
</tr>
<tr>
<td>oC#</td>
<td>spot</td>
<td>[spɔɾ]</td>
</tr>
</tbody>
</table>
5.2.4 Analysis

There were two different analyses performed to investigate the gestural composition of the Brazilian Portuguese rhotics: (1) a smoothing spline (SS)ANOVA (Gu, 2002; Davidson, 2006) was performed at the point of maximum constriction; (2) a dynamic analysis was performed with the average tongue contours; and (3), an RMS analysis was performed as a measure of resistance to coarticulatory effects.

The point of maximum constriction was determined through acoustic and visual inspection for the peak displacement of the tongue root/tip articulator during articulation of each target phoneme in each of the environmental conditions. Statistical comparisons were performed across all target phonemes and experimental conditions (12 conditions: e.g. #_a, a_a, a_; 8 conditions for each rhotic, due to positional restrictions). Comparisons also included comparisons of positional and environmental variation within a single phoneme (e.g. /ɾ/ in i_i, e_e, a_a, o_o). Tongue traces were performed in AAA software (Articulate Instruments Ltd., 2012). Coordinates were exported and SSANOVA (Gu, 2002; Davidson, 2006) plots were created using script written by Mielke (2013) in R (R Development Core Team, 2017). Mielke’s (2013) script calculates average tongue contours and 95% confidence intervals.

The dynamic analysis was performed by calculating the average tongue contour for each frame from the 1st to 11th frame. The point of maximum constriction was considered the 6th frame and measurements were taken 5 frames before and after this point. Plots for each phoneme in each of the environmental conditions were created. Two plots for each environment were created: the 1st plot shows the average tongue contours for each frame from the 1st frame to the point of maximum constriction and the 2nd plot shows each frame from the point of maximum constriction to the 11th frame. This means there are a total of 480 plots (80 per speaker).

Figure 42. Example of the AAA interface and with tongue trace (left) and the division of the tongue into different regions as discussed in this paper (right, modified from Catford, 1988).
AAA software (Articulate Instruments Ltd., 2012) was also used to preform a comparison of the difference in coarticulatory effects for each phoneme. The average tongue contours calculated for the previous SSANOVA analysis were used for a Root Mean Square (RMS; Kenney & Keeping, 1962) measure of the entire tongue contour and was compared statistically for the factors phoneme (5 levels: /t/, /ɾ/, /ɾ/, /k/, and /R/) and environment (3 levels: CV, VCV, VC) with multiple linear regression. To test the modelled linear regression, Satterthwaite approximation for degrees of freedom was used to perform an ANOVA in R (R Development Core Team, 2017). Post-hoc t-tests were performed using Bonferroni correction.

5.3 Results

The results are presented in three sections: (1) the comparison of the tongue contours at the point of maximum constriction (Section 4.3.1); (2) the dynamic results (Section 4.3.2); and the coarticulatory examination (Section 4.3.3).

5.3.1 SSANOVA Results

In this section, the results for a representative speaker, BP5, will be presented and discussed first and then general cross-speaker results will be discussed. First, the two control phonemes, /k/ and /t/, are discussed, followed by the phonemes of interest, /ɾ/, /ɾ/, and coda /R/.

The SSANOVA results for BP5 for /k/ are presented in Figure 43 and indicated that /k/ was highly influenced by phonetic environment in the word-initial and intervocalic regions. Place of articulation was the most retracted when a back vowel, /o/, was in the surrounding phonetic environment and most advanced when front vowel /i/ or /e/ were present; in the phonetic environment for /a/, the place of articulation for /k/ was between its place of articulation for the /o/ and /e i/ environments. This was consistent for both word-initial and intervocalic regions. The word-final position showed significant fronting in all environments (Figure 43, right panel), consistent with previous descriptions that Brazilian Portuguese obstruents are often palatalized in word-final position (e.g. Bettoni-Techio & Koerich, 2006).
Figure 43. SSANOVA results for BP5’s articulation of /k/ in all vocalic environments for the word-initial position (left), intervocalic position (center), and word-final position (right).

The results for the production of /t/ had a similar pattern as /k/: The tongue root position and fronting of the tongue body was relative to the phonetic environment. /o/ environments produced the most tongue root retraction during the articulation of /t/, while /i/ environments produced the most advancement of the tongue root, accompanied with tongue body raising. This is consistent with descriptions of palatalization in Brazilian Portuguese (Barbosa & Albano, 2004). /a e/ environments did not differ significantly in root retraction or tongue body raising and fronting for the realization of /t/ in the word-initial position, but in the intervocalic position, /e/ environments produced comparable palatalization for the realization of /t/ to /i/ environments. In word-final position, all realizations of /t/ showed palatalization, except for the /o/ environment. Figure 44 shows /t/ in all positions and environments.

Figure 44. SSANOVA for BP5’s articulation of /t/ in all vocalic environments for the word-initial position (left), intervocalic position (center), and word-final position (right).

The results for BP5 for /ʁ/ showed a similar pattern for both onset and intervocalic positions. However, the effects were much more extreme in the intervocalic position. In the onset position, there was significant resistance of both the tongue root and tongue body to phonetic environments. There was a small amount of tongue root advancement in the /i/ environment and small amount of retraction in the /o/ environment. There was also a small but significant
difference in tongue root advancement between the /a/ and /e/ environments, with /e/ being slightly more advanced. In the intervocalic position, the tongue root and body shape and position for /ʁ/ varied greatly based on the vocalic context. The most retraction was produced in the /o/ environment, which was accompanied by tongue body raising and retraction. /i/ and /e/ both produced tongue root advancement and tongue body fronting and raising. /ɾ/ exhibited a relatively invariant tongue root in the intervocalic position, but there was some advancement in the /i/ environment and some retraction in the /o/ environment. In the /o/ and /i/ environments, the tongue body was also raised. /a/ and /e/ environments differed slightly in degree of tongue body raising, but there was only a small – yet significant – difference in the tongue root advancement. In the word-final position, all of the articulations for coda /ʁ/ were realized as a bunched approximant [ʃ]. There was almost no variation in the tongue root or body in each environment; however, there was a small, but significant amount of tongue root advancement in the /i/ environment. Figure 45 and 46 show the SSANOVA results for /ʁ/, /ɾ/ and coda /ʁ/.

Figure 45. SSANOVA results for BP5’s articulation of /ʁ/ in word-initial position (left) and intervocalic position (right).

Figure 46. SSANOVA results for BP5’s articulation of /ɾ/ in intervocalic position (left) and /ʁ/ in word-final position (right).
The overall SSANOVA results indicated that /k/ consisted of a closure made with the posterior region of the tongue body. However, it was highly susceptible to phonetic environments and underwent significant fronting in the /e/ and /i/ environments and was articulated in this case with the anterior region of the tongue body. In the /o/ environment, there was also significant retraction of the tongue root for /k/. The influence of phonetic environment was present in word-initial and intervocalic environments and the degree to which /k/ retracted/advanced did not particularly vary. In word-final environments, /k/ was palatalized by all speakers.

The results for /t/ indicate that it was primarily produced with the tongue tip.blade. The tongue body and root were highly susceptible to phonetic environment. In /o/ environments, the tongue root was retracted for the realization of /t/ and in /e/ and /i/ environments, the tongue root was advanced. In all word-final positions, the tongue blade tip underwent a slight degree of fronting and the resulting articulation is slightly palatalized, producing an articulation closer to [ʃ]. This happens irrespective of whether the surrounding phonetic environment is a high front vowel or not. In summary, there was no major variation between the speakers for the articulation of /t/ and /k/. See Figures 96 – 119 the Appendix for plots of each of the phonemes in each environment for every speaker.

All speakers (BP1-BP6) produced an [h] in intervocalic position for /ɾ/. This was evidenced by the lack of a consistent tongue root gesture and a tongue position that closely resembled the phonetic environment. Intervocally, /t/ was consistently realized with retraction of the tongue root coordinated with a tongue tip gesture. There was minor variation across speakers in the articulation. All speakers exhibited high degrees of resistance to coarticulatory effects for /ɾ/.

In the word-final position, the coda /R/ had substantial speaker variation. It was realized either [ɾ ɻ χ]. BP1 and BP5 both produced an approximant [ɾ]; however, there was further variation in the realization of [ɾ]. BP1 produced a much more “retroflex” articulation, with the tongue tip/blade raised and retracted to a more post-alveolar place of articulation, similar to [ɾ]. Unlike canonical retroflex productions, the tongue tip/blade was not facing upwards toward the palate but was rather perpendicular to it. BP5 produced a “bunched” articulation, with the tongue tip/blade raised towards the post-alveolar region but facing downwards. The tongue root also had a distinct retraction into the lower pharyngeal cavity. BP2 and BP4 produced an alveolar tap in
word-final position. Similar to the intervocalic position, [ɾ] showed little variation in the position of the tongue root, showing some advancement for the [i_] environment. Finally, BP3 produced a uvular fricative articulation in the word-final position. An observable degree of tongue retraction took place in each vocalic environment with noticeable frication. All speakers came from the same region of Brazil, São Paulo, and likely from a similar socio-economic class, so the variation is ideolectal and is not due to differences in regional or social dialect. Table 15 summarizes each speakers realization of coda /R/.

Table 15. Summary of the Realizations of /R/

<table>
<thead>
<tr>
<th></th>
<th>/R/</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP1</td>
<td>retroflex approximant [ɻ]</td>
</tr>
<tr>
<td>BP2</td>
<td>alveolar tap [ɾ]</td>
</tr>
<tr>
<td>BP3</td>
<td>uvular fricative [χ]</td>
</tr>
<tr>
<td>BP4</td>
<td>alveolar tap [ɾ]</td>
</tr>
<tr>
<td>BP5</td>
<td>bunched approximant [ɻ]</td>
</tr>
<tr>
<td>BP6</td>
<td>alveolar tap [ɾ]</td>
</tr>
</tbody>
</table>

5.3.2 Dynamic Results

As with the SSANOVA, the results for the controls /k/ and /t/ are presented first, followed by presentation of /ɾ/, /ɾ/, and coda /R/.

