PHONETICS AND PHONOLOGY OF REGRESSIVE VOICING ASSIMILATION IN RUSSIAN NATIVE AND NON-NATIVE SPEECH

by

Natalya Y. Samokhina

Copyright © Natalya Y. Samokhina 2010

A Dissertation Submitted to the Faculty of the

GRADUATE INTERDISCIPLINARY PROGRAM IN SECOND LANGUAGE ACQUISITION AND TEACHING

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2010
THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Natalya Y. Samokhina

entitled Phonetics and Phonology of Regressive Voicing Assimilation in Russian Native and Non-native Speech

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

___________________________________________________ ____________________
Date: 
Diane Ohala

___________________________________________________ ____________________
Date: 
Natasha Warner

___________________________________________________ ____________________
Date: 
Michael Hammond

___________________________________________________ ____________________
Date: 

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

___________________________________________________ ____________________
Date: 
Dissertation Director: Diane Ohala

___________________________________________________ ____________________
Date: 
Dissertation Director: Natasha Warner
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Natalya Y. Samokhina
ACKNOWLEDGMENTS

First and foremost I would like to thank my dissertation committee chairs, Natasha Warner and Diane Ohala for their support, encouragement and generosity of time. I would also like to thank Mike Hammond for his insightful and thought-provoking comments. My thanks also go to John Leafgren for being available to answer my numerous questions about the Russian language. I would also like to take this opportunity to thank Naomi Ogasawara and Ben Tucker for their help with statistics and Praat. I would also like to thank all my native and non-native speaking subjects for participating in my study.

In addition, I would like to thank my dearest friends at Trinity College. Thank you Julia Goesser for your unconditional support and help with multiple rounds of proofreading. Thank you Andrea Scapolo for listening to my continuous whining and being there for me in times of crisis.
DEDICATION

To my parents, Nadezhda and Yuriy Samokhin. This work would have not been possible without your love, support and confidence in me.

Моим родителям, Надежде и Юрию Самохиным, посвящается. Эта работа не состоялась бы без вас.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>14</td>
</tr>
<tr>
<td>LIST OF TABLES AND TABLEAUX</td>
<td>18</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>21</td>
</tr>
<tr>
<td>CHAPTER ONE: INTRODUCTION</td>
<td>22</td>
</tr>
<tr>
<td>1.0. Introduction</td>
<td>22</td>
</tr>
<tr>
<td>1.1. Regressive voicing assimilation as a general phenomenon</td>
<td>24</td>
</tr>
<tr>
<td>1.2. Voice contrast in Russian</td>
<td>25</td>
</tr>
<tr>
<td>1.2.1. Basic facts</td>
<td>25</td>
</tr>
<tr>
<td>1.2.2. RVA in Russian L2</td>
<td>26</td>
</tr>
<tr>
<td>1.3. Research objectives</td>
<td>27</td>
</tr>
<tr>
<td>1.4. Previous research</td>
<td>31</td>
</tr>
<tr>
<td>1.4.1. Introduction</td>
<td>31</td>
</tr>
<tr>
<td>1.4.2. Temporal cues to [voice]</td>
<td>32</td>
</tr>
<tr>
<td>1.4.3. Voice neutralization</td>
<td>35</td>
</tr>
<tr>
<td>1.4.3.1. Voice neutralization in word-final position</td>
<td>35</td>
</tr>
<tr>
<td>1.4.3.2. Voice neutralization in word-internal obstruent clusters</td>
<td>37</td>
</tr>
<tr>
<td>1.4.3.3. Summary</td>
<td>39</td>
</tr>
<tr>
<td>1.4.4. L2 phonetics</td>
<td>40</td>
</tr>
<tr>
<td>1.4.4.1. L1-L2 phonetic transfer</td>
<td>41</td>
</tr>
<tr>
<td>1.4.4.2. Summary</td>
<td>45</td>
</tr>
<tr>
<td>1.5. Conclusion</td>
<td>48</td>
</tr>
</tbody>
</table>
CHAPTER TWO: REAL-WORD PRODUCTION EXPERIMENT

2.0. Introduction

2.1. Methods

2.1.1. Subjects

2.1.2. Materials

2.1.3. Procedure

2.2. Analysis

2.2.1. Measurements

2.2.1.1. Segment duration

2.2.1.2. Voicing proportion

2.2.1.3. VOT

2.2.1.4. Data analysis

2.3. Results

2.3.1. Duration of the first consonant in the cluster (C1)

2.3.1.1. C1 duration in assimilating-to-voiceless vs. fully-voiced environments

2.3.1.2. C1 duration in assimilating-to-voiceless vs. fully-voiceless environments

2.3.1.3. Summary

2.3.2. C1 voicing proportion

2.3.2.1. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments

2.3.2.2. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environments

2.3.2.3. Summary

2.3.3. Duration of the second consonant in a cluster (C2)
2.3.3.1. C2 duration in assimilating-to-voiceless vs. fully-voiced environments

2.3.3.2. C2 duration in assimilating-to-voiceless vs. fully-voiceless environments

2.3.3.3. Summary

2.3.4. Voicing proportion of the second consonant in a cluster (C2)

2.3.4.1. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments

2.3.4.2. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environment

2.3.4.3. Summary

2.3.5. Voice Onset Time (VOT) of the second consonant in a cluster (C2)

2.3.5.1. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiced environments

2.3.5.2. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiceless environments

2.3.5.3. Summary

2.3.6. Duration of the vowel preceding a cluster (V1)

2.3.6.1. V1 duration in assimilating-to-voiceless vs. fully-voiced environments

2.3.6.2. V1 duration in assimilating-to-voiceless vs. fully-voiceless environments

2.3.6.3. Summary

2.3.7. Duration of the vowel following a cluster (V2)

2.3.7.1. V2 duration in assimilating-to-voiceless vs. fully-voiced environments

2.3.7.2. V2 duration in assimilating-to-voiceless vs. fully-voiceless environments

2.3.7.3. Summary
2.4. Discussion

2.4.1. Conclusion

CHAPTER THREE: NONSENSE-WORD PRODUCTION EXPERIMENT

3.0. Introduction

3.1. Methods

3.1.1. Subjects

3.1.2. Materials

3.1.3. Procedure

3.2. Analysis

3.2.1. Measurements

3.3. Results

3.3.1. Duration of the first consonant in the cluster (C1)

3.3.1.1. C1 duration in assimilating-to-voiceless vs. fully-voiced environments

3.3.1.2. C1 duration in assimilating-to-voiceless vs. fully-voiceless environments

3.3.1.3. Summary

3.3.2. C1 voicing proportion

3.3.2.1. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments

3.3.2.2. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environments

3.3.2.3. Summary

3.3.3. Duration of the second consonant in a cluster (C2)

3.3.3.1. C2 duration in assimilating-to-voiceless vs. fully-voiced environments
3.3.3.2. C2 duration in assimilating-to-voiceless vs. fully-voiceless environments .................................................................125

3.3.3.3. Summary ........................................................................................................................................127

3.3.4. Voicing proportion of the second consonant in a cluster (C2) .................................................................127

3.3.4.1. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments .................................................................128

3.3.4.2. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environments .................................................................130

3.3.4.3. Summary ........................................................................................................................................131

3.3.5. Voice Onset Time (VOT) of the second consonant in a cluster (C2) .................................................................132

3.3.5.1. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiced environments .................................................................132

3.3.5.2. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiceless environments .................................................................134

3.3.5.3. Summary ........................................................................................................................................135

3.3.6. Duration of the vowel preceding the cluster (V1) .........................................................................................135

3.3.6.1. V1 duration in assimilating-to-voiceless vs. fully-voiced environments .................................................................135

3.3.6.2. V1 duration in assimilating-to-voiceless vs. fully-voiceless environments .................................................................137

3.3.6.3. Summary ........................................................................................................................................139

3.3.7. Duration of a vowel following a cluster (V2) .........................................................................................140

3.3.7.1. V2 duration in assimilating-to-voiceless vs. fully-voiced environments .................................................................141

3.3.7.2. V2 duration in assimilating-to-voiceless vs. fully-voiceless environments .................................................................142

3.3.7.3. Summary ........................................................................................................................................144

3.4. Discussion ........................................................................................................................................145
3.4.1. Conclusion

CHAPTER FOUR: PHONOLOGY OF REGRESSIVE VOICING
ASSIMILATION

4.0. Introduction

4.1. Optimality Theory

4.1.1. Learnability Theories in OT: Error Driven Constraint Demotion (EDCD)
(Tesar & Smolensky 1993,1996)

4.1.2. Gradual Learning Algorithm (Boersma & Hayes 2001)

4.1.3. Summary

4.2. Analysis of voicing typology in languages other than Russian from an L1
perspective

4.2.1. Word-medial obstruent clusters and word-final obstruents

4.2.2. Complex codas

4.2.3. Summary

4.3. Analysis of RVA in L2 phonology

4.3.1. Polish speakers learning L2 English

4.3.2. English speakers learning L2 Polish

4.3.3. Summary

4.4. Analysis of RVA in Russian L1

4.4.1. Basic facts about Russian voicing as outlined in Padgett (2002)

4.4.2. OT analysis of Russian RVA

4.4.3. Summary

4.5. RVA in Russian L1 and L2

4.5.1. Constraint interaction and acquisition of RVA

4.5.1.1. Categorical RVA and categorical L2 acquisition
4.5.1.2. Variable ranking of AGREE(VOICE)_{CG} and IDENT(VOICE) in L2

4.5.2. Gradient RVA and gradual L2 acquisition of RVA

4.5.2.1. Frequency distribution of C1 voicing proportion in the assimilating environment

4.5.2.2. Partially-devoiced obstruents in RussianL2

4.5.3. Phonetically detailed constraint families

4.5.3.1. A family of constraints in the perception grammar
(Boersma 1997, 1998)

4.5.3.2. A family of constraints in the production grammar
(Jun 1995, 2004)

4.5.4. Gradient RVA and gradual L2 acquisition

4.5.4.1. A continuous constraint family approach

4.5.4.2. Incorporating the continuous constraint family
\{AGREE_{C1}^{n}C2(VOICE)_{CG}\} in the analysis

4.5.5. Summary

4.6. Conclusion

CHAPTER FIVE: CONCLUSIONS AND DISCUSSION

5.0. Introduction

5.1. Summary of experimental results

5.1.1. Findings of acoustic experiments

5.1.2. Partial voicing neutralization

5.2. Implications for phonetic theory

5.2.1. Phonetic L1 – L2 transfer: evidence from the analysis of voiced stops and fricatives

5.2.2. Approximation of L2 phonetic properties: evidence from the analysis of voiceless and voiced stops
5.3. Implications for phonological theory: learning of phonological patterns rather than categories .......................................................... 255

5.3.1. Implications for phonological acquisition .......................... 255

5.3.2. Gradual acquisition of L2 phonological patterns ................. 257

5.4. Implications for L2 teaching .................................................. 258

5.4.1. Russian L2 learners’ production ......................................... 258

5.4.2. Does formal L2 instruction matter? ................................... 259

5.4.3. Suggestions for Russian L2 instruction ............................ 261

5.5. Conclusion and further research .......................................... 264

5.5.1. Overall conclusion .......................................................... 264

5.5.2. Further research questions ............................................ 265

5.5.3. Conclusion .................................................................. 269

APPENDIX A: TEST REAL WORDS FOR EXPERIMENT ONE .......... 271

APPENDIX B: TEST NONSENSE WORDS FOR EXPERIMENT TWO .... 275

REFERENCES ......................................................................... 281
LIST OF FIGURES

Figure 2.1. Sample waveform and spectrogram of the V1-C1-C1-V2 sequence of the word lobzik –/lobzik/ ‘fret-saw’ produced by a native speaker.........................58

Figure 2.2. Sample waveform and of the C1[+vce]C2[-vce]V1 sequence in the word lozhka – /loʒka/ ‘spoon’ produced by a native subject……………………59

Figure 2.3. Comparisons of the duration of the first consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments.........................63

Figure 2.4. Comparisons of the duration of the first consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments.................................................................66

Figure 2.5. Comparisons of voicing proportion of the first consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments.................................................................70

Figure 2.6. Comparisons of voicing proportion of the first consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments.................................................................73

Figure 2.7. Comparisons of the duration of the second consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments.................................................................77

Figure 2.8. Comparisons of the duration of the second consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments.................................................................80

Figure 2.9. Comparisons of voicing proportion of the second consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments.................................................................83

Figure 2.10. Comparisons of voicing proportion of the second consonant in the stop-fricative cluster in the assimilating vs. fully-voiceless environments......................86

Figure 2.11. VOT of the second consonant in the fricative-stop clusters in the assimilating vs. fully-voiced environments.................................................................88

Figure 2.12. VOT of the second consonant in the fricative-stop clusters in the assimilating vs. fully-voiceless environments.................................................................90

Figure 2.13. Comparisons of the duration of the vowel preceding a stop-fricative and fricative-stop cluster in the assimilating vs. fully-voiced environments.........................92
Figure 2.14. Comparisons of the duration of the vowel preceding a stop-fricative and fricative-stop cluster in the assimilating vs. fully-voiceless environments………..…95

Figure 2.15. Comparisons of duration of the vowel following a stop-fricative and fricative-stop cluster in the assimilating vs. fully-voiced environments…………………..98

Figure 2.16. Comparisons of duration of the vowel following a stop-fricative and fricative-stop cluster in the assimilating vs. fully-voiceless environments……………….99

Figure 3.1. Comparisons of the duration of the first consonant in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments…………..112

Figure 3.2. Comparisons of the duration of the first consonant that is expected to assimilate in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments……………………………………………………………..114

Figure 3.3. Comparisons of voicing proportion of the first consonant that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments………………………………………………………………….…….117

Figure 3.4. Comparisons of voicing proportion of the first consonant that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments………………………………………………………………………..120

Figure 3.5. Comparisons of the duration of the second consonant in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments…………….…124

Figure 3.6. Comparisons of the duration of the second consonant in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments………………..126

Figure 3.7. Comparisons of voicing proportion of the second consonant that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments……………………………………………………..……....129

Figure 3.8. Comparisons of voicing proportion of the second consonant that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments……………………………………………………………..131

Figure 3.9. Comparison of VOT in the second consonant in the fricative-stop cluster in the assimilating vs. fully-voiced environments…………………………………..133

Figure 3.10. Comparison of VOT in the second consonant in the fricative-stop cluster in the assimilating vs. fully-voiceless environments……………………………..…134

Figure 3.11. Comparisons of the duration of the vowel preceding a stop-fricative and a fricative-stop cluster in the assimilating vs. fully-voiced environments………..136
Figure 3.12. Comparisons of the duration of the vowel preceding a stop-fricative and a fricative-stop cluster in the assimilating vs. fully-voiceless environments  

Figure 3.13. Comparisons of duration of the vowel following a stop-fricative and a fricative-stop cluster in the assimilating vs. fully-voiced environments  

Figure 3.14. Comparisons of duration of the vowel following a stop-fricative and a fricative-stop cluster in the assimilating vs. fully-voiceless environments  

Figure 4.1. Categorical ranking of $\text{AGREE(VOICE)}_{CG} \gg \text{IDENT(VOICE)}$: Russian L1  

Figure 4.2. Categorical ranking of $\text{IDENT(VOICE)} \gg \text{AGREE(VOICE)}_{CG}$: English L1  

Figure 4.3. Variable ranking of $\text{IDENT(VOICE)}$ and $\text{AGREE(VOICE)}_{CG}$ in English L1- Russian L2. Common result: $\text{IDENT(VOICE)} \gg \text{AGREE(VOICE)}_{CG}$  

Figure 4.4. Variable ranking of $\text{IDENT(VOICE)}$ and $\text{AGREE(VOICE)}_{CG}$ in English L1-Russian L2. Less common result: $\text{AGREE(VOICE)}_{CG} \gg \text{IDENT(VOICE)}$  

Figure 4.5. The amount of voicing within token in C1 in three voicing environments in a real-word experiment  

Figure 4.6. The amount of voicing within token in C1 in three voicing environments in a non-word experiment  

Figure 4.7. The frequency distribution of the voicing proportion of C1 in three voicing environments in all language groups in the real-word experiment  

Figure 4.8. The frequency distribution of the voicing proportion of C1 in three voicing environments in all language groups in the nonsense-word experiment  

Figure 4.9. The frequency distribution of the voicing proportion of assimilated C1 in the L1 production in the real-word experiment  

Figure 4.10. The frequency distribution of the voicing proportion of assimilated C1 in advanced L2 production in the real-word experiment  

Figure 4.11. The frequency distribution of the voicing proportion of assimilated C1 in intermediate L2 production in the real-word experiment  

Figure 4.12. The frequency distribution of the voicing proportion of assimilated C1 in L1 in the nonsense-word experiment  

Figure 4.13. The frequency distribution of the voicing proportion of assimilated C1 in the advanced L2 production in the nonsense-word experiment
Figure 4.14. The frequency distribution of the voicing proportion of assimilated C1 in the intermediate L2 production in the nonsense-word experiment………………….213

Figure 4.15. The *WARP family on the continuous auditory dimension: a possible initial state…………………………………………………………………………..220

Figure 4.16. Production distributions of the three categories /30/, /50/, and /70/…..222

Figure 4.17. Variable ranking of IDENT(VOICE) and AGREE(VOICE)$_{CG}$ in English L1- Russian L2: IDENT(VOICE) >> AGREE(VOICE)$_{CG}$………………………….230

Figure 4.18. Variable ranking of IDENT(VOICE) and AGREE(VOICE)$_{CG}$ in English L1-Russian L2: AGREE(VOICE)$_{CG}$ >> IDENT(VOICE)……………………………………..231
LIST OF TABLES AND TABLEAUX

Table 1.1. Russian obstruent inventory..............................................................26
Table 2.1. Example stimuli containing Russian obstruent combinations in the three voicing environments.................................................................53
Table 2.2. Summary of the C1 measurements......................................................75
Table 3.1. Sample nonsense items in the three voicing environments.................108
Table 4.1. Voicing typology.................................................................................167
Table 4.2. Voicing typology and corresponding constraint rankings....................177
Table 4.3. Summary of predicted paths & stages in acquisition of L2 English voicing contrast.................................................................182
Table 4.4. Summary of predicted paths and stages in English L1 – Polish L2........188
Table 4.5. The highest percentage of residual voicing in C1 in the three voicing environments in the three language groups..............................................205
Table 4.6. The interaction of the WEAKENING and PRESERVE constraints which results in various reduction patterns..................................................226
Table 4.7. Summary of production patterns and corresponding constraint rankings...........................................................................................................239

Tableau 4.1. Mark cancellation in the loser > winner data pair..........................157
Tableau 4.2. Grammar that requires several rounds of demotion.......................157
Tableau 4.3. LINEARITY >> (*|?σ) >> IDENT-IO([SYLLABIC])..........................164
Tableau 4.4. (*|?σ) >> LINEARITY >> IDENT-IO([SYLLABIC]).........................164
Tableau 4.5. Syllable-final devoicing in German................................................169
Tableau 4.6. Syllable-final devoicing in a word-internal cluster in German: *Lar >> IDLar........................................................................................................170
Tableau 4.7. Syllable-final devoicing in a word-internal cluster in German: IDOnsLar >> *Lar >> IDLar.................................................................170
Tableau 4.8. Syllable-final devoicing in a word-internal cluster: German.............171
Tableau 4.9. Unrestricted voice: English.........................................................172
Tableau 4.10. Voicing assimilation and word-final faithfulness: Yiddish..........172
Tableau 4.11. Voicing assimilation and word-final neutralization: Polish.............173
Tableau 4.12. Word-final complex coda neutralization: German.......................174
Tableau 4.13. Word-final complex coda neutralization: Polish..........................175
Tableau 4.15. RVA and word-final devoicing: Polish L1..................................179
Tableau 4.16. Stage 1: RVA and FD in Polish L1 - English L2........................180
Tableau 4.17. Stage 2 - Path 1: No RVA and FD – target-like English...............182
Tableau 4.18. Stage 2 - Path 2: Regressive devoicing and FD.........................183
Tableau 4.19. Stage 3 - Path 2b: Regressive devoicing..................................184
Tableau 4.20. Stage 2 - Path 3: RVA but no FD.............................................185
Tableau 4.21. Target-like English....................................................................186
Tableau 4.22. English L2 – Polish L1, Stage 1: No RVA and no FD...................186
Tableau 4.23. Voicing assimilation in onset clusters: Russian..........................193
Tableau 4.24. Voicing assimilation in word-final clusters: Russian....................193
Tableau 4.25. RVA in word-internal obstruent clusters: Russian L1...................195
Tableau 4.26. RVA in word-medial clusters in Russian L1: categorical ranking of 
AGREE(VOICE)_{CG} >> IDENT(VOICE)..............................................................196
Tableau 4.27. Unrestricted voice: English.......................................................197
Tableau 4.28. Categorical acquisition: No RVA in Russian L2, at the starting point of 
English L1 speakers’ acquisition .................................................................198
Tableau 4.29. Categorical acquisition: RVA in Russian L2...............................199
Tableau 4.30. The interaction of the *WARP constraints in the perception grammar, with category /50/ as the optimal candidate……………………………………………221

Tableau 4.31. The interaction of the *WARP constraints in the perception grammar, with category [30] as the optimal candidate…………………………………………223

Tableau 4.32. RVA in Russian: {AGREE\textsubscript{1}\textsubscript{n}C2(VOICE)\textsubscript{CG}} >> IDENT(VOICE)…………………………………………………………….229

Tableau 4.33. The lack of RVA: {AGREE\textsubscript{1}\textsubscript{n}C2(VOICE)\textsubscript{CG}} and IDENT(VOICE) ranking ranges overlap………………………………………………………….231

Tableau 4.34. The interaction of the constraints within {AGREE\textsubscript{1}\textsubscript{n}C2(VOICE)\textsubscript{CG}} family which results in various degrees of assimilation……………………………………………………………………………….233

Tableau 4.35. RVA in Russian L2: {AGREE\textsubscript{40\%}(VOICE)\textsubscript{CG} >> AGREE\textsubscript{35\%}(VOICE)\textsubscript{CG} >> IDENT(VOICE)}……………………………………………………………………………………………………………………………………………………………………234

Tableau 4.36. Target Russian L2 ranking: {AGREE\textsubscript{35\%}(VOICE)\textsubscript{CG} >> AGREE\textsubscript{40\%}(VOICE)\textsubscript{CG} >> IDENT(VOICE)}……………………………………………………………………………………………………………………………………………………………………235

Tableau 4.37. RVA in Russian L2: AGREE\textsubscript{35\%}(VOICE)\textsubscript{CG} >> IDENT(VOICE) >> AGREE\textsubscript{70\%}(VOICE)\textsubscript{CG}………………………………………………………………………………………………………………………………………………………………………………………………………………236

Tableau 4.38. RVA in Russian L2: AGREE\textsubscript{70\%}(VOICE)\textsubscript{CG} >> IDENT(VOICE) >> AGREE\textsubscript{35\%}(VOICE)\textsubscript{CG}……………………………………………………………………………………………………………………………………………………………………………………………………………………………………236
ABSTRACT

In recent years, a great deal of research on second language (L2) phonetics and phonological acquisition has focused on non-target production patterns of L2 learners, addressing issues such as native language (L1) transfer into L2 and the nature and source of developmental errors. Previous studies of phonological processes in L2 acquisition have traditionally focused on the analysis of discrete L2 segments (Flege 1987, 1999; Major & Kim 1996), rather than on L2 production patterns. This study, however, examines the production of sequences of sounds in Russian L1 and L2 from both the phonetic and phonological perspectives.

Two acoustic experiments in this dissertation analyze native and non-native subjects’ production of real and nonsense words containing obstruent clusters in which regressive voicing assimilation is required. For example, the form *lodka* ‘boat’ is rendered orthographically with a voiced stop /d/ which is, however, pronounced as a voiceless /t/ when followed by a voiceless obstruent. Both L1 and L2 speakers were found to exhibit a great deal of variability in their production of the devoiced obstruents in question, thus demonstrating gradience in devoicing. Non-native production also suggests gradual acquisition of L2 phonological patterns.

The findings of these experiments are analyzed within the stochastic Optimality Theory framework. Gradual L2 acquisition is accounted for by the re-ranking of L1 constraints; whereas, gradience in production is viewed as a result of the re-ranking of constraints within phonetically detailed constraint families.
CHAPTER ONE
INTRODUCTION

1.0. Introduction
The acquisition of a second language (L2) presents adult learners with many challenges. In addition to learning new grammar structures and lexical items, L2 learners must acquire both a new sound system and new phonological patterns. Despite all of their efforts, most adult L2 learners rarely succeed in mastering the L2 sound system due to the conflict between their native (L1) and L2 phonetic and phonological systems. The issue of L2 phonetic and phonological learning is addressed in the present work by investigating the acquisition of Russian pronunciation patterns by native speakers of English. More specifically, I investigate the acquisition of voicing assimilation within consonant clusters in which, for example, a voiced fricative /z/ is pronounced as [s] when followed by a voiceless consonant. It has been frequently observed that in the earliest stages of acquisition, L2 learners identify an L2 sound with a phonetically similar L1 sound, and use articulatory patterns and phonetic categories established in L1 to produce L2 sounds (Flege 1984, 1986; Lado 1957; Weinreich 1953). However, as L2 learners gain experience, they modify previously established patterns of segmental production. As a result, they begin approximating the phonetic norms of the L2, and produce similar L2 phones differently from their L1 counterparts, albeit not authentically (Flege 1984, 1987; Flege & Port 1981; MacKay et al., 2001).

For the past several decades, phonetic category restructuring and new category formation have been the focus of extensive research in the field of L2 phonetic and
phonological acquisition. More specifically, L2 production errors are believed to be
due to the language learners’ inability to establish new categories for L2 sounds. For
example, Japanese learners of English have difficulties producing English /r/ and /l/
correctly because they lack separate categories for these sounds. Factors affecting the
accuracy of L2 speech production have been studied along with such variables as the
age of the learner, their length of residence (in the country of the target language),
amount of L1 and L2 used, etc.

Flege and colleagues (1984, 1987, 1991, 1993) have carried out multiple
studies of the production of L2 sounds by speakers of various language backgrounds.
One of such studies is the investigation of the acquisition of French vowels /y/ and /u/
by native English speakers, just to give one example. English has a number of high
vowels /i, ɪ, u, u/ but does not have any high front rounded vowels like French /y/.
The researcheres investigated whether L2 learners of French form a new category for
this sound or merge English high vowels into one category to accommodate a new L2
sound and how such factors as age of arrival in the L2 country, length of residence in
that country, and the amount of L1 spoken affected L2 acquisition. The majority of
such studies investigated the acquisition of discrete segments exclusively (Caramazza
2001; Major 1987; Munro et al., 2000; Port & Mitleb 1983). However, as Flege
(1992:595) observes:

Just as L2 segments may or may not be identified with segments in the L1, learners may (or
may not) identify strings of segments in the L2 as meeting the structural descriptions for the
application of an L1 phonological constraint.
For example, Hammarberg (1990) observed that beginning German learners of Swedish pronounced the L2 word ‘sang’ (‘bed’) with an /s/ as appropriate for Swedish, when prompted, but with a /z/, typical for German, when produced spontaneously. The issue of whether adult L2 learners do indeed apply L1 phonological patterns to the production of L2 sound sequences is addressed in the present work from the phonetic and phonological perspectives, through the investigation of the acquisition of a phonological process known as regressive voicing assimilation by L2 learners of Russian.

1.1. Regressive voicing assimilation as a general phenomenon

The phenomenon of voice neutralization, which eliminates the voiced-voiceless contrast, is shared by a number of languages. In some, such as German and Catalan, it is exemplified by the rule of word/syllable-final devoicing only. For example, German form Alb (‘elf’) is pronounced with a voiceless stop /p/, despite the orthography. In others, voice neutralization also takes place in word-internal obstruent clusters as the examples in (1) – (4) below demonstrate. This process, known as regressive voicing assimilation (RVA), has been attested to in a wide range of world languages belonging to different language families, both Indo-European (Afrikaans, Dutch, Polish, Russian, etc.) and non-Indo-European (Hebrew, Hungarian). RVA requires two adjacent obstruents in a cluster to agree in voicing, with the feature [voice] spreading leftwards, as in the Russian form lodka – lo[t]ka (‘boat’).

Depending on the voicing specification of the second member in a cluster (C2), the first member of a cluster (C1) either devoices or becomes voiced1.

---

1 All examples are given in the Roman orthography.
Modern Hebrew (reproduced from Barkai & Horvath 1978):

1. (1) ahavti - aha[ft]i ‘I loved’ vs. ahava - aha[v]a ‘love’

2. lisgor – li[zg]or ‘to close’ vs. sagar – [s]agar ‘he closed’

Russian:


4. prosjba – pro[z^n]ba ‘request’ vs. prosit – pro[s]it ‘to ask’

In (1) and (3) voiced obstruents /v/ and /d/ devoice when followed by voiceless stops /t/ and /k/ respectively; whereas in (2) and (4) voiceless fricative /s/ becomes voiced when followed by voiced stops /g/ or /b/.

1.2. Voice contrast in Russian

1.2.1. Basic facts

In Russian, obstruent voicing is contrastive within a word but non-contrastive word-finally. That is, Russian obstruents devoice word-finally without exception (Avanesov 1956):

5. knig - kni[k] ‘books’ (Gen. Pl.) vs. kni[g]a (Nom. Sing.)

6. raz – ra[s] ‘occasion, time’ (Nom. Sing.) vs. ra[z]a (Gen. Pl.)

As a result of RVA, voicing contrast is also eliminated in obstruent clusters which occur both across morpheme boundaries, as in (7), and within a phonological word, as in (8):

7. podpisat – po[t]pisatj ‘to sign’

8. ot babushki – o[db]abushki ‘from grandma’

(Padgett 2002)

---

1 I use the superscript symbol / to indicate palatalized (‘soft’) consonants in Russian.
Combinations of fricative and fricative, fricative and stop, stop and fricative, and stop and stop (see Table 1.1 for the Russian obstruent inventory) all participate in RVA in Russian, as shown in examples (9) - (12) in the morpheme boundary environment.

(9) uehavshii - ueha[fʃ]ii ‘the one who has left’
(10) lozhka – lo[ʃk]a ‘spoon’
(11) podsesť – po[tʃ]esť ‘to sit down next to’
(12) lodka - lo[tʃk]a ‘boat’

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labio-dental</th>
<th>Dental</th>
<th>Post-alveolar</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p b</td>
<td></td>
<td>t d</td>
<td></td>
<td>k g</td>
</tr>
<tr>
<td>Fricative</td>
<td>f v</td>
<td></td>
<td>s z</td>
<td>§ ž</td>
<td>X</td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td>Tš</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1. Russian obstruent inventory

1.2.2. RVA in Russian L2

It is a well-known fact that almost all adult learners of second/foreign languages maintain foreign accents throughout their lives (Flege 1990). Informal observation of English-speaking students learning Russian as an L2 has suggested that one of the difficulties they encounter is an accurate production of RVA in word-internal obstruent clusters. It can be argued that since English does not exhibit a similar pattern of voicing assimilation within a phonological word (See point (2) below), less proficient L2 learners’ production will vary with respect to RVA. There are at least two conceivable production patterns:

(1) As English has unrestricted [voice] (Lombardi 1999), and obstruent clusters in English can disagree in voice specifications (e.g., o[bt]ain), L2 learners might produce each segment in an obstruent cluster with differing [voice]
specifications (*ve[zt]i for vezti ‘to drive’ rather than ve[st]i). This assumption is compatible with the fact that RVA is not reflected in Russian orthography. Therefore, it is not unreasonable to assume that less proficient L2 learners might be led by L2 orthography rather than L2 phonology to erroneously produce voice specification of the segment as implied by the grapheme, rather than following the phonological patterns of the L2, since their L1 phonology allows for this.

(2) Yet another possibility suggested by research in L2 phonetics is that Russian L2 speakers might indeed apply RVA in obstruent clusters; but, not to the same extent as native speakers. It could be hypothesized that L2 speakers do not completely devoice a fully-voiced obstruent in neutralizing (assimilating) environments, which results in a partially devoiced obstruent that could be perceived by Russian native speakers as voiced rather than as intentionally voiceless.

Finally, the production patterns described above could include variability and errors in L2 speech, as well as both phonological and phonetic L1 transfer. All of these issues will be taken up in this chapter; whereas a review of the formal L1 and L2 phonological acquisition literature is provided in Chapter Four which is where a phonological analysis RVA will be presented.

1.3. Research objectives

In recent years a great deal of research in L2 phonetics and phonological acquisition has focused on non-target production patterns of L2 learners, addressing such issues as L1 transfer, the nature and source of developmental errors, and factors affecting their retention. It has been shown fairly conclusively that adult learners perceive and produce L2 sounds in terms of their native language (Flege 1984, 1987, 1995, 1999,
2001; Flege & Port 1981; MacKay et al., 2001). For example, English L1-French L2 speakers mispronounce French /y/ as /u/; whereas, Portuguese L1-French L2 speakers mispronounce the same target sounds as /i/ (Flege 1995). Consequently, it is claimed that difficulties in L2 pronunciation are due to the inability of adult learners to modify sound categories previously established for their native language and/or establish new ones for the L2.

Such L2 studies have traditionally focused on the production of individual segments and phonetic category restructuring (Aoyama et al., 2004; Flege 1987; Flege et al., 1999; Major & Kim 1996). For example, much acoustic research has examined the production of voiceless stops in terms of their voice onset time (VOT). (Flege & Eefting 1987; Flege & Hillenbrand 1984; Port & Mitleb 1983). It was found that L2 speakers fail to produce English /p t k/ authentically, either producing VOT with L1 values, or with values intermediate between their L1 and L2 norms. However, L2 research has neglected the issues pertaining to the acquisition of L2 phonetic norms and phonological patterns as opposed to the acquisition of segmental categories. Therefore, one aim of this study is to fill in gaps in existing phonetic and phonological research on L2 speech, such as the elimination of voice contrast in obstruent clusters in Russian (as in lodka – lo[tk]a ‘boat’), by analyzing sequences of sounds, rather than discrete segments. Another aim is to analyze the phonological processes involved in the production of such sequences of sounds. This study also investigates the time course of learning and demonstrates that at least some L2 learners succeed in modifying previously established pronunciation patterns and achieve an approximated L2 pronunciation norm.
Although numerous phonological studies have examined various phonological processes in world languages, including voicing assimilation, with a few exceptions the discussions were based on auditory impressionistic evidence rather than acoustic phonetic evidence. Very few studies investigated the phonetic properties of a language that could be a result of the application of phonological rules (e.g., voice neutralization in word-final position or in obstruent clusters word-internally). Therefore, another aim of this study is to analyze a phonological process, namely regressive voicing assimilation (RVA), from both phonetic and phonological perspectives, in order to expand upon previous studies. This work is part of a general attempt to systematically study the phenomenon of voicing assimilation by examining its phonetic effects (Burton & Robblee 1997; Ringen & Helgason 2004; Slowiaszek & Dinnsen 1985; Wissing & Roux 1995).

This research also extends previous analyses of RVA to include additional contexts and acoustic measures, and it contributes to our general knowledge of voicing assimilation systems in the languages of the world. Moreover, this study addresses these issues through the investigation of RVA not only in L1 but also L2 speech. While some L2 research has used Optimality Theory (OT) to account for the phenomenon of RVA in various L1’s and L2’s (Boersma et al., 2003; Hancin-Bhatt 2000; Swanson 2001), such work has not been performed for Russian L2. This study, therefore, in addition to the phonetic studies, investigates RVA in Russian L2 within the OT framework. Specifically, it will be shown that the acoustic details of L2 acquisition obtained in production cannot be captured by any non-stochastic version of OT. Rather, a model that allows for both gradience and gradual acquisition, such as the stochastic models proposed by Boersma & Hayes (2001), is adapted for L2
acquisition. None of these issues (i.e., gradience and gradual acquisition) is addressed or argued for in the previous analyses of RVA; therefore, this research aims to close the existing gap in L2 phonological research.

To summarize, the objectives of the proposed study are: (1) to investigate the production of RVA in obstruent clusters in Russian native speakers from a phonetic perspective in order to determine the acoustic characteristics of RVA in the Russian language; (2) to investigate the production of RVA in obstruent clusters produced by L2 speakers from a phonetic perspective in order to determine the acoustic details of RVA in Russian L2; (3) to compare non-native production to native production for the purpose of determining which, if any, acoustic characteristics of RVA L2 learners fail to produce correctly; (4) to provide a phonological analysis of RVA in Russian L1 and L2 speech based on the results of the phonetic analysis, and (5) to offer an interpretation of the emerging L2 production patterns in both L1 and L2 by modeling them within a stochastic model of acquisition (i.e., the Gradual Learning Algorithm by Boersma & Hayes 2001).

Thus, the present dissertation has a twofold goal: one is to analyze RVA from a phonetic perspective using acoustic measurements; the second is to account for non-native production of RVA from the phonological perspective. The dissertation is organized as follows: Chapter One is a literature review of previous phonetic research on L2 acquisition and on the phonetics of voicing and voicing assimilation as general phenomena. Chapter Two describes a real-word production experiment and provides an acoustic analysis of RVA in Russian L1 and L2. Chapter Three describes a nonsense-word production experiment. In Chapter Four, I provide a literature review of relevant L2 phonological research and analyze the phonology of RVA in L1 and L2...
within the revised Optimality Theory framework of Boersma & Hayes (2001).

Chapter Five concludes with a general discussion and implications for future research.

1.4. Previous research

1.4.1. Introduction

As mentioned previously, there are a significant number of phonetic studies that have analyzed word-final or syllable-final voicing neutralization in various languages such as Catalan, German, and Polish (Dinnsen & Charles-Luce 1984; Fourakis & Iverson 1984; Port et al., 1981; Slowiaczek & Dinnsen 1985). However, to date very few studies have investigated the acoustic characteristics of voicing neutralization through assimilation in an obstruent cluster, rather than because of syllable-final devoicing. Jassem & Richter (1989) have done so for Polish, and Burton and colleagues (1994; 1997) have done so for Russian. Furthermore, this phenomenon has received little attention from the perspective of L2 phonetics. To my knowledge, there are no acoustic studies that have addressed the specifics of RVA in L2. Moreover, very few studies have provided a thorough analysis of RVA from both the phonetic and phonological perspectives (Swedish: Ringen & Helgason 2004; Afrikaans: Wissing & Roux 1995).

In the following sections, I will first review acoustic correlates of voice. I will then turn to relevant phonetic literature on voice neutralization in word-final obstruents and word-internal obstruent clusters, concluding with a review of L2 phonetic studies discussing L1 phonetic transfer and L1 category restructuring.

---

3 Hayes (1984) has also analyzed voicing assimilation in Russian; however, his analysis is restricted to sonorant clusters and the labiodentals fricative /v/ which undergoes devoicing but does not trigger it itself.
1.4.2. Temporal cues to [voice]

I will begin by briefly reviewing the feature [voice] and its correlates in obstruents as presented in the literature. English uses voiceless aspirated (e.g., cat – [kʰæt]), voiceless unaspirated (e.g., stick – [stʰk]), and voiced stops (e.g., dog – [dɡ]), the distribution of which is dependent on the environment. Slavic languages, on the other hand, distinguish between fully-voiced or prevoiced stops and voiceless unaspirated stops (Jessen 2001; Keating 1984; Lisker & Abramson 1964). Therefore, English L1 - Russian L2 learners have to learn a new distribution pattern of the feature [voice], different from English L1.

Because languages realize aspirated/unaspirated and voiced/voiceless obstruents in various manners depending on the environment in which they occur, and because Voice Onset Time (VOT) relates aspiration to voicing phonetically (Keating 1984), the measurements of voicing duration, VOT, and consonant duration should be taken in order to see how these phonetic categories are implemented in a specific environment for a specific language. In addition, the duration of the preceding and following vowels has been identified as an acoustic correlate of [voice] in obstruents in the studies of English obstruents (Crystal & House 1988; Lisker & Abramson 1964) and in the work on incomplete neutralization in various languages (Port & Crawford 1989; Port & O’Dell 1985; Warner et al., 2004). I will briefly describe each of these temporal characteristics in turn (for more details see Chapter Two).

Closure voicing. Jakobson et al., (1952) define voicing in stops acoustically as the presence of a periodic vocal fold vibration manifested by the voice bar in the spectrogram. The presence of vocal fold vibration is also taken as a sign of voicing in fricatives; whereas, there is no vocal fold vibration during the interval in which
frication noise is generated in voiceless fricatives. In voiced fricatives, the vocal folds vibrate for at least a portion of the interval when frication noise is present (Stevens et al., 1992). Although even intervocalic voiceless stops usually have some voicing into the closure as a result of the carryover effect of the vocal cord vibration in the preceding vowel, the presence of vocal fold vibration during closure/frication is thus generally considered the basic correlate of [voice]. For example, Port & O’Dell (1985) in their study of word-final devoicing in German reported somewhat more voicing into closure in underlyingly voiced stops.

Closure duration. In intervocalic stops, closure duration is defined as the time interval between termination of the vowel formant structure preceding the stop and onset of the following vowel. The closure duration of English intervocalic voiceless stops was found to be greater than that of voiced stops (Crystal & House 1988; Lisker 1957). A similar pattern was confirmed for French by Laeufer (1992) and for Russian, by Barry (1988).

Frication duration. Fricative duration is usually measured from the offset of F2 to onset of the following F2. A number of studies have found that the duration of frication in voiceless fricatives is greater than in voiced fricatives (Cole & Cooper 1975; Crystal & House 1988).

VOT. VOT is the temporal interval from stop release to the onset of voicing in the following vowel. In Polish and Russian, voiceless unaspirated stops are characterized by short-lag or zero VOT, while voiced stops are realized with negative VOT (prevoicing) (Jessen 2001; Keating 1984; Lisker & Abramson 1964). In English, on the other hand, voiceless aspirated stops are produced with long lag VOT, and
phonemically ‘voiced’ stops with short lag in some environments and negative VOT in others (Keating 1984; Kingston & Diehl 1994).

**Preceding and following vowels.** It has long been observed that vowels preceding voiced consonants are generally longer than vowels preceding voiceless consonants (Crystal & House 1988; Lehiste 1969; Port & Dalby 1982; Port & O’Dell 1985). In a study by Chen (1970) it was found that in English, Russian, French, and Korean vowels are longer before voiced consonants than before voiceless ones. The effects of consonants on the duration of following vowels are ambiguous. However, Allen & Miller (1999) claim that English stops /p t k/ decrease duration as compared to /b d g/ both in the preceding and following vowels.

**Additional measurements.** F0 onset and F1 onset are also listed as additional correlates of [voice]. F0 is the fundamental frequency at the onset of the vowel, which has been shown to be higher in languages with a voiced/voiceless distinction (no aspiration) after voiceless than voiced stops (Ladefoged 1977 on French; Shimizu 1989, 1996 on Japanese, cited in Jessen 2001). F1 frequency at the onset of the following vowel has also been reported to start higher after voiceless unaspirated stops of French than after voiced stops (Jessen 2001).

Although all of these acoustic properties have been investigated in phonetic research on the implementation of voice contrast, to my knowledge, none of the studies analyzing L2 production have examined *all* of the temporal cues to voicing. The present work takes an all-inclusive approach and investigates all the temporal cues to voice mentioned above, which allows for a more in-depth analysis of the

---

4 However, Keating (1984:292) notes that “some languages do not show this effect (Flege 1979; Keating 1979)” and points to the study of English and French word-final /b d g p t k/ by Mack (1982), who found that vowels before voiceless stops were 53% as long as those before voiced stops in English, and 74% in French.
factors affecting the maintenance of the underlying [voice] distinction (See Chapters Two and Three).

1.4.3. Voice neutralization

1.4.3.1. Voice neutralization in word-final position

A great deal of phonetic research has focused on the examination of word-final devoicing, a process by which the contrast between voiced and voiceless obstruents is neutralized word-finally. The underlying purpose of this research has been to determine whether and to what degree the underlying [voice] contrast is preserved, if at all. This phenomenon is attested in a number of languages, including Afrikaans, Catalan, Dutch, German, Polish and Russian (Barry 1988; Charles-Luce & Dinnsen 1984; Fourakis & Iverson 1984; Jassem & Richter 1989; Port & O’Dell 1985; Slowiaczek & Dinnsen 1985; Warner et al., 2004). A number of studies have shown that although speakers devoice final voiced consonants, voice contrast is not neutralized completely, at least in many languages. For example, in German, only voiceless syllable-final obstruents are allowed, so that both forms Alp (‘mountain pasture’) and Alb (‘elf’) should be pronounced with voiceless /p/, despite the orthography. However, in their study Port & O’Dell (1985) found that a small degree of difference was maintained between word-final /p/ and /b/.

Generally, incomplete neutralization is manifested by durational differences, such as the preceding vowel duration, duration of closure, duration of voicing, etc. For example, the results of Barry’s (1988) study of word-final obstruents in Russian suggested that obstruent duration and preceding vowel duration served as cues to maintain the voice contrast word-finally. Barry used eight native Russian speakers and 11 minimal pairs, which differed in voicing underlingly and orthographically, as
in luk - lu/k/ (‘onion’ Nom.Sg.) and lug – lu/k/ (‘meadow’ Nom.Sg.) vs. luka – lu/k/a (‘onion’ Gen.Sg) and luga – lu/g/a (‘meadow’ Gen.Sg.). The subjects in this study showed a large amount of variability, with some speakers using either of these two properties to signal the distinction between underlying voiced and voiceless obstruents, and some using both.

The observed production pattern in Barry’s data is similar to that found in Polish word-final obstruents (Slowiaczek & Dinnsen 1985). There was considerable variability in which acoustic measures seemed to show the effect, but the preceding vowel duration seemed to make a relatively consistent difference. Overall, the incomplete neutralization effect appeared to be reliable, if small.

Other phonetic studies of word-final devoicing have also reported incomplete voice neutralization (Charles-Luce & Dinnsen 1984; Port & O’Dell 1985; Port & Crawford 1989; Warner et al., 2004). That is, although subjects in these studies devoiced final obstruents, they preserved the underlying voice specification in other parts of the syllable. This, for example, was manifested by longer duration of a preceding vowel, which is characteristic of voiced obstruents.

Some researchers, however, ascribe the results of the studies that reported incomplete neutralization to the effects of orthography or hypercorrection (Fourakis & Iverson 1984; Jassem & Richter 1989), arguing that voicing neutralization would be complete in more natural speech. In a verb conjugation task given to four native speakers of German in the study by Fourakis & Iverson (1984), no effects of underlying voicing in terms of closure duration or the preceding vowel duration were found. In the reading task, however, the authors found vowel duration differences and concluded that these results were due to hypercorrection under linguistically artificial
conditions. Jassem & Richter (1989) obtained similar results in their analysis of final devoicing and RVA in Polish. The authors elicited 232 tokens from four Polish native speakers produced in a dialogue. They attributed an observed slight tendency towards longer vowel duration and shorter obstruent duration in the voiced condition to hypercorrectness, and argued that voice neutralization was phonetically complete.

In a more recent study, Warner et al., (2006) presented more evidence that orthography could cause incomplete neutralization. It was found that word pairs in Dutch which differed at the underlying morphological string, but did not differ in orthography exhibited complete neutralization. This finding in conjunction with the results of Warner et al., (2004) study suggested that it is orthographic differences that can lead to incomplete neutralization not the differences at the underlying level.

1.4.3.2. Voice neutralization in word-internal obstruent clusters

The results reported in the only acoustic study of RVA (regressive voicing assimilation) in Russian, by Burton and Robblee (1997), are in accord with those of Fourakis & Iverson (1984) and Jassem & Richter (1989). Burton & Robblee examined the phonetic effects of RVA across prepositional boundaries in stop-fricative and fricative-stop clusters produced by five Russian L1 - English L2 speakers. The target items consisted of five minimal pairs, each preceded by a preposition that ended in a voiced or voiceless stop /d/ or /t/ or a voiced or voiceless fricative /z/ or /s/ (e.g., nad sadom ‘above the garden’ and bez takta ‘lacking tact’), embedded in frame sentences.

It was found that in the fricative-stop clusters, the mean duration of the same fricative phonemes preceding /d/ was 41ms shorter than that preceding /t/. This finding was taken as evidence of RVA in this type of cluster. In the stop-fricative
clusters, where the fricative is the second consonant of the cluster and is not expected to be changed by assimilation, voiced fricatives were found to be 51ms shorter than voiceless fricatives. This pattern fits with a general cross-linguistic pattern of voiced obstruents being shorter than voiceless obstruents. However, an analysis of stops in the fricative-stop clusters revealed no significant difference in the durations of voiced and voiceless stops. Burton & Robblee concluded that this result was consistent with other studies showing that prevoicing, not duration, serves as a cue to voicing in Russian and Polish stops (Barry 1988, 1991; Keating 1979, 1984).

Thus, in contrast to overall segment duration, duration of voicing was found to play a more significant role in distinguishing voiced from voiceless obstruents. For both fricatives and stops in the first position in both types of clusters in the assimilating environment, duration of voicing in C1 was longer before voiced C2. Likewise, in obstruents in the second position, duration of voicing was longer for underlying voiced consonants than for underlying voiceless consonants. Overall, despite some speaker differences in the production of voicing, the authors concluded that acoustic analyses of the obstruent clusters occurring across word boundaries supported a process of RVA in Russian.

On the other hand, Barry (1988) in her investigation of obstruents in Russian, found that intervocalic voiced and voiceless obstruents were distinguished by the amount of voicing into closure and segment duration; whereas, the contrast between underlying voiced and voiceless obstruents was manifested by the segment duration only. Further evidence that voicing into closure is the most important cue to the voicing contrast comes from Barry’s (1991) perception study of intervocalic stops in Russian. In the forced identification task, native speakers of Russian identified a fully-
voiced stops /d/ as such at almost a 100% rate; likewise, “absence of voicing leads to 100% /t/ responses” (Barry 1991:59).

The voicing pattern in Russian obstruents in Barry’s studies is similar to the pattern described by Keating (1979, 1984) in her studies of Polish. In Polish, phonemically voiced stops, both initial and medial, have voicing during closure, and sometimes through the burst, making them fully-voiced. Phonemically voiceless stops, on the other hand, “do not have voicing into closure, with voicing always beginning after the burst” (Keating 1984:301).

1.4.3.3. Summary

The review of literature on word-final obstruent devoicing and RVA in word-internal obstruent clusters has demonstrated that a combination of different acoustic properties, such as voicing into closure, segment duration, VOT, and preceding and following vowel duration, rather than one specific property, is responsible for the voiced vs. voiceless distinction across languages, with sometimes one or the other cue being the most prominent indicator of RVA or sometimes several together.

Despite the significant number of studies on voice neutralization, such research has been restricted to native speakers of the various languages studied. The question as to whether and to what degree L2 speakers neutralize the distinction between underlying voiced and voiceless obstruents in appropriate contexts has not been addressed in phonetic research, and this is the issue the current work addresses in subsequent chapters.

In the next section, however, I present a review of L2 phonetic research that pertains to the more general question of why L2 speakers retain foreign accents. More
specifically, these studies investigate how L2 speakers acquire new phonemes and establish new sound categories.

1.4.4. L2 phonetics

Human languages differ according to the sound inventories and phonetic categories present within each language. Adult learners of an L2 are, therefore, expected to be aware of the new sounds in the L2 and learn how to produce them. In addition, they must learn to modify previously established sound categories and production patterns. In early stages of L2 acquisition, adult learners have been shown to judge an L2 sound category as an instance of an L1 category, substituting an L2 sound for a similar L1 sound (Flege 2001). However, in later stages of acquisition, adult learners’ pronunciation of L2 sounds begins gradually to change so as to resemble target sounds. There are a number of factors that are believed to influence the degree of perceived accent in L2 speakers.

Literature on L2 acquisition has shown quite conclusively that early L2 learners fare better than late learners. Some studies have suggested that foreign accent retention is due to the age-related loss of neural plasticity (Lenneberg 1967; Scovel 1969). An alternative hypothesis states that age-related changes in the degree of foreign accent are due to the interaction between the L1 and L2 systems, and the more developed L1 system exerts more influence on the L2 system (Flege 1987, 1995; Oyama 1979).

The age at learning is usually confounded with such variables as length of residence in the L2-speaking community and the amount of L1 and L2 use. The general agreement is that length of residence is a less important factor than the age at learning, if the effects of both length of residence and age at learning are reported. It
is believed that highly experienced L2 speakers are unlikely to improve their pronunciation significantly as the result of additional years of experience in the L2 environment. Less proficient speakers, on the other hand, may benefit from additional experience in the early phases. However, as Piske et al., (2001) point out, it is not possible to define precisely what constitutes ‘early phases’ of acquisition.

A number of studies have also looked at the influence of language use patterns on L2 pronunciation. The results of Thompson’s (1991) study of Russian-English bilinguals have suggested that the high amount of Russian L1 used by the subjects on a regular basis might have been responsible for their detectable foreign accent in English. Likewise, Flege (1995) found that early Italian-English bilinguals who spoke Italian frequently had stronger foreign accents than a matched group of subjects who used Italian infrequently. An effect of L1 use was also reported, among others, by Guion et al., (2000), who found that the amount of L1 use influenced the L2 speech in Quichua-Spanish bilinguals.

1.4.4.1. L1-L2 phonetic transfer

In order to account for L1 phonetic transfer into L2, Flege (1985, 1995) proposed the Speech Learning Model (SLM). According to the model, native and non-native sounds exist in a ‘common phonological space,’ and interact through the mechanism of category assimilation when new category formation has been blocked. Furthermore, if adult learners continue to perceive and judge L2 sounds as instances of already-established L1 categories, they may initially produce a non-native sound as a corresponding native sound without any modification. However, as more input becomes available, learners establish a merged category that reflects both native and non-native input. In such instances, “the perceived phonetic similarity of the L1 and
L2 speech sounds would be too great for phonetic category formation to occur, yet the cross-language phonetic differences would be auditorily detectable” (Flege 2001:226).

Thus, a merged category is used to produce and perceive instances of “the perceptually equated L1 and L2 speech sounds” (ibid.). If a new sound category is established, the native and non-native sound systems interact through the mechanism of category dissimilation when a new category is established for a speech sound that is found in the L2 but not the L1. The SLM posits that in order to maintain contrast within and between phonetic categories, a new L2 category moves away from the nearest L1 category. In this case, neither the L1 category nor the new L2 category is identical to the categories established by monolinguals (Flege 2001).

The interaction between the L1 and L2 phonetic subsystems has been looked at in a number of studies, which have shown that even fairly advanced L2 speakers do not produce L2 sounds authentically (Flege 1987, 1990; Flege & Eefting 1987; Flege & Port 1981; MacKay et al., 2001; for a detailed review see Piske et al., 2001). This issue has been addressed through the investigation of the acquisition of L2 acoustic properties. The findings of such studies are consistent with Flege’s (1990) Speech Learning Model. For example, Flege (1987) examined the production of VOT in /t/ in French and English words by French subjects who were highly experienced in English, and English speakers who were also experienced in French, and English and French monolinguals. As previously mentioned, English voiceless stops are produced with long-lag VOT whereas voiced stops have short-lag VOT values. The English monolinguals produced /t/ with long-lag VOT values and French monolinguals produced /t/ with short-lag values, as appropriate. The French-English bilinguals
produced English /t/ with longer average VOT than the French monolinguals produced in French /t/, but not as long as VOT produced by English monolinguals. English-French bilinguals, on the other hand, produced French /t/ with shorter VOT than the English monolinguals produced in English stops, but longer VOT in French /t/ than the French monolinguals did. Likewise, Caramazza et al., (1973) found that French L1 speakers produced English stops /p t k/ with shorter VOT than native English speakers did, yet VOT was longer than is typical for French stops produced by French monolinguals.

Similar results were reported by Flege (1990) in his study of stop VOT values in Spanish – English bilinguals. It was found that the control group of Spanish monolinguals produced Spanish /t/ with considerably shorter VOT values than the English monolinguals. The late L2 learners produced English /t/ with VOT values intermediate to Spanish L1 and English L1 values. The early bilinguals, however, did not differ from English native speakers in their production of English VOT values. Thus, Flege concluded that learners who begin learning their L2 in early childhood, but not those who begin learning an L2 as adults, are able to establish separate phonetic categories for similar L1 and L2 sounds.

According to Flege’s Speech Learning Model (1990), a Spanish L1 child learning English as L2 at the age of 5-6 years will be able to establish separate phonetic categories for each English stop, at the same time maintaining phonetic categories already established for Spanish stops. Adult learners of English are not expected to establish new categories for ‘similar’ L2 sounds because they fail to notice the acoustic differences between short-lag Spanish stops and long-lag English stops, and thus perceive and produce these sounds as belonging to the same category.
The findings of these studies have been taken as evidence that phonetic learning has taken place in the L2. Late L2 speakers phonetically approximated, but did not achieve, the target VOT value for /p t k/ in English L2. Given these results, Flege (1987, 1990) claimed that although adult L2 speakers are aware of the phonetic differences between VOT values in their L1 and English L2, they nevertheless fail to establish new phonetic categories for the L2 sounds.

A similar conclusion was reached by MacKay, Flege, Piske & Schirru (2001), who tested phonetic category assimilation posited by the Speech Learning Model. In four experiments, the authors analyzed the production of English /b/ and the perception of English /b d g/ in Italian-English bilinguals. In English, phonemically voiced stops /b d g/ are usually realized with short-lag VOTs that are similar to Italian voiceless stops /p t k/; whereas, Italian /b d g/ are realized as prevoiced stops. As in the previous studies by Flege (1987, 1990), the bilinguals were found to approximate the English phonetic norm for /b d g/. Early Italian-English bilinguals were shown to resemble the natives to a greater extent than late learners, yet both groups of bilinguals differed from Italian monolinguals. More importantly, bilinguals were shown to differ from both Italian monolinguals and English monolinguals, and the more experienced bilinguals showed the greatest effect of English on their production of Italian stops. This result has been taken as evidence that the Italian-English bilinguals had not established separate phonetic categories for the L2, but instead developed merged English and Italian categories for /b d g/, which embraced the properties of corresponding L1 and L2 stops. As the authors argue, had the L2 speakers established a new category for an L2 sound, their production of L1 stops
would not have changed because two categories would have been kept separate and intact.

Overall, the findings of this study led MacKay and colleagues to claim that phonetic learning was possible in the absence of new category formation, and the category formation was unlikely even by early bilinguals because of the perceptual similarity of English and Spanish stops. The early bilinguals in their study performed better than the late bilinguals because they were more likely to be exposed to more native realizations of English short-lag stops and receive more authentic input from native speakers, not because they were able to establish independent phonetic categories for English stops.

1.4.4.2. Summary

The findings of phonetic research on L1-L2 transfer and category restructuring could be extended to the present study. In light of the Speech Learning Model (Flege 1995), L2 learners are believed to be able to restructure existing sound categories to encompass both L1 and L2 sounds. According to Flege, partial approximation to L2 phonetic norms indicates that L2 learners note acoustic differences between similar L1 and L2 sounds. Although an L2 sound is not produced in a native manner, it nevertheless differs acoustically from its counterpart in L1. English L1-Russian L2 speakers in this study are thus expected to be able to modify existing L1 phonetic categories at least somewhat in order to incorporate new, acoustically different, L2 obstruents, such as prevoiced stops as opposed to English short-lag VOT.

It should be recalled that Flege’s (1995) SLM has been proposed as a model of acquisition of L2 segments and sound categories. This research, however, is primarily concerned with the acquisition of L2 phonological patterns. Assuming that the same
mechanism that applies to the acquisition of L2 sounds can apply to phonological learning, there are several possible scenarios. First, one can argue that similarly to how adult learners create a merged L1 - L2 category for a specific sound, they merge L1 and L2 constraint rankings, too. During the course of acquisition the rankings are not stable and the constraints are in flux. As a result, learners might switch from one ranking to another producing either the target L1 pattern or the L2 pattern.

Second, given the overall progress L2 learners usually make over a period of time while acquiring segmental categories, it is possible that non-natives exhibit gradual acquisition of L2 phonological patterns as well. Since the application of phonological patterns is argued to be categorical, it is not unconceivable that L2 learners might successfully apply them on certain occasions but fail to do so on other occasions over the course of learning.

Yet another possibility is that L2 learners learn to produce the target pattern; yet, its production is dependent on the realization of the phoneme string in question. That is, the realization of the phonological pattern is turned over to the learners’ segmental categories, which may be L1 – L2 merged ones, or transferred L1 ones, or separate L1 and L2 ones for early bilinguals. In other words, learners might be aware of RVA in Russian, but produce it a non-native manner due to L1 phonetic transfer.

Recall that previous research has shown that only early but not later bilinguals have been shown to establish new, separate, categories for L2 sounds. Since this work investigates late bilinguals, the subjects are not expected to produce Russian stops and fricatives in a native way. Rather, advanced learners are predicted to approximate L2 sounds; whereas, less proficient, intermediate, learners are predicted to maintain existing L1 categories and interpret L2 sounds as instances of L1 sounds.
It is not unreasonable to assume that L2 subjects’ inability to produce Russian obstruents authentically might interfere with their overall production of RVA in obstruent clusters. That is, although L2 speakers might be aware of RVA in consonant clusters in Russian, it could be applied in their production to various extents, and differently from Russian speakers, due to the fact that L2 speakers use English acoustic values when producing Russian obstruents.

The predicted inability of adult L2 learners to produce L2 sounds acoustically similar to natives raises the question of whether L2 learners are also unable to acquire L2 phonological patterns. Therefore, the experiments reported in Chapters Two and Three will not only present the empirical data but will also provide information about the details of L2 phonological acquisition. More specifically, the exact facts concerning the production of RVA in the native and non-native speakers will inform my analysis about the mechanism L2 learners employ in order to modify their L1 phonological patterns, and will allow me to propose a model that captures acoustic differences in the L1 and L2 production of RVA and its gradual acquisition.

To sum up, there are two conceivable reasons why Russian L2 productions of RVA might differ from L1 production. One is learners’ failure to acquire the L2 phonological pattern. As a consequence, they fail to produce RVA altogether and use the English L1 production pattern. The second reason is that L2 learners might acquire the L2 pattern but may still transfer the L1 category for the particular segments into Russian L2 when attempting to produce RVA. In this case, they still sound like a non-native because they produce a wrong category.
1.5. Conclusion

In this chapter, I discussed temporal cues to [voice], and acoustics of voice neutralization in word-final position and obstruent clusters in word-internal position, as presented in current phonetic research. I also reviewed L2 learning models proposed by Flege (1985, 1995), which account for the L2 learners’ inability to produce L2 sounds authentically. In Chapters Two and Three, I will explore these claims as they relate to a real-word production experiment and a nonsense-word experiment, on RVA in native Russian and English-Russian L2 learners, respectively. Chapter Four will turn to the formal phonological analysis of L2 acquisition of RVA, as informed by the studies in the previous two chapters. Chapter Five will conclude with implications for phonetic and phonological theories and for L2 teaching.
CHAPTER TWO

REAL-WORD PRODUCTION EXPERIMENT

2.0. Introduction

Although the acoustics of regressive voicing assimilation (RVA) in Russian have been studied before (Burton et al., 1994; Burton & Robblee 1997), the focus was exclusively on RVA at the clitic-word boundaries in native speakers' Russian. In order to address the question of RVA in a different phonological environment, and in learners of Russian, two experiments were carried out. Experiment One investigated RVA in Russian L2 learners in a word-medial environment using real-words and Experiment Two investigated the same phenomena in nonsense stimuli. This chapter provides an account of the real-word production experiment, Experiment One, which was designed in order to test and compare native and non-native productions of word-medial obstruent clusters with RVA.

Recall that RVA is said to occur when voicing properties of the first obstruent in the obstruent cluster are consistent with the voicing properties of the second obstruent. If RVA does not take place, the voicing properties of the first obstruent in the cluster remain the same and are not affected by the following obstruent. Experiment One thus tested whether and to what extent native and non-native speakers of Russian applied RVA in obstruent clusters.

Previous research on L2 phonetic acquisition has demonstrated that even proficient L2 speakers produce L2 sounds acoustically differently from L1 speakers (Flege et al., 1997, 1999; Flege & Liu, 2001; MacKay et al., 2001). As amount of exposure to the target language increases, L2 speakers attain more authentic
pronunciation; albeit not completely nativelike pronunciation. As mentioned previously, RVA production in L2 has not been studied at all; however, it is not unreasonable to assume that L2 speakers in this study will behave similarly to L2 subjects in earlier literature on acquisition of other sounds. The specific purpose of these experiments was to detect which correlates of RVA (e.g., voicing duration or segment duration) natives and non-natives produce differently, and how learners acquire these properties. It was predicted that in both experiments, L2 learners would produce obstruent clusters, in which RVA applies, less authentically than native speakers, if at all. Recall that in Section 1.4.4.2., it was hypothesized that L2 learners acquire phonological patterns categorically; therefore, some learners were predicted to fail noticing RVA and fail to produce it altogether, and some were predicted to be able to acquire RVA. However, it was also hypothesized that although L2 learners had acquired the L2 phonological pattern, they might not realize it phonetically in a native manner. That is, L2 speakers might employ L1 sound categories when producing devoiced obstruents in the clusters where RVA is required, or use merged L1 – L2 categories which resulted in the non-native production of RVA. Overall, more proficient L2 learners were expected to perform better (devoice more consistently) than less proficient speakers, who have had less exposure to the target language.

2.1. Methods

2.1.1. Subjects

Two groups of ten non-native Russian speakers were recruited to participate in the experiment: advanced English L1-Russian L2 learners and intermediate English L1-
Russian L2 learners. In addition, a group of monolingual native Russian speakers (N=10, mean age 46) was recruited as a control group, for a total of 30 subjects. The subjects had no known history of speech or hearing disorders. All the intermediate learners were recruited among the third-semester Russian students (RSSS 201a) of the Department of Russian and Slavic Studies at the University of Arizona and were given extra credit by the instructor for participation in the experiment. None of the intermediate students had studied or traveled to a Russian-speaking country and had very little exposure to the target language outside classroom.

All the advanced learners were recruited among Master’s students (1st and 2nd year) enrolled at the Department of Russian and Slavic Studies at the time of the experiment. The mean age of the participants was 28 years. Only those advanced students who had studied or lived in Russia or some other Russian-speaking country (e.g., Kazakhstan) were recruited for the study. Advanced learners were paid $10 for participation. All the non-native subjects were recorded during the second half of the semester.

Native speakers were monolingual Russian speakers living in Kazakhstan who did not have any foreign language exposure. All of the L1 speakers were college graduates, with two of them holding doctoral degrees. Recruitment was done through word-of-mouth. As with the advanced learners, the monolingual Russian speakers were paid a small amount of money for their participation.

The latter subjects were tested in Kazakhstan in order to avoid any influence of English on subjects’ speech\(^5\) that might cause phonetic/phonological interference.

\(^5\) Research in second language acquisition has discovered that a second language might exert an influence on the first language, especially as far as pronunciation is concerned (Flege 1988). It has been found that the quality of consonants and vowels of the first language could change after speakers have
In addition to Kazakh, Russian is an official language of the Republic of Kazakhstan and remains the language of everyday communication with the native Russian-speaking population being about 30% of the total population in the country. All the subjects were monolingual Russian speakers who did not have any knowledge of the Kazakh language.

2.1.2. Materials

The stimuli consisted of 36 real words of Russian, each containing a medial obstruent cluster, representing two orders of obstruents: (1) a stop-fricative cluster, as in

\[ \text{obsharit} \rightarrow /\Lambda p\text{ʃarit}/ \ 'to rummage', \]

and (2) a fricative-stop cluster, as in

\[ \text{vezti} \rightarrow /\text{vist}/ \ 'to drive, transport'. \]

The stop-fricative condition included four obstruent combinations where RVA applies: \( bsh, bs, ds, \) and \( dsh \). Each cluster was represented in two words in three different environments: an assimilating environment where the cluster was voiced-voiceless and was expected to undergo RVA (obshyt \( \rightarrow o/p\text{ʃ}/yt \ 'to saw around')

\[ /\text{obzhyt} /yt \ 'to settle down'), \]

and a second non-assimilating environment where both members of the cluster were fully-voiced (obzhyt \( \rightarrow o/b\text{ʃ}/yt \ 'to settle down')

\[ /\text{lapsha} /\text{ʃ}/a 'noodles'). \]

Thus, the first obstruent of the cluster represents an obstruent made voiceless by assimilation, one that is underlying voiced and remains so, and one that is underlyingly voiceless and remains so. The latter two conditions provide a comparison to determine to what extent the assimilated obstruent has been devoiced. The total number of tokens in this spent a significant amount of time speaking their second language. This might be the case with native speakers of Russian residing in the USA; therefore, in order to obtain valid results, Russian native speakers were recorded in their home country.
condition was 24. Both high and low frequency words were included in the experiment. However, due to the limited number of relevant real-word data, the tokens were not controlled for frequency (Although I return to this issue in Chapter Five). (See Table 2.1 for examples, and Appendix A for the complete list of stimuli).

<table>
<thead>
<tr>
<th>Environment</th>
<th>Assimilating ( C_1 [+\text{vce}]C_2 [-\text{vce}] )</th>
<th>Voiced ( C_1 [+\text{vce}]C_2 [+\text{vce}] )</th>
<th>Voiceless ( C_1 [-\text{vce}]C_2 [-\text{vce}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop-Fricative</td>
<td>podsest-( pa/\text{ts}/\text{est} )( to \text{ sit next to} )</td>
<td>podzaniat-( po/\text{dz}/\text{an/at} )( to \text{ borrow} )</td>
<td>otsudit-( ot/\text{ts}/\text{udit} )( to \text{ judge} )</td>
</tr>
<tr>
<td>Fricative-Stop</td>
<td>vezi-( vi/st/i )( to \text{ drive} )</td>
<td>razdat-( ra/\text{zd}/\text{at} )( to \text{ hand out} )</td>
<td>vesti-( vi/st/i )( to \text{ lead} )</td>
</tr>
</tbody>
</table>

Table 2.1. Example stimuli containing Russian obstruent combinations in the three voicing environments

The fricative-stop condition included three obstruent combinations where RVA applies: \( zk, zhk, \) and \( zt \). The \( zk \) cluster was represented in two words in three different environments: an assimilating environment where the cluster was voiced-voiceless and was expected to undergo RVA (\( \text{salazki} – /\text{s}\Lambda\text{lask}\text{a}/ ‘sledge’ \)), a control environment where both members of the cluster were underlyingly voiced (\( \text{rozgi} – /\text{rozg}\text{i}/ ‘birch-tree rods’ \)), and a second control environment where both consonants were underlyingly voiceless (\( \text{kraska} – /\text{krask}\text{a}/ ‘paint’ \)). (cf. Table 2.1 for examples, and Appendix A for the complete list of real-word stimuli). The \( zhk \) and \( zt \) clusters were each represented in one word only in three different environments (assimilating, voiced and voiceless), due to the difficulty finding items containing word-medial \( zhg \) and \( zt \) clusters. The total number of tokens in this condition was 12.

Labio-dental fricatives \( [v] \) and \( [f] \) occurring in medial clusters were excluded from the analysis altogether due to the fact that with the exception of a \( [v] \) followed by a voiceless stop (\( \text{lavka} – /\text{lafk}\text{a}/ ‘bench’ \)), word-medial combinations of a \( [v] \)
followed by a voiced stop and an [f] followed by a stop are not attested in the language.

Also, items containing clusters where the first member is voiceless and the second member is voiced as in *prosba* - */prosba/ ‘request’ were excluded. Although RVA does occur in this environment in Russian, words with regressive devoicing are much more plentiful. Thus, only obstruent clusters in which neutralizing of the voicing contrast takes place were compared with the clusters in which both members of a cluster are either voiced or voiceless.

As mentioned, all the clusters in the real-word stimuli were word-medial. Out of 36 items, RVA occurred across morpheme boundaries in 25 of them: 18 items were of a prefix + stem shape (e.g., *pod-sest* - */pʌtsesˈt/ ‘to sit down next to’) and seven were of a stem + suffix shape (e.g., *knish-ka* – */knifʃka/ ‘small book’). The remaining 11 items were non-derived forms with the cluster appearing root-internally (e.g., *luzga* – */luzga/ ‘sunflower shells’). (See Appendix A for more details.)

Within each set, the items were controlled for stress and surrounding vowel quality whenever possible: the clusters were either preceded or followed by a stressed vowel. If a stressed vowel preceding or following an obstruent cluster in the assimilating environment was a low-back vowel [a], then items containing an analogous vowel in the same position in the other voicing conditions were chosen whenever possible; e.g., *obshárit* - */ʌpʃarit/ (‘to rummage’) - *obzhárit* /ʌbʒarit/ (‘to deep fry’) - *lapshá* - /ʌpʃa/ (‘noodles’). However, as this set of tokens demonstrates, the vowels, as well as the shape of a syllable they occurred in, are not identical. In the first two tokens, the vowel preceding the consonant cluster is in an open syllable;
whereas, the vowel following the cluster, is in the closed syllable. In the third token in the set, the preceding vowel is in the closed syllable, and the following vowel is in the open syllable. Therefore, it is possible that inherent differences in vowel quality might affect some measurements.

2.1.3. Procedure

Subjects were presented with the stimuli list, which was typed in Cyrillic, and asked to read the items on the list. The instructions were presented verbally in the native language of the subjects. The pronunciation and meanings of each of the stimuli were double checked by the author, who is a native speaker of Russian, and another Russian native speaker. In order to control for intonation and speech rate, the stimuli were presented in a frame sentence typed in Russian: “Ya govoriu slovo ___ seychas” (‘I am saying the word ___ now’). The subjects were asked to read the stimuli at a regular rate with a normal intonation. The subjects were not told about the purpose of the experiment.

Each item was recorded twice to ensure that a usable token of each item was obtained. English-speaking subjects were recorded using a high-quality head-mounted microphone and a CD-recorder at 44.1 Hz sampling rate in a sound-protected booth at the phonetics lab at the Department of Linguistics at the University of Arizona. The native speakers were recorded in Kazakhstan at the participant’s or the investigator’s home. The recordings were done in a quiet room avoiding traffic noise and other types of noise. The subjects were recorded digitally straight to hard-drive of a laptop using a microphone which was also used for recording non-native subjects in the lab. Both real and nonsense words, the latter to be discussed in Chapter Three, were recorded during one recording session with a break in between. The real stimuli were
recorded first. Thirty filler tokens were included at the beginning of the token list as practice items. This was done in order to allow the participants to familiarize themselves with the task requirements and get used to speaking to the microphone. On average, the experiment took 30 minutes.

Since some low frequency words were included in the real word stimulus list, it was necessary to ensure that the non-native speakers were at least somewhat familiar with these words and did not treat these items as nonsense words. To ascertain this, non-native participants’ language knowledge (both intermediate and advanced) was tested prior to the production experiment. First, the non-native speakers were given a survey containing the items from the stimulus list, which they were asked to define. They were instructed that the surveys would not be graded and were taken for research purposes only. Afterwards the participants were familiarized with the stimulus definitions but were not tested on their knowledge.

2.2. Analysis

2.2.1. Measurements

The methodology and measurements utilized in this experiment to analyze the degree of RVA are based on the ones used by Burton & Robblee (1997), previously discussed in Chapter One. As in their study, the following measurements were taken: overall duration of each obstruent in a cluster, amount of voicing in each obstruent in the cluster, duration of the vowel preceding and following the obstruent cluster, and VOT of the stops in the fricative-stop combination only (since VOT is not measurable otherwise). Details of how these were measured appear below.
2.2.1.1. Segment duration

Stop. In the stop-fricative clusters, stop duration was measured from the end of the preceding vowel, signaled by the end of F2, to the onset of frication noise. In the fricative-stop clusters, stop duration was measured from the offset of frication noise to the beginning of a vowel, signaled by the beginning of F2. (See Figure 2.1 for an illustration.)

Fricative. In the stop-fricative clusters, fricative duration was measured from the onset of frication noise to its offset. The end of frication was indicated by the emergence of vocalic formant (F2) structure in the following vowel. In the fricative-stop clusters, fricative duration was measured from the end of vocalic formant (F2) structure to the offset of frication noise. The end of frication prior to the stop was determined by the emergence of a periodic wave pattern for voiced stops and the absence of sound louder than background noise for voiceless stops. (See Figure 2.1 for illustration.)

Vowel. The beginning of the vowels preceding and following the cluster was determined by the emergence of peaks in the spectrogram and by the emergence of vocalic formant (F2) structure in the waveform. The end of the vowels was determined by the disappearance of vocalic structure and peaks. (See Figure 2.1.)
Figure 2.1. Sample waveform and spectrogram of the V1-C1-C1-V2 sequence of the word *lobzik* –/lobzık/ ‘fret-saw’ produced by a native speaker. Vertical lines show duration of each segment in the sequence. Voicing duration is the same as segment duration in both obstruents, as both are voiced throughout.

2.2.1.2. Voicing proportion

Voicing proportion for each segment was calculated by dividing voicing duration by segment duration. Voicing duration for stops and fricatives was measured as periodic portions of the segments as displayed in the repeating wave and/or as a voice bar on the bottom part of a spectrogram from the onset of stop/fricative as described above to the end of voicing. If voicing did not continue throughout the entire segment, the voicing portion of the segment was measured. For fully-voiced obstruents, voicing duration was the same as segment duration. (cf. Figure 2.1.)

2.2.1.3. VOT
Voice Onset Time (VOT) is known to play a role in distinguishing voiced and voiceless stops. Voiced stops have either negative VOT or very short positive VOT, and voiceless stops have positive VOT (Ladefoged 2001; Lisker & Abramson 1964).

Since VOT could only be measured for stops, in this experiment, this measurement was taken for C2 in the fricative-stop clusters only. In cases where there is no voicing during the stop closure, such a stop is considered voiceless, and its VOT is measured from the stop burst (the closure release) to the onset of vocalic formant (F2) structure. (See Figure 2.2. for an illustration). If voicing begins before the closure release, the stop is considered fully-voiced (prevoiced), and has a negative VOT value.

![Figure 2.2. Sample waveform and of the C1[+vce]C2[-vce]V1 sequence in the word lozhka – /loʃkə/ ‘spoon’ produced by a native subject. Vertical lines indicate the segment boundaries.](image-url)
2.2.1.4. Data analysis

Waveforms and spectrograms of the segments under scrutiny were analyzed using the speech analysis software Praat (Boersma & Weenink 1993). The measurements were analyzed using statistical software (SPSS Inc., Chicago IL). As mentioned previously, there were three factors in the experiment: cluster type (stop-fricative and fricative-stop), voicing environment (assimilating, voiced and voiceless) and language level (native, non-native advanced and intermediate). The language level factor was a between-subjects factor, while the other two factors were within-subjects. Data was averaged over items to give one value for each speaker for each condition (by-subjects testing). Most dependent variables (duration of each consonant, proportion of voicing of each consonant, duration of the preceding and following vowels) were analyzed using all three factors. However, since VOT could only be measured in the fricative-stop clusters, cluster-type in that analysis is moot; therefore only the following two factors were used for the VOT analysis: language level and voicing environment. However, instead of an overall ANOVA (exclusive of VOT), two interaction comparisons were used for each dependent variable in order to test the hypotheses in a targeted way, as described below.

2.3. Results

The important issue in this experiment was to determine whether speakers, both L1 and L2, realize consonants that should assimilate to voiceless as similar to the fully-voiceless consonants (i.e. successfully assimilated), similar to the fully-voiced consonants (i.e. not assimilated at all), or somewhere in between (i.e. partially assimilated). In order to determine which was the case, the data were analyzed by
comparing the assimilation environment to the fully-voiceless environment, and separately, the assimilation environment to the fully-voiced environment. For each dependent measure, two interaction comparisons were used to do this. Each interaction comparison was in itself a 2x2x3 ANOVA, with cluster type (stop-fricative and fricative-stop) and language level (native, non-native advanced, and non-native intermediate) as two of the factors. The third factor, with two levels, was the voicing environment. As mentioned, one interaction comparison included only the assimilation environment and the fully-voiceless environment in order to determine whether assimilated consonants differ from fully-voiceless ones; such a difference would indicate a failure to assimilate fully on the part of the speaker. The other interaction comparison used only the assimilation environment and the fully-voiced environment data, in order to determine whether assimilated consonants differed from fully-voiced ones; such a difference would indicate at least some assimilation on the part of the speaker.

Russian native speakers served as controls and were expected to produce RVA in the assimilated-to-voiceless condition consistently. They were expected to devoice a fully-voiced obstruent when followed by a voiceless obstruent, thus rendering such an obstruent similar to a fully-voiceless obstruent in the voiceless environment in terms of segment and voicing duration. Using the two levels of language learners allowed me to determine whether non-native speakers devoiced fully-voiced obstruents to a lesser degree than native speakers, thus rendering such obstruents more similar to fully-voiced obstruents in terms of segment and voicing duration. In previous L2 phonetic research (Flege et al., 1997, 1999; MacKay et al., 2001), L2 speakers were shown to produce L2 sounds acoustically differently from native
speakers, with proficient speakers performing better than beginners. Therefore, in this study, assuming that language learners acquire RVA categorically, it was predicted that those L2 learners who had noticed RVA devoiced C1 in a cluster but to a lesser degree than natives because they used merged L1 – L2 phonetic categories or L1 categories to produce L2 obstruents. In addition, it was also hypothesized that less proficient learners might fail to acquire RVA and fail to apply it altogether. I will now examine the results for each acoustic measure in turn.

2.3.1. Duration of the first consonant in the cluster (C1)

In this section, I analyze the duration of the first obstruent in a cluster (C1), which is a stop in the stop-fricative clusters, and a fricative in the fricative-stop clusters. The working hypothesis was that native speakers applied RVA in the voiced-voiceless consonant clusters, devoicing voiced consonants and producing them similarly (in duration as well as other characteristics) to fully-voiceless ones. Devoiced and voiceless obstruents were expected to be longer than fully-voiced. Both groups of L2 speakers were predicted to apply RVA less consistently than L1 speakers producing voiced consonants in the assimilating environment only partially devoiced. Therefore, L2 consonants in the assimilating environment were predicted to be shorter in duration than fully-voiceless ones but longer than fully-voiced. I will first analyze C1 duration in the assimilating vs. fully-voiced environments and then turn to the analysis of C1 duration in the assimilating vs. fully-voiceless environments.

2.3.1.1. C1 duration in assimilating-to-voiceless vs. fully-voiced environments

Figure 2.3. below presents an analysis of C1 duration in the assimilating vs. fully-voiced environments. In all groups of subjects, the duration of C1 stops presented in the top panel of the graph, was similar in both environments. However, as the height
of the bars in the bottom panel shows, all subjects produced C1 fricatives in the assimilating longer than in the fully-voiced environment.