The gestural results for /k/ revealed that the tongue body was raised toward the palate in a straight up motion. This means that /k/ was overall highly susceptible to the surrounding phonetic environments. In the /e/ and /i/ environments, the fronting of the tongue body/root for the preceding or following high-front vowel impacted the target area for the tongue-palate contact by causing significant fronting. The opposite was true for the case of /a/ and /o/, the more retracted tongue body resulted in a more retracted tongue-palate contact. See Figure 47 – 49 for dynamic plots of /k/ in the environments /CV/, /VCV/, and /VC/.
Figure 47. Gestural dynamics for /k/ for BP1 in the environment /#_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 48. Gestural dynamics for /k/ for BP1 in the environment /e_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
/t/ was made with a tongue tip/blade constriction, but the tongue body/root was highly susceptible to vocalic context. There was some tongue root advancement moving towards the point of maximum constriction, but the relative position at the point of maximum was either more retracted or more advanced based on the surrounding phonetic environment. /t/ also underwent palatalization in the /e/ and /i/ environments, which results in a more raised tongue body/blade and a closure made primarily with the tongue blade. In word-final position, the tongue blade was always raised to produce a more palatalized articulation. This occurred regardless of the phonetic environment. See Figure 50 – 52 for dynamic plots of /t/ in the environments /CV/, /VCV/, and /VC/. 

Figure 50. Gestural dynamics for /t/ for BP1 in the environment /#_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
The gestural results for /ʁ/ showed two different patterns depending on word position. Word-initially, there was a consistent tongue root gesture, involving retraction towards a uvular place of articulation. The dynamics also suggest some degree of tongue tip/blade activation. This was particularly visible in the /#_a/, /#_e/, /#_i/ environments, where there was clear raising of the tip/blade. In the intervocalic position, the root gesture appeared to be weakened significantly. The entire tongue shape was influenced strongly by the flanking vowels; this was most strongly observed in the /o_o/ and /i_i/ environment. For some speakers, in this position, there was no difference between the onset, offset and the point of maximum constriction. See Figure 53 – 54 for dynamic plots of /ʁ/ in the environments /CV/ and /VCV/.
Figure 53. Gestural dynamics for /ʁ/ for BP1 in the environment /#_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 54. Gestural dynamics for /ʁ/ for BP1 in the environment /e_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

The dynamic results for /ɾ/ showed both a consistent tongue root gesture and a tongue tip/blade gesture. In the intervocalic position, there was a consistent gesture moving the root into the pharyngeal cavity, observed by the advancement/retraction of the tongue. In the case of /o_o/, clear advancement from the previously retracted position was observed, while in the other positions retraction occurs to achieve the gestural target. In every position, there was a short duration tongue tip/blade raising to create the brief closure associated with a tap. There was significant variation of the specific gestural dynamics of coda /R/ in the word-final position depending on the speaker. However, all speakers showed the same general tendency, the coordination of a tongue root gesture and a tongue tip/blade gesture. For the speakers that
produced and approximant, there were both “retroflex” (BP1) and “bunched” (BP5) articulations. The “retroflex” involved a raising of the tongue tip/blade into the post-alveolar region. There was clear tongue root retraction for this articulation; this was most evident in the lower areas of the root. “Bunched” articulations showed the same tongue root retraction, but instead the tongue blade/body raised while the tongue tip remained facing down. See Figure 55 for dynamic plots of /ɾ/ in the environment /VCV/ and figure 56 for dynamic plots of coda /R/ in the /VC/ environment.

Figure 55. Gestural dynamics for /ɾ/ for BP1 in the environment /e_e/, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.

Figure 56. Gestural dynamics for coda /R/ for BP1 in the environment /e_/#, including the gestures towards the maximum constriction (left) and away from it (right). Tongue tip is on the right.
5.3.3 Coarticulatory Results

The results of the multiple linear regression, after Satterthwaite approximation for degrees of freedom, indicated a significant result for the factors phoneme \([F(4, 57.75) = 14.12, p < 0.0001]\) and environment \([F(2, 38.71) = 40.31, p < 0.0001]\). Post-hoc \(t\)-tests were performed with a Bonferroni correction and indicated a significant difference between \(/t/\) (RMS: 4.06) and both \(/k/\) (RMS: 5.56, \(p = 0.0027\)), \(/ʁ/\) (RMS: 6.37, \(p < 0.0001\)), and \(/R/\) (RMS: 2.18, \(p = 0.017\)), but not \(/ɾ/\) (RMS: 3.89, \(p = 0.1818\)). \(/t/\) had a lower RMS than \(/k/\) and \(/ʁ/\), but a higher RMS than \(/R/\). \(/k/\) was significantly different from \(/t/\) (\(p < 0.0522\)) and \(/R/\) (\(p < 0.0001\)), but not \(/ʁ/\) (\(p = 0.8285\)). \(/k/\) was had a higher RMS than \(/ɾ/\) and \(/R/\). There was a significant difference between \(/t/\) and \(/ʁ/\) (\(p < 0.010\)), but not \(/R/\) (\(p = 0.1948\)). \(/t/\) had a lower RMS than \(/R/\). Finally, there was a significant difference between \(/R/\) and \(/ʁ/\) (\(p < 0.0001\)), where \(/R/\) had a lower RMS than \(/ʁ/\). Figure 57 shows a bar plot of each phoneme in each environment and Table 16 shows the means and standard deviation of each phoneme in each environment and the means and standard deviation of each phoneme and each environment. Table 17 presents the results of the mixed effects model.

![Bar plot of the mean root mean square for each phoneme in each environment.](image)

*Figure 57. Bar plot of the mean root mean square for each phoneme in each environment.*

76
Table 16. Means and standard deviations of each phoneme in each environment.

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>VCV</th>
<th>VC</th>
<th>means</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>3.77 (1.91)</td>
<td>5.46 (1.79)</td>
<td>2.96 (2.71)</td>
<td>4.06 (2.40)</td>
</tr>
<tr>
<td>r</td>
<td>-</td>
<td>3.89 (1.56)</td>
<td>-</td>
<td>3.89 (1.56)</td>
</tr>
<tr>
<td>R</td>
<td>-</td>
<td>-</td>
<td>2.18 (1.14)</td>
<td>2.18 (1.14)</td>
</tr>
<tr>
<td>ʁ</td>
<td>4.24 (2.08)</td>
<td>8.51 (4.11)</td>
<td>-</td>
<td>6.37 (3.88)</td>
</tr>
<tr>
<td>k</td>
<td>6.91 (3.81)</td>
<td>6.97 (3.93)</td>
<td>2.79 (1.64)</td>
<td>5.56 (3.81)</td>
</tr>
<tr>
<td>means</td>
<td>4.97 (3.04)</td>
<td>6.21 (3.64)</td>
<td>2.64 (1.57)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 17. Summary of the mixed effects model (***(*** indicates significant p-values))

|        | Estimate | Std. Error | df | t-value | Pr (>|t|) |
|--------|----------|------------|----|---------|----------|
| (Intercept) | 6.9063 | 0.4950 | 27.20 | 13.952 | < 0.0001 *** |
| r | -3.0854 | 0.6439 | 35.10 | -4.792 | < 0.0001 *** |
| R | -0.6035 | 0.6219 | 173.30 | -0.970 | 0.3332 |
| ʁ | -2.6699 | 0.6216 | 218.90 | -4.295 | < 0.0001 *** |
| t | -3.1412 | 0.6194 | 332.70 | -5.072 | < 0.0001 *** |
| VC | -4.1185 | 0.6257 | 152.00 | -6.582 | < 0.0001 *** |
| VCV | 0.0683 | 0.6219 | 252.20 | 0.110 | 0.9126 |
| t:VC | 3.3151 | 0.8757 | 343.40 | 3.786 | 0.0002 *** |
| ʁ:VCV | 4.2026 | 0.8757 | 343.40 | 4.799 | < 0.0001 *** |
| t:VCV | 1.6301 | 0.8757 | 343.40 | 1.861 | 0.0635 |

5.4 Conclusions

This chapter examined the articulatory properties of the Brazilian Portuguese rhotics with the goal of determining the gestural composition of the uvular fricative, alveolar tap, and coda /R/ for Brazilian Portuguese and how that effects their synchronic and diachronic patterns. The study revealed /ʁ/ in onset positions had consistent tongue root retraction and uvular-pharyngeal constriction. There was a high degree of coarticulatory resistance, indicated by a low RMS score. In intervocalic positions, /ʁ/ had a highly variable place of articulation, dependent wholly on phonetic context. This suggests the speakers tested here have an [h]-like articulation intervocally. The weak resistance to coarticulatory effects was reflected in a high RMS score.
in the intervocalic position. In short, /r/ is best described as a consistent tongue root and body gesture constriction in the uvular-pharyngeal region.

The data presented here suggests that the gestural composition of /r/ involves the coordination of a tongue root, tongue body, and tongue tip gesture. This is best observed by the relative resistance of the tongue root and body to phonetic environments. The tongue root and body position were relatively stable, showing some retraction in the /o_o/ environment and some advancement in the /i_i/ environment. This was accompanied by a consistent tongue tip gesture, clearly observed by a motion towards the palate.

Coda /R/ had a variety of realizations dependent on speaker. Three speakers produced an alveolar tap, two speakers produced an approximant, and one speaker produced a uvular fricative. In all cases, there was a consistent tongue root and tongue body gesture. The target varied by articulation. The alveolar taps and uvular fricative all exhibited a retracted tongue root and tongue body. The approximants had a retracted tongue root, but the tongue body was more advanced, making a constriction in the palatal region. The tongue tip/blade was, in both cases, making an alveolar constriction.

The finding is interpreted that tongue tip rhotics are comprised of complex articulations involving the whole tongue. This assertion also accounts for the high coarticulatory resistance of these segments. The notion that uvular rhotics are composed of a tongue root and tongue body gesture is also supported in the evidence presented here. This may also explain the ease of diachronic changes from alveolar to uvular place of articulation among rhotics. Loss of the tongue tip gesture results in a simplified coordination between the tongue root and tongue body. This leaves a uvular-pharyngeal constriction as the primary place of articulation. Theoretical implications are discussed in detail in Chapter 7, Section 7.2.2.
Chapter 6
Rhotic perception provides evidence for classhood and accounts for cross-linguistic distribution of rhotics in phonemic inventories

6.1 Introduction
Chapter 2 provided evidence for a consistent formant frequency and trajectory for F2 among Sorbian alveolar and uvular rhotics, and chapters 3 and 4 provided ultrasound evidence for a consistent tongue root gesture for uvular rhotics. In chapter 5, the results indicated there was also evidence for a tongue root gesture in both uvular and alveolar rhotics in Brazilian Portuguese. The consistent acoustic-articulatory properties might suggest a perceptual correlate to rhotics as a class. Therefore, this chapter is aimed at examining exactly that question.

5.1.1 Hypothesis
At the center of this research is the hypothesis that there is a perceptual correlate to rhotic classhood. This hypothesis is expected to manifest itself in the research with speakers having low perceptual sensitivity for rhotic/rhotic comparisons, indicated by low d-prime scores, and for relatively small perceptual areas (see 5.2.4 for explanation of what the perceptual area is) for rhotics as a group. It is also anticipated that the rhotics will group together in the perceptual space, as will other groups, but they will all form separate, distinct groupings in the perceptual space.