Figure 2.3. Comparisons of the duration of the first consonant (C1) in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments

The results of the analysis revealed that there were significant main effects of voicing environment: (F(1,27) = 79.550, p < .001) and cluster type: (F(1,27) = 8.152, p < .001); whereas, the main effect of language level was not significant (F(2,27) = 1.164, p > .05). Of the three two-way interactions, only the interaction of cluster type by environment was significant: (F(2,27) = 30.374, p < .001); cluster type by language level: (F(2,27) = 1.538, p > .05); environment by language level: (F<1). The three-way interaction (level by environment by cluster type) was significant (F(2,27) =
4.519, p < .05). Because of the significance of the three-way interaction the data were analyzed further in order to test the effects of language level and environment for each cluster separately.

For the stop-fricative clusters, there were no significant effects of environment: (F(1,27) = 3.533, p > .05) or language level: (F < 1). The two-way interaction (level by environment) was not significant either (F(2,27) = 1.461, p > .05). Thus, the stop (C1) duration in stop-fricative clusters was statistically the same for all environments and language levels.

For the fricative-stop clusters however, there were significant main effects of environment: (F(1,27) = 72.009, p < .001) and language level: (F(2,27) = 1.258, p < .05). The two-way interaction (level by environment) for the fricative-stop cluster was also significant (F(2,27) = 72.009, p < .001). Because of the significance of the two-way interaction, the data were further analyzed in order to test the effect of environment for each level separately. The C1 fricative was significantly shorter in fully-voiced clusters than in assimilated-to-voiceless clusters for each speaker level (natives: F(1,9) = 27.019; advanced: F(1,9) = 38.397; intermediate: F(1,9) = 8.555; p < .001).

The fact that C1 fricatives were significantly longer in the assimilating environment, for all language levels, suggested that all levels of speakers treated assimilated-to-voiceless fricatives differently than fully-voiced fricatives. However, the interaction with language level was significant, reflecting the smaller effect of

---

6 For all 3-factor analyses, even when only one interaction was significant, I followed a standard procedure of testing the two-factor design of language level by environment for each cluster type separately, regardless of which interactions were significant. This was done because environment was the factor of primary interest. Also it was necessary to keep the interpretation of results standardized across all the measures in the two experiments in order to facilitate the interpretation of the results.
environment for intermediate learners. This suggested that the intermediate learners, more than the other two groups, made the assimilated-to-voiceless fricatives relatively similar to the fully-voiced fricatives.

2.3.1.2. C1 duration in assimilating-to-voiceless vs. fully-voiceless environments

The same three-factor interaction comparison as was used for the assimilating vs. fully-voiced environments above (language level, cluster type, and voicing environment) was used here to compare the assimilating-to-voiceless environment to the fully-voiceless environment. As in the previous analysis, the C1 stop duration was similar in both environments in all groups of subjects, and so was the C1 fricative duration (bottom panel) but in the native and advanced groups only. These data appear in Figure 2.4. below (with assimilating data repeated from Figure 2.3 for comparison).
Figure 2.4. Comparisons of the duration of the first consonant (C1) in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments

As in the previous analysis, there were main effects of environment: \( F(1,27) = 94.439, p < .001 \) and cluster type: \( F(1,27) = 6.543, p < .05 \). The main effect of language level was not significant \( F<1 \). Unlike the previous analysis, the two-way interaction of language level and cluster type was significant \( F(2,27) = 3.980, p < .05 \), whereas the two-way interactions of language level by environment and environment by cluster type were not significant \( F<1 \). The three-way interaction (environment by cluster type by language level) was also not significant \( F(2,27) = 32.655, p > .05 \).

Because the two-way interaction of language level and cluster type was significant, a two-way ANOVA with language level (native, advanced non-native and
intermediate non-native) and environment (assimilating and fully-voiceless) was performed for each cluster (stop-fricative and fricative-stop). Contrary to the standard procedure, the data were split across the consonant cluster type factor rather than being collapsed over the factor in order to analyze the obstruent cluster production in each cluster type separately. This type of statistical design allowed me to see whether the production of RVA in obstruents in a cluster varied as a result of a cluster type.

In the stop-fricative clusters, there was no significant main effect of environment (F(1,27) = 3.476, p > .05) or language level (F(2,27) = 1.057, p > .05). The two-way interaction (language level by environment) for the stop-fricative cluster was also not significant (F(2,27) = 1.607, p > .05). Thus, as in the previous analysis, the stop (C1) duration in stop-fricative clusters was the same for all voicing environments and language levels.

In the fricative-stop clusters, however, there were significant effects of environment: (F(1,27) = 34.658, p < .05) and language level: (F(2,27) = 5.062, p < .01). The two-way interaction (language level by environment) was also significant (F(2,27) = 4.658, p < .01). Thus, the data were analyzed further in order to test the effect of environment for each level separately. The C1 fricative duration was significantly longer in fully-voiceless clusters than in assimilated-to-voiceless clusters for intermediate subjects only (natives: F(1,9) = 2.425, p > .05); advanced: F < 1; intermediate: (F(1,9) = 5.808, p < .05).

2.3.1.3. Summary

The significant effects of environment and language level in the fricative-stop clusters demonstrated that subjects’ production of fricatives in the C1 position varied with regard to the segment length produced in each fricative. The fact that the duration of
fricatives in an assimilating environment produced by intermediate L2 speakers was similar to the duration of fully-voiced fricatives indicated that intermediate L2 speakers did not apply the RVA pattern to the same extent as native-speaking and advanced L2 subjects. However, the intermediate learners’ C1 fricative duration differed significantly between fully-voiceless and assimilated, too. Intermediate learners made their assimilated C1 fricatives an intermediate length between their fully-voiced and their fully-voiceless ones. Thus, it could be concluded that they assimilated, but not completely.

In addition, the lack of a significant simple effect for native speakers and advanced learners in the comparison of the assimilating vs. fully-voiceless environments, and the presence of an effect for the assimilating vs. fully-voiced environments showed that these two groups both assimilated, and in the same way, on this measure at least. Unlike C1 fricatives, the duration of C1 stops was not dependent on the environment or language level. This result was in accord with Burton & Robblee’s (1997) results, who following Keating (1984) claim that prevoicing, not duration, is the primary cue to voicing in stops in Slavic languages.

2.3.2. C1 voicing proportion

Previous research has identified voicing duration as one of acoustic parameters for the implementation of the voice contrast (Slowiaczek & Dinnsen 1985), which, in turn has been shown to be an important correlate of voicing assimilation (Burton & Robblee 1997). Moreover, Burton & Robblee’s (1997) argued that overall stop duration was inappropriate for assessing assimilation effects in their data, and that voicing duration was shown to be a more reliable assessment tool.
In this section I present an analysis of the voicing proportion of C1, which was calculated with the purpose of determining the amount of voicing present or lack thereof. The voicing proportion was calculated by dividing segment voicing duration by segment duration, as discussed in 2.2.1.2 above.

Recall that if RVA occurred, the voicing proportion of C1 in the assimilating environment was expected to be similar to the C1 voicing proportion in the fully-voiceless environment. If not, the C1 voicing proportions in the assimilating and fully-voiced environments were expected to be similar.

2.3.2.1. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments

The same three-factor interaction comparison with language level, environment and cluster type as variables as was used to analyze C1 duration, was performed on C1 mean voicing proportion. L1 subjects were expected to produce shorter voicing duration in the assimilating than in fully-voiced environments; whereas, L2 subjects were predicted to produce longer voicing duration than L1 subjects in at least the assimilating environment. As Figure 2.5. demonstrates, this prediction was born out but for C1 stops only, which are presented in the top panel of the graph. Natives (the light bar) produced the shortest voicing duration in assimilating C1 stops. Although both intermediate and advanced subjects produced longer C1 voicing duration than natives did in the assimilating environment, it was nevertheless shorter than their values in the fully-voiced environment (compare the height of the leftmost and rightmost graphs in the top panel). All groups produced little voicing in assimilating C1 fricatives, as seen in the bottom panel of the graph.
Figure 2.5. Comparisons of voicing proportion of the first consonant (C1) in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments

There were significant main effects of environment: (F(1,27) = 224.583, p < .001) and cluster type: (F(1,27) = 313.191, p < .001). The main effect of language level was not significant (F<1). All three two-way interactions were significant (environment by language level: F(2,27) = 21.216; cluster type by environment: F(2,27) = 22.269; cluster type by language level: F(2,27) = 25.674; all p’s < .001). The three-way interaction (level by environment by cluster type) was not significant (F<1). Since all the two-way interactions were significant, the data were analyzed further in order to test the effects of level and environment for each cluster type separately.
In the stop-fricative clusters, there were significant main effects of environment: (F(1,27) = 82.553, p < .001) and language level: (F(2,27) = 4.537, p < .05). The two-way interaction (language level by environment) was significant (F(2,27) = 14.948, p < .001), and the data were further analyzed in order to test the effect of environment for each level separately. The C1 voicing proportion was significantly smaller in the assimilating environment than in the fully-voiced environment for each language level (natives: F(1,9) = 82.932, p < .001; advanced: F(1,9) = 9.786, p < .02; intermediate: F(1,9) = 10.507, p < .02).

The significant main effect of environment for each group of subjects indicated that there was more voicing produced by the subjects in all levels in the fully-voiced stops than in assimilated-to-voiceless stops. In other words, all subjects at least partially devoiced the stops in the assimilating environment as required by RVA. The native speakers produced less voicing in assimilated-to-voiceless environment than any of the speakers in either of the remaining two language levels. The significant interaction of level by environment showed that the language levels differed significantly in how strong the effect of devoicing was, with natives devoicing more thoroughly than learners.

In the fricative-stop clusters, there were also significant main effects of environment: (F(1,27) = 258.896, p < .001) and language level: (F(2,27) = 8.718, p < .05). The two-way interaction (language level by environment) was again significant (F(2,27) = 11.194, p < .001), and the data were analyzed further in order to test the effect of environment for each level separately. C1 voicing proportion was significantly shorter in the assimilating condition than in the fully-voiced for each
speaker level (natives: $F(1,9) = 1384.877$; advanced: $F(1,9) = 27.450$, intermediate: $F(1,9) = 64.687$, all $p$’s < .001).

The significant effect of environment for all subjects’ levels indicated that the amount of voicing in the fricatives produced by all subjects in the assimilating environment was significantly less than the amount of voicing in the fully-voiced fricatives. The voicing proportion in the devoiced fricatives appeared to be similar for all groups of subjects; that is, all subjects regardless of their level devoiced fully-voiced fricatives in the assimilating environment in compliance with RVA. Interestingly, both groups of non-native speakers produced less voicing in the fricatives in the fully-voiced environment than native speakers, and the significant interaction of level by environment showed that learners had a smaller effect of environment than natives. However, this was a difference between native and non-native subjects in a non-assimilating (fully-voiced) environment where no application of RVA was required.

2.3.2.2. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environments

The same three-factor interaction comparison as was used for the assimilating vs. fully-voiced environments above was used here to compare the assimilating to fully-voiceless environments (see Figure 2.6.). Both groups of L2 subjects (the leftmost and center bars in the graph) were found to maintain more voicing in C1 stops in the assimilating environment than L1 speakers, represented by a rightmost bar in the graph. However, all groups of subjects performed in a similar manner in the fricative-stop clusters producing equal amount of voicing in the assimilating vs. fully-voiceless environments.
As in the previous analysis, there were significant main effects of environment: (F(1,27) = 99.091, p < .001) and cluster type: (F(1,27) = 120.848, p < .001). The main effect of language level was not significant (F(2,27) = 2.407, p > .05). Unlike the previous analysis, the interaction of environment and language level was not significant (F(2,27) = 2.344, p > .05), although the interactions of environment and cluster type (F(2,27) = 109.231, p < .001) and of language level and cluster type were significant (F(2,27) = 24.186, p < .001). As the three-way interaction was also significant (F(2,27) = 24.119, p < .001), the data were analyzed further in order to test the effects of level and environment for each cluster type separately.
In the stop-fricative clusters, there was a significant main effect of environment ($F(1,27) = 132.358, p < .001$) but no significant main effect of language level ($F(2,27) = 2.512, p > .05$). The two-way interaction (language level by environment) was significant ($F(2,27) = 27.470, p < .001$). The data were analyzed further in order to test the effect of environment for each level separately.

There was no significant effect of voicing environment for native speakers ($F(1,9) = 1.220, p > .05$), which demonstrates that natives make stops equally voiceless for fully-voiceless and assimilated-to-voiceless stops. However, both groups of non-native speakers produced more voicing in stops in the assimilating environment than in stops produced in the voiceless environment, which is different from the voicing pattern obtained from the native speakers (advanced: $F(1,9) = 55.064$; intermediate: $F(1,9) = 103.150; p < .001$).

Similarly, in the fricative-stop clusters, there was a significant main effect of environment, with assimilated-to-voiceless fricatives slightly more voiced than fully-voiceless ones: ($F(1,27) = 6.574, p < .02$). Like in the stop-fricative clusters, there was no main effect of language level: ($F < 1$). The two-way interaction (level by environment) was not significant ($F(2,27) = 2.242, p > .05$). This result could be interpreted as demonstrating that all groups of subjects assimilated C1 fricatives but not enough to render them similar to the fully-voiceless fricatives in terms of voicing proportion. However, in English, voiced fricatives tend to be produced with very little voicing (Docherty 1992; Haggard 1978); therefore, non-native subjects produced even voiced Russian fricatives with less voicing than native speakers and had less voicing to lose when devoicing the fricatives. Thus, although L2 subjects seemed to assimilate
similarly to natives, the obtained result could simply reflect L2 speakers’ general lack of voicing during fricatives rather than their knowledge of the rule of RVA.

2.3.2.3. Summary

The results obtained for C1 voicing proportion demonstrate that non-native speakers maintained more voicing in the stops in the assimilating environment compared to the voicing amount of the stops in fully-voiceless environment. However, there also was a significant effect of voicing environment for assimilating vs. fully-voiced stops for L2 learners, which suggests that non-native subjects devoice stops in the assimilating environment but only partially. In the fricative-stop clusters, both groups of non-native subjects produced less voicing than natives in the fully-voiced environment but performed similarly to natives in the assimilating and fully-voiceless environments. The results of the measurements obtained for C1 (duration and voicing proportion) are briefly summarized in the table below.

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Stop-fricative</th>
<th>Fricative-stop</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 duration</td>
<td>no significant difference among speakers in all environments</td>
<td>L1 &amp; L2: shorter C1 duration in voiced environment</td>
</tr>
<tr>
<td>C1 voicing Proportion</td>
<td>L1: same voicing proportion in assimilated &amp; voiceless environments; L2: more voicing in assimilating environment than in voiceless</td>
<td>L1 &amp; L2: same voicing proportion in assimilated &amp; voiceless environments</td>
</tr>
</tbody>
</table>

Table 2.2. Summary of the C1 measurements (C1 duration & C1 voicing proportion)

2.3.3. Duration of the second consonant in a cluster (C2)

In this section I will analyze duration and proportion of voicing of the second member (C2) in a consonant cluster, using the same statistical design as was used in the previous sections. Since duration and voicing of one segment was shown to affect
duration of an adjacent segment (Crystal & House 1987; 1988), analysis of C2
duration and voicing proportion was necessary for the overall analysis of voicing
assimilation. C2 provides the environment to which the C1 in the assimilating
environment assimilates (cf. Section 1.4.2.). Recall that in the assimilating condition,
the first obstruent of a cluster is underlingly voiced and the second obstruent is
underlyingly voiceless; therefore, C1 assimilates to voiceless C2 and surfaces as
voiceless (e.g., lozhka ‘spoon’ – /loʃkə/). However, no assimilation to the voice
specification of C2 occurs in fully-voiced (both consonants are underlying voiced,
e.g., zazhgi ‘set on fire’ – /zaʒgɨ/ and fully-voiceless (both consonants are underlying
voiceless, e.g., moshka ‘midge’- /moʃkə/) conditions.

Since C2 did not participate in the process of RVA and did not undergo
assimilation itself, its duration and voicing duration were not expected to undergo
large changes. In general, the duration of voiceless obstruents is longer than that of
voiced obstruents, and this was an expected pattern for this measure: in both
assimilating and fully-voiceless environments C2 was fully voiceless and therefore
was expected to be longer than fully-voiced obstruents in the fully-voiced
environment.

2.3.3.1. C2 duration in assimilating-to-voiceless vs. fully-voiced environments

Figure 2.7. below demonstrates the results of the C2 duration analysis. The prediction
for this measure was that C2 in the assimilating environment would be longer than
fully-voiced C2 because voiceless obstruents are inherently longer than voiced ones.
This is indeed what was found, but for C2 fricatives only, shown in the top panel of
the graph.
Figure 2.7. Comparisons of the duration of the second consonant (C2) in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments

An overall three-factor ANOVA with language level, voicing environment and cluster type as variables was performed to compare C2 duration in the assimilating to the fully-voiced environment. (cf. Figure 2.7. for the results.) There were significant main effects of environment: (F(1,27) = 108.810, p < .001) and cluster type: (F(1,27) = 129.678, p < .001). The main effect of language level was not significant (F(2, 27) = 1.005, p > .05). The interaction of cluster type and language level was significant (F(2,27) = 4.199, p < .05), as was the interaction of environment and cluster type (F(2,27) = 29.996, p < .001). The interaction of environment and language level was not significant (F < 1). The three-way interaction (level by environment by cluster type) was also not significant (F<1). Because all factors participated in at least one
significant interaction, a two-way ANOVA with language level and environment as variables was performed for each cluster type (stop-fricative and fricative-stop).

It was predicted that voiceless obstruents in the C2 position (and thus assimilating C2 as well) would be longer than voiced, and that is what was found in the stop-fricative clusters: the main effect of environment was significant (F(1,27) = 182.999, p < .001). This finding verified the direction of voicing assimilation as regressive rather than progressive. If progressive assimilation had taken place in the assimilating environment, a voiceless C2 would be affected by a voiced C1 and would shorten. The fact that C2 duration remained unchanged in this study serves as a proof that in Russian voicing assimilation is indeed regressive (Burton & Robblee, 1997).

The main effect of language level and the two-way interaction (language level by environment) were not significant (F<1), which indicated that all groups of subjects produced C2 fricatives in a similar manner, with voiceless C2 being longer and voiced C2 being shorter.

In the fricative-stop clusters, there were no significant main effects of environment: (F<1) or language level: (F(2,27) = 1.088, p > .05). The two-way interaction (level by environment) was not significant (F(2,27) = 2.080, p > .05) either. The results revealed that there was no or little difference in the duration of the assimilated-to-voiceless vs. fully-voiced C2 stops. The absence of any significant effects in this subset of the data indicates that voicing does not influence duration of a C2 stop (in such clusters) for any language level. This finding contradicted my prediction that fully-voiced obstruents were shorter than fully-voiceless. However, this pattern was in accord with Burton & Robblee’s (1997) study. They found little
difference between voiced and voiceless stop duration and argued that overall stop duration did not provide a cue to voicing in either C1 or C2 position.

2.3.3.2. C2 duration in assimilating-to-voiceless vs. fully-voiceless environments

The same three-factor interaction comparison as was used in the assimilating vs. fully-voiced environments above (language learner, cluster type, voicing environment) was used here to compare C2 duration in the assimilating environment to the fully-voiceless environment. The results of the analysis are presented in Figure 2.8. Assimilating and fully-voiceless C2 stops and fricatives were expected to be similar in duration. This prediction was born out only partially. Only the duration of C2 stops produced by native speakers (the rightmost bar in the graph, bottom panel) in the assimilating and fully-voiceless environments was similar. The duration of C2 fricatives was found to be longer than the learners’ in all conditions.
As in the previous analysis, there were significant main effects of environment: (F(1,27) = 117.973, p < .001) and cluster type: (F(1,27) = 9.287, p < .05), and there was no main effect of language level (F(2,27) = 1.978, p > .05). The two-way interactions of language level and environment, and cluster type and language level were not significant (F<1). The interaction of cluster type and environment was significant (F(2,27) = 23.051, p < .001). The three-way interaction (level by environment by cluster type) was not significant (F(2,27) = 1.712, p > .05). The data
were analyzed further in order to test the effects of level and environment for each cluster type separately.

In the stop-fricative clusters, the main effect of environment was significant (F(1,27) = 35.770, p < .001) with C2 fricatives shorter when preceded by fully-voiceless stops in the voiceless environment than when preceded by assimilated-to-voiceless stops in the assimilating environment, for all language levels. The main effect of language level was not significant (F < 1). The two-way interaction (language level by environment) was not significant (F(2,27) = 1.241, p > .05), which demonstrated that regardless of the language level, all subjects produced shorter C2 fricatives in the fully-voiceless environment than in the assimilating one. Although the native speakers appear to produce longer C2 fricatives than the learners, this effect was not significant.

In the fricative-stop clusters, there was no significant main effect of environment (F<1) but there was a significant main effect of language level (F(2,27) = 7.185, p < .05). As in the stop-fricative clusters, the two-way interaction of language level and environment was not significant (F<1), which demonstrated that in both environments, both groups of L2 learners produced somewhat shorter voiceless C2 stops than native speakers.

2.3.3.3. Summary

The absence of significant main effects of environment in the fricative-stop clusters in all conditions indicated that C2 stop duration was not dependent on the voicing environment in which the cluster occurred but rather on the language level (C2 stops were shorter in both groups of L2 subjects); whereas duration of C2 fricatives was found to be dependent on the environment in which it occurred (but did not interact
with language level). Therefore, my prediction that voiceless C2 stops and fricatives would be longer than voiced was born out only partially. Only fully-voiceless C2 fricatives were found to be longer than fully-voiced ones, indicating that voice contrast was maintained. However, as discussed previously, it is prevoicing, not overall duration that is used to distinguish phonetic voicing in Russian stops (Burton & Robblee 1997); whereas, overall stop duration was shown to be an unreliable and less important cue to voicing.

2.3.4. Voicing proportion of the second consonant in a cluster (C2)

As mentioned earlier, the amount of voicing was measured for C2 in order to determine whether C1 voicing affects C2 voicing. Recall that establishing this fact helps verify the direction of voicing assimilation in a cluster. If, in an assimilating environment, C2 assumes the voice specification of the preceding obstruent, progressive assimilation takes place. If C2 voicing remains unaffected, RVA occurs, which is an expected pattern in this experiment. An analysis of C2 voicing proportion also verified voicing of the conditioning environment for obstruents undergoing assimilation. A small voicing proportion of C2 in the assimilating-to-voiceless environment would confirm that C2 is indeed fully-voiceless and thus conditions voicing neutralization of C1. All groups of subjects were expected to produce voiceless C2 obstruents in the assimilating and fully-voiceless environments and voiced C2 in the fully-voiced environment.

2.3.4.1. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments

Figure 2.9. demonstrates the results of the C2 voicing proportion analysis. Since C2 does not undergo any changes but conditions RVA itself, C2 voicing was expected to
be the shortest or virtually non-existent in the assimilating environment; whereas, voicing duration of fully-voiced C2 was expected to be significantly greater. This prediction was confirmed for both clusters and all groups of subjects (compare the height of the bars in the assimilating vs. fully-voiced environments).

Figure 2.9. Comparisons of voicing proportion of the second consonant (C2) in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments

An overall three-factor ANOVA with language level, voicing environment, and cluster type as variables was performed on C2 voicing proportion (See Figure 2.4 for the results.) There were significant main effects of environment: (F(1,27) = 32.151, p < .001); cluster type: (F(1,27) = 312.461, p < .001) and language level (F(2,27) = 16.526, p < .001). All three two-way interactions were significant (cluster type by
environment: (F(2,27) = 17.420, p < .05); environment by language level: (F(2,27) = 12.776, p < .001); cluster type by language level: (F(2,27) = 20.904, p < .001). The three-way interaction (level by environment by cluster type) was also significant (F(2,27) = 7.538, p < .05), and the data were analyzed further in order to test the effects of level and environment for each cluster type separately.

In the stop-fricative clusters, there were significant main effects of environment (F(1,27) = 557.742, p < .001) and language level (F(2,27) = 9.295, p < .05). The two-way interaction (level by environment) was also significant (F(2,27) = 15.731, p < .001). The data were analyzed further in order to test the effect of environment for each level separately. There was significantly more voicing in fully-voiced fricatives than in assimilated-to-voiceless fricatives for each language level (natives: (F(1,9) = 6062.239; advanced: (F(1,9) = 51.749; intermediate: (F(1,9) = 273.838; all p’s < .001). The significant effect of environment for all language levels showed that all groups of subjects distinguished voiced from voiceless fricatives, correctly producing C2 fricatives with no or very little voicing in the assimilating environment (C1_[+voice] C2_[-voice]) and with significantly more voicing in the fully-voiced environment (C1_[+voice] C2_[+voice]). The significant interaction of level and environment demonstrated that this effect was significantly smaller for the learners than the natives, because although the learners produced voiced fricatives, they did not voice them as much as native speakers. This result can be taken as evidence of L1 phonetic transfer into L2 because previous phonetic research has shown that voiced fricatives in English are generally produced with very little voicing (Dochezty 1992; Haggard 1978), and this is what was found in this experiment in the learners' Russian productions as well.
Since much of the data for this measure had values of ‘0’ (no voicing present) in the fricative-stop clusters, the data was not appropriate for ANOVA because it would violate the assumption of normal distribution of data. Although no statistical analysis was conducted for the fricative-stop clusters, Figure 2.9 clearly demonstrated that there was more voicing in C2 in the fully-voiced environment than in the assimilating environment for all language levels. The test of simple effects of language level demonstrated that native speakers produced significantly more voicing in C2 in the fully-voiced environment than both groups of L2 subjects (F(2,27) = 13.069, p < .001), as also evidenced by the graphs.

2.3.4.2. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environment

The analysis of C2 voicing in the assimilating vs. fully-voiceless environments revealed that all groups of subjects successfully produced C2 stops as completely voiceless. The measurement values were equal to ‘0’ and therefore only C2 fricative voicing values were included in the graph. Figure 2.10 demonstrates that L2 speakers appeared to maintain some voicing in C2 fricatives in both environments. Native speakers, however, did not maintain any voicing at all, as the absence of the ‘native speakers’ bar in the graph demonstrated. Since the data showed extremely small amounts of voicing, if any, for all conditions, no ANOVAs were performed for this measure.
The analysis of C2 voicing proportion comparison showed that L2 learners produced less voicing than native speakers both in C2 fricatives and stops in the fully-voiced environment. However, this finding was not immediately relevant to the study since this was a control condition. More importantly, both groups of L2 speakers, as well as natives, successfully produced fully-voiceless C2 fricatives and stops in both assimilating and voiceless environments, as predicted.

2.3.5. Voice Onset Time (VOT) of the second consonant in a cluster (C2)

In the previous sections, I have analyzed two acoustic parameters of C2, duration and voicing proportion. C2 provides conditioning environment for RVA. Fully-voiceless
C2 affects C1 in the C1 [+voice] C2 [-voice] (assimilating) environment, which results in C1 devoicing. However, C1 does not undergo any changes in the C1 [+voice] C2 [+voice] and C1 [-voice] C2 [-voice] environments.

VOT is another phonetic parameter that serves to implement the voice contrast in stops (Burton & Robblee 1997). In this experiment, C2 VOT was analyzed in order to confirm that voicing assimilation did not influence this part of the signal; i.e., VOT remained negative in fully-voiced and positive in voiceless stops. Since VOT can only be measured for stops before a vowel, an analysis was conducted only for C2 in the fricative-stop clusters.

Russian contrasts prevoiced and fully-voiceless unaspirated stops (Keating 1984; Petrova et al., 2006) in all positions; therefore, all groups of subjects were expected to produce C2 stops in the assimilating and fully-voiceless environments with positive VOT and in the fully-voiced environment with negative VOT.

2.3.5.1. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiced environments

Due to the fact that VOT can only be measured for stops, not fricatives, only the fricative-stop clusters were included in this analysis. Within each environment, little effect of language group is expected. Although this prediction was confirmed in the assimilating environment, L1 and L2 speakers behaved quite differently in the fully-voiced environment, as it is apparent from Figure 2.11. below. In the fully-voiced environment, the largest downward bar represents VOT values for the native speakers (with substantial voicing during closure); the short upward bar represents intermediate L2 speakers (with voicing beginning just slightly after release), and advanced speakers are in between (with some voicing during closure).
A two-factor ANOVA with language level and voicing environment as variables was performed on C2 VOT in the assimilating vs. fully-voiced environments. There were significant main effects of environment: (F(1,27) = 176.298, p < .001) and language level: (F(2,27) = 17.234, p < .05). The two-way interaction (language level by environment) was significant (F(2,27) = 21.783, p < .05), and the data were analyzed further in order to test for effects of environment for each level separately. VOT was significantly greater in assimilating environment than in the fully-voiced environment for each language level (natives: (F(1,9) = 424,575; advanced: (F(1,9) = 18,057; intermediate: (F(1,9) = 23,922, all p’s < .002).

Figure 2.11. VOT of the second consonant (C2) in the fricative-stop clusters in the assimilating vs. fully-voiced environments
The significant effect of environment for all language levels showed that all groups of subjects used VOT to distinguish between voiced and voiceless C2 stops, with greater VOT for voiceless stops, as was appropriate. However, the significant interaction of environment and language level showed that the VOT difference between fully-voiced and fully-voiceless stops was significantly smaller for learners than for natives. That is, unlike native speakers, L2 learners produced Russian stops with less or no prevoicing: mean VOT duration was -90 ms. for natives, -25 ms., for advanced, and 10 ms. for the intermediate L2 speakers.

2.3.5.2. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiceless environments

The same design two-factor ANOVA with language level and environment as variables was performed on C2 VOT in assimilating vs. fully-voiceless environment. Figure 2.11. demonstrates the results of this interaction comparison. There were no significant main effects of environment (F(2,27) = 4.373, p < .05) and language level (F<1). The two-way interaction (level by environment) was not significant (F<1). The absence of significant main effects indicated that C2 VOT was similar for all language levels in both environments. Therefore, an assimilated-to-voiceless vs. fully-voiceless C1 did not influence the C2 VOT values for any speakers, which confirmed once again the regressive direction of voicing assimilation in obstruent clusters in Russian.
Figure 2.12. VOT of the second consonant (C2) in the fricative-stop clusters in the assimilating vs. fully-voiceless environments

2.3.5.3. Summary

In the assimilating vs. fully-voiced environments, both groups of L2 learners failed to lower their VOT far enough for the voiced stops to approximate native VOT values, although they successfully made the voicing distinction. Advanced L2 learners produced C2 stops with negative VOT; albeit still significantly different from native speakers (mean VOT of -25 ms. for advanced L2 speakers vs. -90 ms. for L1 speakers).

The results of the VOT analysis in the assimilating vs. fully-voiceless environments indicated that all groups of subjects produced C2 stops with similar
VOT values. However, no difference was predicted in this condition, and this comparison was used to confirm that all groups of subjects produced voiceless stops with positive VOT, whether in the assimilating or fully-voiceless environment, and thus are capable of making voiced-voiceless distinction in Russian.

2.3.6. Duration of the vowel preceding a cluster (V1)

Recall that universally vowels are somewhat longer before voiced obstruents than before voiceless (Keating 1984). This effect also often holds even if the voicing distinction is neutralized (Slowiaszek & Dinnsen 1985, Warner et al., 2003). Generally, vowels tend to be longer before fully-voiced obstruents and shorter before fully-voiceless obstruents. This was also found for Russian (Chen 1970). Therefore, V1 duration serves as yet another test for verifying voicing distinction in C1 stops and fricatives. If assimilation does not apply, V1 duration in assimilating environment should be similar to V1 duration in the fully-voiced environment. In this set of data, V1 duration was expected to be shorter before both fully-voiceless and assimilated-to-voiceless obstruents. However, it should be noted that although the tokens were controlled for stress and vowel quality, they were not identical. Therefore, it is possible that this difference in vowel quality could have an effect on the measurements, and the prediction could not be confirmed. Furthermore, based on the past results on incomplete neutralization, speakers might produce a longer V1 in assimilated-to-voiceless environment than in fully voiceless environment, as the assimilated environment has an underlyingly voiced consonant, and is thus similar to the incomplete neutralization situations tested in the past. However, incomplete neutralization effects tend to be quite small (Warner et al. 2004), so this effect might not appear in the current slightly less controlled materials.
2.3.6.1. V1 duration in assimilating-to-voiceless vs. fully-voiced environments

Figure 2.12. below demonstrates the results of the V1 duration comparison in the assimilated vs. fully-voiced environments. Recall that in this analysis the V1 duration in the assimilating environment was expected to be shorter than the V1 duration in the fully-voiced environment. This prediction was born out for the vowels preceding the stop-fricative clusters only (top panel). Vowels preceding the fricative-stop clusters, on the other hand, were found to be slightly longer in the assimilating environment. Intermediate speakers produced the longest V1 in all conditions, as demonstrated by the leftmost bar in the graph.

![Figure 2.13. Comparisons of the duration of the vowel preceding a stop-fricative and fricative-stop cluster (V1) in the assimilating vs. fully-voiced environments](image-url)
An overall three-factor ANOVA with language level, environment, and cluster type as variables was performed on V1 duration in the assimilating vs. fully-voiced environments. The main effect of cluster type was significant ($F(1,27)= 83.450, p < .001$). The main effects of language level: ($F(2,27) = 2.593, p > .05$) and environment: ($F<1$) were not significant. Of the three two-way interactions, only the interaction of cluster type and environment was significant: ($F(2,27) = 18.387, p < .001$). The interactions of cluster type and language level and environment and language level were not significant ($F<1$) nor was the three-way interaction ($F< 1$). Because of the significant interaction of cluster type and environment, the data were analyzed further in order to test the effects of language level and environment for each cluster type separately.

In the stop-fricative clusters, the significant main effects of environment ($F(1,27) = 25.236$) and language level ($F(2,27) = 3.524, p < .05$) indicated that V1 duration was longer before a cluster in the voiced environment than in the assimilating environment in all language level groups. This result also indicated that intermediate subjects produced longer preceding vowels in all environments compared to two other groups of subjects. The two-way interaction (language level by environment) was not significant ($F<1$).

In the fricative-stop clusters, there was a significant main effect of environment ($F(1,27) = 4.330, p < .05$), and no significant main effect of language level ($F(2,27) = 1.566, p > .05$). The two-way interaction (language level by environment) was not significant ($F<1$). The significant main effect of environment indicated that V1 duration was somewhat longer before assimilated-to-voiceless obstruents than before fully-voiced obstruents, contrary to the expected direction of
the effect. Since previous research has shown that vowels are shorter before voiceless obstruents, V1 was expected to be shorter preceding an assimilated obstruent in this analysis as well. One of the explanations for such an unexpected result could be the fact that although some of the stimuli were matched for stress and vowel quality, whenever possible, they were not minimal triplets. Therefore, any differences in vowel quality itself or any differences in acoustic properties of the segments in general could have affected V1 duration (cf. Section 2.1.2.).

2.3.6.2. V1 duration in assimilating-to-voiceless vs. fully-voiceless environments

In this analysis, V1 duration was predicted to be similar in both assimilating and fully-voiceless environments, although if incomplete neutralization has an effect, it would result in V1 duration being longer in assimilating environment (underlyingly voiced) than in fully-voiceless environment. Any such effect would be small at most, however. Similar to the previous analysis, this prediction was not born out consistently. V1 was found to be shorter before assimilating C1 stops and longer before fully-voiceless C1 fricatives. Figure 2.14. below demonstrates the results of this interaction comparison. Interestingly, similar to the previous analysis, intermediate speakers produced the longest V1 in all conditions (the leftmost bar in the graph).
Figure 2.14. Comparisons of the duration of the vowel preceding a stop-fricative and fricative-stop cluster (V1) in the assimilating vs. fully-voiceless environments

The same three-factor comparison as was used in the assimilating vs. fully-voiced environments above (language learner, cluster type, voicing environment) was used here to compare V1 duration in the assimilating vs. fully-voiceless environments. The main effect of cluster type was significant ($F(1,27) = 80.572, p < .001$); whereas main effects of environment ($F<1$) and language level were not significant either ($F(2,27) = 2.016, p > .05$). Of the three two-way interactions tested, only one was significant (cluster type by environment: $F(2,27) = 30.991, p < .001$); the other two were not significant (environment by language level and cluster type by language level: $F<1$). The three-way interaction (level by environment by cluster type) was not significant either ($F<1$). However, because the interaction of cluster type and environment was
significant, the data were analyzed further in order to test the effects of level and environment for each cluster type separately.

In the stop-fricative clusters, the significant main effect of environment (F(1,27) = 18.771, p < .001) and no significant main effect of language level (F(2,27) = 2.665, p > .05) indicated that contrary to the expected direction, V1’s were longer before a fully-voiceless cluster than in the assimilating environment in all language level groups. The two-way interaction (language level by environment) was not significant (F<1).

In the fricative-stop clusters, the two-way interaction (language level by environment) was not significant (F(2,27) = 1.654, p > .05). A significant main effect of environment (F(1,27) = 18.071, p < .001) and no significant main effect of language level (F(2,27) = 1.155, p > .05) indicated that for all language levels V1 duration was longer in the assimilating environment than in the fully-voiceless environment. This finding could be taken as evidence for incomplete voicing neutralization in the clusters for this measure, with C1 retaining some voicing, which affected V1 duration (Port & O’Dell 1985; Port & Crawford 1989; Slowiaczek & Dinnsen 1985). However, the analyses of C1 duration and voicing proportion, at least for native speakers, supported the process of voicing assimilation. Also, as discussed in Section 2.2.6.2, the stimuli were not minimal triplets, and phonetic differences between the stimuli could be responsible for the obtained result.

2.3.6.3. Summary

The significant main effect of environment in all conditions and all language level groups indicated that V1 duration differed depending on voicing environment. Based on the vowel duration results reported in other studies (Barry 1991; Chen 1970), it
was predicted that vowels were longer in the fully-voiced environment than in the assimilating environment. However, this was found in the stop-fricative clusters only. Contrary to the expected direction of effect, vowels preceding the stop-fricative clusters were found to be longer in the fully-voiceless environment than in the assimilating environment. In the fricative-stop clusters, the shortest V1 was produced in the voiceless environment, as was appropriate. However, V1’s were slightly longer in the assimilating environment than in the fully-voiceless, which can be taken as evidence of incomplete obstruent neutralization (to be discussed in Section 2.4.).

Overall, in this analysis, only two of the predictions were born out: longer V1 was found in the fully-voiced vs. assimilating environments in the stop-fricative clusters, and shortest V1 was found in the voiceless environment in the fricative-stop clusters. As previously discussed, the tokens used in the experiment were not minimal triplets, and inherent differences in vowel quality and other acoustic properties could be responsible for the obtained results.

2.3.7. Duration of the vowel following a cluster (V2)

In addition to duration of a vowel preceding a cluster, duration of a vowel following the cluster was also measured. Previous research has demonstrated that V1 duration is dependent on the voicing specification of the following obstruent. However, little research has been performed on vowels following an obstruent cluster (V2) but a few studies have shown that voiceless obstruents decrease duration of vowels following them (Campbell 1991; Allen & Miller 1999). In this section, I present an analysis of V2, which was included with the purpose of determining whether C2 voicing distinction had any effect on V2 duration at all. Overall, V2 was expected to be longer after fully-voiced obstruents and shorter after fully-voiceless and assimilated.
2.3.7.1. V2 duration in assimilating-to-voiceless vs. fully-voiced environments

The results of the V2 duration interaction comparison are shown in Figure 2.15.7. All groups of subjects produced slightly longer V2 in the fully-voiced environment than in the assimilating environment, as expected.

![Figure 2.15](image)

Figure 2.15. Comparisons of duration of the vowel following a stop-fricative and fricative-stop cluster (V2) in the assimilating vs. fully-voiced environments

The standard three-factor interaction comparison with language level, environment, and cluster type as variables was performed on V2 duration. The main effect of environment was marginally significant (F(1,27) = 3.944, p = .056); whereas, the main effect of language level was not significant (F<1). The main effect of cluster type was significant, with vowels being longer after the fricative-stop clusters (F(1,27)= 35.537, p < .001). None of the two-way interactions were significant.
(cluster type by environment: (F<1); environment by language level: (F<1); cluster type by language level: F(2,27) = 1.283, p > .05). The three-way interaction (level by environment by cluster type) was not significant (F<1) either. Since none of the interactions was significant, no further analysis was conducted.

2.3.7.2. V2 duration in assimilating-to-voiceless vs. fully-voiceless environments

In this analysis, the V2 duration was predicted to be similar in both environments. However, contrary to the prediction, all groups of subjects were found to produce longer V2 in the assimilating environment. (See Figure 2.16. for the results).

![Figure 2.16. Comparisons of duration of the vowel following a stop-fricative and fricative-stop cluster (V2) in the assimilating vs. fully-voiceless environments](image)

The three-factor interaction comparison with language learner, cluster type, voicing environment as factors, was used here to compare V2 duration in the assimilating vs.
voiceless environments. As in the previous analysis, there was no significant main 
effect of language level (F<1). The significant main effect of cluster type (F(1,27) = 
43.312, p< .001) and the marginal effect of environment (F(1,27) = 4.174, p = 0.51) 
indicated that all groups of subjects produced longer V2 after the fricative-stop 
clusters and somewhat longer V2 in the assimilating environment than in the fully-
voiceless environment in both clusters. None of the interactions (two-way: cluster 
type by environment; environment by language level; cluster type by language level; 
three-way: level by environment by cluster type) was significant (F<1), and no further 
analysis was conducted.

2.3.7.3. Summary

The absence of a significant main effect of language level suggested that the post-
cluster vowel had similar duration for all groups of subjects in each relevant 
environment. The significant main effect of cluster type and marginal main effect of 
environment indicated that in all language level groups, V2 duration was dependent 
on both the cluster type after which it occurred and voicing environment. More 
specifically, V2 was longer after the fricative-stop clusters than after the stop-fricative 
clusters in all environments, and V2 was the longest after voiced C2 and shortest after 
voiceless C2, as predicted. The lack of significant interactions with language level 
indicates that learners are not in the process of acquiring a distinction natives have, 
but rather that the effects that are significant are properties of the words used or of the 
inherent durations of segments, not a language-specific pattern to be acquired.
2.4. Discussion

As discussed in the literature (Burton & Robblee 1997; Keating 1984), in Slavic languages, duration of voicing, not the segment duration, serves as the most important cue to voicing contrast. The difference in the RVA production between native and non-native speakers in Experiment One was most clearly demonstrated precisely on this measure: C1 voicing proportion. In the C1 voicing proportion analysis, L1 speakers were found to produce the same amount of voicing in both C1 fricatives and stops respectively in the assimilating and fully-voiceless environments. L2 speakers performed similarly to natives producing very little voicing, if any, in assimilated and fully-voiceless C1 fricatives but they maintained significantly more voicing in C1 stops in the assimilating environment compared to the fully-voiceless environment. One can argue that non-natives behaved similarly to natives on this measure because they have acquired RVA and successfully applied it to practice. However, L2 subjects did not exhibit a similar pattern for the stops in the assimilating environment. They produced less voicing in the assimilated stops than in fully-voiced ones but more than in fully-voiceless, which indicated gradual acquisition of RVA in stop-fricative clusters. Overall, it could be speculated that Russian stops are generally more difficult for English speakers to acquire and produce because unlike English, Russian voiceless stops are not aspirated, and voiced stops are prevoiced (Jessen 2001; Petrova et al., 2006).

The finding that L2 speakers produced assimilated and fully-voiceless fricatives similarly to L1 speakers could be accounted for by acoustically-motivated factors. For example, since English speakers normally do not completely voice even fully-voiced fricatives (Haggard 1978; Stevens at al., 1992), it is possible that they
simply transferred an L1 phonetic pattern into the L2. As a result, they produced Russian fully-voiced fricatives with much less voicing than Russian natives (cf. Figure 2.5.), and had less voicing to lose in the assimilating environment to begin with. Therefore, although L2 subjects were found to treat the assimilated-to-voiceless fricatives as fully-voiceless in this measure, this could partly reflect their general lack of voicing during fricatives. Overall, since voiced English fricatives tend to contain less voicing than Russian voiced fricatives, the results of this analysis could simply indicate that L2 learners have not yet acquired the pronunciation norms of target Russian fricatives, which require more voicing than English ones. As a result, Russian fricatives were produced with English-like voicing, which nevertheless resulted in the targeted pronunciation in the assimilating environment.

The analysis of C2 stop and fricative voicing proportion supported the latter claim. Although L2 speakers successfully produced voiceless C2 in the assimilating and fully-voiceless environments, they were found to produce fully-voiced C2 with significantly less voicing than native speakers, which could be taken as an indication of L1 phonetic transfer into L2. Interestingly, advanced learners produced more voicing in C2 than intermediate learners. This observation could serve as evidence of gradual acquisition, with more proficient learners, who have received more exposure to the language, performing better than less proficient learners.

An L1 phonetic effect on L2 speech was also observed in the analysis of C2 VOT. Although L2 speakers produced stops in the assimilating and fully-voiceless environments with the VOT values similar to the L1 values, L2 voiced stops were not nearly as prevoiced as L1 voiced stops. Previous research has shown that English speakers do not generally prevoice voiced stops (Keating 1984), and this is what was
found in this experiment as well. The analysis of C2 VOT demonstrated that intermediate speakers produced L2 voiced stops virtually without any prevoicing, which could suggest that this group of learners applied English short-lag VOT values when producing Russian stops. However, advanced L2 speakers produced voiced obstruents with some prevoicing, albeit with less prevoicing than in natives. Although this finding was not immediately relevant to the process of RVA and was simply used to confirm voice specification of C2 in a cluster, it pointed to L2 speakers’ gradual approximation of the target VOT values in voiced stops.

The measurements of other acoustic correlates of RVA did not demonstrate any significant differences between native and non-native speakers. The analysis of the segment (C1 and C2) duration revealed that all groups of subjects produced C1 and C2 stops with the same duration in all the three environments. This finding supports Burton & Robblee’s (1997) claim that overall stop duration did not serve as a cue to voicing in Russian; otherwise, C1 and C2 stops were longer in the fully-voiced environment. C1 and C2 fully-voiced fricatives were shown to be shorter than assimilated and fully-voiceless, in all groups of subjects, which confirmed the cross-linguistic pattern of voiced obstruents being longer than voiced. The duration of fully-voiceless and assimilating C1 fricatives was the same in the native and advanced subjects. Intermediate subjects, on the other hand, produced assimilated C1 fricatives shorter than fully-voiceless, which indicated that this group of subjects retained more voicing in the assimilating fricatives.

As discussed in Section 2.3.6, vowels preceding voiced obstruents were expected to be longer than when preceding voiceless obstruents. However, this cross-linguistic pattern for V1 was confirmed only partially. In the fricative-stop clusters,
vowels in the assimilating environment were found to be slightly longer than in the voiceless environments, which could be taken as evidence of incomplete voice neutralization. Voicing neutralization processes eliminate surface voice contrasts and render underlying different obstruents similar in terms of their voice specifications. If neutralization is complete, all the phonetic properties, including vowel duration, employed for the implementation of the voice contrast are expected to be similar in assimilating and fully-voiceless environments (Slowiaszek & Dinnsen 1985, Warner et al., 2003). If neutralization is incomplete, assimilated obstruents differ from truly voiceless ones in terms of one or more acoustic properties. In this experiment, V1 being longer in the assimilating than in the voiceless environment could suggest that fully-voiced fricatives were not completely devoiced, and, therefore, affected the vowel duration. However, the results of the C1 voicing duration analysis in the fricative-stop clusters demonstrated that fricatives were almost completely devoiced in all subjects in the assimilating environment; therefore, the process of incomplete neutralization is not supported by the results of other analyses.

Also, contrary to the expected pattern, in the stop-fricative clusters, V1 was the longest in the fully-voiceless environment. This finding did not support the presence of incomplete neutralization in the stop-fricative clusters in all groups of subjects. If incomplete neutralization had occurred in this condition, V1 would have been longer before assimilated obstruents than before fully-voiceless ones, as the V1 duration in the fricative-stop clusters demonstrated. Overall, although these findings did not provide evidence against incomplete neutralization in this condition (because voicing environments are a between-item condition and the stimuli were not minimal pairs), they did not support it either. Since the stimuli were not minimal triplets, it is
not unreasonable to assume that the differences in V1 quality, as well as suprasegmental differences, might be responsible for such a result.

The results of the V2 analysis did not parallel the results of the V1 analysis. However, V2 results were in accord with Allen & Miller’s (1999) claim that V2 are longer after voiced obstruents and shorter after voiceless ones. In this experiment, it was found that in all conditions, the V2 duration was the longest in the fully-voiced and shortest in fully-voiceless environments, with the V2 duration in the assimilating environment being intermediate between the two. Vowels following fully-voiceless obstruents in the assimilating and fully-voiceless environments were expected to be similar in duration. However, this is not what was found. Like in the analysis of V1 duration, the absence of an expected could perhaps be accounted for by the difference in vowel quality and suprasegmental properties of the stimuli because they were not minimal triplets.

2.4.1. Conclusion

In this chapter I have presented a real-word production experiment and analyzed the following acoustic correlates of RVA: C1 and C2 duration, C1 and C2 voicing proportion, C2 VOT, and V1 and V2 duration. C1 voicing proportion analysis demonstrated an unambiguous difference in the RVA production between native and non-native subjects. The original hypothesis that L2 speakers applied RVA to a lesser degree than L1 speakers was thus confirmed, at least on this measure.

On other measures, no such clear difference between the L1 and L2 speakers’ production was found. All groups exhibited much variability, and disparity among the groups of subjects led to inconclusive results. In the following chapter, I turn to the analysis of the nonsense-word experiment.
3.0. Introduction

In Chapter Two I presented the real-word experiment, in which native and non-native speakers of Russian were tested on their production of regressive voicing assimilation (RVA) in obstruent clusters in real words. In this chapter, I turn to the analysis of the nonsense stimuli. Nonsense stimuli were included for two reasons. The first is that it was necessary to verify the results of the real-word experiment with a larger number of items, because the number of attested words containing word-internal voiced-voiceless obstruent clusters in Russian is limited. The second reason nonsense stimuli were included was to test whether both native and non-native speakers produced RVA in obstruent clusters in such words at all. It could be argued that both native and non-native subjects produced RVA in real words solely due to their familiarity with the words. Using non-words ensured the absence of lexical knowledge on the part of the experiment participants, and tested whether they were capable of generalizing the RVA production pattern to ‘unfamiliar’ words. Also, due to the nature of nonsense words, it was possible to have minimal triplets which differed only in the conditioning environment.

Like in the real-word experiment presented in Chapter Three, the goal of the nonsense-word experiment was to determine whether all groups of speakers produced voiced consonants in the assimilating environment as similar to the fully-voiceless consonants (as appropriate), similar to the fully-voiced consonants (not assimilated at all), or only partially assimilated.
3.1. Methods

3.1.1. Subjects

The same subjects used in the real-word experiment participated in the nonsense-word experiment. The subjects were recorded during one recording session with a break in between the two experiments (with the real words recorded first). As above, there were two groups of L2 learners, intermediate (N=10) and advanced (N=10), and a control group of monolingual Russian speakers (N=10).

3.1.2. Materials

The stimulus list originally consisted of 105 nonsense words; however, seven sets of items were discarded, as it was discovered post-experiment that the words in the assimilating position in these sets resembled existing Russian words (nouns and proper names). The final number of nonsense items was, therefore, 81.

The words represented the same obstruent cluster types (1) stop-fricative, as in kabsa - /kΛpsa/ and (2) fricative-stop, as in kazta - /kΛsta/. The stimuli were assigned to the same conditions as in the real-word experiment (see Appendix B for a complete list of nonsense stimuli). The stop-fricative clusters included the same combinations as attested in real words: bs, bsh, ds, and dsh. Each cluster was represented in four or five words in three voicing environments: an assimilating environment where the cluster was voiced-voiceless and was expected to undergo RVA (kabsa - /kΛpsa/); a control environment where both members of the cluster were fully-voiced (kabza - /kΛbza/); and a second control environment where both obstruents in the cluster were fully-voiceless (kapsa- /kΛpsa). Initially, there were 57 items (9 triplets) total; however, one triplet was discarded because post-experiment it was discovered that items sounded similar to a proper name. Therefore, the total number of tokens in this
condition was 54 (18 triplets). (See Table 3.1. for examples, and Appendix A for the complete list of nonsense stimuli).

The fricative-stop condition included the same obstruent combinations as attested in real words: zk, zhk, and zt. The zk and zt clusters were represented in four tokens in each of the three voicing environments: an assimilating environment where the cluster was voiced-voiceless and was expected to undergo RVA (e.g., kazta /kΛsta/; a control environment where both members of the cluster were fully-voiced (kazda /kΛzda/); and a second control environment where both obstruents in the cluster were fully-voiceless (kasta /kΛsta). The sequence zhk was represented in one set of tokens only. The original number of tokens was 39 (11 triplets); however, due to the same reason as above, 4 triplets were discarded post-experiment. The total number of tokens in this condition was thus 27 (9 triplets). There were fewer items in this condition because the number of possible nonsense words containing the fricative-stop combination in question which were not reminiscent of existing Russian words was limited.

Table 3.1. below provides examples of the nonsense stimuli in the stop-fricative and fricative-stop clusters and in the three environments.

<table>
<thead>
<tr>
<th>Environment</th>
<th>CC type</th>
<th>Environment</th>
<th>CC type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop-Fricative</td>
<td>padsa-/pAtsa/</td>
<td>padza-/pΛdza/</td>
<td>patsa-/pΛtsa/</td>
</tr>
<tr>
<td>Fricative-Stop</td>
<td>kazta-/kΛsta/</td>
<td>kazda-/kΛzda/</td>
<td>kasta-/kΛsta/</td>
</tr>
</tbody>
</table>

Table 3.1. Sample nonsense items in the three voicing environments: assimilating-to-voiceless, fully-voiced, and fully-voiceless

The nonsense stimuli, although conforming to Russian phonotactic rules, were devised in such a manner as to avoid any influence of lexical knowledge on the part of the participants. All the nonsense stimuli were bisyllabic non-words consisting of
both a closed and an open syllable. The items were controlled for stress and vowel quality: obstruent clusters were placed in the pre-stressed position. Unlike real-word stimuli, all nonsense stimuli were minimal triplets. In all nonsense items, the final syllable was stressed with a low-back vowel [a] in the stressed syllable and a reduced vowel [A] in the pre-stress syllable.

3.1.3. Procedure

The same procedure employed in the real-word experiment was used in this experiment. Both real and nonsense words were recorded during one recording session, with a break in between. The real stimuli were recorded first. Subjects were familiarized with the nonsense words prior to the whole recording. In order to demonstrate how to read nonsense words, the subjects were given a number of practice nonsense words that did not contain obstruent clusters (e.g., pulok). The subjects were informed that the stimuli were not real Russian words, but were instructed to read them as real words. All groups of subjects, but especially native speakers who were generally older than non-native participants and less used to performing metalinguistic tasks than university students, seemed to be slightly confused and very much aware of the fact that the items were made-up words. This was manifested by the subjects’ hesitation while reading and exaggerated enunciation of the stimuli. Intermediate subjects were less hesitant about reading the nonsense stimuli, perhaps due to their limited L2 vocabulary, and treated the tokens as real but unfamiliar Russian words.
3.2. Analysis

3.2.1. Measurements

The methodology and measurements used in the real-word experiment were used here (cf. Section 2.2.1. for details). All the measurements were performed according to the same criteria used in the real-word experiment. The measurements included: C1 and C2 overall segment duration, C1 and C2 voicing duration, and C2 VOT.

3.3. Results

In order to determine how both native and non-native speakers realized RVA in obstruent clusters, and, more specifically, how they realized fully-voiced obstruents in the assimilating environment, the data were analyzed by comparing the assimilating to fully-voiceless environments, and the assimilating to fully-voiced environments, as was done for real words above. This experiment drew upon the same design, 2x2x3 ANOVA, as was used in the real-word experiment. The within-subjects factors were cluster type, in which a consonant cluster occurred (stop-fricative and fricative-stop), and language level (native, non-native advanced, and non-native intermediate). The third factor, with two levels, was the voicing environment. For each measure, with the exception of VOT, two interaction comparisons were performed. Since VOT could only be measured in the fricative-stop clusters, only one interaction comparison with language level and voicing environment as factors was performed. The same procedure was utilized as a standard throughout the analysis.

As in the real-word experiment, Russian native speakers served as controls, and were expected to consistently assimilate fully-voiced consonants when followed by a fully-voiceless obstruent. Therefore, native speakers were expected to produce
such a fully-voiced obstruent similar to a fully-voiceless obstruent in terms of segment and voicing duration. Both groups of non-native speakers, on the other hand, were predicted to devoice fully-voiced obstruents to a lesser degree than native speakers, thus producing such obstruents similar to fully-voiced obstruents in terms of segment and voicing duration. I will now examine the results for each acoustic measure in turn.

3.3.1. Duration of the first consonant in the cluster (C1)

I will first analyze the duration of the first obstruent in a cluster, a stop in the stop-fricative clusters, and a fricative in the fricative-stop clusters. L1 subjects were expected to devoice fully-voiced consonants in the voiced-voiceless clusters, and to produce them with duration similar to fully-voiceless consonants. Voiceless consonants are usually longer than voiced, and both L1 fully-voiceless and devoiced consonants were expected to be longer than their fully-voiced counterparts. L2 subjects, on the other hand, were predicted to devoice fully-voiced consonants to a lesser degree than natives. L2 consonants in the assimilating environment were predicted to be shorter than fully-voiceless. I will analyze the C1 duration in the assimilating vs. fully-voiced environments first, and will then present an analysis of the C1 duration in the assimilating vs. fully-voiceless environments.

3.3.1.1. C1 duration in assimilating-to-voiceless vs. fully-voiced environments

An interaction comparison with voicing environment (assimilating and fully-voiced), cluster type (stop-fricative and fricative-stop), and language level (native, advanced, and intermediate) was performed in order to determine whether or not obstruent clusters in the assimilating environment were produced similarly to the clusters in the fully-voiced environment. Figure 3.1. below shows the results of the analysis of the
C1 duration in the assimilating vs. fully-voiced environments. Like in Experiment Two in Chapter Three, C1 fricative duration was found to be longer than C1 stop duration, with natives producing the longest fricatives and stops in all environments.

![Figure 3.1 Comparisons of the duration of the first consonant (C1) in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments](image)

The main effects of environment: (F(1,27) = 44.839, p < .001), cluster type: (F(1,27) = 32.693, p < .001), and language level: (F(2,27) = 3.660, p < .05) were significant. All the two-way interactions (language level by cluster type: (F(2,27) = 4.608, p < .05); language level by environment: (F(2,27) = 3.584, p < .05); and environment by cluster type: (F(2,27) = 55.001, p < .001) were significant. The three-way interaction (environment by cluster type by language level) was not significant (F(2,27) = 2.186,
Due to the significant interactions, a two-way ANOVA with learner level and environment was performed for each cluster type. In the stop-fricative clusters, there were no significant effects (environment: \(F<1\); language level: \(F(2,27) = 1.933, p > .05\); interaction: \(F(2,27) = 1.463, p > .05\)). This indicated that C1 stop duration in the stop-fricative clusters was the same for all voicing environments and language levels, even though assimilated-to-voiceless and fully-voiced clusters were being compared. When a stop was the first consonant of the cluster, its duration was the same for these cluster types.

In the fricative-stop clusters, there were significant effects of environment: \((F(1,27) = 55.943, p < .001)\) and language level: \((F(2,27) = 4.581, p < .02)\). The two-way interaction (language level by environment) was also significant \((F(2,27) = 4.461, p < .03)\), and the data were further analyzed in order to test the effect of environment for each level separately. The C1 fricative duration was significantly shorter in fully-voiced clusters than in assimilated-to-voiceless clusters for native and advanced subjects only \((natives: F(1,9) = 55.467, p < .001);\) advanced: \((F(1,9) = 17.649, p < .01)\); intermediate: \((F(1,9) = 4.472, p > .05)\). Thus, by lengthening their C1 fricatives in the assimilating environment, native speakers and advanced learners showed an effect of assimilation that intermediate learners did not. However, the intermediate learners did show a non-significant trend in the same direction.

### 3.3.1.2. Assimilating-to-voiceless vs. fully-voiceless environments

The same three-factor interaction comparison as was utilized for the assimilating vs. fully-voiced environments above \((learner\ level,\ cluster\ type,\ and\ voicing)\).

---

7 As explained in Chapter 2, the data were split across the cluster type rather than being collapsed over the factor because this procedure allowed analyzing the production of obstruent clusters in the two types of clusters individually.
environment) was used here to compare the assimilating-to-voiceless environment to the fully-voiceless environment. These data appear below in Figure 3.2. Similar to the assimilating vs. fully-voiced environment analysis, natives produced the longest C1 stops and fricatives in all environments.

![Figure 3.2. Comparisons of the duration of the first consonant (C1) that is expected to assimilate in the stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments](image)

There were significant main effects of voicing environment: (F(1,27) = 105.225, p < .001), cluster type: (F(1,27) = 9.248, p < .01), and language level: (F(1,27) = 4.326, p < .03). The two-way interaction of language level by environment was also significant (F(2,27) = 4.246, p < .03). The interactions of cluster type and language level (F(2,27) = 2.364, p > .05) and of environment and cluster type (F(2,27) = 2.631, p > .05) were
not significant. The three-way interaction (level by environment by cluster type) was not significant either (F<1). Because of the significance of a two-way interaction, the data were analyzed further in order to test the effects of language level and environment for each cluster separately.

There were no significant effects of environment in the stop-fricative clusters: (F(1,27) = 3.511, p > .05) or learner level: (F(2,27) = 2.097, p > .05). These results indicated that the C1 stop duration in stop-fricative clusters was the same for all environments and language levels, even though the figure indicated that the advanced learners produced slightly (but not significantly) shorter C1 stops. The two-way interaction (level by environment) was not significant either (F<1), and no further analysis was performed.

The same design two-way ANOVA with level and environment as independent variables was performed for the fricative-stop clusters. Fricatives in the fully-voiceless environment were slightly longer than fricatives in the assimilating environment (F(1,27) = 8.198, p < .01). The significant main effect of language level (F(2,27) = 5.437, p < .01) showed that natives produced longer fricatives than the two groups of learners, both in the assimilating and fully-voiceless environments. The two-way interaction (level by environment) was not significant (F(2,27) = 2.155, p > .05), and no further analysis was performed.

3.3.1.3. Summary
The absence of significant two-way interactions in the stop-fricative clusters in the assimilating vs. fully-voiceless environments, and in the assimilating vs. fully-voiced environments, suggested that all groups of subjects produced similar C1 stop duration.
This finding showed that, like in the real-word experiment, the duration of C1 stops was not affected by the process of voicing assimilation.

The significant main effect of environment and language level in the assimilating vs. fully-voiced environments indicated that assimilated C1 fricatives produced by native and advanced L2 subjects were longer than fully-voiced fricatives, which suggests RVA. Russian L1 voiced fricatives were found to be longer than those produced by advanced L2 speakers; whereas the duration of C1 fricatives produced by intermediate L2 subjects in the assimilating vs. fully-voiced environments was not statistically different, suggesting that they produced assimilation less successfully, if at all.

3.3.2. C1 voicing proportion

In this section, I present an analysis of voicing proportion of C1, which was used to measure the amount of voicing present, or lack thereof, in C1 stops and fricatives. The prediction was that as in the real-word experiment, L1 speakers would assimilate fully-voiced obstruents in the assimilating environment more thoroughly than non-natives, thus producing C1 assimilated obstruents resembling the fully-voiceless ones in terms of voicing duration. Non-natives, on the other hand, were predicted to differ from each other in their production of C1. Advanced L2 learners were expected to assimilate C1 obstruents to a greater extent than intermediate learners. However, all groups of subjects were expected to produce fully-voiced C1 stops and fricatives equally well.
3.3.2.1. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments

The same three-factor interaction comparison used to analyze the C1 duration was also performed on C1 mean voicing proportion. Figure 3.3 below demonstrates the results of the interaction comparisons for this measure in the assimilating vs. fully-voiced environment. As predicted, L2 subjects maintained more voicing in C1 in the assimilating environment than L1 subjects. However, this prediction was borne out for C1 stops only. As Graph 3.3 demonstrates, all groups of subjects devoiced C1 fricatives equally well (see the bottom panel). Furthermore, all groups of subjects produced fully-voiced C1 obstruents as such, albeit native C1 fully-voiced obstruents contained more voicing in both types of clusters (rightmost bars in the graph).

![Figure 3.3. Comparisons of voicing proportion of the first consonant (C1) that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments](image-url)
The results of this analysis were as follows. There were significant main effects of environment: \( (F(1,27) = 123.918, p < .001) \) and cluster type: \( (F(1,27) = 481.884, p < .001) \); whereas the main effect of language level was not significant \( (F<1) \). All the two-way interactions were significant: environment and language level: \( (F(2,27) = 6.563, p < .01) \); environment and cluster type: \( (F(2,27) = 19.176, p < .001) \); language level and cluster type: \( (F(2,27) = 25.777, p < .001) \). The three-way interaction (level by environment by cluster type) was not significant \( (F<1) \). Since all the two-way interactions were significant, the data were analyzed further in order to test the effects of level and environment for each cluster type separately.

In the stop-fricative clusters, there was a significant main effect of environment: \( (F(1,27) = 114.193, p < .001) \), and no significant main effect of language level: \( (F(2,27) = 2.512, p > .05) \). The two-way interaction (language level by environment) was significant \( (F(2,27) = 15.818, p < .001) \), and the data were further analyzed in order to test for effects of environment for each level separately. The C1 voicing proportion was significantly larger in the fully-voiced environment than in assimilating environment for each language level (natives: \( F(1,9) = 109.932, p < .001 \); advanced: \( F(1,9) = 26.860, p < .01 \); intermediate: \( F(1,9) = 13.028, p < .01 \)). Although all groups of subjects devoiced obstruents in the assimilating environment, native speakers did so to a greater degree than each group of L2 learners.

There was a significant main effect of environment in the fricative-stop clusters: \( (F(1,27) = 219.526, p < .001) \), and no main effect of language level: \( (F(2,27) = 1.268, p > .05) \). Since the two-way interaction (language level by environment) was significant \( (F(2,27) = 5.015, p < .02) \), the data were further analyzed to test for effects
of environment for each level separately. All groups of subjects produced more voicing in C1 fricatives in the fully-voiced environment than in the assimilating environment (natives: $F(1,9) = 519.846, p < .001$; advanced: $F(1,9) = 31.145, p < .01$; intermediate: $F(1,9) = 47.462, p < .001$), with native speakers producing more voicing in fully-voiced C1 than each group of learners.

3.3.2.2. C1 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environments

The same three-factor interaction comparison with language level, environment and cluster type as variables as was used for the assimilating vs. fully-voiceless environments above, was used here to compare the assimilating to fully-voiced environment. These data appear in Figure 3.4. below.

Native assimilated C1 obstruents were expected to be similar to the fully-voiceless ones; yet, this prediction was borne out for C1 fricatives only (see bottom panel of the graph). Voicing duration of C1 stops produced by natives in the assimilating environment was significantly greater than in the fully-voiceless environment. It was also predicted that L2 subjects would produce more voicing in the assimilating environment than in the fully-voiceless one, which was indeed the case, but for C1 stops only.
Figure 3.4. Comparisons of voicing proportion of the first consonant (C1) that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments

There were significant main effects of environment: (F(1,27) = 18.844, p < .001) and cluster type: (F(1,27) = 26.312, p < .001); whereas the main effect of language level was not significant (F<1). All three two-way interactions were significant (environment by language level: F(2,27) = 9.191; cluster type by environment: F(2,27) = 263.699; cluster type by language level: F(2,27) = 27.066; all p’s < .001). The three-way interaction (level by environment by cluster type) was not significant (F<1). Because all the two-way interactions were significant, the data were analyzed further in order to test the effects of level and environment for each cluster type separately. As above, two-way ANOVA with level and environment as the
independent variables was performed for each cluster (stop-fricative and fricative-stop) separately.

In the stop-fricative clusters, there were significant main effects of environment: (F(1,27) = 185.390, p < .001) and language level: (F(2,27) = 3.871, p < .05). The two-way interaction (language level by environment) was also significant (F(2,27) = 10.279, p < .001). The data were further analyzed in order to test the effect of environment for each language group separately. The C1 voicing proportion was significantly larger in the assimilating environment than in the fully-voiceless environment for each language level (natives: F(1,9) = 25.232, p < .05; advanced: F(1,9) = 85.764, p < .001; intermediate: F(1,9) = 81.731, p < .001). The significant simple effect of environment for each group of subjects indicated that there was more voicing produced by subjects of all levels in the assimilated-to-voiceless stops than in the fully-voiceless stops. The significant interaction of language level by environment showed that the subjects differ significantly in how strong the effect of devoicing was, with natives devoicing to a greater extent than either group of learners.

In the fricative-stop clusters, there was a significant main effect of environment for all subjects’ levels (F(1,27) = 47.278, p < .001). This indicated that the assimilated-to-voiceless C1 fricatives had significantly more voicing than the fully-voiceless ones, even though they were still less than 20% voiced. There was no significant main effect of language level: (F(2,27) = 1.044, p > .05). The two-way interaction (language level by environment) was not significant (F(2,27) = 1.163, p > .05), and the data were not analyzed further. Thus, the pattern described here applied to both learner groups, as well as natives.
3.3.2.3. Summary

The results obtained for the C1 voicing proportion analysis demonstrated that all groups of speakers maintained more voicing in C1 in the assimilating environment than in the stops in the fully-voiceless environment. This result combined with the significant effect of voicing environment for assimilating vs. fully-voiced stops for all groups of subjects suggested that both native and non-native subjects devoiced stops in the assimilating environment, but only partially.

In the fricative-stop clusters, the significant effect of environment indicates that the amount of voicing produced in the fricatives by all groups of subjects was significantly larger in the assimilating than in fully-voiceless environments. It was also found that both groups of non-native subjects produced less voicing than natives in the fully-voiced environment (however, this was a difference in the environment where no voicing assimilation was required). Previous phonetic research has shown that voiced English fricatives tend to be produced with little voicing. Several studies reported that in some phonetically-voiced fricatives in intervocalic position, glottal vibration occurred only near the boundary with the vowel (Haggard 1978; Stevens et al., 1992). Therefore, it is not unreasonable to assume that as a result of L1 phonetic transfer, L2 learners produced voiced Russian fricatives with less voicing than native speakers.

In the assimilating environment, neither L1 nor L2 speakers devoiced C1 completely. However, non-natives were shown to assimilate to a lesser degree than natives, which suggests incomplete learning of RVA. That is, L2 speakers produced their assimilated C1 differently from voiced C1, but did not obtain the same voicing proportion values in the assimilating environment as natives.
3.3.3. *Duration of the second consonant in a cluster (C2)*

In this and the following sections, two measures, duration and voicing proportion of the second member (C2) in a consonant cluster, are analyzed. The same statistical design as was applied in the previous section was used again. As discussed in the real-word experiment, C2 was voiced in fully-voiced environment, and voiceless otherwise. The analysis of C2 duration and voicing proportion was necessary for the overall analysis of voicing assimilation, as C2 conditioned the environment for C1 assimilation. However, C2 itself did not participate in the process of RVA and did not undergo assimilation itself. Therefore, its duration and voicing proportion were not expected to undergo large, if any, changes in the assimilation environment. In general, the duration of voiceless obstruents is longer than the duration of voiced obstruents. Therefore, C2 in both assimilating and fully-voiceless environments were expected to be longer than voiced obstruents in the fully-voiced environment.

3.3.3.1. *C2 duration in assimilating-to-voiceless vs. fully-voiced environments*

Figure 3.5. below demonstrates the results of the interaction comparisons of C2 duration in the assimilating vs. fully-voiced environments. As expected, all subjects produced longer C2 in the assimilating environment, as in the real-word experiment. In addition, L1 speakers were found to produce the longest C2 in all conditions, which is apparent from the height of the rightmost bars.
An overall three-factor ANOVA with language level, voicing environment, and cluster type as variables was performed to compare C2 duration in the assimilating to fully-voiced environment. (cf. Figure 3.5, for the results.) There was a significant main effect of environment (F(1, 27) = 66.345, p < .001), with voiced C2 being longer than assimilated ones. The significant main effect of cluster type (F(1, 27) = 112.273, p < .05) suggests that C2 fricatives were longer than C2 stops. There was a significant main effect of language level (F(2, 27) = 4.189, p < .05), with natives producing longer C2 in all conditions than non-natives. None of the two-way interactions was significant (language level by environment: (F(2, 27) = 1.299, p > .05); cluster type by
environment \((F(2,27) = 1.026, p > .05)\); and cluster type by language level: \((F<1)\). The three-way interaction (level by environment by cluster type) was not significant either \((F<1)\). Since neither the three-way nor the two-way interactions were significant, no further analysis was performed.

3.3.3.2. C2 duration in assimilating-to-voiceless vs. fully-voiceless environments

The same three-factor interaction comparison as was used in the assimilating vs. fully-voiced environments above (language learner, cluster type, voicing environment) was employed here to compare C2 duration in the assimilating environment to the fully-voiceless environment. The results of this comparison are shown in Figure 3.6. The C2 duration was expected to be the same in both environments. This prediction was not borne out for the fricative-stop clusters: C2 stops were found to be longer in the assimilating than in the fully-voiceless environments (see the bottom panel of the graph).
Figure 3.6. Comparisons of the duration of the second consonant (C2) in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments

In this ANOVA, there were significant main effects of environment: \(F(1,27) = 112.229, p < .001\) and cluster type: \(F(1,27) = 6.724, p < .02\), and language level: \(F(2,27) = 3.826, p < .05\). The interactions of cluster type and language level: \(F(2,27) = 1.124, p > .05\) and environment and language level: \(F(2,27) = 1.855, p > .05\) were not significant. The interaction of environment and cluster type was significant \(F(2,27) = 6.032, p < .05\). The three-way interaction (level by environment by cluster type) was not significant \(F<1\). Because the two-way interaction of cluster type and environment was significant, a two-way ANOVA with language level and environment as variables was performed for each cluster type (stop-fricative and fricative-stop).
In the stop-fricative clusters, the main effect of environment was not significant (F<1). The main effect of language level was significant (F(2,27) = 4.822, p < .05), with native speakers producing longer C2 fricatives than each group of non-native speakers. The two-way interaction (level by environment) was not significant (F(2,27) = 1.681, p > .05), and no further analysis was performed.

There was a significant main effect of environment in the fricative-stop clusters: (F(1,27) = 7.093, p < .05), with C2 stops in the assimilating environment being somewhat longer than the ones in the fully-voiceless environment for all groups of subjects. There was no significant main effect of language level: (F(2,27) = 2.143, p > .05), although the learners showed a trend toward shorter C2s than natives, which was the same direction as the significant effect above for stop-fricative clusters. Since the two-way interaction (level by environment) was not significant (F<1), no further analysis was performed.

3.3.3.3. Summary

The analysis revealed that native speakers, overall, produced longer C2 than learners. The C2 duration comparison in the assimilating vs. fully-voiceless and assimilating vs. fully-voiced environments in the nonsense data confirmed a cross-linguistic pattern of voiceless fricatives being longer than voiced fricatives. The duration of C2 stops was similar in the assimilating and fully-voiceless environments; whereas C2 fricatives were found to be slightly longer in the assimilating environment.

3.3.4. Voicing proportion of the second consonant in a cluster (C2)

In this section, the amount of voicing in C2 is analyzed in order to determine whether C1 voicing affected C2 voicing. C2 was fully-voiced in the fully-voiced environment, and fully-voiceless in the assimilating and voiceless environments. If C2 voicing
remained unaffected, the directionality of regressive assimilation would be confirmed, and this was an expected pattern in this experiment. An analysis of C2 voicing proportion also verified voicing of the conditioning environment for the C1 obstruents undergoing assimilation. A small voicing proportion of C2 in the assimilating environment confirmed that C2 was indeed fully-voiceless, and was thus able to condition devoicing in C1. Overall, the amount of C2 voicing was expected to be the same in the assimilating and fully-voiceless environments, and smaller in the fully-voiced environment.

3.3.4.1. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiced environments.

Figure 3.7. below demonstrates the results of the C2 voicing proportion analysis in the assimilating vs. fully-voiced environments. It was predicted that the amount of voicing in fully-voiced C2 obstruents would be significantly greater than in the assimilating one regardless of subjects’ level. This prediction was borne out for both C2 stops and C2 fricatives (compare the bar heights on the left to the ones on the right, in the both panels).
Figure 3.7. Comparisons of voicing proportion of the second consonant (C2) that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiced environments

The same overall three-factor ANOVA as was run in the previous sections, was run here in order to analyze C2 voicing proportion in the assimilating vs. fully-voiced environments. There were significant main effects of language level: (F(2,27) = 6.528, p < .01), cluster type: (F(1,27) = 7.095, p < .02), and environment (F(1,27) = 250.772, p < .001). All the two-way interactions were significant (cluster type by language level: (F(2,27) = 4.880, p < .02); environment by language level: (F(2,27) = 9.503, p < .01); cluster type by environment (F(2,27) = 4.689, p < .05). Since the three-way interaction (cluster by language level by environment) was significant (F(2,27) = 4.489, p < .05), the data were analyzed further in order to test for the effects of level and environment for each cluster type separately.
In the stop-fricative clusters, there were significant main effects of environment: \((F(1,27) = 300.998, p < .001)\) and language level: \((F(2,27) = 5.554, p < .05)\). The two-way interaction (level by environment) was significant \((F(2,27) = 10.040, p < .01)\), and the data were analyzed further to test the effects of environment for each language level separately. There was significantly more voicing in fully-voiced C2 fricatives than in assimilating C2 fricatives for each language level (natives: \((F(1,9) = 1128.554)\); advanced: \((F(1,9) = 35.412)\); intermediate: \((F(1,9) = 66.496)\); all \(p’s < .001\)). The significant interaction of language level and environment indicated that L2 speakers produced less voicing in fully-voiced C2 fricatives than native speakers; that is, that they did not make as great a distinction between fully-voiced and fully-voiceless C2 fricatives.

In the fricative-stop clusters, there were significant main effects of environment: \((F(1,27) = 112.769, p < .001)\) and language level: \((F(2,27) = 6.437, p < .01)\). The two-way interaction (level by environment) was significant \((F(2,27) = 6.951, p < .01)\), and the data were analyzed further to test the effects of environment for each language level separately. Like in the stop-fricative clusters, there was significantly more voicing in fully-voiced C2 than in assimilated C2 for each language level (natives: \((F(1,9) = 116.751)\); advanced: \((F(1,9) = 18.550)\); intermediate: \((F(1,9) = 21.129)\); all \(p’s < .01\)), although the interaction again indicated that this difference in voicing was smaller for learners than for natives.

### 3.3.4.2. C2 voicing proportion in assimilating-to-voiceless vs. fully-voiceless environments

Voicing proportion of assimilated and voiceless C2 obstruents is shown below in Figure 3.8. Native subjects produced their C2 completely voiceless in the
assimilating and fully-voiceless environments, as demonstrated by the absence of light grey bars in the graph. Overall, the data in this subset were ‘0’ or close to ‘0’. Running an ANOVA with these data would incur violations of the assumption of normal distribution; therefore, no interaction comparison of C2 voicing proportion in the assimilating vs. fully-voiceless environments was performed.

![Figure 3.8. Comparisons of voicing proportion of the second consonant (C2) that had voicing in stop-fricative and fricative-stop clusters in the assimilating vs. fully-voiceless environments](image)

3.3.4.3. Summary

Although both groups of L2 learners produced less voicing than native speakers in fully-voiced C2 obstruents, they produced significantly more voicing in the fully-
voiced C2 obstruents than in the assimilating ones. Together with the fact that L2 learners successfully produced voiceless C2 in the assimilating and fully-voiceless environments, this demonstrated that they distinguished between fully-voiceless and fully-voiced obstruents, and were able to produce them as such.

3.3.5. Voice Onset Time (VOT) of the second consonant in a cluster (C2)

VOT can only be measured for stops before a vowel, and an analysis was conducted only for C2 stops in the fricative-stop clusters. C2 VOT was analyzed in order to confirm that voicing assimilation is not influencing this part of the signal. Voice Onset occurring before stop release is described as having a negative VOT value, and as having a positive value when Voice Onset occurs after stop release (Ladefoged 2001; Keating 1984). Russian was shown to have a true voicing distinction rather than an aspiration distinction in stops (Jessen 2001; Keating 1984; Petrova et al., 2006); that is, it contrasts voiceless unaspirated stops (in every position) with fully-voiced stops. Therefore, in all groups of subjects in this experiment, fully-voiceless C2 stops were expected to have positive VOT values in the assimilating and fully-voiceless environments, and negative VOT in the fully-voiced environment.

3.3.5.1. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiced environments

A two-factor ANOVA with language level and voicing environment as variables was performed on C2 VOT in the assimilating vs. fully-voiced environment (See Figure 3.9. for the results). As predicted, all groups of subjects distinguished between fully-voiceless and fully-voiced stops, albeit L2 learners’ VOT values, both positive and negative, were smaller than in natives.
Figure 3.9. Comparison of VOT in the second consonant (C2) in the fricative-stop cluster in the assimilating vs. fully-voiced environments

The results of the VOT comparison in the assimilating vs. fully-voiced environments were as follows. There were significant main effects of environment: (F(2,27) = 117.610, p < .001) and language level: (F(2,27) = 3.799, p <. 05). The two-way interaction (level by environment) was also significant (F(2,27) = 16.389, p < .001), and the data were further analyzed to test the effects of environment for each language level separately. Voiced stops were produced with significantly lower (negative) VOT values in the fully-voiced environment than in the assimilating environment for each language level (natives: (F(1,9) = 104.090, p < .001; advanced: (F(1,9) = 21.951, p < .02; intermediate: (F(1,9) = 12.086, p < .02). The significant interaction showed that this effect was smaller for L2 speakers, although it was still significant.
3.3.5.2. C2 VOT in fricative-stop clusters in assimilating-to-voiceless vs. fully-voiceless environments

The same design two-factor ANOVA as used above was performed on C2 VOT in the assimilating vs. fully-voiceless environments. As expected, all groups of subjects produced C2 VOT values similarly in the assimilating and fully-voiceless environments, although, yet again, L1 speakers’ VOT values were greater than in L2 learners (see Figure 3.10. below).

![Figure 3.10. Comparison of VOT in the second consonant (C2) in the fricative-stop cluster in the assimilating vs. fully-voiceless environments](image)

The comparison of VOT values in the assimilating vs. fully-voiceless environments showed that there was no significant main effect of environment ($F(1,27) = 1.435, p > .05$). The main effect of language level was significant ($F(2,27) = 10.416, p < .001$).
with native speakers producing longer VOT. The two-way interaction (language level by environment) was not significant (F<1), and no further analysis was performed.

3.3.5.3. Summary

The comparison of the VOT values in all the three environments (assimilating, fully-voiced, and fully-voiceless) indicated that native speakers produced C2 stops with more extreme VOT (negative in fully-voiced and positive otherwise) than each group of non-native subjects. Intermediate learners, on the other hand, produced the least difference in VOT among environments.