The perceptual magnet effect (e.g. Grieser & Kuhl, 1989, Kuhl, 1991, Iverson & Kuhl, 1995) is at the core of the experimental design. The perceptual magnet effect suggests the adults’ perceptual map is altered or distorted relative to the linguistic input of their native language. Kuhl, Williams, Lacerda, Stevens, and Lindblom (1992) performed an identification task comparing [i] and [y] tokens with less prototypical members of each class. This was achieved by manipulating F1 and F2 for [i] and [y] tokens to get them further away in formant frequencies from the prototypical ones. American and Swedish native speakers then performed an identification task. Listeners heard a token and had to identify it as either [i] or [y]. The results revealed that vowel identification was influenced by a perceptual magnet effect: tokens close to a prototype token were tightly clustered and received high ratings among listeners; less prototypical members fell further away from the prototype in the perceptual space and were also rated lower among listeners. Ultimately, the data suggests that perceptually similar sounds are
attracted to an internalized prototype phoneme caused by a distortion of the perceptual space due to native language input. The results presented in Kuhl et al. (1992) are consistent with a perceptual magnet model of speech category perception.

Thus, the underlying assumption of this research is that non-native sounds with similar acoustic/perceptual characteristics will map on to the same internalized phoneme. For this reason, native speakers of English were chosen for a perception experiment involving all non-native English sounds. The hypothesis is that the warping of the perceptual space due to being a native speaker of English will cause perceptually similar sounds to be attracted to a perceptual node for a native sound. The goal is to observe whether non-native rhotics are all attracted to an internalized rhotic phoneme. If the rhotics all cluster closely together, it will indicate that they are similar in perceptual characteristics and that they have all been attracted to a single internalized phoneme on listeners’ the perceptual map.

6.2 Methods
6.2.1 Participants
Forty-four native speakers of Canadian English were recruited through the Linguistics participant pool at the University of Toronto. All participants were undergraduate students aged 18-24 years enrolled in LIN228H1F (Phonetics). Thirty-eight speakers were female and 6 were male. All participants provided informed consent and had no self-reported speaking or hearing problems.

6.2.2 Stimuli
The stimuli were composed of three rhotics, /r/, /ɻ/, and /ʀ/, three stops, /d/ɼ/, /ɖ/ɼ/, and /ɟ/ɼ/, three nasals, /n/ɼ/, /ɳ/ɼ/, and /ɲ/ɼ/, three fricatives, /z̪/ɼ/, /ʐ/ɼ/, and /ʑ/ɼ/, and three laterals /l/ɼ/, /ɭ/ɼ/, and /ɭ/ɼ/. These tokens were selected on the basis that they represent 5 different natural classes (rhotics, stops, nasals, fricatives, and laterals) and 5 different places of articulation (alveolar, retroflex, alveolopalatal, palatal, and uvular). For example, the rhotics provide a range of acoustic differences within each natural class because of the inclusion of an alveolar, /r/, retroflex, /ɻ/, and uvular, /ʀ/, place of articulation. Each of the stimuli were produced by one male and one female native speaker of a language which contains the target phoneme and was the basis for selecting each of the languages (/r/, /ɻ/, and /ʀ/: Czech; /d/ɼ/, /n/ɼ/, /ɭ/ɼ/, and /ɭ/ɼ/: Russian; /ʀ/: Upper Sorbian; /z̪/ɼ/, /ʐ/ɼ/, and /ʑ/ɼ/: Polish; /ɭ/, /ɭ/ɼ/, /ɭ/ɼ/, and /ɭ/ɼ/: Malayalam). All speakers were aged 20-24 years. Stimuli were produced in a /Ca/ sequence 6 times. All tokens were recorded in a sound
attenuated booth at the Phonetics Lab at the University of Toronto with a sampling frequency of 48,000 Hz and 16-bit, except the Upper Sorbian speakers who were recorded at the Phonetics Lab at the Max Planck Institute for Evolutionary Anthropology at the same sampling frequency and bit depth.

All tokens were normalized to an average of 70 dB SPL and a duration of 220 ms. A mean of 70 dB SPL was set for each token in Praat (Boersma & Weenink, 2014) using the scale intensity function. Duration normalization was achieved by clipping the end of the vowel so that the target phoneme and vowel sequence was 220 ms in duration. The full target phoneme was always present and no less than 75% of the original duration of the vowel was present. The target phoneme was always aligned to the left edge of the audio file and the end of the vowel was aligned to the right edge. The beginning of the trills, /ɾ/ and /ɻ/, was taken to be the initial contact made in production. This was clearly visible by a spectral occlusion in the acoustic signal (Figure 58). The onset of the retroflex rhotic, /ɻ/, was taken to be the beginning of the steady state portion of F2 and F3 on the spectrogram. Auditory confirmation of r-fullness was also carried out to ensure the normalized tokens were suitable for the experiment. An SSANOVA plot was also created to compare the formant structure of the tokens used for the perceptual experiment (Figure 59) and the formant frequencies at the mid-point of each rhotic is presented in Table 18.

![Figure 58. Spectrogram for the token /ɾ/ as produced by the female native Russian speaker. The left edge is aligned to the onset of the tongue-palate contact. The occlusion for trill contact can be clearly observed on the left edge.](image-url)
Figure 59. F1-F3 of the rhotics used in the perceptual experiment. The left edge is the onset of the rhotic, the mid-point is the end of the rhotic, and the right edge is the mid-point of the vowel.

Table 18. Mean formant frequencies (Hz) for the rhotics used in the experiment

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>624</td>
<td>1639</td>
<td>2315</td>
</tr>
<tr>
<td>l</td>
<td>453</td>
<td>1817</td>
<td>2561</td>
</tr>
<tr>
<td>R</td>
<td>670</td>
<td>1402</td>
<td>2865</td>
</tr>
</tbody>
</table>

6.2.3 Procedure

Each participant took part in two AX discrimination tasks out of a possible four (rhotics vs. stops, rhotics vs. nasals, rhotics vs. fricatives, or rhotics vs. laterals). See Table 19 for a list of each of the phonemes used for comparison in each experimental condition. The experiments that the participants took part in were randomized, but they were counterbalanced to control for which experiment each person participated in and which order the experiments were done in. Twenty-two speakers participated in the rhotics vs. stops experimental condition, 21 speakers participated in the rhotics vs. nasals experimental condition, 25 speakers participated in the rhotics vs. fricatives experimental condition, and 20 speakers participated in the rhotics vs. lateral experimental condition. Each token was heard 10 times, for a total of 300 tokens per listener per condition, and a total of 600 tokens per listener (44 listeners x 600 tokens = 26,400 total tokens). Stimuli was presented in a randomized order and each comparison was balanced for which segment came first (e.g. 5 tokens of r vs. d and 5 tokens of d vs. r).
Participants sat in front of a computer and wore a pair of Sony MDR-7506 headphones. The experiment was presented using DMDX (Forster & Forster, 2003). Participants were prompted with instructions at the beginning of the experiment. Participants were instructed to press $F$ for pairs they perceived as the same and $J$ for pairs they perceived as different. Participants were given two breaks during each of the experimental blocks, which lasted until they were ready to progress by pressing the spacebar (this was never longer than 3 minutes). A fixation point appeared at the beginning of each trial for a randomized time between 835 ms and 1,250 ms. After the prompt, the pair would be played with a 500 ms duration between each sound. Participants were given a maximum of 2500 ms to answer before the trial timed-out and the next trial began. After the experiment was completed, a short break was again given while the second experiment was loaded. Participants were again prompted with instructions at the beginning of the main experiment. Instructions, breaks, and format were the same for both experimental blocks.

### 6.2.4 Analysis

Responses that were faster than 300 ms or slower than 2300 ms were discarded from the analysis due to high probability of listener error in those ranges of response times (8.5% of 26,400 tokens). $d$-prime scores (Macmillan & Creelman, 2004) were calculated for each pair in each condition. $d$-prime scores of 2.0 are considered typical. Below 2.0 is generally considered poor. Scores above 3.0 are generally considered to indicate good discriminability. 4.5 and above are exceptional scores which indicate near perfect discriminability.

Four repeated measures ANOVA were performed (1 ANOVA for each experiment) in R (R Development Core Team, 2017) with the factor Comparison (composed of each possible $d$-prime comparison), which was comprised of the $d$-prime results for each pair (e.g. $d$-prime scores for $r$ vs. $ʀ$ were compared against the $d$-prime score for $r$ vs. $ɻ$; this was done for all
possible comparisons in each experimental condition). Post-hoc analyses were performed with pairwise *t*-tests with a Bonferroni correction for multiple comparisons.

Multidimensional scaling (MDS; Johnson, 2008) was also performed on each individual participant and on the mean data using d-primes. The individual MDS data coordinates were used to calculate the perceptual area in *perceptual units (pu)* for each group of sounds (rhotics, stops, nasals, fricatives, laterals). This was done using the Surveyor’s Area Formula (Braden, 1986) to calculate the area of the triangle formed by the three *x*, *y*-coordinates from each phoneme from each natural class:

\[
\text{area} = \frac{1}{2} |x_1y_2 + x_2y_3 + x_3y_1 - x_2y_1 - x_3y_2 - x_1y_3|
\]

The areas for each group was compared using a Welch two sample *t*-test for each of the four experiments. A multiple linear regression was also used to compare the perceptual areas formed by the rhotics in each of the experimental conditions. The model compared the area formed by the rhotics by the factor experimental condition as the *fixed effect* (4 levels: stops, nasals, fricatives, laterals). Participant (44 levels, 1 for each participant) was set as the *random effect*. The mean data was used for data visualization, which was done using the sci-plot package (Morales, 2015) in R (R Development Core Team, 2017).

A kmeans clustering analysis (Kanungo, Mount, Netanyahu, Piatko, Silverman, & Wu, 2002) was performed on the coordinates of the MDS solutions in order to determine how well the data fit two distinct clusters forming two perceptual categories for each of the pairs of natural classes examined in each experiment.

## 6.3 Results

In this section, first the d-prime analysis is presented (Section 5.3.1), followed by the presentation of the MDS results (Section 5.3.2) and the kmeans results (Section 5.3.3).

### 6.3.1 d-prime Results

Comparing rhotics and stops, Mauchly’s Test for Sphericity revealed that the data was not spherical \([W(2) < 0.01, p < 0.0001]\). After Greenhouse-Geisser correction \([\text{GGGe} = 0.3836]\), there was a main effect for Comparison \([F(2, 14) = 34.31, p < 0.0001]\). Crucially, post-hoc tests with a Bonferroni correction revealed that the rhotic comparisons (r vs. r, r vs. l, l vs. r) all had a lower d-primes \((p < 0.0001)\) than every other comparison, except for the comparison between d vs. j \((p\)
= 1.0000). Figure 60 shows a bar plot of the results and Table 20 shows mean d-primes with standard deviations presented in parentheses.