3.3.6. Duration of the vowel preceding the cluster (V1)

The duration of the vowel preceding an obstruent is usually dependant on the voicing specification of that obstruent. Cross-linguistically, vowels were shown to be shorter before voiceless obstruents and longer before voiced (Keating 1984). Chen (1970) confirmed this observation for Russian. In this experiment, vowels were expected to be shorter before both assimilating-to-voiceless and fully-voiceless obstruents, and longer before fully-voiced ones. If assimilation did not apply, V1 duration in the assimilating environment was similar to V1 duration in the fully-voiced environment. If it applied, the duration of V1 would be similar in the assimilating and fully-voiceless environments.

3.3.6.1. V1 duration in assimilating-to-voiceless vs. fully-voiced environments

Figure 3.11. below demonstrates the V1 duration comparisons in the assimilating vs. fully-voiced environments. The duration of V1 was predicted to be longer when preceding fully-voiced C1, and shorter when preceding assimilated C1. This prediction was not, however, borne out. As shown by the graph, in each respective
group of subjects V1 duration before assimilated C1 was similar to V1 duration before fully-voiced C1.

An overall three-factor ANOVA comparison with language level, environment, and cluster type as variables was performed on V1 duration in the assimilating vs. fully-voiced environments. The main effect of was cluster type significant (F(1,27) = 106.867, p < .001). The main effects of environment (F<1) and language level (F(2,27) = 2.788, p > .05) were not significant. The two-way interactions of cluster type and environment, and environment and language level were not significant: (F<1). The interaction of cluster type and language level was significant: (F(2,27) =
4.844, \( p < .02 \)). The three-way interaction (level by environment by cluster type) was not significant either (\( F<1 \)). Because the two-way interaction of cluster type and language level was significant, the data were analyzed further in order to test the effects of language level and environment for each cluster type separately.

In the stop-fricative clusters, the main effects of environment: (\( F(1,27) = 2.305, p > .05 \)) and language level: (\( F(2,27) = 1.735, p > .05 \)) were not significant, which indicated that V1 was produced similarly with regard to duration in the assimilating vs. fully-voiced environments in each group of subjects. The interaction of environment and language level was not significant either (\( F<1 \)), and no further analysis was performed.

In the fricative-stop clusters, the main effect of language level was significant (\( F(2,27) = 3.614, p < .05 \)), with native and intermediate subjects producing overall longer V1 than advanced speakers. The main effect of environment and the two-way interaction of environment and language level were not significant either (\( F<1 \)), and no further analysis was performed.

3.3.6.2. V1 duration in assimilating-to-voiceless vs. fully-voiceless environments

The same three-factor interaction comparison used in the assimilating vs. fully-voiced environments above (language learner, cluster type, voicing environment) was used here to compare V1 duration in the assimilating vs. fully-voiceless environments (see Figure 3.12. for the results). As in the previous analysis, the prediction was not borne out. Contrary to the expected result, V1 duration was somewhat longer in the assimilating environment. Also, as the graph demonstrates, V1 duration was longest for the intermediate speakers (leftmost bars in the graph), and shortest for the advanced ones (center bars in the graph).
Figure 3.12. Comparisons of the duration of the vowel preceding a stop-fricative and a fricative-stop cluster (V1) in the assimilating vs. fully-voiceless environments

In this analysis, the main effects of cluster type ($F(1, 27) = 103.954, p < .001$) and environment ($F(1, 27) = 14.816, p < .01$) were significant. The main effect of language level was not significant ($F(2, 27) = 2.433, p > .05$). The two-way interactions of cluster type and environment ($F(2, 27) = 11.297, p < .02$) and cluster type and language level ($F(2, 27) = 5.398, p < .02$) were significant; whereas the two-way interaction of environment and language level was not ($F(2, 27) = 2.486, p > .05$). The three-way interaction (level by environment by cluster type) was not significant ($F < 1$). Because of the significance of two two-way interactions, the data were analyzed further in order to test the effects of language level and environment for each cluster type separately.
In the stop-fricative clusters, there was a significant main effect of environment: (F(1,27) = 12.377, p < .01) and no significant main effect of language level: (F(2,27) = 1.248, p > .05). Because the two-way interaction (language level by environment) was significant (F(2,27) = 4.992, p < .02), the effect of environment was tested for each language group separately. V1 duration was significantly longer in the assimilating environment for intermediate learners only (natives and advanced: (F<1); intermediate: (F(1,9) = 3.164, p < .02).

In the fricative-stop clusters, there was a significant main effect of environment (F(1,27) = 14.244, p < .01), with longer V1 duration before assimilating obstruents than before fully-voiceless ones. There was a significant main effect of language level (F(2,27) = 3.619, p < .05), with natives and intermediate subjects producing overall longer V1. The two-way interaction (language level by environment) was not significant (F(2,27) = 1.698, p > .05), and no further analysis was performed.

3.3.6.3. Summary

The significant main effect of cluster type in both V1 duration analyses indicated that V1 duration was longer preceding a C1 fricative, and shorter before a C1 stop. As expected, vowels were longer in the fully-voiced than in the fully-voiceless environments. Contrary to the expected pattern, however, V1 duration was similar in the assimilating vs. fully-voiced environments, and longer preceding assimilated obstruents in the assimilating vs. fully-voiceless environments, even in the native speaking group. Furthermore, in all voicing environments and cluster types, intermediate speakers produced the longest vowels, and advanced speakers produced the shortest. In the cross-linguistic study of vowel length by Chen (1970), English
speakers were shown to produce overall longer vowels than Russian speakers, regardless of the voicing specification of the following consonant. The finding that intermediate speakers consistently produced longer V1 than each other group of subjects could simply reflect this fact. More importantly, this finding suggests that the results of L1 speakers producing the longest C1 and C2 do not mean that the natives talked slowly during the experiment because, for example, they felt uncomfortable with the nonsense words. It could be argued that intermediate speakers applied acoustic properties of English vowels when producing Russian vowels.

3.3.7. Duration of a vowel following a cluster (V2)

In addition to the duration of a vowel preceding a cluster, the duration of the vowel following the cluster was also measured. There was no strong reason to expect voicing environment and cluster type to affect the duration of the following vowel in important ways. However, several studies (Campbell 1991; Allen & Miller 1999) have shown that sometimes segments could influence the durations of other segments, even at quite a distance. Therefore, an analysis of V2 duration was performed with the purpose of determining whether V2 duration was affected by C2 voicing. V2 duration was expected to be longer following a fully-voiced C2, and shorter following a fully-voiceless C2 in all groups of subjects.

3.3.7.1. V2 duration in assimilating-to-voiceless vs. fully-voiced environments

The results of the interaction comparison for this measure are shown below in Figure 3.13. Contrary to the prediction, V2 duration was longer after a fully-voiceless C2 in the assimilating environment than after a fully-voiced C2. Also, as the black bars in the graph demonstrate, advanced L2 subjects produced the longest V2 in all conditions.
Figure 3.13. Comparisons of duration of the vowel following a stop-fricative and a fricative-stop cluster (V2) in the assimilating vs. fully-voiced environments

The standard three-factor interaction comparison with language level, environment, and cluster type as variables was performed on V2 duration. Unlike the V1 duration analysis, in this interaction comparison there was no significant main effect of cluster type (F<1). The significant main effect of language level (F(2,27) = 4.093, p < .05) and the main effect of environment (F(1,27) = 16.991, p < .001) indicated that all groups of subjects produced longer V2 in the assimilating environment, with advanced speakers producing the longest V2.

None of the two-way interactions were significant (cluster type by environment: (F<1); environment by language level: (F(2,27) = 1.641, p > .05; cluster type by language level: (F(2,27) = 1.714, p > .05)). The three-way interaction (level...
by environment by cluster type) was not significant either (F<1). Since none of the interactions were significant, no further analysis was conducted.

3.3.7.2. V2 duration in assimilating-to-voiceless vs. fully-voiceless environments
The same three-factor interaction comparison with language learner, cluster type, and voicing environment as factors used in the previous section, was used to compare the V2 duration in the assimilating vs. fully-voiced environments (see Figure 3.14. for the results). As in the previous analysis, advanced L2 speakers were shown to produce the longest V2 in all conditions. Contrary to the prediction, V2 duration in the assimilating environment was slightly longer than in the fully-voiceless one.

![Figure 3.14. Comparisons of duration of the vowel following a stop-fricative and a fricative-stop cluster (V2) in the assimilating vs. fully-voiceless environments](image-url)
As in the previous analysis, there were significant main effects of environment (F(1,27) = 25.525, p < .001) and language level (F(2,27) = 6.045, p < .01); whereas, the main effect of cluster type was not significant (F<1). The two-way interaction of environment and language level was significant (F(2,27) = 10.951, p < .001), and the interaction of cluster type and language level was only marginally significant (F(2,27) = 3.106, p = .061). The interaction of environment and cluster type was not significant (F(2,27) = 1.240 p > .05), and the three-way interaction (level by environment by cluster type) was not significant (F<1) either. Since two two-way interactions were (nearly) significant, the data were further analyzed to test the effects of language level and environment for each cluster separately.

In the stop-fricative clusters, there were significant main effects of environment (F(1,27) = 13.118, p < .01) and language level (F(2,27) = 6.067, p < .01). The interaction of environment and language level was also significant (F(2,27) = 6.404, p < .01), and the data were analyzed further to test the effect of environment for each language level separately. The significant simple effect of environment for native speakers (F(1,9) = 10.722, p < .05) demonstrated that natives produced longer V2 in the assimilating environment than in the fully-voiceless environment. Both groups of L2 speakers produced similar V2 duration-wise in both environments (advanced: (F(1,9) = 1.233); intermediate: (F(1,9) = 2.055; both p’s > .05).

In the fricative-stop clusters, there were significant main effects of environment (F(1,27) = 15.650, p < .01) and language level (F(2,27) = 6.434, p < .01). The interaction of environment and language level was also significant (F(2,27) = 6.135, p < .01), and the data were analyzed further to test the effect of environment for each language level separately. The significant simple effect of environment for
natives (F(1,9) = 10.270, p < .05) and advanced speakers (F(1,9) = 5.612, p < .05) demonstrated that native and advanced subjects produced longer V2 in the assimilating environment than in the fully-voiceless environment. Intermediate speakers produced similar V2 duration in both environments (F(1,9) = 1.462, p > .05).

3.3.7.3. Summary

The absence of a significant main effect of environment and a significant main effect of cluster type suggested that for all groups of subjects, the post-cluster vowel duration was dependent on the type of the preceding obstruent rather than on its voicing.

The observed V2 duration production pattern represented an inverse correlation with V1 duration (cf. Figures 3.11. and 3.12.). Advanced speakers produced the longest V2 and the shortest V1; whereas intermediate speakers produced the longest V1 and the shortest V2, with the natives in between. Conceivably, intermediate subjects talked slower than advanced speakers; however, the fact that intermediate speakers produced the shortest V2 in all conditions demonstrated that the vowel duration differences were not due to the overall speech rate. It should be also recalled that the native participants produced the longest C1 and C2 duration which was not a speech rate effect either. It is unclear as to why advanced speakers produced much longer V2 than each of two other groups of subjects. One could argue that a stress shift could be responsible for such a result. All nonsense stimuli were of the CVCCV shape, with the word-final stress. Advanced speakers might have over-emphasized the stressed vowel, producing it longer, as stressed vowels are usually longer than unstressed ones (Crystal & House 1988).
Intermediate subjects, on the other hand, produced the longest V1. At least some of the intermediate L2 learners were heard to shift the stress from the ultimate syllable to the penultimate one. As a result, they produced V2 as unstressed and, thus, shorter than V1; consequently, duration of their V1 was longer than expected.

3.4. Discussion

Similar to the results of Experiment One, the crucial difference in the RVA production in nonsense stimuli between native and non-native speakers was demonstrated in the most important correlate of RVA: C1 voicing proportion. Although all subjects were shown to produce less voicing in C1 in the assimilating environment than in the voiced environment, native speakers devoiced C1 obstruents more thoroughly than non-native speakers. However, none of the groups devoiced C1 obstruents in the assimilating environment to such an extent as to render them similar to fully-voiceless ones. This pattern was especially noticeable in the production of C1 stops (cf. Figures 3.1. & 3.2.). It is interesting to note that both native and non-native speakers devoiced fully-voiced C1 fricatives in the assimilating environment more than C1 stops, although non-natives retained more voicing than natives. The finding that natives did not assimilate stops as much as fricatives could perhaps be accounted for by the differences in the aerodynamic properties of stops and fricatives, which make fricatives more susceptible to devoicing (J. Ohala 1997).

The finding that non-natives assimilated C1 fricatives more than stops could be taken as evidence that L2 learners have learnt to produce RVA, which they did successfully in the fricative-stop clusters. However, had L2 learners completely acquired RVA, they would have applied it to the C1 stop production as well, and
devoiced C1 as much. It seems possible then, that factors other than the knowledge of RVA could be responsible for the native-like production of C1 fricatives by L2 learners. As discussed in Section 2.3.2.2., English speakers normally do not completely voice even fully-voiced fricatives (Haggard 1978; Stevens at al., 1992); therefore, it could be argued that L2 speakers produced Russian voiced fricatives with less voicing to begin with, and when RVA applied, there was less voicing to lose. Although non-native subjects appeared to devoice C1 fully-voiced fricatives more than C1 stops, this could partly reflect their general lack of voicing during fricatives. This finding could indicate that L2 learners had not yet acquired the voicing values of target Russian obstruents and, as a result, produced C1 fricatives with English-like voicing values. This claim was supported by the results of the C2 voicing proportion analysis, in which L2 speakers also produced fully-voiced C2 fricatives with significantly less voicing than native speakers.

Native speakers not only produced more voicing in voiced fricatives, but made their C1 and C2 stops and fricatives longer than non-natives in all conditions. Overall, as in Experiment One, the C1 stop duration was the same for each respective group of subjects in all environments. This pattern was different from the one obtained for C2 fricatives and stops. Fully-voiceless C2 stops and fricatives, in both assimilating and fully-voiceless environments, were longer than fully-voiced in all groups of subjects. Assimilating and fully-voiceless C1 fricatives were likewise longer than fully-voiced in all groups of subjects.

On the whole, the results of the C1 stop analyses were inconclusive as to the application of RVA. However, previous research on Russian RVA has shown that prevoicing in stops, rather than stop duration, serves as a cue to voicing in Slavic
stops (Burton & Robblee 1997; Keating 1984), and the findings of this study confirmed this claim. The results of the C2 fully-voiceless stops and C1 and C2 fricatives, on the other hand, conformed to the cross-linguistic pattern of voiceless obstruents being longer than voiced.

An L1 phonetic effect on L2 speech, which was observed in the analysis of C1 and C2 voicing duration, was also observed in the analysis of C2 VOT. As in the real-word experiment, L1 fully-voiceless and fully-voiced stops were produced with larger VOT than L2 stops. In voiceless stops, intermediate VOT was close to 0 milliseconds, and advanced VOT was a bit larger; whereas both intermediate and advanced voiced stops were not nearly as prevoiced as L1 fully-voiced stops. English voiced stops are shown to have little, if any, prevoicing (Keating 1984; Ladefoged 2001); and the results of the VOT analysis suggested that L2 learners applied English short-lag VOT values when producing Russian stops. Likewise, L2 speakers transferred English VOT values into the production of Russian voiceless stops. The VOT values obtained in this analysis could be represented on a continuum, with the intermediate speakers having the shortest, and the native the longest VOT, and with the advanced L2 speakers in between. This pattern suggested that L2 speakers were in the process of gradually acquiring the target VOT values. However, this finding was not immediately relevant to the process of RVA, and was used simply to confirm the voice specification of C2 in a cluster.

Vowel duration did not serve as a reliable indication of RVA. Generally, vowels preceding or following fully-voiced obstruents are longer than when preceding or following voiceless obstruents (Allen & Miller 1999; Chen 1970). In this experiment, if a fully-voiced obstruent devoiced in the assimilating environment, a
preceding vowel was then expected to shorten. However, the duration of vowels preceding an obstruent in the assimilating environment was similar to the vowel duration in the fully-voiced environment, and slightly longer than those in the voiceless environments.

This finding could be taken as evidence of incomplete voice neutralization. The voice neutralization process eliminates surface voice contrasts, and renders underlying different obstruents similar in terms of their voice specifications. If neutralization is complete, all the phonetic properties, including vowel duration, employed for the implementation of the voice contrast are expected to be similar in assimilating and fully-voiceless environments (Slowiaszek & Dinnsen 1985, Warner et al., 2003). If neutralization is incomplete, assimilated obstruents differ from truly voiceless ones in terms of one or more acoustic properties. In this experiment, V1 being longer in the assimilating than in the fully-voiceless environments could suggest that fully-voiced fricatives were not completely devoiced, and, therefore, affected the vowel duration. However, the results of the C1 voicing proportion analysis, at least in the fricative-stop clusters, demonstrated that fricatives were almost completely devoiced in all subjects in the assimilating environment; therefore, the process of incomplete neutralization was not supported. Overall, although these findings did not provide evidence against incomplete neutralization in this condition (because voicing environment was a between-item condition), they did not support it either.

Although all subjects produced longer V1 in the assimilating and fully-voiced environments than in fully-voiceless one, the longest vowels were produced by
intermediate learners, and the shortest by advanced learners. On this measure, non-natives performed exactly the same as in they did in real-word stimuli.

The results of the V2 analysis did not parallel the results of the V1 analysis, and contradicted Allen & Miller’s (1999) claim that vowels are longer after fully-voiced obstruents and shorter after fully-voiceless ones. In this experiment, in all conditions the V2 duration was the shortest after the fully-voiced and longest after fully-voiceless environments, with advanced L2 speakers producing the longest V2, and intermediate, the shortest, in all conditions. The absence of an expected result could not be accounted for by the difference in vowel quality, because unlike the real-word stimuli, all nonsense stimuli were matched for V1 and V2 quality (cf. Section 3.3.7.3. for the discussion of possible reasons for such an outcome).

3.4.1. Conclusion

In this chapter I have presented a nonsense-word production experiment and analyzed the following acoustic correlates of RVA: C1 and C2 duration, C1 and C2 voicing proportion, C2 VOT, and V1 and V2 duration. C1 voicing proportion analysis demonstrated an unambiguous difference between native and non-native production in the assimilating vs. fully-voiced environments. The original hypothesis that L2 speakers apply RVA to a lesser degree than L1 speakers, at least on this measure, was thus confirmed. As to other RVA correlates, no such clear difference between L1 and L2 speakers’ production was found. However, instances of L1 phonetic transfer into L2 were manifested in the analysis of C2 VOT, overall obstruent duration, and amount of voicing in fully-voiced fricatives. Although these measurements were not immediately relevant to the process of RVA and were used to verify the directionality of RVA, their analysis demonstrated that in addition to gradual learning of the
phonological RVA pattern, there was also sub-phonemic L1-L2 influence at the phonetic level. More specifically, the results of the C1 voicing proportion analysis showed the phonological learning in L2 speakers; whereas, the C1 fricative voicing proportion and VOT results showed L1-L2 influence at the phonetic level.

More specifically, the comparison of the real and nonsense words revealed that the C1 stop duration was the same in all conditions; whereas the duration of voiced C1 fricative was shorter in all subjects. The C2 fricative duration was the same in all conditions in the real tokens and in natives and advanced, in the nonsense tokens. Voiced C2 stops in the nonsense words were found to be longer than voiceless ones in all language levels but L2 voiced stops were found to be shorter than the natives’ in the real words.

As far as segment voicing duration is concerned, in both experiments, there was less voicing in the assimilated C1 than in the voiced ones, with native speakers generally maintaining the least amount of voicing. Also, the voicing proportion of assimilated L1 and L2 obstruents in the nonsense words was greater than in the real words. The amount of voicing of C1 fricatives in the real-word experiment did not differ among all language groups; whereas in the nonsense word experiment, assimilated fricatives were shown to maintain more voicing than voiceless ones. As for C2 voicing, it was found that all subjects distinguished between voiced and voiceless obstruents, with natives producing more voicing in voiced C2 obstruents than both groups of L2 learners.

The results of the VOT analysis showed that in both experiments, in voiced stops, L1 subjects had more prevoicing than L2 subjects. However, the VOT values of voiceless stops were found to be the same in all groups of participants.
The analysis of the V1 and V2 duration did not reveal any consistent patterns. In both experiments, V1 preceding a fricative in the assimilating environment was shown to be longer than in other environments, for all language groups. V1 preceding a stop in the assimilating environment was shorter than V1 in other environments, in real tokens only. In the nonsense tokens, on the other hand, V1 before a stop in the assimilating and voiced environments was longer than V1 before a stop in the voiceless environment. The duration of V2 in nonsense tokens was the same when following a voiceless fricative, in native and advanced speakers; whereas, V2 following a stop in the assimilating and voiceless environments did not differ in native speech only. Such V2 was longer in the assimilating environment in non-native production. In the real data, on the other hand, the duration of V2 following a fricative was found to be the same in all conditions but shorter than V2 following a stop.

Overall, the phonetic analysis of RVA in real and nonsense stimuli in Russian L1 and L2 has demonstrated that L2 speakers have not acquired the target pronunciation yet. More specifically, advanced L2 learners were found to perform slightly better on several measurements than intermediate learners (cf. Figures 3.3., 3.9. – 3.10). This finding thus suggests a gradual pattern of L2 acquisition, with more proficient subjects performing better than their less proficient counterparts, which needs to be accounted for from a theoretical perspective.

It should be also recalled that the amount of (de)voicing in C1 in the assimilating-to-voiceless environment was established to be the most reliable indication of voicing in both experiments. All groups of subjects were found to devoice C1 to various degrees. Natives maintained less voicing in the appropriate
contexts than L2 speakers, and advanced subjects fared somewhat better than intermediate ones. More specifically, some of the lower level learners were shown to transfer L1 phonetic properties into L2. Overall, the variability found in the native production suggests that the notion of what counts as a voiceless/devoiced obstruent in Russian should be revised and accounted for in the formal analysis. On the other hand, the variability in non-native production needs to be included in the analysis as well. In the following chapter, I will provide a phonological analysis of RVA in Russian native and non-native speech within the Optimality Theory framework which accounts for the findings of the acoustic experiments.

CHAPTER FOUR

PHONOLOGY OF REGRESSIVE VOICING ASSIMILATION

4.0. Introduction

In Chapters Two and Three I presented an acoustic analysis of regressive voicing assimilation (RVA) in native and non-native Russian speech in both real-word and nonsense-word experiments. At this point it should be recalled that RVA requires two adjacent obstruents to agree in voicing as the following examples taken from the nonsense word experiment demonstrate: bazta – ba/st/a vs. basta – ba/st/a vs. bazda – ba/zd/a. In the first item bazta, a voiced fricative /z/ is followed by a voiceless stop /t/ which conditions the devoicing of /z/, as in ba/st/a. The experiments confirmed that RVA takes place in voiced-voiceless obstruent clusters in Russian L1, which was demonstrated by the amount of voicing present in a voiced obstruent in an assimilating environment vs. a fully-voiced obstruent in a fully-voiced environment.
The experiments also demonstrated that L2 Russian learners did not devoice fully-voiced obstruents in an assimilating environment to the same extent as native speakers, although advanced L2 speakers applied RVA more thoroughly than intermediate L2 speakers. These findings were in accordance with the original hypothesis that L2 speakers would perform less authentically than natives in regard to RVA, and that advanced speakers would perform better than intermediate speakers. In conclusion, the analyses showed that consonant voicing duration was the single most important indicator of RVA in all groups of subjects.

This chapter addresses the theoretical implications of the acoustic analysis of RVA done in the previous chapters. The process of RVA was empirically tested and verified in the two production experiments analyzed in Chapters Two and Three. However, a phonological analysis of the data is offered to complement the phonetic experiments. The phonological theory to be employed for this type of analysis should offer a uniform and straightforward account of the empirical findings of the phonetic experiments, and it should be able to explain the acquisition of RVA in L1 and L2 in general. The phonetic experiments showed that the process of RVA is gradient, which was demonstrated by different degrees of devoicing of C1 in an obstruent cluster, even in native speech. It was also found that the L2 acquisition of RVA is gradual, with intermediate L2 speakers performing better than beginners on at least two measures: the amount of voicing and VOT values. These findings have obvious implications for theories of phonological acquisition. In particular, the correct theory must allow for both the gradual acquisition and the gradient nature of RVA shown by the phonetic data. The phonological analysis of the phonetic data will be performed within the Optimality Theory (OT) framework, more specifically, the stochastic OT
model proposed by Boersma (1997) and Boersma & Hayes (2001) because this model has been shown to fulfill these conditions (See Section 4.1.2.).

As discussed in Chapter One, up to now only a relatively small number of studies in L2 phonology have treated L2 phenomena within OT (Broselow et al., 1998; Hancin-Bhatt 2000; Swanson 2001), and very few are based on empirical phonetic data. Although RVA in Russian has not been completely ignored, no phonological studies of RVA in Russian L2 have been performed. In this chapter, I will make an attempt to fill in the gap in existing L1-L2 research by providing an OT analysis of RVA in Russian L1 and L2, which will incorporate the concept of gradual acquisition and gradient devoicing.

The structure of this chapter is as follows: I will begin by briefly reviewing the basic facts about OT, which are followed by a review of optimality-theoretic works on learning L1 and L2. I will then present a review of literature on voicing typology in OT and RVA in Russian L1, and move on to the analysis of the phonetic data presented in Chapters Two and Three for Russian L2 speakers. In particular, I will show how the standard OT analysis of L2 phonological acquisition, which uses categorical constraints, is incapable of accounting for gradient RVA and gradual RVA acquisition. I will then propose an analysis based on the Gradual Learning Algorithm, proposed by Boersma & Hayes (2001), which encompasses both the notion of incomplete RVA in L1 and L2 and gradient acquisition.

4.1. Optimality Theory

Optimality Theory (OT) was developed by Prince and Smolensky (1993) and McCarthy and Prince (1995) as a categorical theory of language knowledge and
language acquisition. In OT, knowledge of a language consists of knowing a universal set of constraints that allows for the production of all and only those phonological forms that appear in languages, as well as a language-particular constraint-ranking (Prince and Smolensky 1993). The set of constraints is proposed to be innate, and acquisition thus becomes a matter of learning the constraint rankings of the target language. In the early stages of acquisition, individual constraints are not ranked with respect to each other, with the family of Faithfulness constraints ranked below Markedness constraints (Boersma 1998; Hayes 1999; Smolensky 1996), and the grammar is unstable. This allows multiple forms to emerge given the same input. However, as the learner gains more and more exposure to the target language, the constraint rankings are stabilized and the target language grammar is acquired.

In recent years, a number of constraint-based learnability theories have been developed to specifically address how OT might account for first language learning (Hayes 1997; Prince & 1999). What follows below is a review of the two most relevant theories, the Error Driven Constraint Demotion algorithm (EDCD) proposed by Tesar & Smolensky (1993, 1996) and the Gradual Learning Algorithm proposed by Boersma (1997) and Boersma & Hayes (2001). Both of these algorithms were put forth to account for the acquisition of L1 phonology through the process of constraint ranking on the basis of the input data. The result of such a ranking is a complete grammar. This process is not unlike the acquisition of L2 phonology, wherein L2 learners have to determine the correct L2 constraint hierarchy in order to achieve the target grammar, a point that will be taken up in a later section.

Recall that in OT language knowledge equals the knowledge of a set of universal constraints and language-specific rankings. Originally it was assumed that in the learner’s initial state all constraints were equally ranked (Tesar & Smolensky 1993). However, in a later version of EDCD (Tesar & Smolensky 1996), it was proposed that markedness constraints dominated faithfulness constraints, which is consistent with other work on child language acquisition and OT (Gnanadesikan 2004; D. K. Ohala 1996). This latter approach made it possible to capture the fact that a learner’s early forms tend to reflect universally unmarked structures. The learner’s task is to infer the correct grammar (constraint hierarchy) from the available data (the actual realization of an utterance), which is termed overt structure in Tesar & Smolensky.

Since OT is by nature a comparative theory, a learner is presented with two competing structural descriptions: a data pair. One, called the winner, is the grammatical structural description assigned by the target grammar (the adult form)\(^8\). The other one, the loser, is an alternative parse of the same input, which is suboptimal under the target ranking. In Tesar & Smolensky (1993, 1996), such a suboptimal form is generated by the learner’s grammar. The learner’s task is then to find a ranking in which each winner is more harmonic than its corresponding loser. A mismatch between the two descriptions is called an error. Learning, which is error-driven, thus begins when learners notice the discrepancy between their output and target forms.

The EDCD algorithm learns the target grammar in the following way. The learner perceives an overt form and analyzes it using interpretive parsing, providing a

---

\(^8\) Adult forms, thus, provide positive evidence about the nature of the optimal output, as well as indirect negative evidence about impossible optimal forms.
full structural description that includes an underlying form. The perceived overt form is the optimal output, and it is checked against the current learner’s ranking. If the perceived form matches what the learner would have produced, then the ranking is consistent with the learning data. However, if the learner’s grammar selects a different ‘optimal’ output, to the grammar must be changed. In EDCD, this is done by applying Constraint Demotion, which demotes the constraints violated by the winner (the target form) in the hierarchy so that they are below the constraints violated by the loser (the learner’s output). In the learner’s initial ranking all constraints are assumed to be unordered and placed in a single stratum, as in \{C_1, C_2, C_3, \ldots, C_n\}. The learner perceives the data and assigns full structural descriptions to the overt forms. The learner forms mark data pairs and compare the optimal candidate together with the constraints which it violates and a suboptimal candidate together with the constraints which it violates. The relationship between them is expressed as loser < winner, shown in Tableau 4.1. below.

<table>
<thead>
<tr>
<th>loser &lt; winner</th>
<th>marks (loser)</th>
<th>marks (winner)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a &lt; b</td>
<td>C_1, C_2</td>
<td>C_2, C_3</td>
</tr>
</tbody>
</table>

**Tableau 4.1. Mark cancellation in the loser > winner data pair**

Since one of the constraints (C_2) is violated by both candidates, it has to be deleted. Since under the learner’s current ranking the winner (b) violates some constraints, such constraints need to be moved below the constraints violated by the loser and down to the next stratum. A new stratum is created, and the demoted constraints are placed there, as in \{C_1, C_2\} >> \{C_3\}. However, it may happen that after mark cancellation, the winner marks are not dominated by the loser marks. This case could be illustrated by assuming the following constraint ranking: \{C_3, C_4\} >> \{C_1\} >> \{C_2\}, presented in Tableau 4.2. below.
Tableau 4.2. Grammar that requires several rounds of demotion

<table>
<thead>
<tr>
<th>loser &lt; winner</th>
<th>marks (loser)</th>
<th>marks (winner)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a &lt; b</td>
<td>C1, C2</td>
<td>C3, C4</td>
</tr>
</tbody>
</table>

The highest-ranked loser mark is C1 which does not dominate the winner marks in the provisional grammar. In order to fix this problem, the learner demotes these constraints (C3, C2, C4) to the stratum immediately below C1:

First demotion: \{C1\} >> \{C3, C2, C4\}

Then the learner checks if the remaining loser mark C2 dominates the winner marks. As C2 is in the same stratum as C3 and C4, the learner applies constraint demotion once again:

Second demotion: \{C1\} >> \{C2\} >> \{C3, C4\}

That is, the process can be repeated if needed until the target ranking is achieved.

Tesar & Smolensky (1996:42) claim that given the potentially large number of possible human grammars, EDCD achieves its goal with “modest time and data requirements.” EDCD also successfully exploits the implicit negative evidence about what cannot be an optimal candidate, which becomes available to the learner through the process of comparing loser-winner pairs.

However, as has been discussed in literature (Cutillas-Espinosa 2001), EDCD cannot easily explain cases of free variation. Under EDCD, constraint demotion is drastic and cannot be undone easily. Also, as Cutillas-Espinosa (2001) points out, if a ranking is changed as the result of a slip of the tongue, the constraint demotion process must begin anew in order to restore the initial state. Furthermore, only the most harmonic structural descriptions are fully grammatical, whereas all suboptimal forms are ungrammatical. This does not leave room for free variation, which presents a problem for the interpretation of the data in this work. Past evidence shows a great
deal of token-to-token variability, and the results of the two acoustic experiments described in Chapters Two and Three, when examined for frequency of various degrees of devoicing, supports this claim as well (see Section 4.5.1.). Therefore, the EDCD model is unlikely to capture these results.

An alternative, stochastic-OT approach, namely the Gradual Learning Algorithm (GLA), developed in the works of Boersma (1998, 2003) and Hayes (2000) and Boersma & Hayes (2001), handles these problems better. In the next section I will review the GLA in some detail.

4.1.2. Gradual Learning Algorithm (Boersma & Hayes 2001)

The GLA could be viewed as a development of Tesar & Smolensky’s EDCD (1993, 1996), because it is error-driven and changes constraint rankings when presented with conflicting data. However, unlike EDCD, which presupposes discrete rankings, GLA assumes a continuous scale of constraint strictness; that is, each constraint has a ranking value along the scale. The distance between constraints can vary. Some are located relatively close to each other; others are further away.

The ranking scale is arranged in arbitrary units, and higher values correspond to higher-ranked constraints. Strict constraint ranking, C1 >> C2 >> C3, could be presented on such a scale as in (1) below. The short distance between constraints C2 and C3 is taken to mean that the relative ranking of this constraint pair is less fixed than that of C1 and C2. This could lead to an overlap, as shown in (3) below.

(1) Categorical ranking of three constraints along a continuous scale (reproduced from Boersma & Hayes 2001:47)

<table>
<thead>
<tr>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
</table>

---
Boersma & Hayes suggest that in the process of speaking, the position of each constraint is temporarily perturbed by a negative or positive value. The constraints are thus associated with ranges of values along a normal curve called *selection points*, rather than single points. The value more permanently associated with the constraint, such as the center of the range, is called the *ranking value*. Therefore, constraints are associated with groups of values that can overlap with the values of other constraints. If the constraints’ selection points do not overlap, the scale represents categorical ranking, as in (2), reproduced from Boersma & Hayes 2001:47.

(2) C1 >> C2

\[
\begin{array}{c}
\text{C1} \\
\text{strict} \\
\hline
\text{C2} \\
\text{lax}
\end{array}
\]

However, if the constraint ranges overlap, variable ranking is obtained, as shown in figures (3) and (4), reproduced from Boersma & Hayes 2001:48.

(3) Variable ranking, common result: C2 >> C3

\[
\begin{array}{c}
2 \\
\text{C2} \\
\hline
3 \\
\text{C3} \\
\hline
\text{strict} \\
\text{lax}
\end{array}
\]

Variable ranking is explained as follows. At evaluation time, the selection points could be chosen from anywhere within the ranges of the two constraints. In (3), C2 would normally outrank C3. However, if the selection points are taken from the
overlapping part, more specifically, from the upper part of C3’s range and the lower part of C2’s range, then C3 would outrank C2, as shown in (4). As a result, free variation, wherein the same underlying form produces more than one possible output, would occur.

(4) Rare result: C3 >> C2

```
C2 3 2 C3
strict
```

The GLA views grammar as stochastic: “at every evaluation of the candidate set, a small noise component is temporarily added to the ranking value of each constraint, so that the grammar can produce variable outputs if some constraint rankings are closer to each other” (Boersma & Hayes 2001:46). Furthermore, the constraint ranges are interpreted as probability distributions. For each constraint, the GLA assumes “a function that specifies the probability that the selection point will occur at any given distance above or below the constraint’s ranking value at evaluation time” (ibid.). Probability distributions make predictions about relative frequencies of outputs generated by the grammar.

According to the GLA, the selection points for constraints are distributed normally, and the mean of the distribution occurs at the ranking value. It should be noted that there are no endpoints to the curve, which predicts that regardless of the distance between the two constraints, there exists probability that they can be re-ranked. The normal distributions are assumed to have the same standard deviation for every constraint; e.g., an arbitrary value of 2.0. The term *evaluation noise* is used to designate the standard deviation of the distribution (Boersma & Hayes 2001:49). At
evaluation time, a small amount of noise is added to the ranking of each constraint. The result of this procedure is that two constraints that are ranked at approximately the same height can be ranked in either order during an evaluation.

The overlapping ranking distribution on an arbitrary ranking scale reproduced from Boersma & Hayes (2001:49) is shown in Figure 5 below.

(5) Overlapping ranking distribution

\[
\begin{array}{cc}
\text{C1} & \text{C2} \\
\end{array}
\]

In the overlapping distribution, the ranking values for C1 and C2 are assigned hypothetical values of 87.7 and 83.1, with the evaluation noise of 2.0. Therefore, the constraints will overlap substantially. C1 will outrank C2 at evaluation time in most cases, but the opposite ranking will occasionally hold\(^9\).

The GLA assumes that initially all constraints are assigned the same arbitrary ranking value of 100. The GLA basic mechanism is similar to EDCD in that the conflict between a learning datum (adult surface forms), and the learner’s current grammar triggers some changes in the ranking of constraints. The algorithm compares the constraint violations of the learning datum with the learner’s form. Unlike EDCD, however, the mismatch between the target and sub-optimal candidates does not lead

\(^9\) Although the authors do not provide the actual calculations, they note that “the percentages for these outcomes will tend towards the values 94.8% (C1 >> C2) and 5.2% (C2 >> C1)” (Boersma & Hayes, 2001:49).
to immediate constraint demotion. The algorithm makes small adjustments of the
ranking values of constraints by increasing the ranking value for the learner’s
candidate, and decreasing the ranking value for the learning datum candidate (that is,
both constraint demotion and promotion are allowed). The algorithm is repeated until
the target grammar is achieved.

In order to illustrate how the GLA works, Boersma & Hayes (2001) put the
GLA to use in the analysis of several phonological phenomena in Ilokano (an
Austronesian language of the Northern Philippines) and several other languages.
What follows is a representative selection of their analysis of Ilokano data pertaining
to free variation and multiple output generation.

Ilokano exhibits an optional process of metathesis; for example, the sequence
\(/\text{tw}/\) optionally becomes \(/\text{wl}/\), as in \text{da\\text{o}\\text{o} ‘kind of tree’ vs. pagd\text{a\\text{o}}\\text{wan} or pagd\text{a\\text{o}}\\text{wan} ‘place where da\\text{o}\\text{o}’s are planted’. In all cases [w] is derived from underlying [o].

Glottal stops are not generally permitted in Ilokano syllable codas, and the only glottal
stop codas are those that are accompanied by a glide formation.

Boersma & Hayes posit two constraints that account for metathesis: a
constraint that bans glottal stops in coda (*\(\text{?}\sigma\)) and a constraint that requires
faithfulness to underlying linear order (LINEARITY). The form \text{ta\\text{?}\\text{en}} avoids coda [\text{?}],
whereas the form \text{ta\\text{?}\\text{wen}} preserves the order of \(/\text{\text{r}}/) and \(/\text{lo}/\) (which becomes \(/\text{wl}/\).
Both candidates alter the syllabicity of \(/\text{lo}/\) and thus violate the constraint IDENT-
IO(SYLLABIC) is also proposed. IDENT-IO([SYLLABIC]). The interaction of the proposed constraints derives the two alternative outputs: \textit{taʔ.wen} and \textit{taw.ʔen}. The form \textit{taw.ʔen} avoids a coda with a glottal stop; whereas the form \textit{taʔ.wen} preserves the relative order of /ʔ/. To account for the forms like \textit{taʔ.wen}, the following ranking, LINEARITY $>> (\text{syllabic})_\sigma$ $>>$ IDENT-IO([SYLLABIC]), is assumed, as shown in Tableau 4.3. below. The winning candidate is indicated by a pointing finger.

<table>
<thead>
<tr>
<th>/taʔo-en/</th>
<th>LINEARITY</th>
<th>(\text{syllabic})_\sigma</th>
<th>IDENT-IO([SYLLABIC])</th>
</tr>
</thead>
<tbody>
<tr>
<td>taʔ.wen</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>taw.ʔen</td>
<td>!</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.3. reproduced from Boersma & Hayes (2001:57)

As the data suggest, when LINEARITY is ranked above (\text{syllabic})_\sigma, one of the possible outputs, \textit{taʔ.wen}, surfaces. Tableau 4.4. below demonstrates a reversed ranking:

$\text{syllabic})_\sigma$ $>>$ LINEARITY $>>$ IDENT-IO([SYLLABIC]).

<table>
<thead>
<tr>
<th>/taʔo-en/</th>
<th>(\text{syllabic})_\sigma</th>
<th>LINEARITY</th>
<th>IDENT-IO([SYLLABIC])</th>
</tr>
</thead>
<tbody>
<tr>
<td>taʔ.wen</td>
<td>!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>taw.ʔen</td>
<td>*</td>
<td>!</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.4. reproduced from Boersma & Hayes (2001:57)

The constraint \text{syllabic})_\sigma bans glottal stops in coda position and, therefore, an alternative output, \textit{taw.ʔen} derived under this ranking surface as optimal.

---

10 Although no formal definition is provided in Boersma & Hayes (2001), the constraint IDENT-IO([SYLLABIC]) could be interpreted as follows: the syllabicity of an input segment must be preserved in its output correspondent.
The data exhibiting free variation shown in Tableaux (a) and (b) can be represented in the form of a ranking scale as in (6) and (7).

(6) Variable ranking, Variant 1: LINEARITY $>> \ (*\?\sigma)$.

\[
\begin{array}{c}
\text{LINEARITY (C1)} \\
\text{C1} \\
\text{strict} \\
\end{array} \quad \begin{array}{c}
\text{(*\?\sigma) (C2)} \\
\text{C2} \\
\text{lax} \\
\end{array}
\]

In (6), LINEARITY outranks (*\?\sigma). However, if the selection points are taken from the upper part of LINEARITY’s range and the lower part of (*\?\sigma) range, then C3 would outrank C2, as in (6).