![Bar plot of d-prime results for rhotics and stops.](image)

**Figure 60. Bar plot of the d-prime results for the comparison between rhotics and stops.**

**Table 20. d-prime means and standard deviations for rhotic and stop comparisons**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Comparison</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>r vs. r</td>
<td>0.2161 (1.10)</td>
<td>r vs. l</td>
</tr>
<tr>
<td>r vs. l</td>
<td>-0.0809 (0.88)</td>
<td>r vs. d̂</td>
</tr>
<tr>
<td>r vs. d̂</td>
<td>3.1078 (0.88)</td>
<td>r vs. ɻ</td>
</tr>
<tr>
<td>r vs. ɻ</td>
<td>2.9502 (1.01)</td>
<td>r vs. ɭ</td>
</tr>
<tr>
<td>r vs. ɭ</td>
<td>2.9458 (1.08)</td>
<td>ɭ vs. d̂</td>
</tr>
<tr>
<td>ɭ vs. d̂</td>
<td>3.1135 (0.97)</td>
<td>ɭ vs. ɭ</td>
</tr>
<tr>
<td>ɭ vs. ɭ</td>
<td>2.9724 (1.01)</td>
<td>d̂ vs. ɭ</td>
</tr>
<tr>
<td>d̂ vs. ɭ</td>
<td>2.9458 (1.08)</td>
<td>ɭ vs. ɭ</td>
</tr>
</tbody>
</table>

Comparing rhotics and nasals, Mauchly’s Test for Sphericity revealed the data was not spherical [W(2) < 0.01, p < 0.0001]. After Greenhouse-Geisser correction [GGe = 0.4287], there was a main effect of Comparison [F(2, 14) = 37.69, p < 0.0001]. Post-hoc tests with Bonferroni correction revealed that, crucially, the rhotic comparisons (r vs. r, r vs. ɭ, ɭ vs. r) all had lower d-primes (p < 0.0001) than every other comparison, except in this case, all the rhotic comparisons were not significantly different from ñ vs. ɭ (p = 1.0000). Figure 61 shows a bar plot of the results and Table 21 shows mean d-primes with standard deviations presented in brackets.
Figure 61. Bar plot of the d-prime results for the comparison between rhotics and nasals.

Table 21. d-prime means and standard deviations for rhotic and nasals comparisons

<table>
<thead>
<tr>
<th>Comparison</th>
<th>d-prime mean (std dev)</th>
<th>Comparison</th>
<th>d-prime mean (std dev)</th>
<th>Comparison</th>
<th>d-prime mean (std dev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>r vs. r</td>
<td>0.1852 (1.15)</td>
<td>r vs. ṭ</td>
<td>0.4432 (1.18)</td>
<td>ṭ vs. ɳ</td>
<td>2.6833 (0.89)</td>
</tr>
<tr>
<td>r vs. ṭ</td>
<td>0.0081 (1.31)</td>
<td>r vs. ɲ̊</td>
<td>2.8078 (1.13)</td>
<td>ɲ̊ vs. ɳ</td>
<td>2.7472 (1.04)</td>
</tr>
<tr>
<td>r vs. ɲ̊</td>
<td>2.2801 (1.02)</td>
<td>r vs. ɳ</td>
<td>2.8090 (1.01)</td>
<td>ɳ vs. ɲ</td>
<td>2.3110 (1.49)</td>
</tr>
<tr>
<td>r vs. ɳ</td>
<td>2.9988 (0.95)</td>
<td>r vs. ɲ̊</td>
<td>2.6155 (0.97)</td>
<td>ɲ̊ vs. ɳ</td>
<td>0.4017 (0.93)</td>
</tr>
<tr>
<td>r vs. ɲ</td>
<td>2.8681 (0.96)</td>
<td>ɲ vs. ɲ̊</td>
<td>2.8378 (0.97)</td>
<td>ɳ vs. ɲ</td>
<td>1.8942 (1.72)</td>
</tr>
</tbody>
</table>

Comparing rhotics and fricatives, Mauchly’s Test for Sphericity revealed data was not spherical \( W(2) < 0.01, p < 0.0001 \). After Greenhouse-Geisser correction \( GGe = 0.4549 \), there was a main effect of Comparison \( F(2, 14) = 60.32, p < 0.0001 \). Post-hoc tests with Bonferroni corrections revealed that, crucially, the rhotic comparisons (\( r \) vs. \( r \), \( r \) vs. \( ɭ \), \( ɭ \) vs. \( ɭ \)) all had lower d-primes \( p < 0.0001 \) than every other comparison, except in this case, all the rhotic comparisons were not significantly different from \( ʂ \) vs. \( ʐ \) \( (p = 1.0000) \). Figure 62 shows a bar plot of the results and Table 22 shows mean d-primes with standard deviations presented in brackets.
Comparing rhotics and laterals, Mauchly’s Test for Sphericity revealed data was not spherical \([W(2) < 0.01, p < 0.0001]\). After Greenhouse-Geisser correction \([GGe = 0.3588]\), there was a main effect of Comparison \([F(2, 14) = 21.79, p < 0.0001]\). Post-hoc comparisons with Bonferroni corrections revealed that the rhotic comparisons (\(r\) vs. \(r\), \(r\) vs. \(l\), \(r\) vs. \(z\)) had a lower \(d\)-prime value \((p < 0.0001)\) than the lateral/lateral or rhotic/lateral comparisons. However, the comparison between the rhotics was not found to be significantly different from \(l\) vs. \(l\) \((p = 1.0000)\). \(r\) vs. \(r\) was also not found to be significantly different from \(l\) vs. \(l\) \((p = 0.0866)\) and \(l\) vs. \(z\) was not found to be significantly different from \(l\) vs. \(l\) \((p = 0.5552)\). Figure 63 shows a bar plot of the results and Table 23 shows mean \(d\)-primes with standard deviations presented in brackets.

**Figure 62. Bar plot of the \(d\)-prime results for the comparison between rhotics and fricatives.**

**Table 22. \(d\)-prime means and standard deviations for rhotic and fricative comparisons**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Comparison</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r) vs. (r)</td>
<td>0.5425 (0.92)</td>
<td>(r) vs. (l)</td>
</tr>
<tr>
<td>(r) vs. (l)</td>
<td>-0.4132 (0.91)</td>
<td>(r) vs. (z)</td>
</tr>
<tr>
<td>(r) vs. (z)</td>
<td>2.8527 (1.03)</td>
<td>(r) vs. (l)</td>
</tr>
<tr>
<td>(r) vs. (l)</td>
<td>3.2070 (0.92)</td>
<td>(r) vs. (z)</td>
</tr>
<tr>
<td>(r) vs. (z)</td>
<td>3.1630 (0.98)</td>
<td>(l) vs. (l)</td>
</tr>
</tbody>
</table>
88

Figure 63. Bar plot of the d-prime results for the comparison between rhotics and laterals.

Table 23. d-prime means and standard deviations for rhotic and laterals comparisons

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Comparison</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>r vs. r</td>
<td>r vs. l</td>
<td>r vs. l</td>
</tr>
<tr>
<td>0.6042 (1.44)</td>
<td>2.6620 (0.85)</td>
<td>2.4097 (1.03)</td>
</tr>
<tr>
<td>r vs. l</td>
<td>r vs. l</td>
<td>r vs. l</td>
</tr>
<tr>
<td>-0.5674 (0.68)</td>
<td>2.4630 (1.18)</td>
<td>1.9257 (1.24)</td>
</tr>
<tr>
<td>r vs. l</td>
<td>r vs. l</td>
<td>r vs. l</td>
</tr>
<tr>
<td>2.6120 (1.08)</td>
<td>3.1042 (0.87)</td>
<td>1.8061 (1.48)</td>
</tr>
<tr>
<td>r vs. l</td>
<td>r vs. l</td>
<td>r vs. l</td>
</tr>
<tr>
<td>2.7229 (1.04)</td>
<td>0.9264 (1.32)</td>
<td>0.3297 (1.09)</td>
</tr>
<tr>
<td>r vs. l</td>
<td>r vs. l</td>
<td>r vs. l</td>
</tr>
<tr>
<td>2.7734 (0.89)</td>
<td>2.3069 (0.86)</td>
<td>2.0926 (1.45)</td>
</tr>
</tbody>
</table>

6.3.2 MDS Results

The comparison of the perceptual area of rhotics and stops revealed a significant difference
[t(27.90) = 3.24, p = 0.0031] between the two. The rhotics had a mean perceptual area of 0.4258
pu² (SD, 0.7230 pu²) and the stops had a mean perceptual area of 1.7404 pu² (SD, 1.7582 pu²).
The MDS indicates clear perceptual separation between the two phonological classes, rhotics and
stops, and that the overall perceptual space for stops is larger. The results are presented in Figure
63.
Figure 64. a. Bar plot of the mean perceptual units for the rhotics vs. stops comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and stops.

The comparison of the perceptual area of rhotics and nasals revealed a significant difference \( t(34.71) = 2.53, p = 0.0157 \) between the two. The mean perceptual area for the rhotics was 0.5288 pu\(^2\) (SD, 0.7865 pu\(^2\)) and the nasals has a mean perceptual area of 1.0685 pu\(^2\) (SD, 0.5775 pu\(^2\)). The MDS indicate clear perceptual separation between the two phonological classes, rhotics and nasals, and that the overall perceptual space for nasals is larger. The results are presented in Figure 65.

Figure 65. a. Bar plot of the mean perceptual units for the rhotics vs. nasals comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and nasals.
The comparison of the perceptual area of the rhotics and the fricatives revealed a significant difference \([t(24.93) = -4.72, p < 0.0001]\) between the two. The mean perceptual area for the rhotics was 0.3836 \(pu^2\) (SD, 0.4132 \(pu^2\)) and fricatives had a mean perceptual area of 3.2118 \(pu^2\) (SD, 2.9701 \(pu^2\)). The MDS indicate clear perceptual separation between the two phonological classes, rhotics and fricatives, and that the overall perceptual space for fricatives is larger. The results are presented in Figure 66.

Figure 66. a. Bar plot of the mean perceptual units for the rhotics vs. fricatives comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and fricatives.