(7) Variable ranking, Variant 2: (*\?\sigma) $>>$ LINEARITY

\[
\begin{array}{c}
\text{LINEARITY (C1)} \\
\text{C2 C1} \\
\text{strict} \\
\end{array} \quad \begin{array}{c}
\text{(*\?\sigma) (C2)} \\
\text{lax} \\
\end{array}
\]

The analysis of the Ilokano data is of course only one example of how this theory can handle variability in native speakers and a frequency of occurrence in addition to L1 learning and is included here for expository purposes only. Other cases of variability which are modeled by using the GLA include Albright & Hayes (to appear), Arbisi-Kelm (2002), Hayes & Londe (2006).

4.1.3. Summary

The GLA has been shown to have certain advantages over EDCD. The continuous ranking scale allows small changes to the constraints’ locations along the scale. Therefore, the resulting changes are not as drastic as those occurring under EDCD.
Boersma & Hayes (2001) argue that the algorithm is robust against occasional errors in the input data, meaning that such errors do not change the ranking dramatically. In addition, by employing the concept of the continuous ranking scale, the GLA is able to handle grammars that generate multiple outputs, which cannot be generated under EDCD. As the data of the acoustic experiments have demonstrated, the RVA production of both native and non-native speakers exhibited a large amount of variability (see Section 4.5.1.). More importantly, both optimal and sub-optimal candidates were shown to surface in L2 production. This finding can be best accounted for within the GLA framework, which, unlike EDCD, performs the re-ranking minimally, thus allowing for variation in the data. To my knowledge, no studies have applied the GLA to the analysis of L2 phonological acquisition and production, and only one study (Swanson 2001) has specifically tested EDCD in L2 acquisition of RVA.

In her study, Swanson (2001) analyzed the process of obtaining the L2 grammar through the application of Tesar & Smolensky’s (1993, 1996) learning algorithm. She specifically looked at the acquisition of RVA and word-final devoicing in Polish L2 – English L1 and lack thereof in Polish L1 – English L2. I will turn to the review of Swanson’s work in Section 4.3.1. I will first, however, review the literature on voicing typology and voicing assimilation as a general phenomenon in L1 within the OT framework, specifically work by Lombardi (1999), which establishes some primary facts about voicing in various languages and its treatment within traditional OT. The constraints introduced in her analysis are utilized in Swanson’s (2001) study of Polish and English L2; thus Lombardi’s investigation serves as the basis for Swanson’s work.
4.2. Analysis of voicing typology in languages other than Russian from an L1 perspective

RVA has been studied in the phonological literature for many years (e.g., Barkai & Horvath 1978; Hayes 1984; J. Ohala 1990; Wissing & Roux 1995). The following discussion will focus on accounts of RVA that have utilized the OT framework. Grijzenhout (2000: Afrikaans, English, German), Lombardi (1999: English, Polish, Yiddish, and Serbo-Croatian), Padgett (2003: Russian), and Petrova & Szentgyorgyi (2001: Hungarian, Russian) have used OT to account for voicing assimilation as a general phenomenon in various languages. Lombardi (1999) offers the most comprehensive OT account of voicing typology and RVA in obstruent clusters observed to pattern differently in languages. This work is reviewed below.

4.2.1. Word-medial obstruent clusters and word-final obstruents

It should be recalled that RVA in Russian requires two adjacent obstruents to Agree in voice, with voice spreading leftwards; e.g., vezti - /vıstı/ ‘to drive’ vs. vezu - /vızu/ ‘I drive’. Within the OT framework, voicing assimilation is viewed as result of the interaction between positional faithfulness and markedness constraints. The table presented in 4.1. below summarizes all the cross-linguistic patterns.

<table>
<thead>
<tr>
<th>Language</th>
<th>Voicing typology</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>Syllable-final neutralization</td>
<td>sagte - sa[k.t]e ‘said, 1st person, Sg.’</td>
</tr>
<tr>
<td>English</td>
<td>Unrestricted voice¹²</td>
<td>pigpen –pi[g.p]en</td>
</tr>
<tr>
<td>Yiddish</td>
<td>VA in obstruent clusters &amp; word-final contrast</td>
<td>ayzkastn - ay[sk]astn ‘icebox’</td>
</tr>
<tr>
<td>Maori</td>
<td>No voicing in obstruents</td>
<td>(No examples provided by the author)</td>
</tr>
</tbody>
</table>

¹¹ Recall that only the cases of the assimilation to [-voice] are addressed in the analysis because the cases of the assimilation to [+voice] are sparse and thus have been omitted.

¹² However, see Hammond (1999) for an opposing view of restricted [voice] in monomorphemic sequences.
Table 4.1. Voicing typology (based on Lombardi 1999:268)

As already noted, treatments of voicing in these languages are argued to be the result of conflicts between relevant faithfulness and markedness constraints. Given that languages rank constraints differently, such conflicts are resolved in various ways that yield the typology shown above. To account for variation in how [voice] in obstruents is employed in these languages, Lombardi proposes a set of faithfulness and markedness constraints (reviewed in the order and form presented by the author), the first one being the constraint *Lar.

(1) *Lar – do not have Laryngeal features

Lombardi assumes that [voice] is a privative feature; that is, one-valued, with the contrast best represented by the presence or absence of [voice]. Therefore, in her analysis, a voiced obstruent receives one mark for *Lar, and voiceless obstruents receive no marks. The constraint *Lar accounts for German syllable-final neutralization, as in rad - ra[t] ‘advise’ or sagt - sa[kt]e ‘said, 1st person, Sg.’, since voiced obstruents in the output will always violate this constraint. The concurrent interaction of *Lar and positional faithfulness constraints (discussed below) allows voiced obstruents to surface in the onset but not in the coda, as an example from German demonstrates: gut - *kut ‘good’ (See Tableau 4.1. below).

Traditionally the directionality of assimilation (regressive vs. progressive) has been explained in terms of positional faithfulness (Lombardi 1999, following Beckman 1997). It is believed that faithfulness to the underlying feature is preserved

---

13 Lombardi (199:271) argues that although the constraint *Lar can be defined as “Obstruents should not have laryngeal features,” a more general form of the constraint would extend it to sonorants. However, she doesn’t pursue this issue any further.
in phonetically and psycholinguistically salient positions, such as the syllable onset position; hence, Lombardi’s faithfulness constraint IDentOnsetLaryngeal.

(2) IDentOnsetLaryngeal (IDOnsLar): consonants in the onset position should be faithful to underlying laryngeal specification.

The interaction of * Lar and IDentOnsetLaryngeal is shown in Tableau 4.5.

<table>
<thead>
<tr>
<th>/rad/</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. rad</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. rat</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>/gut/</td>
<td>IDOnsLar</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. gut</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d. kut</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.5. Syllable-final devoicing: German (reproduced from Lombardi 1999:274)

As seen in this tableau, this ranking gets devoicing syllable-finally, but ensures faithfulness in the onset position. For an onset consonant, it is more important to obey IDOnsLarl than to obey *Lar. Candidate (a) violates IDOnsLar and is therefore eliminated; whereas, candidate (b) satisfies IDOnsLar and is chosen as the optimal candidate.

As mentioned, the constraint IDOnsLar allows underlying voicing specifications to surface in the onset position exclusively. However, in order to militate against unwanted changes in voicing specification, another faithfulness constraint, which ensures that input and output voice specifications are the same, is needed. Lombardi proposes:

(3) IDent(Laryngeal) (IDLar) – Consonants should be faithful to underlying laryngeal specification.

Since syllable-final devoicing is required in German, the constraint IDent(Laryngeal) (IDLar) is assumed to be ranked lower than *Lar, which militates
against voiced segments. Tableau 4.6. demonstrates the ranking of IDLar over *Lar in a word-internal cluster.

<table>
<thead>
<tr>
<th>/sagte/ ‘I said’</th>
<th>*Lar</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sa[gt]e</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. sa[kt]e</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.6. Syllable-final devoicing in a word-internal cluster in German: (reproduced from Lombardi 1999:274&276)

In this dataset, candidate (a) contains a voiced consonant [g] and thus violates high-ranked *Lar. Candidate (b) violates only the low-ranked constraint IDLar and thus emerges as the optimal candidate.

However, recall that in German the constraint *Lar is ranked below IDOnsLar (cf. Tableau 4.5.). Therefore, the IDOnsLar, which targets syllable-final devoicing in obstruent clusters, should be added to the current ranking *Lar >> IDLar. The ranking is thus: IDOnsLar >> *Lar >> IDLar, which is demonstrated in Tableau 4.7.

<table>
<thead>
<tr>
<th>/sagte/ ‘I said’</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sa[gt]e</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. sa[kt]e</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>c. sa[gd]e</td>
<td></td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

Tableau 4.7. Syllable-final devoicing in a word-internal cluster in German (reproduced from Lombardi 1999:274&276)

In this tableau, candidate (a) violated *Lar and is ruled out. Candidate (c) violates high-ranked IDOnsLar, because it contains a voiced obstruent in the onset of the second syllable. A voiceless syllable-initial consonant in (b) remains voiceless, thus satisfying the high-ranked IDOnsLar, and thus this candidate emerges as optimal wins.

The last constraint Lombardi proposes is AGREE, which requires adjacent obstruents to agree in voicing, and thus militates against clusters with differing voice specifications.
(4) Agree – obstruent clusters should agree in voicing