The comparison of the perceptual area of the rhotics and the laterals revealed no significant difference \([t(37.70) = -0.02, p = 0.9881]\) between the two. The mean perceptual area for the rhotics was 1.3847 \(pu^2\) (SD, 1.8819 \(pu^2\)) and laterals had a mean perceptual area of 1.3762 \(pu^2\) (SD, 1.7203 \(pu^2\)). The MDS results indicated clear perceptual distinction between the rhotics and the laterals, but that the overall perceptual space was not significantly different in size. The results are presented in Figure 67.
Figure 67. a. Bar plot of the mean perceptual units for the rhotics vs. lateral comparison. The whiskers indicate the standard error. b. MDS of the mean coordinates across all speakers for the comparisons between the rhotics and laterals.

The linear mixed effects model of the perceptual area of the rhotics across experimental conditions revealed a significant difference between conditions. The perceptual area of the rhotics in the stop condition \((p = 0.2666)\), nasal condition \((p = 0.7524)\), and fricative condition \((p = 0.9574)\) all had no significant difference from each other. However, the comparison with the lateral condition revealed a significant difference \((p = 0.0056)\), when compared to the other conditions. Table 24 gives a summary of the statistics for the linear regression model and Figure 68 shows a bar plot of each comparison.

Table 24. Results of the linear regression model (** indicates a significant result)

|                | Estimate | Std. Error | t-value | Pr(>|t|) |
|----------------|----------|------------|---------|----------|
| (Intercept)    | 0.3637   | 0.3251     | 1.119   | 0.2666   |
| Nasal Condition| 0.1057   | 0.3340     | 0.317   | 0.7524   |
| Fricative Condition | -0.0179 | 0.3340     | -0.054  | 0.9574   |
| Lateral Condition | 0.9643  | 0.3386     | 2.848   | 0.0056 **|
6.3.3 kmeans Results

The results of the kmeans analysis indicated a clear distinction in the perceptual space between the rhotics and stops (within cluster sum of squares by cluster: 230.018, 329.551; \( SS_b: 1354.60 / SS_{tot}: 1914.16 = 70.8\% \)), the rhotics and the nasals (within cluster sum of squares by cluster: 274.265, 180.536; \( SS_b: 1146.94 / SS_{tot}: 1601.74 = 71.6\% \)), the rhotics and the fricatives (within cluster sum of squares by cluster: 256.601, 548.788; \( SS_b: 1556.01 / SS_{tot}: 2361.39 = 65.9\% \)), and the rhotics and the laterals (within cluster sum of squares by cluster: 287.401, 233.461; \( SS_b: 1006.68 / SS_{tot}: 1527.54 = 65.9\% \)). The kmeans results can be interpreted by examining the difference between the between groups sum of squares and the total sum of squares. The ratio of those two measures indicates the significance of the fit. The rhotics and stops and rhotics and nasals comparisons were both above 70%, and the rhotics and fricatives and rhotics and laterals were both above 65%. In both cases, this indicates that the data is fit nicely into two clusters. In all cases, the clusters were distributed into two categories, one on the left of each plot and one on the right side of the plot. Table 25 presents the original values of the original points by group and how they were distributed into each group in the kmeans analysis. Each column is labeled with a category name. All the numbers under that column represent how many points in the analysis were originally from that group. Each row represents a cluster. Everything in the row 1 is distributed into the first cluster and everything in row 2 is distributed into the second cluster. By
looking at the intersection of each column and row, it can be seen how many of each of the categories tested fit into cluster 1 and how many were redistributed into cluster 2. The same can be done for the distribution of data into cluster 2 and the redistribution into cluster 1. Figure 69 – 72 presents plots of the kmeans analysis. The plots are accompanied by line graphs of the total within sum of squares for each possible cluster analysis from 1 – 5 clusters for the given data set. In each case, there was a sharp elbow at two clusters, confirming the between/total sum of squares analysis by indicating that the data was best fit with two clusters.

Table 25. Distribution of the original groupings compared to their retribution in the kmeans analysis

<table>
<thead>
<tr>
<th></th>
<th>rhotics</th>
<th>stops</th>
<th>rhotics</th>
<th>nasals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49</td>
<td>19</td>
<td>51</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>47</td>
<td>12</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>rhotics</td>
<td>fricatives</td>
<td>rhotics</td>
<td>laterals</td>
</tr>
<tr>
<td>1</td>
<td>66</td>
<td>13</td>
<td>46</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>62</td>
<td>14</td>
<td>45</td>
</tr>
</tbody>
</table>

Figure 69. kmeans plot for the rhotics and stops (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right).
Figure 70. kmeans plot for the rhotics and nasals (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right).

Figure 71. kmeans plot for the rhotics and fricatives (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right).
Figure 72. kmeans plot for the rhotics and laterals (left); and a plot of the total within sum of squares for each possible number of clusters from 1 – 5 (right).

6.4 Discussion

The goal of this study was to test for a perceptual similarity across different rhotics. English speakers heard 3 non-native rhotics, /r/, /ɻ/, and /ʀ/, compared against 3 non-native stops, nasals, fricatives, and laterals. Listeners had a harder time discriminating rhotics than other non-native contrasts, observed by consistently low d-prime scores. The MDS solutions indicated that the rhotics were grouped in a small area in all conditions and that they were perceptually distinct from the classes of sounds tested here. The kmeans analysis confirmed this result by grouping 2 clusters, optimally, in the perceptual space that divides the rhotics and the other category of segments tested.

The results support the hypothesis that there is an acoustic similarity for rhotics across places and manners of articulation. The acoustics for the rhotics in this experiment suggests that the similarity may be due to comparable F2 formants and trajectories. This suggestion is in line with Hawkins and Slater (1994) who observed a distinct F2 lowering for English rhotics and with the findings in the previous chapters that the rhotics observed in this dissertation have a distinct tongue root retraction into the pharyngeal cavity. Chapter 7, Section 7.3 discusses the broader theoretical implications for rhotic classhood and typological distribution across languages.
Chapter 7

General Discussion and Conclusion

7.1 Discussion

The discussion of the overall results of the previous chapters is separated into 3 separate sections which synthesize the results into the major themes of this research. In Section 7.2, the results of the acoustic, articulatory, and perception experiments are synthesized to discuss characteristics motivating the class of rhotics. In Section 7.3, the findings are applied to the phonology and typology of rhotics. Section 7.4 summarizes the empirical findings and Section 7.5 summarizes the theoretical claims made in this dissertation. Finally, Section 7.6 reviews some limitations of the current thesis and outlines questions for further research.

7.2 Discussion of Acoustic and Articulatory Properties Uniting the Rhotics

In this dissertation, the acoustic, articulatory, and perceptual characteristics of rhotics were examined with the intent to identify correlates to rhotic articulation. There is a great deal of phonological evidence that rhotics are treated in the phonology as a class of sounds (Walsh-Dickey, 1996; Wiese, 2001), but there has not been a clear relationship between the phonetics of rhotics and the phonology. Lindau’s (1985) landmark study showed there was no single acoustic feature present for all members of the class of rhotics. Rather, there were a series of features that connect different members to each other (e.g. /ɾ/ and /ʀ/ both have formants, while /ɾ/ and /ɾ/ share place of articulation). This is problematic because there are clearer acoustic-articulatory correlates to other natural classes of sounds (e.g. nasals involve nasal airflow, fricatives involve turbulent airflow production).

The results suggest a strong relationship between the acoustics of the rhotics as a group, despite the findings of Lindau (1985). Lindau (1985) suggests that approximants and trills are united by the presence of formants. However, the relationship between the two manners of articulation appears to be much closer than just the presence of formants. Rather, the acoustic results revealed F1 and F2 were almost identical in both dynamics and frequencies for both the unpalatalized rhotics, /ɾ/ and /ɾ/ in Upper and Lower Sorbian. F2 was lower than contrasting /l/, suggesting Sorbian rhotics exhibit an F2 “lowering” effect much like the one discussed in Hawkins and Slater (1994) for English rhotics. There was also an elevated F1 for both the rhotics, compared to the contrasting lateral, /l/. The presence of a lowered F2 and increased F1
for both uvular and alveolar rhotics is indicative of a constriction in the pharyngeal cavity (Stevens, 1989). This finding is consistent with Boyce et al.’s (2016), hypothesis that the cross-linguistic feature common to all rhotics is a secondary constriction in the pharyngeal cavity.

The articulatory evidence presented in this dissertation largely corroborates the acoustic results. The comparison of the uvular rhotic and velar stop in Upper Sorbian indicated a significant and consistent tongue root retraction into the pharyngeal cavity coordinated with a tongue body gesture for /r/. In the framework of Articulatory Phonology (Browman & Goldstein, 1986 et seq.), the coordination of the two gestures for uvular rhotics can be represented as a uvular-pharyngeal tongue body constriction in a 0° phase relationship with a pharyngeal constriction produced with the tongue root. This differs from the representation for /ɡ/, which is solely composed of a tongue body closure at the velar place of articulation. Figures 73 and 74 show the gestural score differences between /ɡ/ and /r/ in the /oCo/ sequence. The reader is referred to Tables 1 and 2 in Chapter 1, which define all the tract variables, constriction locations, and constriction degrees.

Figure 73. Gestural score for /oɡo/. The articulators involved include the tongue root (TR) and the tongue body (TB) with the target constriction location (uvular-pharyngeal and pharyngeal).
Figure 74. Gestural score for /oɾo/. The articulators involved include the tongue root (TR) and the tongue body (TB) with the target constriction location (uvular-pharyngeal and pharyngeal).

Figure 73 presents the gestural score for /oɡo/. /ɡ/ completely lacks a tongue root gesture but does have a tongue body gesture which involves a complete closure. The tongue root gestures for the vowels are coordinated with each other, resulting in extensive retraction throughout the articulation of /ɡ/, which can be observed from the continual activation of the tongue root articulator across articulation. The lack of coarticulatory resistance for /ɡ/ also results in gestural undershoot (Lindblom, 1963) causing a more retracted tongue body closure. The undershoot effect also explains the relatively malleable place of articulation observed with /ɡ/: articulatory pressure from different phonetic environments results in differing places of articulation. In the case of Figure 73, instead of reaching the velar target of articulation, a more uvular articulation is produced due to the surrounding phonetic environment.

However, the gestural score for /oɾo/, presented in Figure 74, is different from /ɡ/ because there is a coordination of tongue body and tongue root gesture for /ɾ/ /ɾ/. The vocalic tongue root gestures between /ɾ/ and the vocalic environment are coordinated together and exert some coarticulatory effects on each other as a result. /ɾ/ is not as susceptible to coarticulatory effects because of the complex coordination of the tongue root and tongue body. The evidence presented in chapter 2 and 3 and the gestural score for /ɾ/ are also in line with current models of the Degree of Articulatory Constraint (DAC) (Recasens & Rodríguez, 2016). The tongue root is far more resistant to phonetic environments and as a result it is coordinated across the vowel and the rhotic. The tongue body shows less coarticulatory resistance than the root, as predicted by the DAC model (Recasens & Rodríguez, 2016), but it does have a relatively stable place of articulation, compared to /ɡ/. Thus, the gestural score of /oɾo/ shows the reduced coarticulatory
effects and the coordination of the tongue root and the tongue body for the articulation of the uvular rhotic.

The evidence presented in chapter 5 suggests that uvular fricative rhotics (as observed in Brazilian Portuguese) can be modelled in much the same way as approximant and trill rhotics: as a complex coordination of a tongue root and tongue body gesture. Word-initial position clearly demonstrates a tongue root and tongue body gesture. The retraction of the tongue root and raising of the tongue body suggests that articulation involves both a tongue root and tongue body gesture. The RMS score, compared to /k/, also suggests higher coarticulatory resistance; this is the result of both a tongue root and tongue body gestures involved in the articulation of /ʁ/, while /k/ involves only a tongue body gesture. Given the facts presented here, the assertion that uvular rhotics are composed of a tongue root and tongue body gesture appears to be correct.

Furthermore, there appears to be some activation of the tongue tip/blade in the case of the word-initial /ʁ/ as Silva (2003) suggested. This is evidenced by the small, but observable raising of the tip/blade during the dynamic measures. This may suggest an underlying coordination of the tongue root and the tongue tip/blade. The tongue tip/blade gesture is, however, weak and only realized as a full tongue-palate contact in the case of the correct prosodic conditions. This should be examined in future research. Figure 75 shows the underlying representation of /ʁ/.

\[
/ʁ/ \\
\text{crit}\{\text{TB(uv-phar)}\} \\
0° \\
\text{wide}\{\text{TR(phar)}\}
\]

*Figure 75. Gestural representation for /ʁ/.*

Alveolar rhotics are more complex than uvular rhotics; /ɾ/ is modelled as a tongue root, a tongue body, and a tongue tip gesture. The presence of these gestures can be observed through the dynamic and coarticulatory analysis provided in this chapter 5 of this dissertation. The relatively stable tongue body across phonetic environments suggests the presence of a phonological tongue body gesture; this is accompanied by a tongue tip gesture in the case of /ɾ/, suggesting that Proctor (2011) is on correct path with his analysis. However, there is also a reason to believe that there is a consistent tongue root gesture. The presence of the tongue root gesture is highlighted by the consistent tongue root retraction and high coarticulatory resistance (low RMS). The dynamic results strengthen these findings, showing a movement of the tongue
root into the pharyngeal cavity. The suggestion that the tongue root is involved in the articulation for taps in Brazilian Portuguese is in line with Recasens & Rodriguez’s (2017) observations for taps in Catalan and is consistent with previous assertions (Delattre & Freeman, 1968; Delattre, 1971) that rhotics are composed of the coordination of a tongue root and tongue tip gesture. Figure 76 shows the gestural representation of /ɾ/.

\[
/ɾ/
\]
\[
\text{nar\{TT(alv)\}} \quad 0°
\]
\[
\text{wide\{TB(uv-phar)\}} \quad 0°
\]
\[
\text{wide\{TR(phar)\}}
\]

*Figure 76. Gestural representation for /ɾ/.*

The acoustic and articulatory evidence reinforce the notion that there is a cross-linguistic similarity for rhotics as a class and suggests the presence of an acoustic-perceptual correlate. The primary finding of chapter 5 is that the rhotic segments all share similar acoustic-perceptual characteristics that cause non-native listeners to categorize them as the same sound. Kuhl (1991) posits that language experience alters the language perception of individuals in such a way that acoustically similar sounds are perceived as the same learned sound. This phenomenon was termed the *perceptual magnet effect* and essentially results in a distortion of the perceptual space such that L2 sounds are perceived within the category of a learned L1 sound. The fact that non-native rhotics all had extremely poor discriminability, especially compared to the discrimination of other non-native sounds, suggests that there is an underlying acoustic-perceptual similarity between the rhotics tested. This suggests that Lindau’s (1985) analysis of the acoustics of rhotics is not completely correct. The observed perceptual effects suggest that there is some acoustic feature or features that cause listeners to confuse the approximant rhotics tested here as the same category. The acoustic analysis of the rhotics used in the perceptual experiment and of the rhotics in chapter 2 suggests that F2 is a point of striking similarity between rhotics as a class. Therefore, the perceptual results are consistent with the expectation of Hawkins and Slater’s (1994) findings of an F2 “lowering” effect for English. However, it should be noted that Lindau’s (1985) study examined more than the three rhotics tested here. Therefore, further research including larger rhotic inventories that tests both perceptual and acoustic/articulatory
characteristics is necessary. Further research into the perception of rhotics as a class is also
necessary to not only pin down specific class features, but also to determine how far the
perception of rhotics as a class extends. Table 26 summarizes the proposed vocalic tract
variables for each of the rhotics discussed in this dissertation. The gestural representation for
trills and taps is the same given the below gestural representation. However, the tongue stiffness,
duration, and glottal aperture settings are theorized to be difference. These factors are argued to
contribute to the aerodynamic conditions necessary for trilling to occur.

Table 26. Hypothesized tract variable specifications for the rhotics examined in this dissertation

<table>
<thead>
<tr>
<th>Tract Variable</th>
<th>/r/</th>
<th>/ɾ/</th>
<th>/ɾ/</th>
<th>/ɾ/</th>
<th>/ɾ/</th>
<th>/ɾ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT – CL</td>
<td>alveolar</td>
<td>alveolar</td>
<td>alveolar</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TT – CD</td>
<td>nar</td>
<td>nar</td>
<td>nar</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TB – CL</td>
<td>uvu-phar</td>
<td>uvu-phar</td>
<td>uvu-phar</td>
<td>uvu-phar</td>
<td>uvu-phar</td>
<td>uvu-phar</td>
</tr>
<tr>
<td>TB – CD</td>
<td>wide</td>
<td>wide</td>
<td>wide</td>
<td>nar</td>
<td>nar</td>
<td>crit</td>
</tr>
<tr>
<td>TB – CL</td>
<td>-</td>
<td>palatal</td>
<td>-</td>
<td>-</td>
<td>palatal</td>
<td>-</td>
</tr>
<tr>
<td>TB – CD</td>
<td>-</td>
<td>wide</td>
<td>-</td>
<td>-</td>
<td>wide</td>
<td>-</td>
</tr>
<tr>
<td>TR – CL</td>
<td>phar</td>
<td>phar</td>
<td>phar</td>
<td>phar</td>
<td>phar</td>
<td>phar</td>
</tr>
<tr>
<td>TR – CD</td>
<td>nar</td>
<td>nar</td>
<td>nar</td>
<td>nar</td>
<td>nar</td>
<td>nar</td>
</tr>
</tbody>
</table>

Taken together, this thesis presents evidence that rhotics of different places and manners
are linked together by the presence of a secondary pharyngeal constriction produced by tongue
root retraction. The measurable tongue root gesture has a distinct effect on the acoustic output for
rhotics by causing a raised F1 and lowered F2. These formant effects are also perceptually
evident in the difficulty speakers have in discriminating non-native rhotics.

7.3 Discussion of the relationship between the Gestural Composition of Rhotics and their
Phonological Behaviour and Typological distribution

In Section 7.2, the gestural representation for uvular and alveolar rhotics was presented. In this
section, the gestural representation is applied to their phonological behaviour.

The gestural representation for rhotics straightforwardly accounts for rhotic vocalization.
In Upper Sorbian, coda rhotics are vocalized (e.g. /por/ ‘couple, pair’ [po], Upper Sorbian
[Jocz, 2013]). The gestural representation proposed in the previous section handles vocalization
in Upper Sorbian as lenition (Kirchner, 1998; Blevins, 2004) of the tongue body gesture. Figure
77 shows the gestural score for /eɾ/. The tongue body gesture for /ɾ/ is lenited and as a result the
coordination and phase relationship between the tongue body and tongue root is broken. The
tongue root gesture persists into the articulation of the vocalized rhotic, which accounts for the
variation in articulation observed. The tongue body gesture from the preceding vowel persists
into the coda rhotic, causing significant coarticulatory effects and accounting for the fact that the
coda rhotic surfaced in different forms dependant on the preceding vowel context. This occurs in
the coda position more prominently due to the lenition of the tongue body gesture for the rhotic.
This analysis can be extended to lenition of tongue tip rhotics with the only difference being that
the tongue tip gesture instead of the tongue body gesture will lenite in these cases.

![Gestural score for /ɛɾ/](image)

**Figure 77.** Gestural score for /ɛɾ/. The articulators involved include the tongue body (TB) and the tongue root (TR) with the target constriction location (coronal, palatal, uvular-pharyngeal, and pharyngeal).

The data also suggests that what is referred to in the literature on Brazilian Portuguese as
the ‘archiphoneme’ /R/ (Barbosa & Albano, 2004) is actually an underlying tap /ɾ/. I am making
this assertion based on the typology of sound change. Cross-linguistically, the coda position is
well known to produce lenition (Kim, 1979; Kirchner, 1998; Iverson, 2004; Mitterer & Ernestus,
2006). Historical sound changes also most commonly involve lenition (Crowley & Bowern,
2010; Campbell, 2013). If we posit that the underlying form is /ɾ/, all of the tongue tip
allophones, which were the most common allophones in this study, must be the result of tongue
tip strengthening. However, if we posit that the underlying form is /ɛɾ/, we can account for all the
changes with a typologically common process: lenition. Another case of Brazilian Portuguese
allophony – the occurrence of [ɾ] in word-final position is straightforwardly explained by a
lenition of the tongue tip gesture. The relationship between /ɛɾ/ and [ɾ] is also simply explained as
lenition. The tongue tip gesture undergoes complete lenition, while the tongue root and body
gestures persist. Figure 78 presents the gestural score for /ɪɾ/ for speakers who produce word-
The tongue tip gesture is crossed off to represent lenition of this gesture, although it is still present in the underlying gestural representation. This type of analysis also straightforwardly explains the other alternations in the word-final position: /ɾ/ > [ɾ] occurs as a result of lenition of the tongue tip gesture. In the case of approximants, the closure is reduced instead to [mid] approximant constriction. This differs in the case of the [χ] realization only in the degree that lenition occurs. The same process is still driving the phonological alternation between all the related members, creating a unified analysis.

Figure 78. Gestural score for /iɾ/ for speakers who produce word-final [χ].