This set of constraints, pertaining to when a language-specific ranking is applied, accounts for the voicing typology summarized in Table 4.1. above. Lombardi argues that syllable-final laryngeal neutralization in German, then, calls for the following constraint ranking: IDOnsLar >> *Lar >> Agree, IDLar. Tableau 4.8. below shows this ranking at work.

```
<table>
<thead>
<tr>
<th>/sagte/ ‘I said’</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>Agree</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sa[gt]e</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. sa[kt]e</td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. sa[gd]e</td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Tableau 4.8. Syllable-final devoicing in a word-internal cluster: German (based on Lombardi 1999:274&276)

As discussed above, in German it is more important for onset consonants to obey IDOnsLar (Onsets should be faithful to underlying voicing) than to obey *Lar (Do not be voiced). Candidate (c) changes the voicing specification on the onset, and thus fatally violates high-ranked IDOnsLar, in addition to violating lower-ranked *Lar twice and IDLar once. Although candidate (a) obeys IDOnsLar by preserving its underlying voicing, it fatally violates high-ranking *Lar, in addition to lower-ranked Agree. The remaining candidate (b) is the optimal one, because it satisfies high-ranked IDOnsLar and *Lar, violating only low-ranked IDLar.

Tableau 4.9. below demonstrates Lombardi’s analysis of unrestricted voicing, as in English. To allow for unrestricted voicing in obstruent clusters (e.g., /gp/ rather than assimilated /kp/ or /gb/), the faithfulness constraints, IDOnsLar and IDLar, must be ranked above the markedness constraints *Lar and Agree. The ranking is thus: IDOnsLar, IDLar >> *Lar, Agree.
Tableau 4.9. Unrestricted voice: English (reproduced from Lombardi 1999:277)

Since high-ranked constraints IDOnsLar and IDLar are not crucially ranked with respect to each other, they both work together to rule out unfaithful candidates such as those in (b) and (c). The remaining candidate (a), with a word-internal cluster that disagrees in voicing, satisfies these high-ranked constraints and emerges as optimal.

Languages such as Yiddish and Serbo-Croatian show assimilation to the voicing of the last obstruent in a cluster, but preserve contrast in word-final positions. Tableau 4.10. below demonstrates the data from Yiddish. In Yiddish, medial input clusters that disagree in voicing (e.g., /kb/ below) must assimilate to the second member of the cluster (e.g., /gb/). In word-final position, the voicing contrast is preserved. This is ensured by IDLar being ranked higher than *Lar. Lombardi’s suggested ranking for RVA in medial clusters and word-final faithfulness in Yiddish is: Agree, IDOnsLar >> IDLar >> *Lar.

Tableau 4.10. Voicing assimilation in medial clusters and word-final faithfulness in Yiddish (reproduced from Lombardi 1999:283)

As the data for Yiddish suggest, ranking Agree over IDLar ensures voicing assimilation in medial clusters, as in candidates (b) and (c). The ranking of IDOnsLar over IDLar suggests that it is also more important to be faithful to an onset laryngeal
specification than to a coda specification. In the first data subset, the only way to satisfy Agree is for the coda to assimilate to the voicing of the onset. Candidate (b), where the coda is assimilated to the onset, satisfies both Agree and IDOnsLar and surfaces as an optimal candidate. Candidate (a) violates Agree, because it has an obstruent cluster disagreeing in voicing. The cluster in (c) Agrees in voicing and satisfies Agree, but it is unfaithful to the onset laryngeal specification, thereby violating IDOnsLar and IDLar. In the second data subset, where there is no obstruent cluster, Agree and IDOnsLar are irrelevant. Unfaithful candidate (d) is ruled out by IDLar. Candidate (e), which satisfies this constraint, is the winner.

Similarly to the Yiddish ranking, languages such as Polish, with RVA and word-final devoicing, rank Agree and IDOnsLar higher than *Lar and IDLar. Agree ensures that assimilation in obstruent clusters takes place, whereas IDOnsLar is responsible for the regressive directionality of voicing assimilation. Lombardi’s ranking for RVA and word-final neutralization is as follows: Agree, IDOnsLar >> *Lar >> IDLar.

<table>
<thead>
<tr>
<th>/zhabka/ ‘small toad’</th>
<th>Agree</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. zhabka</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. zhapka</td>
<td></td>
<td>*!</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>c. zhabga</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>/klub/ ‘club’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. klub</td>
<td></td>
<td>!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>e. klup</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Table 4.11. Voicing assimilation and word-final neutralization: Polish (reproduced from Lombardi 1999:282)

In the first data set, candidates (a) and (c) fatally violate the higher-ranked constraints Agree and IDOnsLar, and are thus ruled out. The remaining candidate (b), which exhibits RVA in its medial cluster, violates the lowest-ranked IDLar only and surfaces
as the optimal candidate. When there is no obstruent cluster, as in (d) and (e), constraints Agree and IDOnsLar are irrelevant. Since *Lar prohibits voiced obstruents, candidate (d), which violates this constraint, is ruled out; whereas the winning candidate (e) violates the lowest-ranked IDLar only.

4.2.2. Complex codas

All the data reviewed above involved only simple codas. However, Lombardi (1999) addresses the issue of complex codas by showing that the proposed constraints are capable of handling complex coda behavior in word-final positions. For example, the constraint ranking IDOnsLar >> *Lar >> Agree, IDLar proposed for German syllable-final devoicing can account for word-final neutralization as well, as Tableau 4.12 below demonstrates.

<table>
<thead>
<tr>
<th>/jagd/ ‘hunt’</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>Agree</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. jagd</td>
<td></td>
<td>#!#</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. jakt</td>
<td></td>
<td>*!</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>c. jagt</td>
<td></td>
<td>#!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Since all of the candidates contain word-final complex codas, the constraint IDOnsLar, which ensures faithfulness to the voice specification in the onset position, is irrelevant, and it is satisfied vacuously within the candidates shown in the tableau. The decision is made by lower-ranked *Lar, which rules out candidates (a) and (c), containing two and one voiced codas, respectively. Remaining candidate (b) devoices both obstruents in a word-final coda, thus satisfying all high-ranked constraints, and emerges as an optimal candidate.

Polish also exhibits word-final voicing neutralization in complex codas; however, this is achieved through a different ranking because Polish also has voicing
assimilation in consonant clusters. The ranking is thus: Agree, IDOnsLar >> *Lar >> IDLar. The Polish data are demonstrated in Tableau 4.13. below.

<table>
<thead>
<tr>
<th>/muzg/</th>
<th>Agree</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. muzg</td>
<td></td>
<td></td>
<td>⬤!</td>
<td></td>
</tr>
<tr>
<td>b. musk</td>
<td></td>
<td></td>
<td></td>
<td>⬤!*</td>
</tr>
<tr>
<td>c. muzk</td>
<td>⬤!*</td>
<td></td>
<td>⬤*</td>
<td>⬤*</td>
</tr>
</tbody>
</table>


Like in German, constraint IDOnsLar is irrelevant here. Candidate (c), which has a consonant cluster disagreeing in voicing, is ruled out by Agree. Constraint *Lar ensures that a form with a voiced word-final coda, as in (a), does not surface. The remaining candidate (b) satisfies all the high-ranked constraints, and surfaces as the optimal one.

As shown in Section 4.2.1., Yiddish disallows word-final voicing neutralization in simple codas. However, it devoices complex codas, as demonstrated in Tableau 4.14.

<table>
<thead>
<tr>
<th>/zogt/</th>
<th>Agree</th>
<th>IDOnsLar</th>
<th>IDLar</th>
<th>*Lar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. zogt</td>
<td>⬤!*</td>
<td></td>
<td>⬤*</td>
<td>⬤*</td>
</tr>
<tr>
<td>b. zokt</td>
<td></td>
<td></td>
<td>⬤*</td>
<td>⬤*</td>
</tr>
<tr>
<td>c. zogd</td>
<td>⬤*</td>
<td></td>
<td></td>
<td>⬤!*</td>
</tr>
</tbody>
</table>


Unlike in Polish and German, in Yiddish the input cluster contains mixed voicing. Candidate (a), with mixed voicing in the coda, violates high-ranked Agree and is ruled out. IDOnsLar is irrelevant in these data and is satisfied vacuously. Candidates (b) and (c) both satisfy Agree by exhibiting voicing assimilation, but do so at the
expense of violating IDLar. The decision is passed on to lower-ranked *Lar. As a result, candidate (b) containing an unmarked voiceless cluster is chosen.

4.2.3. Summary

Lombardi’s (1999) analysis of voicing typology in word-medial and word-final obstruents and obstruent clusters reviewed above has shown that OT is capable of providing a descriptive framework for phonological phenomena occurring in different, unrelated languages in a uniform manner. It does so by employing a set of universal faithfulness and markedness constraints. Under this approach, the voicing typology is viewed as stemming from a conflict of markedness and faithfulness constraints that is resolved through language-specific rankings of these constraints. Constraints promoting or prohibiting voicing assimilation and/or voicing contrast are shown to be the same across the languages; yet these constraints are ranked differently relative to each other in different languages. The direction of voicing assimilation in consonant clusters has also been accounted for in terms of constraint rankings. Lombardi (1999) proposes this to be a result of the interaction of a constraint requiring obstruents to agree in voicing (Agree), and constraints on positional faithfulness (IDOnsLar and IDLar) preventing any modifications to input forms. Table 4.2. below summarizes the voicing factorial typology, as presented by Lombardi (1999).

<table>
<thead>
<tr>
<th>Language</th>
<th>Voicing typology</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>Voicing in onset but not coda: IDOns&gt;&gt;*Lar&gt;&gt;IDLar</td>
<td>sagte - sa[k,t]e ‘said, 1st person, Sg.’</td>
</tr>
<tr>
<td></td>
<td>No assimilation to [voice] in clusters: IDOns&gt;&gt;*Lar&gt;&gt;Agree, IDLar</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>Faithfulness to underlying voicing in all positions:</td>
<td>pigpen –pi[gp]en</td>
</tr>
</tbody>
</table>
Word-final devoicing:  
IDO�,*Lar>>IDLar
VA in obstruent clusters:  
IDO�, Agree >> *Lar >> IDLar

klub – klu[p] ‘club’

zhabka – zha[p]ka ‘frog’

Word-final contrast:  
IDO�>>IDLar>>*Lar
VA in obstruent clusters:  
Agree, IDO�>>IDLar>>*Lar

ayz ‘ice’

ayzkastn - ay[sk]astn ‘icebox’

Bidirectional VA assimilation:  
Agree, IDLar>>*Lar>>IDO�  
No word-final devoicing

RVA:  
vigsel-vi[k:s]el ‘marriage’  
PVA: tisdag – ti[s:t]a ‘Tuesday’  
dag – da[g] ‘day’

No voicing in obstruents:  
*Lar>>IDO�, IDLar

(No examples provided by the author)

Table 4.2. Voicing typology (based on Lombardi 1999)

Lombardi’s work does not include an analysis of RVA in Russian, but one would assume the typology presented will allow for a similar analysis of Russian RVA given the similarity of Russian to Polish. There are two relevant works to review at this point, one is by Swanson (2001), who applies Lombardi’s constraints to Polish L2 productions, and the other is by Padgett (2002), who applies Lombardi’s constraints to Russian L1 productions. These works will be reviewed in that order in the following sections.

4.3. Analysis of RVA in L2 phonology

As discussed above, OT is not only a theory of language knowledge, but also a theory of language acquisition. In OT, L1 acquisition is a matter of figuring out the L1 constraint rankings responsible for various patterns in the language, and different constraint rankings hold for the same phenomenon depending on how that
phenomenon is realized in the language. By the same token, L2 acquisition is a process of figuring out L2 constraint rankings, but this is achieved by re-ranking L1 constraints with the purpose of achieving the target ranking. That is, unlike in L1 acquisition, the initial state for L2 learners is assumed to be the constraint ranking of L1. This assumption has been proposed and tested in several analyses of different phonological patterns exhibited by L2 learners of various languages (Broselow et al., 1998 and Hancin-Bhatt 2000: English L2 coda acquisition; van Rooy & Wissing 2000 and Swanson 2001: acquisition of voice distinction and voice neutralization).

To my knowledge, only one study, by Swanson (2001), analyzes the process of acquiring RVA in obstruent clusters in an L2. The purpose of Swanson’s paper is to test the L1 learning algorithm proposed by Tesar & Smolensky (1993, 1996; cf. Section 4.1.1) as applied to L2 acquisition, which she does by analyzing RVA and word-final devoicing in L2 Polish – L1 English and L2 English – L1 Polish. It should be noted that Swanson does not provide her own data, but bases the analysis on categorical by-ear data collected by others (Rubach 1983, 1984, cited in Swanson 2001). However, her analysis will be reviewed in this section as her work and the work by Padgett (2002, to be reviewed in Section 4.4) serve as the starting point of my own analysis of L2 RVA for Russian, for which I provide data (cf. Chapters Two and Three).

4.3.1. Polish speakers learning L2 English

In contrast to English, in Polish (and Russian), word-final obstruents devoice (e.g., sad – sa[t] ‘orchard’) and word-medial obstruent clusters agree in voicing (ogrudka – ogru[tk]a ‘garden’). To account for this pattern, Swanson adopts the following constraints posited for Polish by Lombardi (1999) (repeated here from Section 4.2.).
(1) IDOnsLar – onsets should be faithful to underlying laryngeal specification

(2) *Lar – Do not have laryngeal features

(3) IDLar – consonants should be faithful to underlying laryngeal specification

(4) Agree – Adjacent obstruents should agree in voicing

As shown earlier, Lombardi’s proposed ranking for Polish L1 is: Agree, IDOnsLar >> *Lar >> IDLar. Tableau 4.15. (adapted from Swanson 2001) is based on this ranking and shows RVA and word-final devoicing in Polish L1.

<table>
<thead>
<tr>
<th>/ogrudka/ ‘small garden’</th>
<th>Agree</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ogrudka</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>b. ogrutka</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/sad/ ‘orchad’</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>c. sad</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>d. sat</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.15. RVA and word-final devoicing: Polish L1 (reproduced from Swanson 2001:32)

As the tableau demonstrates, ranking Agree over IDLar results in assimilation to voiceless, as in (b). In this form the voiced input coda devoices in the output. Agree is thus satisfied at the cost of a low-ranked IDLar, and candidate (b) emerges as optimal. Faithful candidate (a), on the other hand, satisfies IDLar but violates high-ranked Agree and is ruled out. For word-final obstruents (candidates c-d), the constraints Agree and IDOnsLar are irrelevant. Since candidate (c) violates *Lar, it is ruled out and candidate (d) is an optimal one.

As mentioned earlier, it is a generally accepted view that the learners’ final L1 ranking is the initial state of their L2 ranking (Broselow et al., 1998; Hansin-Bhatt 1999). Therefore, Swanson assumes the Polish ranking Agree, IDOnsLar >> *Lar >> IDLar as the initial state for native speakers of Polish acquiring English as L2.
Therefore, in the earliest stages of acquisition (Stage 1 in Swanson) when they are producing English obstruent clusters disagreeing in voicing, Polish L1 speakers erroneously apply RVA and word-final devoicing (FD), as shown in Tableau 4.16 below. The winning candidates in this tableau, while incorrect from the perspective of the native speaker, are correct for the L2 speaker at this stage. Such candidates are marked with an ‘unhappy’ face in this tableau and tableaux hereafter. The optimal forms in the target language are marked with a pointing finger, whenever appropriate.

<table>
<thead>
<tr>
<th>/obtain/</th>
<th>Agree</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obtain</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>☐b. optain</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. obdain</td>
<td>*!</td>
<td>**</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>/misbehave/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. misbehave</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>☐e. mizbehave</td>
<td></td>
<td></td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>/bob/</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>f. bob</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>☐g. bop</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.16. Stage 1: RVA and FD in Polish L1 - English L2 (reproduced from Swanson 2001:27). The winning candidates in this tableau, while incorrect from the perspective of L1, are correct for L2 speakers at this stage.

In the forms containing medial clusters, Agree rules out candidates (a) and (d), as their clusters disagree in voicing. Candidate (c), although passing Agree, violates IDOnsLar by changing /t/ in the input onset to /d/ in the output onset. Polish L1 – English L2 learners produce forms like obtain as *o[pt]ain rather than o[bt]ain, and mi[zb]ehave rather than mi[zb]ehave (candidates (b) and (e)) because Agree and IDOnsLar are ranked higher than *Lar and IDLar in Polish L1. As a result, the first consonant in a cluster assimilates the voicing specification of the second consonant in the cluster. Therefore, candidates (b) and (e) are erroneously chosen as optimal candidates.
In the pair \textit{bo[b]} – \textit{bo[p]}, the form \textit{bo[p]} in (g) shows final devoicing because \textit{Lar} is ranked higher than IDLar and ensures that final voiced obstruents do not surface. The form \textit{bo[b]} in (f) fatally violates \textit{Lar} and is ruled out. The winning candidate in this pair is thus a non-native form (g).

Following Tesar & Smolensky’s EDCD (1993, 1996), Swanson claims that learning occurs when learners notice a difference between a target form and a form produced by learners’ current grammar. The forms are compared in pairs, and depending on which two forms are compared first, the L2 acquisition paths may differ for learners.

Swanson posits four stages with various paths within these stages, which is based on EDCD. Table 4.3., reproduced from Swanson (2001:31), summarizes all the possible stages and paths in the acquisition of the English voicing contrast. Each path and stage will be reviewed in turn.

<table>
<thead>
<tr>
<th>Path</th>
<th>Learner notices:</th>
<th>Constraint rankings</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No VA to voiceless</td>
<td>Stage 1: IDOnsLar, Agree &gt;&gt; \textit{Lar} &gt;&gt; IDLar</td>
<td>RVA &amp; FD (Polish)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 2: IDOnsLar &gt;&gt; IDLar &gt;&gt; \textit{Lar}, Agree</td>
<td>No VA or FD (English)</td>
</tr>
<tr>
<td>2a</td>
<td>No VA to voiced</td>
<td>Stage 1: Agree, IDOnsLar &gt;&gt; \textit{Lar} &gt;&gt; IDLar</td>
<td>VA &amp; FD (Polish)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 2: IDOnsLar &gt;&gt; \textit{Lar} &gt;&gt; Agree, IDLar</td>
<td>VA to voiceless &amp; FD</td>
</tr>
<tr>
<td></td>
<td>No VA to voiceless</td>
<td>Stage 3: IDOnsLar &gt;&gt; IDLar &gt;&gt; \textit{Lar}, Agree</td>
<td>No VA or FD (English)</td>
</tr>
<tr>
<td>2b</td>
<td></td>
<td>Stage 1: Agree, IDOnsLar &gt;&gt; \textit{Lar} &gt;&gt; IDLar</td>
<td>VA &amp; FD (Polish)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 2: IDOnsLar &gt;&gt; \textit{Lar}</td>
<td>VA to voiceless &amp;</td>
</tr>
<tr>
<td>No VA to voiced</td>
<td>No FD</td>
<td>No VA to voiceless</td>
<td>FD</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------</td>
<td>-------------------</td>
<td>----</td>
</tr>
<tr>
<td>No FD</td>
<td>Stage 3: IDOnsLar &gt;&gt; Agree, IDLar &gt;&gt; *Lar</td>
<td>Stage 4: IDOnsLar &gt;&gt; IDLar &gt;&gt; *Lar, Agree</td>
<td>VA to voiceless but no FD</td>
</tr>
<tr>
<td>No VA to voiced or voiceless</td>
<td>Stage 1: Agree, IDOnsLar &gt;&gt; *Lar &gt;&gt; IDLar</td>
<td>Stage 2: Agree, IDOnsLar &gt;&gt; IDLar &gt;&gt; *Lar</td>
<td>No VA or FD (English)</td>
</tr>
<tr>
<td>3</td>
<td>Stage 3: IDOnsLar &gt;&gt; IDLar &gt;&gt; Agree, *Lar</td>
<td>VA &amp; FD (Polish)</td>
<td></td>
</tr>
<tr>
<td>No FD</td>
<td>VA but no FD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No VA to voiceless or voiceless</td>
<td>VA &amp; FD (Polish)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.3. Summary of predicted paths & stages in acquisition of L2 English voicing contrast (Swanson 2001:31)**

Stage 1 shown in Tableau 4.16. is assumed as a starting point for all L2 learners. However, Stage 2 could have two different paths, (a) and (b), depending on whether learners first notice a lack of both RVA and FD in target forms. If learners first observe that the target parse *o[bt]ain* differs from their parse *o[pt]ain* (cf. Tableau 4.16.), it leads them to notice a lack of regressive devoicing in English. This comparison will lead the learners to demote *Lar and *Agree, which the form *o[bt]ain* violates, below IDOnsLar and IDLar. The resulting (target) ranking IDOnsLar >> IDLar >> *Lar, Agree, which disallows voicing assimilation and final devoicing, is demonstrated in Tableau 4.17. below.

<table>
<thead>
<tr>
<th>/obtain/</th>
<th>IDOnsLar</th>
<th>IDLar</th>
<th>*Lar</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obtain</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. optain</td>
<td>*</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/misbehave/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. misbehave</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d. mizbehave</td>
<td>*</td>
<td>!</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>/bob/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. bob</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. bop</td>
<td>*</td>
<td>!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tableau 4.17. Stage 2 - Path 1: No RVA and FD – target-like English (Swanson 2001:28)**
In the tableau above, the faithfulness constraints IDOnsetLar and IDLar are ranked higher than markedness constraints *Lar and Agree; therefore, the input voicing values surface unaltered as in (a), (c) and (d).

However, if the learners compare forms mi[sb]behave and mi[zb]behave, as in (d) and (e) (cf. Tableau 4.11.), they may notice only the lack of regressive voicing in English. This will lead the learners to demote Agree below *Lar, which will eliminate the assimilation to the following voiced obstruent, which, however, does not have the effect of preventing FD, as shown in Tableau 4.18. below.

<table>
<thead>
<tr>
<th>/obtain/</th>
<th>IDOnsLar</th>
<th>*Lar</th>
<th>Agree</th>
<th>IDLar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obtain</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. optain</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/misbehave/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. misbehave</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. mizbehave</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/bob/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. bob</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| f. bop | | | | *

**Tableau 4.18. Stage 2 - Path 2: Regressive devoicing and FD (Swanson 2001:29)**

The revised re-ranking IDOnsLar >> *Lar >> Agree, IDLar will eliminate the assimilation to the following voiced obstruent, thus preventing candidate (d) from surfacing. However, assimilation to following voiceless obstruents, as in (b) is still allowed because *Lar rules out faithful candidates. Also because *Lar is ranked above IDLar, word-final devoicing will also occur (candidate (f)).

Since the learners produce non-native forms with word-final devoiced obstruents (candidate (f)) and with regressive voicing in word-medial obstruent clusters (candidate (b)), their next step could be either comparing forms o[pt]ain and o[bt]ain, or comparing forms bo[p] and bo[b]. These two comparisons will result in
different re-rankings. The comparison of the *o[p]tain-o[b]tain* pair (Path (2a) in Swanson) will lead to the demotion of *Lar and Agree below IDLar. This demotion results in the same English-like ranking shown in Tableau 4.12., IDOnsLar >> IDLar >> *Lar, Agree.

However, if forms *bo[p]* and *bo[b]* are noticed and compared first (Path (2b)), only *Lar is demoted below IDLar. This would result in a different ranking, and thus different Stage 3 of Path (2b), shown in Tableau 4.19.

<table>
<thead>
<tr>
<th>/obtain/</th>
<th>IDOnsLar</th>
<th>Agree</th>
<th>IDLar</th>
<th>*Lar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obtain</td>
<td>*</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>a. obtan</td>
<td>*</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>/misbehave/</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. misbehave</td>
<td>*</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>d. misbehave</td>
<td>*</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>/bob/</td>
<td>*</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>e. bob</td>
<td>*</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>f. bop</td>
<td>*</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

Tableau 4.19. Stage 3 - Path 2b: Regressive devoicing only (Swanson 2001:30)

This re-ranking results in the elimination of final devoicing, as in (e). However, regressive devoicing still occurs, as demonstrated by the candidate pair (a) and (b). Candidate (a) violates *Lar, whereas (b) does not, and thus (b) wins. However, as Swanson argues, learning still proceeds, because the winner – loser pair (a) - (b) can be compared next. This comparison will lead to the demotion of Agree below IDLar. The resulting Stage 4 is target-like English ranking IDOnsLar >> IDLar >> *Lar, Agree, shown in Tableau 4.17.

The third and final possible path described by Swanson is exhibited when the learners notice the lack of final devoicing in English by comparing forms such as *bo[p]* and *bo[b]* (cf. Tableau 4.17.). As a result, *Lar is demoted below IDLar in Stage 2 of this path, as shown in Tableau 4.20. below.
In this ranking *Lar is ranked below IDLar, and final devoicing is eliminated. As a result, candidate (e) wins. High-ranking Agree prevents candidates containing clusters disagreeing in voicing to surface. Therefore, non-target (b) and (d) emerge as optimal output forms. The comparison of the forms in (a) and (b), or in (c) and (d) force the same re-ranking: Agree being demoted below IDLar. This results in the target English ranking IDOonsLar >> IDLar >> *Lar, Agree, demonstrated in Tableau 4.17.

If the learnability theory proposed by Tesar & Smolensky (1993, 1996) is correct, then, as Swanson (2001) argues, there are four different paths available to Polish L1 – English L2 speakers in eliminating word-final devoicing and RVA. She argues further that Polish speakers learn the lack of word-final devoicing and RVA in English in the same manner as they learned their L1, by demoting markedness constraints alone. She then proceeds to the analysis of English L1 speakers acquiring L2 Polish, which is described in the next section

4.3.2. English speakers learning L2 Polish

Unlike Polish L1 - English L2 speakers, English L1 speakers learning L2 Polish acquire the less-marked properties of Polish. Swanson (2001) assumes again that the initial state of the L2 is the L1 final ranking. The English ranking proposed by

<table>
<thead>
<tr>
<th></th>
<th>Agree</th>
<th>IDOonsLar</th>
<th>IDLar</th>
<th>*Lar</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obtain</td>
<td>*!</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. obtain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. misbehave</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>d. misbehave</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. bob</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>f. bop</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

Tableau 4.20. Stage 2 - Path 3: RVA but no FD (Swanson 2001:30)
Lombardi (1999) accounts for the lack of RVA and word-final devoicing: IDLar >> IDOnsLar >> *Lar, Agree, is repeated in Tableau 4.21. below.

<table>
<thead>
<tr>
<th>/obtain/</th>
<th>IDOnsLar</th>
<th>IDLar</th>
<th>*Lar</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obtain</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. optain</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/misbehave/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. misbehave</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>d. mizbehave</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/bob/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. bob</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>f. bop</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tableau 4.21. Target-like English (Swanson 2001:28)**

Ranking faithfulness constraints IDOnsLar and IDLar over markedness constraints *Lar and Agree ensures that faithful candidates (a), (c), and (e) surface as optimal outputs, and the unfaithful, albeit unmarked, candidates (b), (d), and (f) are banned.

In previous research on L2 acquisition, L2 speakers were shown to transfer L1 processes to L2. More specifically, Polish L1-English L2 speakers were shown to apply RVA in English clusters (Rubach 1983, 1984). Likewise, English speakers were found to erroneously maintain a voicing contrast in L2 Polish (Sussex 1981). It must be remembered that Swanson assumes the learners’ final L1 state to be the initial state of their L2 ranking. Therefore, in the early stages of acquisition, hypothetical Polish L2 speakers are predicted to employ the English L1 constraint ranking when producing Polish obstruent clusters and word-final obstruents, as shown in Tableau 4.22. below.

<table>
<thead>
<tr>
<th>/sad/ ‘orchard’</th>
<th>IDOnsLAR</th>
<th>IDLar</th>
<th>*Lar</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sad</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. sat</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ogrudka/ ‘garden’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ogrudka</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>d. ogrutka</td>
<td></td>
<td>*!</td>
<td>**</td>
<td>*</td>
</tr>
</tbody>
</table>
As the data in the tableau show, high-ranking faithfulness constraints IDOnsLar and IDLar rule out unfaithful candidates (b), (d), (f), and (g), which change voicing specification of the input obstruents. Only faithful, but sub-optimal, candidates (a), (c), and (e) are allowed to surface.

As discussed in the previous section, language learners notice mismatches between the target and the produced form by comparing different winner-loser pairs. Depending on which pair is noticed first, several possible paths will be available to English L1 – Polish L2 learners. Because the re-ranking proceeds in the same manner as for Polish L1 – English L2, Swanson (2001) does not provide the detailed re-ranking, but offers the summary of paths and stages of acquisition, which is repeated in Table 4.4.

Table 4.22. English 2 – Polish 1, Stage 1: No RVA and no FD - English-like (Swanson 2001:32)

<table>
<thead>
<tr>
<th>Path</th>
<th>Learner notices:</th>
<th>Constraint rankings</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2</td>
<td>FD or VA to voiceless</td>
<td>Stage 1: IDOnsLar, IDLar &gt;&gt; *Lar, Agree</td>
<td>No RVA or FD (English)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 2: IDOnsLar &gt;&gt; *Lar, Agree &gt; IDLar</td>
<td>VA to voiceless &amp; FD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 3: IDOnsLar &gt;&gt; Agree &gt;&gt; IDLar, *Lar</td>
<td>VA &amp; variation in FD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 4: IDOnsLar &gt;&gt; Agree &gt;&gt; *Lar &gt;&gt; IDLar</td>
<td>VA &amp; FD (Polish)</td>
</tr>
<tr>
<td>3</td>
<td>VA to voiced</td>
<td>Stage 1: IDOnsLar, IDLar &gt;&gt; *Lar, Agree</td>
<td>No VA or FD (English)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stage 2: IDOnsLar</td>
<td>VA &amp; variation in</td>
</tr>
<tr>
<td>FD</td>
<td>&gt;&gt;Agree&gt;&gt; IDLar, *Lar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 3: IDOnsLar &gt;&gt; Agree</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;&gt; *Lar &gt;&gt; IDLar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FD</td>
<td>VA &amp; FD (Polish)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.4. Summary of predicted paths and stages in English L1 – Polish L2 (Swanson 2001:33)**

In the initial English L1 – Polish L2 ranking IDOnsLar, IDLar >> *Lar, Agree, faithfulness outranks markedness. Therefore, as Swanson points out, if demotion takes place faithfulness constraints will have to be demoted. However, in L1 acquisition, markedness outranks faithfulness, and only demotion of markedness takes place in child acquisition (Smolensky 1996 and others). Given this observation, Swanson argues that L2 acquisition may differ from L1 acquisition in that learners must realize that faithfulness constraints could be demoted, although they never had to do so when learning L1. As re-rankings in Table 4.3. show, on some occasions, the faithfulness constraint IDLar is demoted. Swanson takes this as evidence that English L1 – Polish L2 speakers cannot learn Polish by demoting only markedness constraints.

There are two other paths that English L1 – Polish L2 speakers can take in acquiring word-final devoicing and RVA. Paths 1 and 2 converge: learners may first notice word-final devoicing (Path 1), or VA to a following voiceless obstruent (Path 2). Both of these options will result in the same re-ranking in Stage 2. According to Swanson, however, noticing RVA first (Path 3) can eliminate one stage in learning, although this prediction has not been tested empirically.

### 4.3.3. Summary

In her analysis, Swanson (2001) claims to have demonstrated that Tesar and Smolensky’s EDCD (1993, 1996) learning algorithm can predict different paths and
stages within each path in L2 acquisition. Although Swanson’s analysis appears to be viable, it relied on the Polish data collected by others (Dziubalska-Kolaczyk 1990; Rubach 1983, 1984; Sussex 1981, cited in Swanson 2001), which did not include the fine-grained phonetic analyses of a language such as those presented in Chapters Two and Three of this dissertation for Russian.

Although neither Lombardi (1999) nor Swanson’s (2001) work addresses Russian RVA specifically in L1 or L2, their analyses appear to be excellent starting points for analyzing my own data. The conditioning environment for RVA in Polish is exactly the same as in Russian, and it is not unreasonable to assume that Russian L2 learners will behave similarly to Polish L2 speakers and follow the same learning stages and paths as described by Swanson (2001). It also remains to be seen whether the constraints used in Lombardi’s and Swanson’s analyses are straightforwardly applicable to Russian. As mentioned earlier, Padgett investigated RVA in Russian L1 following Lombardi’s work, so his work will be reviewed next.

Although the main focus of Padgett’s analysis is the problem of the Russian voiced fricative [v], which undergoes RVA but does not cause it itself, Padgett does address word-final neutralization and RVA in Russian obstruent clusters.

I will translate Padgett’s work into my own L1 and L2 data, and show that a straightforward adoption of his analysis into an L2 framework proposed by Swanson (2001) does not capture the L2 learner data. I will instead adopt the GLA (Boersma & Hayes 2001) in order to account for the results obtained in the phonetic experiments.

---

14 Another well-known work on voicing in Russian is by Hayes (1984). It addresses the issue within the generative tradition and focuses on the devoicing of word-final sonorants, which is outside the scope of the present work.

15 As explained in Chapter One, obstruent clusters containing [v] were excluded from the analysis, so this aspect of Padgett’s work will not be addressed here.
4.4. Analysis of RVA in Russian L1

4.4.1. Basic facts about Russian voicing as outlined in Padgett (2002)

Following Lombardi’s (1999) work, Padgett (2002) views RVA in Russian as a result of faithfulness and markedness constraint interaction, and he employs the same set of constraints as utilized by Lombardi with several exceptions and modifications. Below are the basic facts of Russian voicing assimilation reviewed by Padgett (2002).

(a) Obstruents devoice word-finally:

sle[d]a ‘track’ (Gen.Sg.) sled - sle[t] (Nom.Sg.)

kni[g]a ‘book’ (Nom.Sg.) knig - kni[k] (Gen.Pl.)

(b) Obstruents in a word-final cluster devoice:

poezd - poe[st] ‘train’

izb – i[sp] ‘hut’ (Gen.Pl.)

(c) Obstruent clusters within a word Agree in voicing:

podpisat[j] - po[t]pisat[j] ‘to sign’

podzhech – po[d]zhech ‘to set on fire’

podnesti - po[d]nesti ‘to bring to’

As the data in (c) demonstrate, when a voiced obstruent is followed by a voiceless one, both surface as voiceless. When both obstruents in a cluster are voiced, they are produced as such. Finally, when a voiced obstruent is followed by a sonorant, it preserves its voicing. Padgett argues that sonorants in Russian do not participate in voicing processes. Sonorants neither trigger RVA, as shown in (d) below, nor undergo it themselves (e)\(^{16}\).

\(^{16}\) Although some analyses of Russian do suggest that word-final sonorants devoice (Bondarko 1998; Hayes 1984), the work presented in this dissertation does not address this issue.
pis\textsuperscript{\textdagger}mo – pi[s\textsuperscript{\textdagger}]mo ‘letter’

(d) volk – vo[l]k ‘wolf’

4.4.2. *OT* analysis of Russian RVA

It should be recalled that Lombardi (1999) considers an onset to be the most salient position, in which an onset consonant retains its underlying voicing specification.

Padgett (2002) following Steriade (1997), claims that in Russian the key position is not onset position, but rather a position before a sonorant that retains the underlying voice specification. His argument proceeds as follows: in Russian, obstruent clusters must agree in voicing even when all obstruents are in the syllable onset, as in words like *gde* ‘where’ and *kto* ‘who’. Also, in onset clusters derived due to cliticization of prepositions like /k/ as in *k Dime* [gd]ime ‘to Dima’, RVA occurs. Yet, if onset positions in Russian could indeed license distinctive voicing, ill-formed words like *[kde] rather than *gde* ‘where’ should be fine. In order to account for this fact, Padgett (2002) proposes a constraint IDENT\textsubscript{PS} (VOICE) in lieu of Lombardi’s IDENT\textsubscript{OnsetLaryngeal}, which is consistent with the observations shown in (d) and (e) above and includes obstruents before vowels.

\begin{enumerate}
  \item IDENT\textsubscript{PS} (VOICE) – an output obstruent in position before a sonorant and its input correspondent have identical values for the feature [voice].
\end{enumerate}

IDENT\textsubscript{PS} (VOICE) belongs to a more general family of faithfulness constraints

IDENT (VOICE):

\begin{enumerate}
  \item IDENT (VOICE) – output and input obstruents must have identical values for the feature [voice].
\end{enumerate}
Since voiced obstruents are cross-linguistically disfavored in comparison to their voiceless counterparts, Padgett posits a constraint *D (similar to Lombardi’s *Lar) to account for this fact.

(7) *D- voiced obstruents are prohibited.

He also proposes a slightly modified constraint favoring assimilation, AGREE(VOICE)CG, which is made more specific than a general constraint AGREE, as it establishes the domain in which this constraint operates in Russian; namely, a clitic group. A clitic group is a prosodic unit that encompasses a prosodic word and includes enclitics, but is smaller than a phonological phrase. The fact that in Russian voicing assimilation applies across enclitic boundaries (Burton & Robblee 1997) allows Padgett to make such a claim. The examples in (f) and (g), involving a noun followed by an enclitic, are adapted from Padgett (2002:4).

(f) /sok/ ‘juice’ [sog] [ʒe] ‘juice (emph.)’ vs. [sok] [to] ‘juice (topic)’
(g) /sad/ ‘garden’ [sad] [ʒe] ‘garden (emph.)’ vs. [sat] [to] ‘garden (topic)’

In both examples, the word-final obstruents assimilate the voicing specification of the obstruent following the enclitic if the voicing specification differs. However, their voice specification remains the same if adjacent obstruents are either voiced or voiceless. Therefore, the constraint domain is a clitic group.

(8) AGREE(VOICE)CG- within a clitic group obstruents agree in [voice] specification.

Padgett proposes the following constraint ranking for obstruent clusters in Russian: IDENTPS, AGREE(VOICE)CG >> *D >> IDENT(VOICE), which is illustrated in Tableau 4.23. below. Although the initial obstruent in the input form /gto/ is specified for the wrong voicing value, as allowed by Richness of the Base
(proposed by Prince & Smolensky 1993), it surfaces with the correct value as in /kto/.

This is a result of undominated IDENT<sub>PS</sub> and AGREE(VOICE)<sub>CG</sub>, because the underlying specification of a presonorant obstruent must be preserved, and a preceding obstruent must agree with this voicing specification.

<table>
<thead>
<tr>
<th>/gto/ ‘who’</th>
<th>IDENT&lt;sub&gt;PS&lt;/sub&gt;</th>
<th>AGREE(VOICE)&lt;sub&gt;CG&lt;/sub&gt;</th>
<th>*D</th>
<th>IDENT(VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. gto</td>
<td></td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. kto</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. gdo</td>
<td>*!</td>
<td></td>
<td>**</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.23. Voicing assimilation in onset clusters: Russian (Padgett 2002:9)

In this ranking, AGREE(VOICE)<sub>CG</sub> dominates IDENT(VOICE), as assimilation overrides the underlying voice specifications. Undominated IDENT<sub>PS</sub> and AGREE(VOICE)<sub>CG</sub> ensure that the underlying voicing specification of a pre-sonorant obstruent is preserved, which is true of candidates (a) and (b), whose voicing specification before the vowel is maintained. The same constraints also ensure that adjacent obstruents within a clitic group agree in voicing (true of candidates (b) and (c)). Candidate (b) is the only one that satisfies both of these constraints, and thus surfaces as the optimal candidate. Candidates (a) and (c) each violate IDENT<sub>PS</sub> and AGREE(VOICE)<sub>CG</sub> respectively, and are ruled out.

Crucial ranking of *D and IDENT(VOICE) with respect to each other becomes apparent when devoicing of word-final clusters in forms like /pojezd/ ‘train’ are considered (See Tableau 4.24. below).

<table>
<thead>
<tr>
<th>/pojezd/ ‘train’</th>
<th>IDENT&lt;sub&gt;PS&lt;/sub&gt;</th>
<th>AGREE(VOICE)&lt;sub&gt;CG&lt;/sub&gt;</th>
<th>*D</th>
<th>IDENT(VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. pojezd</td>
<td></td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>b. pojest</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>c. pojezt</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.24. Voicing assimilation in word-final clusters: Russian (Padgett 2002:10)
As this dataset demonstrates, the optimal candidate (b), which contains a devoiced word-final cluster, violates only the lowest-ranked constraint IDENT(VOICE). Candidate (a), although satisfying AGREE, has two violations of *D, because it contains two voiced obstruents and is ruled out. Candidate (c) does not show assimilation in a word-final cluster and is also ruled out.

4.4.3. Summary

In this section, I have reviewed work by Padgett (2002) on L1 Russian RVA in OT. Padgett employs the same set of categorically ranked constraints as proposed by Lombardi (1998), albeit with some modifications that allow for the facts of Russian. However, his Russian data include exclusively word-initial and word-final obstruent clusters. In the next section, I extend Padgett (2002) and Swanson’s (2001) analyses to RVA in word-medial clusters in both Russian L1 and L2, showing that such an analysis fails to account for the data here, which is followed by my own analysis based on the model proposed in Boersma (1997) and Boersma & Hayes (2001).

4.5. RVA in Russian L1 and L2

4.5.1. Constraint interaction and acquisition of RVA

Padgett’s (2002) work on voicing assimilation in Russian L1 and Swanson’s (2001) investigation of Polish L1-English L2 reviewed in the previous sections demonstrated how an OT analysis of RVA that uses categorical constraints could proceed. In the following sub-sections, I will demonstrate that the constraints and categorical constraint ranking of the sort proposed in their work is not quite capable of extending to the gradient and gradual L2 acquisition of RVA observed in the acoustic experiments discussed in Chapters Two and Three. I will also present an analysis of
gradient RVA and its gradual acquisition, which is based on the GLA model proposed by Boersma & Hayes (2001). It should be recalled again that the GLA presupposes that multiple outputs and free variation in L1 are triggered by variable constraint ranking. I will follow this postulate in my analysis of variability in L2 production. I will use the constraint set proposed by Padgett (2002).

4.5.1.1. Categorical RVA and categorical L2 acquisition

I will begin by applying the constraints proposed by Padgett (2002) for Russian word-final and word-initial clusters to the form ‘obsharit’ – o/pʃ/arıt ‘to rummage’, which contains a word-internal obstruent cluster. This word was used as a stimulus in Experiment One. In this form, a fully-voiced stop /b/ devoices followed by a voiceless fricative /ʃ/. Recall that Padgett’s proposed ranking is IDENTPS, AGREE(VOICE)CG > *D >> IDENT(VOICE) (cf. Tableaux 4.23. & 4.24.). The data for a native speaker of Russian is presented in Tableau 4.25. below.

<table>
<thead>
<tr>
<th>/obsharit/ ‘rummage’</th>
<th>IDENTPS</th>
<th>AGREE(VOICE)CG</th>
<th>*D</th>
<th>IDENT(VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obsharit</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. opsharit</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. obzhariit</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.25. RVA in word-internal obstruent clusters: Russian L1

I use the constraint AGREE(VOICE)CG, proposed by Padgett (2002), which applies in a clitic group. As Tableau 4.21. demonstrates, constraints AGREE(VOICE) and IDENTPS are not crucially ranked with respect to each other. They both work to rule out two less optimal candidates, (a) and (c). Candidate (a) contains a cluster disagreeing in voicing; whereas candidate (c) contains a voiced cluster. Although candidate (b) has a violation of IDENT(VOICE), it still surfaces as an optimal output in Russian L1, as this constraint is the lower-ranked one in the set. Thus, the RVA
ranking hierarchy proposed by Padgett (2002) for word-initial and word-final clusters works for word-medial clusters, as well, and will be assumed as the Russian L1 adult final ranking in the further analysis of RVA below.

As the tableau above shows, the constraint IDENT<sub>PS</sub> targets forms such as o[bʒ]arit that exhibit fully-voiced obstruent clusters as a result of progressive voicing assimilation. Since forms in which progressive VA has applied were not found in either L1 or L2 production in the phonetic experiments, for the purposes of simplicity, I will omit such forms from further analysis and exclude IDENT<sub>PS</sub> from the ranking, as well. Likewise, the constraint *D which mainly targets word-final voiced obstruents (cf. Tableau 4.18.) is also excluded from the analysis because candidates containing voiced obstruents, as in (a) in the tableau above, are banned by a higher-ranked AGREE(VOICE). Therefore, the crucial ranking for Russian RVA is

\[
\text{AGREE(VOICE)}_{\text{CG}} \gg \text{IDENT(VOICE)}, \text{ shown in Tableau 4.26.}
\]

<table>
<thead>
<tr>
<th>/obsharit/</th>
<th>AGREE(VOICE)&lt;sub&gt;CG&lt;/sub&gt;</th>
<th>IDENT(VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obsharit</td>
<td>*†</td>
<td></td>
</tr>
<tr>
<td>b. opsharit</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.26. RVA in word-medial clusters in Russian L1: categorical ranking of AGREE(VOICE)<sub>CG</sub> >> IDENT(VOICE)

Following Boersma and Hayes (2001), I schematically represent the interaction of these constraints on a ranking scale in Figure 4.1. below.

Figure 4.1. Categorical ranking of AGREE(VOICE)<sub>CG</sub> >> IDENT(VOICE): Russian L1.
This ranking scale demonstrates, that in Russian L1, $\text{AGREE(VOICE)}_{CG}$ and $\text{IDENT(VOICE)}$ are kept apart at a safe distance and are strictly ranked with respect to each other\textsuperscript{17}. As a result, only assimilated obstruents surface in word-medial clusters.

Given Padgett’s focus, his analysis did not provide the English L1 ranking that is required for my data. Although Lombardi (1999) analyzes voicing typology in English (cf. Section 4.2.1), the review of Padgett’s analysis made it clear that Lombardi’s work needed revision to account for the Russian facts. The following tableau presents English L1 data based on Lombardi’s ranking, using, however, Padgett’s constraints for consistency.

The ranking hierarchy for unrestricted voice in English is $\text{IDENT(VOICE)} >> \text{AGREE(VOICE)}_{CG}$, is shown in 4.27. below.

<table>
<thead>
<tr>
<th>/pigpen/</th>
<th>$\text{IDENT(VOICE)}$</th>
<th>$\text{AGREE (VOICE)}_{CG}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. pigpen</td>
<td>$\ast$</td>
<td>$\ast$</td>
</tr>
<tr>
<td>b. pikpen</td>
<td>$\ast$!</td>
<td></td>
</tr>
<tr>
<td>c. pigben</td>
<td>$\ast$!</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.27. Unrestricted voice: English

In the English ranking, faithfulness constraints dominate markedness constraints. Unfaithful candidates (b) and (c), which fatally violate $\text{IDENT(VOICE)}$, are ruled out. The faithful candidate (a), although violating $\text{AGREE(VOICE)}_{CG}$, satisfies high-ranked IDENT (VOICE) and emerges as the surface form.

As was done for the interaction of these constraints in Russian, these constraints are represented on a ranking scale in Figure 4.2.

\textsuperscript{17} Since the GLA scale is arranged in arbitrary units, what counts as a ‘safe’ distance is not made explicit in Boersma & Hayes (2001). However, as they put it: “This distance will be enough that the probability of generating Candidate 2 will become essentially nil, and the resulting grammar will generate Candidate 1 essentially 100% of the time” (Boersma & Hayes 2001:8).
As discussed above, in L2 phonological acquisition research, language learners are assumed to use their native ranking as a starting point in L2 acquisition (Hancin-Bhatt 2000; Swanson 2001). English L1 final ranking is taken then as Russian L2 initial ranking. Therefore, I will consider IDENT(VOICE) >> AGREE(VOICE) as the initial stage in the L2 acquisition of RVA. This is shown in Tableau 4.28. As before, the pointing finger is used to indicate an optimal L1 form; whereas, an unhappy face indicates a sub-optimal form produced by an L2 speaker.

<table>
<thead>
<tr>
<th>/obsharit/ ‘rummage’</th>
<th>IDENT(VOICE)</th>
<th>AGREE(VOICE)_{CG}</th>
</tr>
</thead>
<tbody>
<tr>
<td>☋a. obsharit</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>☋b. opsharit</td>
<td>!</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.28. Categorical acquisition: No RVA in Russian L2, at the starting point of English L1 speakers’ acquisition

In the English ranking, faithfulness over markedness accounts for the lack of RVA. Candidate (b) exhibits RVA, and is in fact the optimal L1 parse. However, it is ruled out by the IDENT(VOICE) constraint which dominates AGREE(VOICE)_{CG}. As a result, a faithful candidate (a), containing a cluster disagreeing in voice, is erroneously chosen as optimal in Russian L2.

In an ideal world, Russian L2 learners would re-rank existing L1 constraints as soon as they have noticed the application of RVA in native speech, which provides evidence for re-ranking. They would immediately demote IDENT(VOICE) which would result in the target Russian ranking: AGREE(VOICE)_{CG} >> IDENT(VOICE). Figure 4.29. shows the results of L1 constraints being re-ranked.
In this tableau, the high-ranked AGREE(VOICE)\textsubscript{CG} rules out the faithful candidate (a) which contains a cluster disagreeing in voicing. The remaining candidate (a'), where RVA applies, surfaces as an optimal one, as in Russian L1.

However, the phonetic experiments presented in Chapters Two and Three demonstrated that RVA was not complete in either native or non-native production. More importantly, learners devoiced obstruents in the assimilating environment to a lesser degree than native speakers (See Section 4.5.2.below for discussion). That is, although some devoicing did take place in L2 production, it was not native-like, with advanced learners performing slightly better than beginners. Therefore, the analysis of RVA in Russian L1 and L2 must be capable of capturing both the notions of gradual acquisition of RVA and the gradient nature of RVA. Although the OT analysis provided in this section demonstrates how the re-ranking of English L1 categorical constraints can proceed, it does not embrace either of these notions. Therefore, the notions of gradient devoicing and gradual acquisition should be introduced to the analysis. I will present such a revised analysis in the next section.

### 4.5.1.2. Variable ranking of AGREE(VOICE)\textsubscript{CG} and IDENT(VOICE) in L2

As argued in the literature on L2 phonological acquisition (Broselow et al. 1998; Hancin-Bhatt 2000), the re-ranking of L1 constraints in L2 takes place when more and more input becomes available to a language learner. Following the same rationale, it could be expected that after having received a sufficient amount of input, Russian L2 learners will demote highly-ranked IDENT(VOICE) below AGREE(VOICE)\textsubscript{CG}.

<table>
<thead>
<tr>
<th>/obsharit/ 'rummage'</th>
<th>AGREE(VOICE)\textsubscript{CG}</th>
<th>IDENT(VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obsharit</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. opsharit</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

**Tableau 4.29. Categorical acquisition: RVA in Russian L2**
It is not unreasonable to assume that during the re-ranking process, when the constraints are pushed along the ranking scale, the distance between them decreases and the strict dominance of one constraint over the other is violated. Boersma & Hayes (2001) conjecture that the short distance between two constraints suggests that the relative ranking of these constraints is less fixed, and at some point they might overlap. Given that in the GLA the grammar is assumed to be stochastic, this will lead to the variable ranking of constraints and thus in variability in the production of C1 in the assimilating environment. The acoustic analysis demonstrated that it was indeed the case that unlike L1 speakers, L2 learners exhibited much more variability in their production of RVA.

Recall that in Boersma & Hayes (2001) constraints are associated with ranges of values, and when the ranges overlap, free ranking takes place. More specifically, a selection point (the value used at evaluation time) could be selected from anywhere within the ranges of the constraints in question. I assume that once L2 learners have noticed RVA in obstruent clusters, re-ranking is initiated, and AGREE(VOICE)CG is pushed along the scale towards IDENT(VOICE) until the two constraints overlap. In the GLA, this process leads to two possible results. First, as shown in Figure 4.3. below, IDENT(VOICE) would outrank AGREE(VOICE)CG, for despite the overlap, the selection point associated with IDENT(VOICE) (represented by number 1 on the scale) is taken from within the IDENT(VOICE) range. (The selection point associated with AGREE(VOICE)CG is assigned number 2.) Therefore, although the process of voice neutralization has been triggered, it has not resulted in any observable effects at this stage. That is, IDENT(VOICE) still outranks
AGREE(VOICE)\textsubscript{CG}, and no noticeable voice neutralization in obstruent clusters in L2 production has taken place.

![Diagram](image)

**Figure 4.3. Variable ranking of IDENT(VOICE) >> AGREE(VOICE)\textsubscript{CG} in English L1-Russian L2, common result**

However, if the selection points are taken from the upper part of the AGREE(VOICE)\textsubscript{CG} range and the lower part of the IDENT(VOICE) range, AGREE(VOICE)\textsubscript{CG} would outrank IDENT(VOICE), as shown in Figure 4.4. As above, number 1 corresponds to IDENT(VOICE), and number 2 corresponds to the lower-ranking constraint AGREE(VOICE)\textsubscript{CG}.

![Diagram](image)

**Figure 4.4. Variable ranking of AGREE(VOICE)\textsubscript{CG} >> IDENT(VOICE) in English L1-Russian L2, less common result**

Under this scenario, although IDENT(VOICE) is still higher up on the ranking scale, constraint AGREE(VOICE)\textsubscript{CG} outranks it. As a result, obstruent clusters agreeing in \([\text{voice}]\) emerge in L2 production.

Following Boersma & Hayes (2001), I assume that the re-ranking process shown in Figures 4.1 – 4.4., refers to a single evaluation time. Over the series of evaluations, the overlapping ranges of values of AGREE(VOICE)\textsubscript{CG} and IDENT(VOICE) will result in variable ranking of these constraints, and multiple outputs for the same underlying form will surface. Therefore, this analysis only shows whether an output is categorically voiced or voiceless. However, the results of phonetic experiments demonstrated a great deal of variability in C1 devoicing both in native or non-native production (see Section 4.5.1.). In the next section, I propose the
analysis of RVA in Russian L1 and L2, whose goal is to account for gradient
devoicing and gradual acquisition.

4.5.2. Gradient RVA and gradual L2 acquisition of RVA

Recall that in the phonetic experiments voicing proportion (duration of voicing during
a segment divided by segment duration) was found to be the most important acoustic
indication of voice contrast. The analysis of voicing duration in both real-word and
nonsense experiments revealed that all groups of subjects produced obstruents within
the assimilating-to-voiceless environment with various degrees of devoicing.

However, the averages presented in Chapters Two and Three do not indicate whether
speakers produce, for example, 30% of tokens in some category with complete
voicing and the rest with no voicing at all, or whether they produce every token with
30% of the consonant’s duration voiced, or something in between. Another possibility
is that 30% of the speakers produce complete voicing nearly consistently, while 70%
produce complete devoicing nearly consistently. That is, is assimilation variable
token-to-token, or even within a token, or both, and how consistent is it across
speakers? Data showing gradient devoicing within single tokens in both L1 and L2
production is shown in Figures 4.5. and 4.6. below. Each bar in the figure represents
the amount of voicing within tokens in C1 stops and C1 fricatives in the three voicing
environments: assimilating-to-voiceless, fully-voiceless, and fully-voiced. The box
plot contains the middle 50% of the data. The upper edge of the box indicates the 75th
percentile of the data set, and the lower edge indicates the 25th percentile. Asterisks
represent outliers, cases with values between 1.5 and 3 box lengths from the upper or
lower edge of the box. Circles represent extreme cases with values more than 3 box
lengths from the upper or lower edge of the box.
Figure 4.5. Real-word experiment. The amount of voicing within token in C1 stops and fricatives in three voicing environments, assimilating-to-voiceless, fully-voiceless, and fully-voiced in a real-word experiment. A horizontal line across the box represents the median value of the data. The box plot contains the middle 50% of the data. The upper edge of the box indicates the 75th percentile of the data set, and the lower edge indicates the 25th percentile. Asterisks and circles represent extreme cases and outliers.
Figure 4.6. Non-word experiment. The amount of voicing within token in C1 stops and fricatives in three voicing environments, assimilating-to-voiceless, fully-voiceless, and fully-voiced in a non-word experiment. A horizontal line across the box represents the median value of the data. The box plot contains the middle 50% of the data. The upper edge of the box indicates the 75th percentile of the data set, and the lower edge indicates the 25th percentile. Asterisks and circles represent extreme cases and outliers.

The results presented in Figures 4.5. and 4.6. are also presented in Table 4.5. The table shows the highest averaged percentage of C1 voicing in both phonetic experiments in all three voicing environments, assimilating-to-voiceless, fully-voiceless, and fully-voiced.
<table>
<thead>
<tr>
<th>Language level</th>
<th>C1 stop</th>
<th></th>
<th>C1 fricative</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assimilating</td>
<td>Voiced</td>
<td>Voiceless</td>
<td>Assimilating</td>
</tr>
<tr>
<td>Native</td>
<td>35%</td>
<td>98%</td>
<td>25%</td>
<td>12%</td>
</tr>
<tr>
<td>Advanced</td>
<td>65%</td>
<td>95%</td>
<td>15%</td>
<td>15%</td>
</tr>
<tr>
<td>Intermediate</td>
<td>70%</td>
<td>98%</td>
<td>22%</td>
<td>15%</td>
</tr>
</tbody>
</table>

Table 4.5. The highest percentage of residual voicing in C1 stops and fricatives in the assimilating-to-voiceless, fully-voiceless, and fully-voiced environments, in the three language groups

The distribution of C1 voicing values within the three voicing environments obtained in the real-word and nonsense-word experiments is also presented in histograms in Figures 4.7. and 4.8., respectively. Of particular interest is the distribution in the assimilated (devoiced) obstruents.

**Figure 4.7.** Histogram showing the frequency distribution of the voicing proportion of C1 stops and fricatives in three voicing environments, assimilating-to-voiceless, fully-voiceless, and fully-voiced in all language groups in the real-word experiment

As Figure 4.7. demonstrates, in the real-word experiment, all groups of subjects cover a wide range of voicing values in their assimilation environment tokens, but with most of their tokens concentrated into relatively small acoustic ranges. In particular,
L1 values are concentrated in the nearly-voiceless area, indicated by a low proportion of voicing. Non-native values are concentrated at both ends, high and low, of the x-axis.

Figure 4.8. The histogram showing the frequency distribution of the voicing proportion of C1 stops and fricatives in three voicing environments, assimilating-to-voiceless, fully-voiceless, and fully-voiced in all language groups in the nonsense-word experiment.