The allophonic realization of /ʁ/ as [h] is explained by two principles: initial onset strengthening (Hock, 1988; Cho, 2004; Recasens & Espinosa, 2005; Recasens & Espinosa, 2007) and lenition. The consistency of the tongue root gesture in the articulation /ʁ/ in the onset is a result of articulatory strengthening in this position, causing the underlying form to surface. However, in the intervocalic position the articulatory demands of the VCV environment coupled with the general weakening when not in the word-initial position results in the loss of the tongue root gesture. The result is an [h]-like articulation with no clear articulatory target; instead the articulation more closely resembles the vocalic environment that it occurs in. Figure 79 shows the gestural score for /oɾo/. It illustrates the lenition of the tongue body and root gestures, and the influence the surrounding phonetic environment has on the articulation of /ʁ/. In the intervocalic position, the place specifications for the tongue root and tongue body are lenited and the constriction degree is also lenited for the tongue root. This results in the coordination of a placeless constriction degree (critical), with the surrounding vocalic environment. This analysis
makes the prediction that intervocalic production of /ʁ/ will be systematically realized in correspondence to the place of articulation of the surrounding environment.

![Diagram](image_url)

Figure 79. Gestural score for /oʁo/, demonstrating intervocalic lenition of /ʁ/ -> [h].

Taken together, the analysis suggested here puts forth a comprehensive view of synchronic rhotic variation and diachronic sound change in Brazilian Portuguese. Synchronic rhotic variation is achieved by the manipulation of underlying tongue gestures. The allophony between rhotics produced with a tongue tip gesture is achieved primarily through a process of lenition. This reconfigures the articulatory setting from [narrow] used to produce /ɾ/ to a [mid], approximant-like setting for [ɹ]. A complete loss of the tongue tip gesture results in uvular articulations like [χ]. Furthermore, the cross-linguistic tendency for one-way sound changes from /ɾ/ to /ʁ/ are captured through the process of manipulation of the gestural settings. The complex representation for /ɾ/ involves the tongue tip, body, and root and the historical change from /ɾ/ to /ʁ/ is captured by lenition of the tongue tip gesture. Historical changes from /ɾ/ to /ʁ/ are frequently, while changes from /ʁ/ to /ɾ/ are less common (Wiese, 2001) because the change requires fortition of the tongue tip gesture, which is considered a cross-linguistically rare historical change (Crowley, 1992). The process Silva (2003) describes where an allophonic change between /ʁ/ and [ɹ] takes place is captured through the representation presented here with a fortition of the tongue tip. The degree to which stress is placed on the phoneme results in varying degrees of allophony, consistent with a positive correlation between stress and fortition (Kirchner, 1998; Gordon, 2011). However, further experimental work on the dynamics of rhotics in Brazilian Portuguese is required to verify this claim. In summary, manipulation of the strength
of the underlying gestural representation of rhotics in Brazilian Portuguese accounts for both the synchronic and diachronic sound changes that occur.

7.3.1 Discussion of Cross-Linguistic Distribution of Rhotics

Cross-linguistically, rhotics tend to pattern adjacent to the syllable nucleus. The representation proposed here also provides direct articulatory evidence for this phenomenon. There is a necessity for the vocalic tongue root gesture for rhotics to be coordinated with the syllable nucleus, resulting in tremendous articulatory pressure to restrict the phonotactics of rhotics to the position adjacent to the syllable nuclei (Gick et al., 2006; Proctor, 2011). This interpretation of the gestural score and the representation of rhotics is also in line with Proctor’s (2011) suggestion that articulatory demands on the tongue root require the close coordination of the vocoid and rhotic articulation, resulting in an adjacent syllable position.

Another implication of the study relates to the cross-linguistic tendency for rhotic inventories to be relatively small. Maddieson (1984) describes inventories with more than one rhotic to be quite rare: The most common phonetic inventory has only one rhotic (57.7% of all surveyed languages). Languages with even two (14.1%) or three (2.5%) rhotics are quite rare and languages with 4 rhotics (0.3%) are extremely uncommon. The cross-linguistic avoidance on large rhotic inventories may be related to an overall perceptual similarity across the class of rhotics. The rhotics tested here cover three places of articulation (alveolar, retroflex, uvular) with different formant structures; however, despite place and manner differences, the overall similarities in formant structures make rhotics (specifically approximants and trills) difficult to distinguish. This potentially contributes to a cross-linguistic avoidance of large rhotic inventories.

The current results also suggest that there is a perceptual enhancement effect when rhotics and laterals are paired together. The perceptual areas for the rhotics tested here were all quite small and did not differ across experimental blocks based on the regression model, except in the condition which compared rhotics and laterals. In that condition, the perceptual space for the rhotics was significantly larger (almost 3 times larger than the perceptual area in any other condition). Maddieson (1984: 73-90) does not describe a language with a large liquid inventory that does not contain laterals; however, large liquid inventories comprised of just laterals do exist (3 laterals: 1.9%; 4 laterals: 0.6%; 5 laterals: 0.6%; 6 laterals: 0.6%). Languages which have only rhotics and no laterals only use 1 rhotic (13.1%) or 2 rhotics (0.9%), and none use 3+ rhotics.
This is quite a striking difference in the distribution of rhotics and laterals across the world’s languages. However, if rhotics are generally difficult to distinguish from each other, this tendency makes sense. Furthermore, if laterals have a perceptual enhancement effect, the overwhelming tendency for rhotics to be paired with laterals – even in systems with only 2 liquids (34.1%) – makes perfect sense. Taken together, this suggests that pairing rhotics with laterals creates a more favourable perceptual environment for rhotics to be correctly identified.

7.3.2 Discussion of the Dispreference for Rhotics and Palatalization

There have been a few conflicting theories on why palatalization disfavors trills. Ladefoged and Maddieson (1996) suggest that the tongue blade is unbraced, which causes a larger vibrating mass. This creates a less favorable aerodynamic environment for trilling. Iskarous and Kavitskaya (2010) propose that there are two possible reasons that trilling and palatalization are incompatible: palatalization may cause a retraction of the tongue tip gesture or decrease the ability to control the stiffness and inertia required to produce trilling. Hall (2000) also notes that the [apical] feature found on alveolar rhotics is not compatible with the [distributed] feature which is accompanied with palatalization. Kavitskaya et al. (2009) and Hall and Hamann (2010) both suggest that there is an inherent conflict in the tongue shapes of rhotics and secondary palatalization. However, it is still not fully clear why palatalized rhotics are typologically marked.

The fact that the palatalized uvular trill exhibits the same formant dynamics as the alveolar trill for F1 to F3 and F2-F1 suggests that the tongue tip gesture is not the main force for the inconsistency between palatalization and trilling due to a complete lack of an apical gesture. However, the apical gesture may still be a factor. Iskarous and Kavitskaya's (2010) suggestion is also contra Bolla’s (1981), Ladefoged and Maddieson's (1996), and Kochetov’s (2005) findings, which show some, albeit moderate, tongue tip advancement for the palatalized trill, relative to the unpalatalized trill. The temporal instability of F2 and F2-F1 is observed in both uvular and alveolar trills and suggests that the inconsistency between palatalization and trills is more likely related to aerodynamic factors, which Ladefoged and Maddieson (1996) also suggest. Iskarous and Kavitskaya (2010) make their claim partially based on the realization of the Russian trills as either full trills or single contact trills. They found that /r/ is realized more frequently as a full trill, while /ɾ/ is more frequently realized as single contact. This also seems to be the case for the trills in Lower Sorbian. The Upper Sorbian trills show a different, but similar pattern: / iterable is most
often realized as a single contact, while /ɾj/ is most often realized as an approximant. This suggests that uvular trills are less likely to be realized as a full trill than alveolar trills are. However, the palatalization gesture still affects the uvular trill in an equivalent way: it reduces the realizations as a trill and in the case of an already reduced number of full trills, it causes more approximant realizations.

The trilling of the tongue is always realized prior to the full realization of the off-glide, even though there is always a constant articulatory gesture towards this target. The gestural target for palatalization likely interferes with the aerodynamic conditions and tongue stiffness required to produce trilling. This suggestion follows from the constantly increasing F2 and F2-F1 over the duration of both the alveolar and uvular palatalized rhotic. The tight constraints on trill production could interfere with the palatalization gesture, which causes a delay in the movement to the palatal target. It is surmised that only after the trilling is complete that the palatalization target can be fully achieved because palatalization causes an increasingly tight constriction in the palatal region, resulting in a Bernoulli effect (Kochetov & Howson, 2015). This would cause a drop in air pressure, which Solé (2002) has shown to cause the loss of trilling. Thus, this suggests an inherent conflict between the physical conditions necessary to produce trilling and the gestural coordination of secondary palatalization. The physical requirement for a strong degree of tongue stiffness and stability has been suggested by previous authors (McGowan, 1992; Ladefoged and Maddieson, 1996; Iskarous & Kavitskaya, 2010), which permits the flexible portion of the tongue to be limited to the tongue tip (in the case of the alveolar trill). This in turn, reduces the vibrating mass, causing trilling to be much easier to achieve. The palatal target likely reduces the ability to achieve proper tongue stiffness and aerodynamics and is thus delayed, permitting trilling.

The ultrasound examination of the palatalized rhotic, /ɾj/, also revealed a consistent degree of tongue root retraction and a high degree of coarticulatory resistance, similar to /ɾ/. The data also revealed that /ɾj/ had a clear uvular target for the tongue body, which was not different from /ɾ/ in place. This suggests a similar tongue root and tongue body gesture involved in the articulation of both /ɾ/ and /ɾj/. However, the results suggest secondary palatalization contributes to a more complex articulation and representation. The complexity of articulation supports Kavitskaya et al.’s (2009) proposal that the reason palatalized rhotics are avoided is because of conflicting demands on the tongue root. The tongue root is advanced because of secondary
palatalization, interfering with the need for tongue root retraction. This conflict is observed directly when the tongue blade/body raises for palatalization, but the tongue root also retracts for vocalic gesture. The dynamic results also revealed a distinct separation between the gestures associated with rhotic articulation and secondary palatalization. Retraction of the tongue root and the tongue body is coordinated together; however, these gestures are sequenced before the tongue body/blade raising and fronting associated with secondary palatalization. This delay in the tongue body gesture has been directly observed through EMA results (Kochetov, 2006; Stoll et al., 2015) and indirectly through acoustic measures (Spajić, et al., 1996; Iskarous & Kavitskaya, 2010). Figure 80 shows the gestural representation for /rj/.

Figure 80. Gestural representation for /rj/. The articulators involved include the tongue body (TB) and the tongue root (TR) with the target constriction location (coronal, palatal, uvular-pharyngeal, and pharyngeal).

The phase relationship between secondary palatalization is 60°. This creates a sequential articulation of tongue root retraction and uvular-pharyngeal tongue body constriction synchronized with an out of phase tongue body gesture and tongue root gesture. This is realized as late achievement of tongue body raising and tongue root advancement. However, the relationship is not the same as two separate phonemes. Rather, there is a tighter degree of overlap and an articulatory trade-off between the tongue root and the tongue body gestures. The phase relationship between the articulatory requirements for the uvular rhotic are in a 0° phase relationship, as are the articulatory requirements for secondary palatalization. In other words, they are not sequenced in the same way as the rhotic and secondary palatalization are. The delicate timing between the gestures creates an unstable articulation (Iskarous & Kavitskaya, 2010), often leading to language specific diachronic changes. Historically, Proto-Slavic is reconstructed as having *rj, but in most of the daughter languages it has been lost (Carlton, 1990). In the case of Slovenian, it has become a sequence of /rj/, while in languages like Belarusian and Slovak the off-glide has been lost completely, with the original palatalized rhotic merging with /r/ (Carlton, 1990). The case of Slovenian is straightforwardly accounted for by the
proposal in presented here. There was a reanalysis of the 60° phase relationship as a sequence of two phonemes, rather than as a single phoneme. In the case of Belarusian and Slovak, the 60° phase relationship resulted in an eventual loss of secondary palatalization due to the articulatory conflict between the two segments. Thus, languages may differ in the repair strategies employed, but the underlying conflict is the same.

The representation presented here contrasts Hall and Hamann’s (2010) analysis which suggests that the conflict rests in the fact that rhotics are articulated with a concave shape while, high front vocoids are convex. The convex shape observed for a number of speakers’ productions of /r/ causes a reason to question that claim. However, Hall and Hamann’s (2010) assertion that the avoidance of sequences of a rhotic and a high front vocoids (and secondary palatalization) is articulatory in nature is certainly on the correct track, given the evidence here. The issue lies not in the entire tongue shape as much as it does in the activity of the tongue root. The tongue root faces an articulatory conflict between wanting to retract for rhotic articulation and advancing for secondary palatalization. The consequence is a cross-linguistic avoidance of secondary palatalization/high front vocoids and rhotics.

7.4 Summary of Primary Findings

1. Upper and Lower Sorbian Acoustics:
   I. The unpalatalized rhotics have a lower F2 and higher F1 than the contrasting lateral.
   II. The unpalatalized rhotics had no significant difference between the frequency and trajectory for F2 and a similar trajectory for F1, but different formant frequencies.
   III. F3 was a marked difference between the uvular and alveolar rhotics (unpalatalized and palatalized), with a higher F3 for uvular rhotics.
   IV. The palatalized rhotics in Upper and Lower Sorbian both had a delayed increase in F2 and F2-F1, indicating similar formant trajectories, but different frequencies.
   V. F1 was lower for the palatalized rhotics, compared to the unpalatalized rhotics.

2. Upper Sorbian Articulation:
   I. Both uvular rhotics, /r/ and /ɾ/, were characterized by a high degree of tongue root resistance to coarticulatory effects, contrasting the lack of tongue root resistance for the voice velar stop, /g/.
II. Both uvular rhotics had a consistent tongue body constriction in the uvular-pharyngeal region, contrasting /ɡ/, which had variable constriction location depending on phonetic environment.

III. The unpalatalized uvular had two distinct tongue shapes across speakers:
   i. A “double bunched” shape characterized by a bunching of the tongue body and root and a slightly raised tongue blade, forming a concave tongue shape.
   ii. A “retracted” shape characterized by a single bunching of the tongue root and body and no noticeable activity of the tongue blade/tip.

IV. A time lag between the uvular constriction and secondary palatalization was observed for /ʁ/. 
   i. Retraction and raising of the tongue body towards a uvular constriction occurred first, followed by a fronting of the tongue body and root characteristic of secondary palatalization.

3. Brazilian Portuguese Articulation:
   I. The uvular fricative rhotic, /ʁ/, had higher coarticulatory resistance (low RMS score) than /ɡ/ in the onset position, but it had less coarticulatory resistance intervocalically, where it was realized as [h].
   II. The alveolar tap, /ɾ/, had the highest coarticulatory resistance (lowest RMS) of all segments intervocalically.
   III. Coda /ɾ/, which only appeared in word-final position in this data set, had the highest coarticulatory resistance (lowest RMS) of all phonemes across all positions.
   IV. Coda /ɾ/ was realized as [ʁ], [ɾ], [χ], and [ɾ]. The variation was related to speaker and not region (all participants were from São Paulo, Brazil).

4. Rhotic Perception:
   I. The non-native rhotics examined, /ɾ/, /ɻ/, and /ʁ/, had low discriminability (indicated by low d-prime scores).
   II. The MDS revealed that rhotics grouped together in the perceptual space, forming a separate category.
i. kmeans analysis confirmed separation of the two groups in the perceptual space.

III. The rhotics occupied a smaller perceptual space than the other natural classes, except for in the lateral condition, where the rhotics and laterals occupied the same amount of perceptual space.

7.5 Summary of Theoretical Implications

1. Tongue root gesture is a uniting articulatory gesture for rhotics as a class.
   i. tongue tip rhotics are also composed of a tongue tip gesture.
   ii. uvular rhotics are also composed of a tongue body gesture.

2. Coda /R/ is best described as /ɾ/, underlyingly.

3. Lenition captures all of the allophony observed in Upper Sorbian and Brazilian Portuguese.
   i. Lenition of the tongue body gesture in Upper Sorbian results in coda vocalization.
   ii. Lenition of the tongue tip gesture in Brazilian Portuguese /ɾ/ captures allophonic changes with [ɾ], [ɾ], and [ɾ].
   iii. Lenition of the tongue body and root gestures for /ɾ/ accounts for the allophony with [ɾ] intervocalically.
   iv. Lenition accounts for historical changes from Galician-Portuguese *r to Brazilian Portuguese /ɾ/.
   v. Lenition explains the high frequency of historical changes from alveolar rhotics to uvular rhotics, but the paucity of historical changes from uvular rhotics to alveolar rhotics.

4. The tongue root component accounts for the distribution of rhotics within the syllable across languages: they are adjacent to vocalic segments in order coordinate tongue root gestures with the vowel.

5. Tongue root gesture has a distinct effect on F2 frequencies for rhotics.

6. Acoustic-perceptual characteristics for rhotics exist across all places of articulation for trill/approximant rhotics.
   i. The acoustic-perceptual characteristic represents a class feature.
   ii. Low discriminability accounts for cross-linguistic tendency for phonetic inventories to have few rhotics.
7.6 Limitations and Future Research

The largest limiting factor of this study was the fact that not all rhotics were tested and only a small number of languages were used. Understanding phonetic and phonological differences amongst rhotics from a cross-linguistic prospective is necessary to get a more complete picture of the acoustics, articulation, and perception of rhotics. This makes the generalizations made in this dissertation subject to further examination. It is also difficult to make generalizations about how the gestural dynamics affect the phonological behaviours that are not present in the languages tested in this dissertation. Therefore, it is important to study individual languages and examine the gestural coordination involved in rhotic articulations to determine most specifically if tongue root retraction is present across the broad range of rhotics observed in natural languages. It is also important to examine languages that have different phonological behaviours for rhotics than the languages tested here to determine if the manipulation of the articulatory specifications proposed here can account for what is observed in natural languages.

A great deal of this research interprets the findings in terms of Articulatory Phonology, which is a task dynamic model of the physical aspects of speech. However, much of this dissertation has focused on acoustic-perceptual correlates of rhotics and how speech perception effects the phonology of rhotics. This is an area that Articulatory Phonology is not currently equipped to handle. Therefore, it is necessary to integrate Articulatory Phonology into cognitive models of speech to get at a more complete and multi-faceted model of speech. This is clearly a promising avenue for future research.
References


116


https://cran.r-project.org/web/packages/sciplot/sciplot.pdf.


Appendix A

Upper Sorbian SSANOVAs

Individual speaker data from Chapter 3.

Figure 81. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S2. Tongue tip is on the right.

Figure 82. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S2. Tongue tip is on the right.

Figure 83. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S2. Tongue tip is on the right.
Figure 84. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S3. Tongue tip is on the right.

Figure 85. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S3. Tongue tip is on the right.

Figure 86. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S3. Tongue tip is on the right.
Figure 87. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S4. Tongue tip is on the right.

Figure 88. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S4. Tongue tip is on the right.

Figure 89. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S4. Tongue tip is on the right.
Figure 90. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S5. Tongue tip is on the right.

Figure 91. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S5. Tongue tip is on the right.

Figure 92. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S5. Tongue tip is on the right.
Figure 93. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), Ce (middle), Co (right) for S6. Tongue tip is on the right.

Figure 94. SSANOVA plots of the point of maximum constriction for each target phoneme in aCa (left), eCe (middle), oCo (right) for S6. Tongue tip is on the right.

Figure 95. SSANOVA plots of the point of maximum constriction for /ɡ/ in aC (left), eC (middle), oC (right) for S6. Tongue tip is on the right.
Appendix B

Brazilian Portuguese SSANOVAs

Individual speaker data from Chapter 5.

Figure 96. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP1. Tongue tip is on the right.

Figure 97. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP1. Tongue tip is on the right.

Figure 98. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP1. Tongue tip is on the right.
Figure 99. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP1. Tongue tip is on the right.

Figure 100. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP2. Tongue tip is on the right.

Figure 101. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP2. Tongue tip is on the right.
Figure 102. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP2. Tongue tip is on the right.

Figure 103. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP2. Tongue tip is on the right.

Figure 104. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP3. Tongue tip is on the right.
Figure 105. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP3. Tongue tip is on the right.

Figure 106. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP3. Tongue tip is on the right.

Figure 107. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP3. Tongue tip is on the right.
Figure 108. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP4. Tongue tip is on the right.

Figure 109. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP4. Tongue tip is on the right.

Figure 110. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP4. Tongue tip is on the right.
Figure 111. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP4. Tongue tip is on the right.

Figure 112. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP5. Tongue tip is on the right.

Figure 113. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP5. Tongue tip is on the right.
Figure 114. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP5. Tongue tip is on the right.

Figure 115. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP5. Tongue tip is on the right.

Figure 116. SSANOVA plots of the point of maximum constriction for each target phoneme in Ca (left), aCa (middle), aC (right) for BP6. Tongue tip is on the right.
Figure 117. SSANOVA plots of the point of maximum constriction for each target phoneme in Ce (left), eCe (middle), eC (right) for BP6. Tongue tip is on the right.

Figure 118. SSANOVA plots of the point of maximum constriction for each target phoneme in Co (left), oCo (middle), oC (right) for BP6. Tongue tip is on the right.

Figure 119. SSANOVA plots of the point of maximum constriction for each target phoneme in Ci (left), iCi (middle), iC (right) for BP6. Tongue tip is on the right.

133