As in the data shown in Figure 4.7., the C1 voicing proportion distribution in the nonsense-word experiment is also characterized by a wide range of voicing values in all groups of subjects. Both native and non-native subjects produced fully-devoiced and fully-voiced obstruents in the assimilating environment. The data in the top row show how speakers realize consonants in the assimilating environment. However, these graphs still do not allow us to determine whether speakers behave consistently relative to each other. Figures 4.9.- 4.11. below address this issue for the data obtained in the real-word experiment; whereas, Figures 4.12.- 4.14. present the nonsense-data. The discussion of the C1 voicing patterns is to follow.
4.5.2.1. Frequency distribution of C1 voicing proportion in the assimilating environment

Figures 4.9. - 4.14. below present the C1 voicing distribution for individual speakers in the assimilating environment in each language group separately, in real and non-word experiments. It should be noted that due to the differing number of items in the real and nonsense experiments, the frequency scale for the nonsense tokens extends up to 20 whereas in real tokens, it goes up to 8 only.
Figure 4.9. Histogram showing the frequency distribution of the voicing proportion of C1 stops and fricatives in the assimilating environment in the native production in the real-word experiment, for each speaker separately (where 1-10 are the native speakers). Within each speaker’s box, the left side of the box is tokens with 0% of the consonant voiced, and the right side of the box is tokens with 100% of the consonant voiced.
Figure 4.10. Histogram showing the frequency distribution of the voicing proportion of C1 stops and fricatives in the assimilating environment in the advanced learners’ production in the real-word experiment (where 11-20 are advanced L2 learners). Within each speaker’s box, the left side of the box is tokens with 0% of the consonant voiced, and the right side of the box is tokens with 100% of the consonant voiced.
Figure 4.11. Histogram showing the frequency distribution of the voicing proportion of C1 stops and fricatives in the assimilating environment in the intermediate learners’ production in the real-word experiment (where 21-30 intermediate L2 learners). Within each speaker’s box, the left side of the box is tokens with 0% of the consonant voiced, and the right side of the box is tokens with 100% of the consonant voiced.
Figure 4.12. Histogram showing the frequency distribution of the voicing proportion of C1 stops and fricatives in the assimilating environment in the native production in the nonsense-word experiment (where 1-10 native speakers). Within each speaker’s box, the left side of the box is tokens with 0% of the consonant voiced, and the right side of the box is tokens with 100% of the consonant voiced.
Figure 4.13. Histogram showing the frequency distribution of the voicing proportion of C1 stops and fricatives in the assimilating environment in the advanced learners’ production in the nonsense-word experiment (where 11-20 advanced L2 learners). Within each speaker’s box, the left side of the box is tokens with 0% of the consonant voiced, and the right side of the box is tokens with 100% of the consonant voiced.
The analysis of C1 voicing proportion in the assimilating-to-voiceless environment in real and nonsense-stimuli has revealed several patterns. First, it was found that some speakers, both native and non-native, produced a relatively large number of completely assimilated (devoiced) tokens. For example, among native speakers, subjects 6, 7 and 10 in the real words, and subjects 2, 4, 5, and 7 in the nonsense

---

18 Because of the scarcity of the real-word data for each speaker, performing a statistical analysis was not possible, and all the production patterns were defined by-eye.
words, assimilated fairly completely (cf. Figures 4.9 & 4.12). Among L2 subjects, speakers 18 and 29 exhibited a similar pattern in the nonsense tokens. A trivial amount of closure voicing observed in some of these subjects’ productions is attributable to the fact that the vocal cords do not stop vibrating instantaneously.

Second, a number of subjects applied assimilation categorically within a token but variably across tokens by either devoicing C1 (nearly) completely or maintaining its underlying voicing (nearly) completely. That is, speakers 15, 17, 19, 20, 21, and 30 in the real-word experiment (cf. Figures 4.10 & 4.11), and speakers 19, 23 and 24, in the nonsense tokens (cf. Figures 4.13. & 4.14.), produced both fairly voiced and fairly devoiced C1, but rarely produced something in between.

Third, some subjects (3, 9, 11, 13, and 16 in the real tokens, and 6, 10, 15, 16, 22, 25, 27, 28, and 30 in the nonsense tokens, cf. Figures 4.9. – 4.14.) demonstrated non-categorical production by not favoring complete assimilation or absence of assimilation. They had voicing values spread out over the whole voicing range, with little concentration of tokens in any range.

Fourth, a number of subjects produced a similar pattern to the speakers described above in that they also had a wide range of voicing values in their production. However, contrary to the group above, they preferred complete C1 assimilation or non-assimilation, while still producing many partially assimilated tokens as well. These were non-native speakers 14, 26, 28, and 29 in the real-word experiment, and 1, 3, 19, 23, and 24 in the nonsense tokens (cf. Figures 4.10 – 4.14.).

To summarize, the analysis of the C1 voicing proportion data obtained in the real and nonsense-word experiments has revealed the following production patterns for the groups numbered above:
(1) speakers who produced the majority of tokens with assimilated (devoiced) C1;
(2) speakers who produced both assimilated and non-assimilated C1, but few tokens in between;
(3) speakers who produced tokens with C1 voicing values spread out over a wide range of voicing values with no preference for assimilation or lack thereof;
(4) speakers who produced tokens with C1 voicing values spread out over a wide range of voicing values, but with a preference for complete assimilation or non-assimilation.

Overall, all language groups demonstrated a great deal of variability within a token and among tokens in their C1 production. Intermediate speakers were more likely to maintain more voicing in the assimilated C1 than the other language levels, with advanced subjects assimilating and failing to produce equally often and, and the natives showing a strong bias toward assimilating. While more of the natives belong to sub-group 1 that favors complete devoicing, and sub-group 2 consists of L2 learners exclusively, with quite a few of the intermediate learners belong to sub-groups with a lot of within-token variability, there are members of each language group in most or all of the sub-groups. This finding suggests the possibility that subjects possess different production (sub)grammars. I will return to this idea and its implications for phonological theories in Chapter Five.

Some speakers, both native and non-native, behaved differently in the nonsense words than in the real words. I assume that this finding are attributable to task effects whereby natives, much like non-natives, were led by Russian orthography when facing an unfamiliar lexical item. Arguably, L1 speakers create an underlying form for a non-word based on their interpretation of orthography. Nonetheless, native
speakers are familiar with real-word tokens having heard them a number of times in the past. Nonsense tokens, on the other hand, were completely unfamiliar and did not have meanings. Although the influence of orthography can explain the production of some L1 speakers in the nonsense-word experiment, it can perhaps account only for some of the L2 production. On the whole, L2 participants were fairly inexperienced with even real words and, therefore, it is debatable whether they were perplexed and uncomfortable with made-up items, as much as native speakers.

These types of questions beg the larger question of whether frequency may play some role in the pattern of results observed (e.g., low frequency words and nonsense words may be treated similarly by some speakers); however, as noted earlier, frequency was not controlled in the real word experiment, so further study is needed to determine the likelihood of this possibility. In the short term, I assume that the L2 production patterns are, in fact, indicative of their acquisition processes as reflected by the process of constraint re-ranking discussed in the next section.

4.5.2.2. Partially-devoiced obstruents in Russian L2

It has been established that all subjects, both native and non-native, exhibited a great deal of variability in their production of devoiced obstruents in the assimilating-to-voiceless environment. Some natives, in some tokens, devoiced C1 almost completely, with a residual voicing proportion of, or slightly higher than, 0%. However, the highest averaged percentage of residual voicing maintained by native speakers is 35% (cf. Table 4.5.). Based on this result, I assume that 0% - 35 % is the target value for a voiceless obstruent in Russian L1. As the analysis of obstruents in the fully-voiceless environment demonstrated, fully-voiceless C1 contained up to 25% voicing (cf. Table 4.5.).
More importantly, no statistical difference was found between native C1 in the fully-voiceless and assimilated C1. Although these percentages of voicing clearly indicate that voicing neutralization was not complete in L1 speech, they are, nevertheless, native values. Hence, I assume that for an obstruent to belong to a voiceless category in Russian L1, it could contain any percentage of voicing between 0% and 35%. Based on the natives’ behavior, I propose to treat the range from 0% to 35% as a target range for L2 speakers. In other words, the percentage of voicing, which is low enough to render the target obstruent as voiceless in native speech, is placed at any point on the scale ranging from 0% to 35%. It is possible that the target range can, in fact, be even lower than 35% if the token produced with this value was a speech error or an outlier. However, using 35% as a target value allows for L2 learners to attend and process any token that they heard from a native speaker. In general, averaged devoiced values for a native subject could be placed at the range, at any point between 0% and 35%.

Likewise, both groups of non-native subjects devoiced C1 to various degrees and on the average maintained up to 70% of voicing in C1 in the assimilating environment. This value, not low enough to approximate native norms of devoicing, is nevertheless significantly lower than voicing proportion produced in fully-voiced C1 (up to 98%), thus indicating partial devoicing in L2 speech. (Note that the native voicing value of fully-voiced C1 is also near 100%.) Therefore, L2 residual voicing values could be represented as occurring at any point between 0% and above 35% (cf. Table 4.5.). However, no devoiced C1 containing up to 70% of voicing was found in L1 production, which can be taken as an indication that in L2, all the occurrences beyond 35% manifest failure to produce RVA in a native-like manner. Therefore, if
L2 learners maintain a voicing value of, or below 35%, such learners are said to have approximated native values, and thus have produced RVA successfully. I also assume that an obstruent produced with the voicing value placed at any point between the ‘no assimilation’ range of 70% - 100% is intended as voiced.

Thus, non-natives were shown to apply a wider range of voicing values, up to 70%, producing C1 which did not belong to either the native voiceless/devoiced class /p/ or voiced class /b/. In order to account for this finding, I propose treating the partially-devoiced obstruents as a non-native phonetic realization of a Russian voiceless category. Similarly to how in the initial stages of L2 acquisition language learners operate L1 constraint rankings, they also employ L1 acoustic values in the production of L2 segments and overall utilize non-native versions of target sounds. That is, I assume that L2 learners are ‘aware’ of the acoustic difference between Russian and English; however, they have not been able yet to modify their L1 phonetic categories in order to accommodate a new L2 category (Flege 1984, 1987; Flege & Port 1981; MacKay et al., 2001).

Therefore, the specific voicing values with which such a partially-devoiced obstruent is made are placed on the acoustic continuum ranging from 70% (the highest C1 voicing proportion in the assimilating environment produced by non-natives) to 0% (fully-voiceless). I assume that a partially-devoiced non-native C1 changes over time as a result of language learners’ increased practice and exposure to the target language. Thus, the L2 phonetic realization of the L1 phonological category depends on the learners’ knowledge of Russian voicing values available through processing input data.
As the experimental results show, gradient within-token values, although not very frequent in the real-word data, were nevertheless present. They were common in the nonsense word data, and common for some speakers (groups 3 and 4 above) in the real-word data. Therefore, in order to correctly account for the production data of both L1 and L2 speakers, the notion of incomplete, or gradient, devoicing must be introduced into any analysis of RVA in Russian in addition to the notion of categorical (de)voicing. In order to account for gradiency in the data, I will employ the concept of continuous constraint families, analogous to Boersma’s (1997, 1998) perceptual *WARP constraints and Jun’s (1995, 2004) PRESERVE sub(place) constraints in the production grammar (to be reviewed in the next sections). In addition, I will employ the GLA by Boersma & Hayes (2001) (cf. Section 4.1.2.) in order to account for gradual L2 acquisition.

4.5.3. Phonetically detailed constraint families

4.5.3.1. A family of constraints in the perception grammar (Boersma 1997, 1998)

In his work, Boersma (1997, 1998) focuses on the issue of the acquisition of a perception grammar whose task “is to abstract raw acoustic material, with its dependence on the age, sex, physiology, and state of the speaker, on room acoustics and the weather, and on some more random-like causes of variation, into a more reproducible (probably discrete) representation that is more suited for lexical access and that can, for purposes of learning, be compared with the output of the listener’s own production grammar” (Boersma 2000:2). According to Boersma, the perception grammar consists of constraints that classify the acoustic input into a finite number of perceptual categories, such as a family of *WARP constraints.
*WARP \( f: x, y \): An acoustic value \( x \) on a perceptual tier \( f \) is not categorized into the category whose centre is at \( y \).

(Boersma 1997:49)

The *WARP constraints are locally ranked: *WARP \( f: x_1, y \) \( >> \) *WARP \( f: x_2, y \) in order to ensure that “a less distorted recognition is preferred over a more distorted recognition” provided that \( x_1 \) and \( x_2 \) are on the same side of the category center \( y \) (ibid.). Boersma further considers perceptual categorization along a continuous auditory dimension whose values range from \([0]\) to \([100]\) in a language which has three contrastive categories /30/, /50/, and /70/ along this dimension. The *WARP constraint family is represented on a continuous auditory dimension as shown in Figure 4.11., adapted from Boersma (1997:49). A simulation of a possible L1 initial state with unbiased categorization shown in the figure below is performed by Boersma by using the Praat script.

![Diagram](image)

**Figure 4.15. The *WARP family on the continuous auditory dimension: a possible initial state (adapted from Boersma 1997:49)**

To show the interaction of these *WARP constraints in the listener’s categorization system, Boersma considers the hypothesized input datum of /44/. The ranking shown
in Tableau 4.30. below is reproduced from Boersma (1997). The listener is assumed to have three categories, /30/, /50/, and /70/, and the perception grammar determines the optimal candidate.

<table>
<thead>
<tr>
<th>[44]</th>
<th>*WARP ([44], /70/)</th>
<th>*WARP ([44], /30/)</th>
<th>*WARP ([44], /50/)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/30/</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>q/50/</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>/70/</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.30. The interaction of the *WARP constraints in the perception grammar (Boersma 1997:50)

Boersma bases his explanation of the ranking of the three *WARP constraints presented in the tableau above on the figure reproduced in 4.11. In the figure, the dotted line at [44] goes from the bottom of the figure up. First, it goes through the *WARP (x, /50/) curve, then through the *WARP (x, /30/) curve, and finally through the *WARP (x, /70/) curve. The *WARP (x, /50/) curve is the lowest of the three curves for x [44], and the acoustic input is categorized into the /50/ category. Given the shape and location of the curves “the discrimination criteria are obviously at [40] and [60], and if evaluation is not stochastic, these criteria are hard: every input above [60] is classified as /70/, every input below [40] as /30/, and every other input as /50/” (Boersma 1997:50).

Boersma further considers variations within and between speakers which result in random distributions of the acoustic input to the listener’s ear. As before, it is assumed that a language has three categories with midpoints at [30], [50], and [70] along a perceptual dimension, and a problematic three-way contrast: the middle category is weaker than the others by virtue of having fewer lexical occurrences. The distribution of a speaker’s productions, which are the inputs to the listener’s
perception grammar, which can be obtained by running a Praat script, is shown in the Figure 4.12. (adapted from Boersma 1997:50).

\[
\text{Produced acoustic value}
\]

**Figure 4.16. Production distributions of the three categories /30/, /50/, and /70/ (adapted from Boersma 1997:50)**

Boersma further argues that the listener will make the fewest mistakes in initial categorization if s/he uses the criterion of *maximum likelihood*. That is, the learner will choose the perceptual category \( y \) that maximizes the a posteriori probability, if the received acoustic input is \( x \).

According to Boersma, if the acoustic input is [44], an optimal listener will choose the /30/ category because the curve of the distribution of the production of /30/ shown in Figure 4.12. is placed above the curve associated with the production of the category /50/, although the value [44] is nearer to the midpoint of the /50/ category than to the midpoint of /30/. As a consequence, the listener will initially categorize all inputs below the criterion [45.5] into the class /30/, all the values between [45.5] and the second criterion [54.5] into the class /50/, and all values above [54.5] into the class /70/.
The categorization continuum shown in Figure 4.12. does not reflect what the speaker’s intended category is. Therefore, if a learner gets to know (in the recognition phase) which category was intended, s/he may take a learning step. For example, the speaker meant to produce the /30/ category but the learners perceived it as the /50/ category (cf. 4.30.). The learner learns that s/he made a categorization error and will have to demote the offending constraint, *WARP ([44], /30/). In the GLA (cf. Section 4.1.2.), constraint demotion is done by lowering the ranking of the offending constraint by a small amount (Boersma suggests an arbitrary amount of 0.01) along the continuous ranking scale. The resulting ranking is shown in 4.31. below.

<table>
<thead>
<tr>
<th>[44]</th>
<th>*WARP ([44], /70/)</th>
<th>*WARP ([44], /50/)</th>
<th>*WARP ([44], /30/)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/30/</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>/50/</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>/70/</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.31. The interaction of the *WARP constraints in the perception grammar (reproduced from Boersma 1997:51)

In this tableau, although the intended input is [30], the actual input is [44]. The previously high-ranked constraint WARP ([44], /30/) is lowered below WARP ([44], /50/). It is now ranked the lowest among the three *WARP constraints corresponding to the three contrastive categories along the auditory dimension. Such a ranking allows the listener to perceive the input [44] in the intended /30/ class.

To sum up, Boersma represents constraints as phonetically detailed constraint families, with the family members placed along a continuum of some phonetic property. In his analysis, similar properties of phonetics and phonology are derived from the same constraints in a perception grammar. In the following section I will
review the work by Jun (1995, 2004) who adopts a similar approach in her analysis of gradient place assimilation and utilizes the concept of a constraint family.

4.5.3.2. A family of constraints in the production grammar (Jun 1995, 2004)

Jun (1995, 2004) analyzes place assimilation in consonant clusters. Place assimilation occurs when one of the two neighboring consonants takes on the place of articulation of another (e.g., /tk/ → /kk/). The stop /t/ is completely assimilated by /k/. However, in gradient assimilation, a gesture corresponding to C1 in a cluster is preserved in the surface form. Thus, the incomplete nature of assimilation is due to the attempt to preserve remnants of the perceptual cues for the original segment. The cue preservation is ensured by the constraint PRESERVE(place), which preserves perceptual cues for input features and can be decomposed into a set of ranked constraints differentiated by the amount of place of articulation information.

\[
\text{PRESERVE}_n(\text{place}): \text{Preserve at least } n \text{ per cent of the perceptual cues for place of articulation, where } 1 \leq n \leq 100.
\]

Jun (2004:80)

The constraint PRESERVE_{100}(place) requires maximum preservation of the perceptual cues for place. A complete closure gesture is produced when this constraint is high-ranked. However, PRESERVE_1(place) … PRESERVE_{99}(place) require preservation of lesser degrees of perceptual cues and, when high-ranked, ensure the production of incomplete closure. In order to guarantee that at least some of the perceptual information is preserved when the maximum preservation is not possible, Jun proposes the following ranking: PRESERVE_1(place) >> … PRESERVE_{50}(place) … >> PRESERVE_{100}(place). Likewise, other constraints could be broken down into families of continuous constraints. For example, the constraint WEAKENING is yet another constraint responsible for place assimilation (Jun 2004:80):
WEAKENING\textsubscript{m}: Do not produce an articulatory gesture whose effort cost is at least \textit{m}.

The universal ranking of the decomposed constraints is: WEAKENING\textsubscript{1x} $> >$ …

WEAKENING\textsubscript{0.5x} … $> >$ …. WEAKENING\textsubscript{0.1x}, where \textit{1x} is the effort cost required for the production of a complete closure gesture. Jun shows how various reduction patterns of the target segment emerge from the proposed families of constraints by considering ten ranked constraints selected from each of the PRESERVE and WEAKENING families: \{PRES\textsubscript{10}(place), PRES\textsubscript{20}(place) … PRES\textsubscript{100}(place)\} and \{WEAK\textsubscript{1x}, WEAK\textsubscript{0.9x} … WEAK\textsubscript{0.1x}\}. The constraints WEAKENING and PRESERVE are in conflict: PRES\textsubscript{100}(place) requires the maximum preservation of the perceptual cues for place, which results in a complete closure; whereas, WEAK\textsubscript{1x} bans the occurrence of a complete closure gesture. Jun assumes that the constraint ranking is dependent on the speech rate and style. That is, the ranking of WEAKENING relative to PRESERVE is increased as a result of faster and more informal speech and is decreased as a result of slower and more formal speech. Table 4.6. below, reproduced from Jun (2004:81) demonstrates how the interaction of WEAKENING and PRESERVE produce different reduced forms of the target segment in place assimilation.
In the table above, each column represents the ranking of WEAKENING and PRESERVE for a given speech style and rate. The higher the constraint in the table, the higher its ranking. For example, in the first column, corresponding to the most formal and slowest speech, all PRESERVE constraints outrank all WEAKENING constraints, which results in no reduction of the target segment. In the second column which represents less formal and slow speech \( W_{1x} \) outranks \( P_{100} \), which prevents a full closure gesture. Yet, \( P_{90} \) is still ranked higher than \( W_{0.9x} \), and the gesture reduction must not result in the loss of more than 10% of perceptual cues. Thus, an optimal output is the form with at most 10% reduction of the target gesture. As the speech rate becomes faster and the style becomes more informal, the ranking of WEAKENING
constraints becomes higher relative to the PRESERVE family constraint. As a result, the degree of reduction in the output forms also increase. The last column (the most informal and fastest speech) shows that all WEAKENING constraints outrank all PRESERVE constraints which results in an output with zero closure.

Thus, the analysis proposed by Jun (1995, 2004) not only produces forms with full and completely reduced closure, but is also capable of generating semi-reduced forms found in gradient place assimilation. More importantly, the proposed analysis is based on phonetic research, and ‘the proposed constraints and their universal rankings are determined by principles and properties that are supported empirically by research in speech perception and production’ (Jun 2004:83).

To sum up, the approach proposed by Boersma (1997, 1998) and Jun (1995, 2004) employing the notion of continuous constraint families which cover multiple values is capable of accounting for gradiency in speech production and perception. Overall, given the nature and empirically-proven basis for such constraint families, it is not unreasonable to assume that such constraint families could be applied to the analysis of the gradient data obtained in the acoustic experiments presented here as well. I will turn to this point in the following section.

4.5.4. Gradient RVA and gradual L2 acquisition

In this section, I will incorporate the notion of phonetically-detailed constraint families into the analysis of RVA production in L1 and L2. More specifically, I will break down the existing AGREE(VOICE)\textsubscript{CG} constraint into a series of constraints which will allow me to account for various degrees of devoicing produced by the subjects.
4.5.4.1. A continuous constraint family approach

Recall that in the analysis of RVA proposed in Section 4.5.1., I adopted constraints AGREE(VOICE)\textsubscript{CG} and IDENT(VOICE) put forward by Padgett (2002). I repeat these constraints below:

\begin{align*}
\text{AGREE(VOICE)}\textsubscript{CG} &- \text{within a clitic group obstruents agree in [voice] specification.} \\
\text{IDENT(VOICE)} &\text{–An output obstruent and its input correspondent have identical value for the feature [voice].}
\end{align*}

Recall that the ranking AGREE(VOICE)\textsubscript{CG} >> IDENT\textsubscript{PS}(VOICE) was shown to account for categorical assimilation. However, as discussed in previous sections, the phonetic experiments presented in Chapters Two and Three demonstrated that RVA was not complete in either native or non-native production. Not only L2 learners devoiced obstruents in the assimilating environment less often than natives, but they did do to a lesser degree as well (cf. Section 4.5.3.), with advanced learners performing somewhat better than beginners. In order to account for gradient devoicing, I will re-formulate AGREE(VOICE)\textsubscript{CG} as:

\begin{align*}
\text{AGREE\textsubscript{C1},C2(VOICE)}\textsubscript{CG} &- \text{within a clitic group obstruents agree in [voice] to within } n \text{ percent.}
\end{align*}

The $n$ value depends on how much acoustic detail a human ear needs to hear in order to discriminate between two obstruents, one of which contains $n\%$ voicing. The exact value of $n$ can be decided by a psychoacoustic study of Just Noticeable Difference for this dimension, but for the purposes of this analysis, I represent $n$ as arbitrary portions of 5%. The constraint \{AGREE\textsubscript{C1},C2(VOICE)\textsubscript{CG}\}\textsuperscript{19} could thus be represented as:

\textsuperscript{19} From hereafter I will enclose AGREE\textsubscript{C1},C2(VOICE)\textsubscript{CG} in the curly brackets in order to distinguish it as a family of continuous constraints from a regular constraint.
AGREEC1_{5\%} C2(VOICE)_{CG} >> \ldots \text{ AGREEC1}_{35\%} C2(VOICE)_{CG} >> \ldots \\
AGREEC1_{70\%} C2(VOICE)_{CG} \ldots >> \text{ AGREEC1}_{100\%} C2(VOICE)_{CG}.

As before, I assume that English L1 – Russian L2 speakers’ initial ranking is:
IDENT(VOICE) >> \{AGREEC1_n C2(VOICE)_{CG}\}. The constraint family
\{AGREEC1_n C2(VOICE)_{CG}\} >> IDENT(VOICE) is the target ranking for L2 learners.
I will analyze each ranking corresponding to a particular sub-group presented in
Section 4.5.2.1. in turn.

Sub-group 1. I assume that in the process of acquisition, L2 learners eventually re-rank individual constraints within the \{AGREEC1_n C2(VOICE)_{CG}\} constraint family, and constraints gradually slide along the acoustic continuum toward the target values.
As a result, their production of RVA changes over the course of acquisition, and after
a series of re-rankings within the \{AGREEC1_n C2(VOICE)_{CG}\} constraint family, the
degree of devoicing of C1 lowers from the initial non-native percentage to the one
that falls within the native percentage range. A series of tableaux below (4.28. – 4.33.)
demonstrates the interaction of the \{AGREEC1_n C2(VOICE)_{CG}\} constraint family and
IDENT(VOICE). Several permutations of the members of the
\{AGREEC1_n C2(VOICE)_{CG}\} family illustrate how various outputs can be obtained
under each respective ranking. Tableau 4.32. represents RVA in Russian L1, with
\{AGREEC1_n C2(VOICE)_{CG}\} ranked above IDENT(VOICE). This ranking is fixed.

<table>
<thead>
<tr>
<th>/obsharit/ ‘to rummage’</th>
<th>{AGREEC1_n C2(VOICE)_{CG}}</th>
<th>IDENT(VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. obsharit</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. opsharit</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.32. RVA in Russian: \{AGREEC1_n C2(VOICE)_{CG}\} >> IDENT(VOICE).
Devoiced C1 is the optimal output.
In the tableau above, candidate (a) which contains a fully-voiced C1 /b/ is ruled out under the current ranking. Candidate (b), the native form, satisfies the high-ranked constraint \{AGREE_{C1,n}C2(VOICE)_{CG}\} which ensures that an assimilated C1, containing the native voiceless value, is produced. This ranking alone accounts for the categorical production resulting in a voiceless output, as in sub-group 1 described in Section 4.5.2.1.

Sub-group 2. The next case I will consider is when the process of constraint re-ranking has been initiated, and the ranking ranges of \{AGREE_{C1,n}C2(VOICE)_{CG}\} and IDENT(VOICE) overlap. As a result, the selection points are chosen from anywhere within the ranges of the two constraints (Boersma & Hayes 2001), as shown in Figures 4.14. and 4.15. below. Figure 4.14. presents the case when at the evaluation time the selection points are taken from the upper part of IDENT(VOICE) and the lower part of \{AGREE_{C1,n}C2(VOICE)_{CG}\}. Number 1 corresponds to the selection point within IDENT(VOICE), and number 2 corresponds to lower-ranked \{AGREE_{C1,n}C2(VOICE)_{CG}\}. As a result, the following ranking holds:

\[
\text{IDENT(VOICE)} \gg \{\text{AGREE}_{C1,n}C2(VOICE)_{CG}\}.
\]

Figure 4.17. Variable ranking of IDENT(VOICE) and AGREE(VOICE)_{CG} in English L1- Russian L2: IDENT(VOICE) \gg AGREE(VOICE)_{CG}

The data presented in Figure 4.14. corresponds to the ranking under which a fully-voiced obstruent surfaces in L2 shown in Tableau 4.33. below.
Tableau 4.33. The lack of RVA: \{AGREE1\textsubscript{a}C2(VOICE)\textsubscript{CG}\} and IDENT(VOICE) ranking ranges overlap. Voiced C1 surfaces as optimal.

In this dataset, a faithful candidate (a) containing a fully-voiced C1 emerges as the optimal form; whereas, candidate (b) containing a voiceless obstruent is banned by IDENT(VOICE). This ranking in conjunction with the ranking presented in Tableau 4.32. accouns for the production in which assimilation applied completely categorically within a token but variably among tokens, as in sub-group 2, discussed in Section 4.5.2.1. That is, due to the constraint overlap, either fully-assimilated C1 (cf. Tableau 4.32.) or fully-voiced C1 (cf. Tableau 4.33.) were shown to surface in the same speaker.

Figure 4.15. presents the case when at the evaluation time the selection points are taken from the upper part of \{AGREE1\textsubscript{a}C2(VOICE)\textsubscript{CG}\} and the lower part of IDENT(VOICE). As above, number 1 corresponds to IDENT(VOICE), and number 2 corresponds to \{AGREE1\textsubscript{a}C2(VOICE)\textsubscript{CG}\}. As a result, \{AGREE1\textsubscript{a}C2(VOICE)\textsubscript{CG}\} >> IDENT(VOICE).

Figure 4.18. Variable ranking of IDENT(VOICE) and AGREE(VOICE)\textsubscript{CG} in English L1-Russian L2: AGREE(VOICE)\textsubscript{CG} >> IDENT(VOICE)

The data in Figure 4.15. correspond to the ranking which results in assimilated C1 shown in Tableau 4.32.
According to the GLA, with learners’ further exposure to learning data and after rounds of re-ranking, constraints will be eventually ranked at a safe distance relative to each other. The desirable outcome of such re-ranking is having categorically ranked constraints that ensure that only devoiced C1 surface. Recall that such a pattern was obtained in predominantly native production (cf. Section 4.5.2.1.). However, even some native speakers were found to produce assimilation non-categorically, with voicing values spread out over the whole voicing range, similarly to L2 speakers. I turn to the discussion of these findings and the analysis of the two remaining production patterns (sub-groups 3 and 4) in the following section.

4.5.4.2. Incorporating the continuous constraint family \{AGREEC1_nC2(VOICE)_{CG}\} in the analysis

I assume that after having been exposed to the ambient language, L2 learners start noticing the discrepancy between the input devoiced C1 in obstruent clusters and a voiced C1 in their own production. They take a learning step and demote IDENT(VOICE) below \{AGREEC1_nC2(VOICE)_{CG}\}. Although this categorical ranking prevents voiced C1 from surfacing, it does not account for variability exhibited by L2 speakers. It should be also recalled that the production of a group of speakers, both native and non-native, was characterized by gradient devoicing in C1 (cf. patterns three and four described in Section 4.5.2.1.). In order to account for these facts, I decompose \{AGREEC1_nC2(VOICE)_{CG}\} into a set of phonetically-detailed constraints, as proposed above: \(AGREEC1_{5\%}C2(VOICE)_{CG} >> \ldots \)

\(AGREEC1_{35\%}C2(VOICE)_{CG} >> \ldots AGREEC1_{70\%}C2(VOICE)_{CG} \ldots >>\)

\(AGREEC1_{100\%}C2(VOICE)_{CG}.

Following Boersma (1997, 1998) and Jun (2004), I present the interaction of \{AGREE_{C1n}C2(VOICE)_{CG}\} constraints in the tableau below. Instead of using a conventional asterisk with an exclamation point to indicate a fatal violation, I will use a check mark instead in order to indicate a possible fatal violation under a current ranking.

<table>
<thead>
<tr>
<th>/obsharit/</th>
<th>AGREE \textsubscript{5%} (VOICE)</th>
<th>....</th>
<th>AGREE \textsubscript{35%} (VOICE)</th>
<th>....</th>
<th>AGREE \textsubscript{70%} (VOICE)</th>
<th>....</th>
<th>AGREE \textsubscript{100%} (VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 100%</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
</tr>
<tr>
<td>......</td>
<td></td>
<td>....</td>
<td></td>
<td>....</td>
<td></td>
<td>....</td>
<td></td>
</tr>
<tr>
<td>n = 70%</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td></td>
</tr>
<tr>
<td>......</td>
<td></td>
<td>....</td>
<td></td>
<td>....</td>
<td></td>
<td>....</td>
<td></td>
</tr>
<tr>
<td>n = 35%</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td></td>
</tr>
<tr>
<td>......</td>
<td></td>
<td>....</td>
<td></td>
<td>....</td>
<td></td>
<td>....</td>
<td></td>
</tr>
<tr>
<td>n = 5%</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td>\checkmark</td>
<td>....</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 4.34. The interaction of the constraints within the \{AGREE_{C1n}C2(VOICE)_{CG}\} family which results in various degrees of assimilation

In the tableau above, the input is a Russian form /obsharit/ which contains a cluster disagreeing in voicing. The leftmost column represents the percentage of voicing in each candidate. Recall that the $n$ is an arbitrary value in increments of 5%. The constraint row across the tableau represents the members of the continuous constraint family, \{AGREE_{C1n}C2(VOICE)_{CG}\}. Depending on the \{AGREE_{C1n}C2(VOICE)_{CG}\} constraint ranking, candidates containing various C1 voicing values could surface as optimal. For example, if AGREE\textsubscript{70\%}C2(VOICE)\textsubscript{CG} dominates all other constraints, only a candidate containing a partially-devoiced non-native C1 will surface. Likewise, if AGREE\textsubscript{35\%}C2(VOICE)\textsubscript{CG} happens to dominate the hierarchy, a candidate containing 35% voicing will emerge as optimal.
Sub-group 3. As previously discussed, in order for learners to obtain the target ranking, they should initiate the re-ranking within the \{AGREE_{C1,C2(VOICE)CG}\} constraint family. Recall that under the GLA model put forth by Boersma & Hayes (2001) the re-ranking of constraints takes place if the constraint ranges overlap, and variable ranking occurs depending on the location of the selection points. Therefore, I conjecture that depending on the location of selection points chosen during the evaluation time, of the two members placed close to each other, either one member of the constraint family or the other will come out on top of the hierarchy. For example, if \(\text{AGREE}_{35\%}(\text{VOICE})_{CG}\) and \(\text{AGREE}_{40\%}(\text{VOICE})_{CG}\) are variably ranked, in some cases \(\text{AGREE}_{35\%}(\text{VOICE})_{CG}\) will surface as optimal and in some other cases \(\text{AGREE}_{40\%}(\text{VOICE})_{CG}\) will be optimal (cf. Section 4.1.2.). I will use values of 35% and 40% as examples in the subsequent relevant tableaux. These two possible cases are presented in Tableaux 4.35. and 4.36. Tableau 4.31. below demonstrates the \{\text{AGREE}_{40\%}(\text{VOICE})_{CG} \gg \text{AGREE}_{35\%}(\text{VOICE})_{CG}\} \gg \text{IDENT(VOICE)}\) ranking.

<table>
<thead>
<tr>
<th>/obsharit/ ‘to rummage’</th>
<th>\text{AGREE}<em>{40%}(\text{VOICE})</em>{CG}</th>
<th>\text{AGREE}<em>{35%}(\text{VOICE})</em>{CG}</th>
<th>\text{IDENT}(\text{VOICE})</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (n = 40%)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. (n = 35%)</td>
<td>*!</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.35. RVA in Russian L2: \{\text{AGREE}_{40\%}(\text{VOICE})_{CG} \gg \text{AGREE}_{35\%}(\text{VOICE})_{CG}\} \gg \text{IDENT(VOICE)}\). Only a partially-devoiced C1 surfaces.

Candidate (a) satisfies \(\text{AGREE}_{40\%}(\text{VOICE})_{CG}\) by virtue of containing a partially-devoiced C1 \((n = 40\%)\) and emerges as an optimal output. Candidates (b) with a voiceless C1 \((n = 35\%)\) violates \(\text{AGREE}_{40\%}(\text{VOICE})_{CG}\) and is ruled out.

Tableau 4.36. presents the case of \{\text{AGREE}_{35\%}(\text{VOICE})_{CG} \gg \text{AGREE}_{40\%}(\text{VOICE})_{CG}\} \gg \text{IDENT(VOICE)}\).
The natively devoiced C1 in (b) which emerges as the optimal output contains no more than 35% of voicing and thus satisfies the highest-ranked constraint AGREE\textsubscript{35\% (VOICE)}\textsubscript{CG}. Candidate (a) contains 40% voicing and is banned by AGREE\textsubscript{35\% (VOICE)}\textsubscript{CG} under the current ranking.

Arguably, the rankings presented in tableaux 4.35. and 4.36. can account for a non-categorical production pattern in sub-group 3 (cf. Section 4.5.2.1.) characterized by a wide spread of voicing values and no preference for assimilation or non-assimilation. It is possible that the detailed constraints, members of the \{AGREE\textsubscript{C1\textsubscript{n},C2(VOICE)}\textsubscript{CG}\} constraint family, are ranked extremely close to each other and overlap to an extensive degree. Depending on the location of selection points, one of these closely-ranked, phonetically-detailed constraints will come out on top at one evaluation time, and a different one, at another evaluation time (cf. Figures 4.14. & 4.15).

Recall, however, that there was another production pattern discussed in Section 4.5.2.1. Sub-group 4 was characterized by both the wide spread of voicing values, as in the pattern discussed above, and the presence of complete assimilation or the absence of assimilation. I will now present the analysis of this production pattern.

Sub-group 4. Tableaux 4.35. and 4.36. demonstrate the ranking interaction with the assumption that IDENT(VOICE) is strictly ranked with respect to \{AGREE\textsubscript{C1\textsubscript{n},C2(VOICE)}\textsubscript{CG}\}. However, one can assume that in the process of re-

---

<table>
<thead>
<tr>
<th>/obsharit/</th>
<th>AGREE\textsubscript{35% (VOICE)}\textsubscript{CG}</th>
<th>AGREE\textsubscript{40% (VOICE)}\textsubscript{CG}</th>
<th>IDENT(VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. n = 40%</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. n = 35%</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 4.36. Target Russian L2 ranking: \{AGREE\textsubscript{35\% (VOICE)}\textsubscript{CG} >> AGREE\textsubscript{40\% (VOICE)}\textsubscript{CG} >> IDENT(VOICE)\}. A devoiced C1 surfaces.
ranking, IDENT(VOICE) can end up somewhere between the 
{AGREE1_nC2(VOICE)cG} constraints. For example, if the ranking is:
AGREE35%(VOICE)cG >> IDENT(VOICE) >> AGREE70%(VOICE)cG, an optimal 
native form in (c) will surface, as long as AGREE35%(VOICE)cG is the highest-ranked 
constraint in the hierarchy banning all faithful forms or forms containing partially-
devoiced obstruents (See Tableau 4.37. below).

<table>
<thead>
<tr>
<th>/obsharit/</th>
<th>AGREE35% (VOICE)cG</th>
<th>IDENT (VOICE)</th>
<th>AGREE70% (VOICE)cG</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. n = 100%</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. n = 70%</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. n = 35%</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.37. RVA in Russian L2: AGREE35%(VOICE)cG >> IDENT(VOICE) >> AGREE70%(VOICE)cG. A devoiced C1, optimal in Russian L1, surfaces.

Likewise, if AGREE70%(VOICE)cG is the highest-ranked constraint, as in 
AGREE70%(VOICE)cG >> IDENT(VOICE) >> AGREE35%(VOICE)cG, shown in 
Tableau 4.38., only partially-devoiced obstruents are allowed to surface. Forms with 
voiceless C1 and faithful forms are ruled out under this ranking. Overall, as long as 
one of the {AGREE1_nC2(VOICE)cG} constraints dominates the hierarchy, faithful 
forms containing fully-voiced obstruents will not surface and only C1 with some 
degree of devoicing will emerge as optimal.

<table>
<thead>
<tr>
<th>/obsharit/</th>
<th>AGREE35% (VOICE)cG</th>
<th>AGREE70% (VOICE)cG</th>
<th>IDENT (VOICE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. n = 100%</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. n = 70%</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. n = 35%</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 4.38. RVA in Russian L2: AGREE70%(VOICE)cG >> IDENT(VOICE) >> AGREE35%(VOICE)cG. A partially-devoiced C1 surfaces.

However, it is not unreasonable to assume that should the ranking in which the 
members of the {AGREE1_nC2(VOICE)cG} constraint family and IDENT(VOICE)
are variably ranked with respect to each other occur, at some evaluation time, a fully-voiced obstruent will emerge as a winning output (cf. Tableau 4.33.). Therefore, the production pattern with both assimilated and non-assimilated C1 and a wide range of voicing values can be accounted for a variable ranking of the
\{AGREEC1_nC2(VOICE)_{CG} \} constraint family and IDENT(VOICE).

To sum up, the proposed target ranking is thus as follows:

\begin{align*}
AGREEC1_nC2(VOICE)_{CG} \{ \text{AGREE}_{5\%}(VOICE)_{CG} \ldots & \gg \text{AGREE}_{35\%}(VOICE)_{CG} \ldots \\
& \gg \text{AGREE}_{70\%}(VOICE)_{CG} \ldots \gg \text{AGREE}_{100\%}(VOICE)_{CG} \} \gg \text{IDENT}(VOICE).
\end{align*}

As proposed by the GLA (Boersma & Hayes 2001), language learners begin the re-ranking of the existing constraint ranking after they have noticed the discrepancy between the forms in their own production and the forms in the ambient language. I assume that L2 learners re-rank their constraints in a variety of ways, according to various native grammars/set of rankings they are exposed to. The four production patterns discovered in the experiments suggest that at any given moment, L2 learners have one of the constraint rankings corresponding to one of those patterns.

Since L2 learners begin acquiring the target language using their native ranking, I assume that at first, L2 learners demote faithfulness constraint IDENT(VOICE), which militates against the changes in the voicing specification and high-ranked in their native grammar, below the family of continuous markedness constraints \{AGREEC1_nC2(VOICE)_{CG}\}.

Following Boersma & Hayes (2001), I also assume that while the re-ranking process is taking place, the two constraints \{AGREEC1_nC2(VOICE)_{CG}\} and IDENT(VOICE) could overlap, and variable ranking could occur which resulted in two possible rankings: \{AGREEC1_nC2(VOICE)_{CG}\} \gg IDENT(VOICE) or
IDENT(VOICE) >> \{AGREE_{1,n}C2(VOICE)_{CG}\}. This ranking accounted for the production pattern in sub-group 2 (cf. Section 4.5.2.1).

Likewise, I consider the case of the overlap of closely-ranked constraints AGREE_{35\%}(VOICE)_{CG} and AGREE_{40\%}(VOICE)_{CG}, with two possible outcomes: the L1 ranking of AGREE_{35\%}(VOICE)_{CG} >> AGREE_{40\%}(VOICE)_{CG} and L2 intermediate ranking of AGREE_{40\%}(VOICE)_{CG} >> AGREE_{35\%}(VOICE)_{CG} which results in the generation of non-natively devoiced C1. This ranking was relevant to the third production pattern, sub-group 3. The last production pattern in sub-group 4 (cf. Section 4.5.2.1.) could be accounted for by the variable ranking of IDENT(VOICE) and various members of the \{AGREE_{1,n}C2(VOICE)_{CG}\} family.

Overall, a unified phonetics-phonology framework employed in such a model frees it from the necessity to posit an additional, language-specific phonetic component. Instead, such phonetic details are included in phonological representations. Although the exact value of \(n\%\) is not known, it can be established by a psychoacoustic study to determine the Just Noticeable Difference for percent voicing of a consonant.

4.5.5. Summary

In this section, I have demonstrated that straightforward categorical ranking of the constraints responsible for RVA and the constraint of the AGREE(VOICE) type presented by Lombardi (1999) and Padgett (2002) fall short of handling variable outputs found in the L2 production of RVA. Following Boersma (1997, 1998) and Jun (1995, 2004), I proposed representing the AGREE(VOICE)_{CG} as a family of continuous constraints which encompass a range of voicing proportion values. These values included native voiceless and voiced categories and the L2 realization of the
Russian L1 voiceless category. This, along with the Boersma & Hayes’ (2001) GLA, incorporated into the analysis of RVA, allowed me to account for both gradiency in L2 production, demonstrated by partial devoicing of C1 in the assimilating environment, and gradual re-ranking and acquisition of L2 constraints.

Table 4.7. below summarizes all the production patterns and the constraint ranking they demonstrate, as discussed in the previous section.

<table>
<thead>
<tr>
<th>Production patterns</th>
<th>Corresponding rankings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-group 1: Mostly assimilated C1</td>
<td>Categorical ranking: {AGREE(C_1,v)C2(VOICE)(_{CG})} &gt;&gt; IDENT(VOICE)</td>
</tr>
<tr>
<td>Sub-group 2: Both assimilated and non-assimilated C1 with few tokens in between</td>
<td>Variable ranking: {AGREE(C_1,v)C2(VOICE)(_{CG})} and IDENT(VOICE)</td>
</tr>
<tr>
<td>Sub-group 3: Spread out C1 voicing values with no preference for assimilation or non-assimilation</td>
<td>Variable ranking of phonetically detailed constraints within {AGREE(C_1,v)C2(VOICE)(<em>{CG})}: AGREE(</em>{35%}(VOICE)<em>{CG}) &amp; AGREE(</em>{40%}(VOICE)_{CG})(^{20})</td>
</tr>
<tr>
<td>Sub-group 4: Spread out C1 voicing values with a preference for assimilation or non-assimilation</td>
<td>Variable ranking of {AGREE(C_1,v)C2(VOICE)(<em>{CG})} &amp; IDENT (VOICE): IDENT(VOICE) is ranked between phonetically detailed constraints within the {AGREE(C_1,v)C2(VOICE)(</em>{CG})} family</td>
</tr>
</tbody>
</table>

Table 4.7. Production patterns and corresponding constraint rankings

The analysis presented in this chapter can be extended to other languages and conditioning environments. The GLA model incorporated into the analysis of L2 production in conjunction with the concept of a continuous constraint {AGREE\(C_1,v\)C2(VOICE)\(_{CG}\)} family can address the issues of gradual acquisition of a specific phonological pattern, as well as gradiency in production. As discussed in Section 4.5.4.1., the lowest percent of voicing to which a C1 in an obstruent cluster has to devoice, \(n\), was assigned an arbitrary value of 5%. This value is an

\(^{20}\) As mentioned above, I use values of 35% and 40% as hypothetical examples. Any other value between 0% and 100% can be used instead.
approximation for however much detail human ears can hear on this specific dimension and should, in principle, be universal.

It is conceivable that this type of an analysis can be applied in a longitudinal study of an L2 acquisition of a particular phonological phenomenon in order to handle learning stages and paths. Furthermore, given the finding that devoicing is not complete, the re-ranking of the members within a detailed constraint family can account for various voicing patterns occurring cross-linguistically.

4.6. Conclusion

This chapter has addressed the phonology of RVA in Russian L1 and L2. I began by reviewing relevant OT literature on voicing neutralization in various L1 and L2, including Russian, as well as constraint learning algorithms, and demonstrated that the analysis of RVA in L2 based on traditional OT could not handle variability, both token-by-token and within token, found in the speech data. Recall that the acoustic experiments revealed two types of variability: variability in the degree of C1 devoicing in the assimilating environment and variability in the production of categorically devoiced (or voiced) C1. The four production patterns indicated by the histograms (cf. Section 4.5.2.1.) were the results of the combinations of these two types of variability. The production of categorically (de)voiced obstruents could be modeled within stochastic OT, which generated outputs that were either voiced or voiceless. Under the GLA, the L2 learners’ production of C1 in the assimilating environment was viewed as a result of the constraint re-ranking. However, this model could not encompass instances of partially-devoiced C1. Following Boersma (1997, 1998) and Jun (1995, 2004), I proposed a revised analysis which incorporated the
concept of a continuous family constraints, such as \{AGREEC1_{n}C2(VOICE)_{CG}\}, and showed how the re-rankings of the constraints within the family, as well as categorical and variable rankings of IDENT(VOICE) and \{AGREEC1_{n}C2(VOICE)_{CG}\} could encompass the four production patterns demonstrated in the acoustic experiments.

Overall, I assumed that after L2 learners have noticed a discrepancy between their own output and the input L1 data, a constraint re-ranking was initiated. After L2 learners obtained a fixed ranking of \{AGREEC1_{n}C2(VOICE)_{CG}\} >> IDENT(VOICE), they proceeded to re-rank the constraints within the \{AGREEC1_{n}C2(VOICE)_{CG}\}. As the amount of exposure to the target language and practice increased, L2 learners were able to attend to the fine phonetic details of the ambient data and modify their own output accordingly.

To conclude, the type of an analysis proposed in this chapter allowed me to account for both gradiently devoiced outputs and gradual L2 acquisition of Russian RVA, as well as for the several patterns of outcomes natives and learners demonstrated. In the following, final, chapter of this dissertation, I will summarize the results of the acoustic experiments, discuss them in the light of L2 acquisition theory, and conclude with the discussion of implications of this research project for L2 teaching.
CHAPTER FIVE

CONCLUSIONS AND DISCUSSION

5.0. Introduction

The goal of this dissertation was to investigate the acquisition of regressive voicing assimilation (RVA) in obstruent clusters in non-native Russian from phonetic and phonological perspectives. This goal was achieved by analyzing and comparing acoustic properties of RVA in Russian L1 and L2 in real and nonsense words (Chapters Two and Three) and accounting for the observed patterns using Optimality Theory (Prince & Smolensky 1993) and Boersma & Hayes’ (2001) Gradual Learning Algorithm (Chapter Four). In this, the final chapter of the dissertation, I will provide a general discussion of the acoustic experiments and phonological analysis. I will begin by summarizing the results of the real-word and non-word experiments and discussing phonetic and phonological effects of the L1 in the L2 speech. Next, I will discuss (incomplete) learning of the L2 phonological patterns and its implications for phonological theory. The conclusion, pedagogical implications, and future research questions will be presented in the final sub-sections.

5.1. Summary of experimental results

5.1.1. Findings of acoustic experiments

In the present investigation of the production of RVA in Russian, I followed methodology employed in the phonetic research on voicing (Burton & Robblee 1997; Jessen 2001; Lisker & Abramson 1961; Slowiaczek & Dinnsen 1985) and analyzed obstruent clusters in terms of C1 (the first obstruent in a cluster) and C2 (the second obstruent in a cluster) duration, C1 and C2 voicing proportion, VOT (Voice Onset
Time) of C2, and V1 (preceding vowel) and V2 (following vowel) duration\textsuperscript{21}. If RVA (regressive voicing assimilation) was present, a devoiced C1 was expected to be similar to a fully-voiceless C1 in terms of voicing and segment duration. Similarly, the duration of vowels was expected to be shorter before devoiced obstruents in the assimilating environment and after fully-voiceless obstruents. However, the results of the acoustic experiments demonstrated that the proportion of voicing in the C1 was the most reliable and consistent indication of RVA (to be discussed below).

Overall, both real and nonsense-word experiments revealed that native speakers of Russian produced RVA in both types of obstruent clusters almost consistently, by substantially reducing the voicing duration in a fully-voiced\textsuperscript{22} C1 in the assimilating-to-voiceless environment. In the real-word experiment, assimilated fully-voiced C1 (e.g., \textit{lozhka} lo[\textepsilon]ka ‘spoon’) was rendered similar to a fully-voiceless C1 in the fully-voiceless environment (\textit{moshka} mo[\textepsilon]ka ‘midge’) in most tokens, which was demonstrated statistically. In the nonsense-word experiment, although native speakers devoiced the C1 in the assimilating environment as well, they maintained more voicing than in C1 in the real-words. Although non-native speakers produced RVA as well, they did so to a lesser degree or less often than natives. In general, L2 performance was much more varied both within tokens and token-by-token, with advanced L2 learners performing better than intermediate speakers (cf. Section 4.5.2., Figures 4.5. and 4.6.).

\textsuperscript{21} The measurements of C2 and V2’s acoustic properties were taken in order to verify that these segments were not affected by the process of voicing assimilation, and thus regressive, rather than progressive, voicing assimilation has taken place. These measures also confirm that C1 is indeed voiceless, and hence can condition devoicing of C1.

\textsuperscript{22} The term ‘fully-voiced/voiceless’ has been used interchangeably with the term ‘underlyingly voiced/voiceless’. The underlying voice specification of a consonant is manifested by its behavior before a sonorant (Padgett 2002).
The analysis of the segment duration of both obstruents and vowels produced mixed results (cf. Sections 2.3 & 3.3). For example, in the nonsense items, the duration of V1 was not statistically different in the assimilating vs. fully-voiced environments, even in the native speakers’ group. The fact that a V1 preceding a C1 in the assimilating environment did not shorten, as expected, could indicate that an assimilated C1 did not devoice and thus was similar to a fully-voiced C1. As a result, a vowel preceding such a non-assimilated obstruent remained of the same duration as a vowel preceding a fully-voiced obstruent. Conceivably, this fact could be accounted for by the subjects’ lack of familiarity with nonsense words because native speakers, much like L2 learners, produced the nonsense tokens at a slower rate, which could be a factor in the overall longer V1 duration. However, this hypothesis is not supported by the results of the voicing proportion analysis of C1 in nonsense tokens, where all groups of subjects devoiced C1 in the assimilating environment and produced it with less voicing than fully-voiced C1.

Another example of variability is some of the ancillary measures is the V2 duration. It was found that in the real-word tokens V2 duration was dependent on both the cluster type after which it occurred and the voicing environment (cf. Sections 2.3.7.2. & 3.3.2.7.). However, in the nonsense-word experiment, the results of the V2 duration analysis revealed that for all groups of subjects the V2 duration depended on the type of the preceding obstruent rather than on its voicing. This pattern is similar to the pattern observed for V1 in real-words.

There are several conceivable factors that could account for the variation in some of the secondary production measures. One explanation for the V1 analysis results might lie in the nature of the stimuli. Although all nonsense stimuli were of the
CVCCV shape (in order to match the majority of the real tokens), with a low-back [a]
as both vowels, some subjects shifted the stress from the final syllable to the initial
syllable. As a result, what was intended to be pronounced as an unstressed vowel [ə]
was produced as a stressed vowel longer in duration than a schwa. Furthermore,
although an attempt was made to control all tokens for stress and vowel quality, in the
real-word tokens, the vowels and the shape of the syllable they occurred in were not
identical. Therefore, it is not unreasonable to assume that inherent differences in
vowel quality in addition to the differences in the supra-segmental properties might be
responsible for the variability of the results.

More importantly, it might be the case that unlike other languages, Russian
does not utilize the vowel duration (and segment duration in general) strongly to
implement the voice contrast\(^\text{23}\). As a result, the vowel duration differences show little
effect of the feature [voice], and other phonetic parameters are used as cues to
voicing\(^\text{24}\). In addition, given that not all stimuli were minimal pairs, it is possible that
the effect of vowel lengthening before voiced obstruents could be too small to be
consistent relative to the variability from other sources.

As far as obstruent duration is concerned, in both experiments, no group of
subjects distinguished among C1 and C2 stops in the three environments
(assimilating, fully-voiceless, and fully-voiced) in terms of segment duration (cf.
Sections 2.3.1.2 & 3.3.1.2.). The only condition wherein all groups of subjects
demonstrated an effect of assimilation, or the lack thereof, was in the nonsense-words.

\(^\text{23}\) Keating (1985) in a discussion of vowel duration argues that two Slavic languages, Polish and
Czech, do not shorten vowels before voiceless consonants.

\(^\text{24}\) Similar results are reported in the study of the production and perception of assimilated stops in
French (Snoeren et al., 2006). The authors argue that only relative voicing duration of a stop closure
provides a robust measure of voicing for French stops in non-initial position; whereas, other cues to
voicing, such as vowel or closure durations, do not provide a reliable cue to voicing.
However, Burton & Robblee (1997) have argued that segment duration in Russian appears to be less indicative of RVA than voicing duration, and duration of voicing serves as the most important cue to voicing contrast in Slavic languages (Keating 1984).

In fact, the proportion of voicing of C1 in the assimilating environment was the only measure pertaining to the RVA production on which statistically significant differences were found among the groups of subjects (cf. Sections 2.3.2.2. & 3.3.2.2.). Compared to other measures, the analysis of this measure elicited the most consistent results in both real and nonsense tokens across all groups and environments. All groups of subjects were found to produce less voicing in C1 in the assimilating environment (as in kazta) compared to a fully-voiced C1 (as in kazda), but more voicing than in truly voiceless obstruents (as in kasta). Notably, even native speakers did not completely devoice C1 in the assimilating environment but devoiced fricatives more than stops. Non-native speakers, on the other hand, whereas devoicing C1 fricatives as much as L1 speakers, were found to devoice C1 stops to a lesser degree than natives. Thus, the current work verifies that proportion voicing of the C1 is the primary, and perhaps only, marker of

The analysis of the same measure, voicing proportion, for C2 confirmed that in both experiments, all groups of subjects distinguished between voiced and voiceless stops and fricatives. It was also found that overall, L2 learners produced less voicing than natives in fully-voiced obstruents, both in C1 and C2 stops and fricatives, as in lu[zg]a ‘sunflower shells’ and lo[bz]ik ‘fret-saw’. This finding was somewhat supported by the results of the C2 stop VOT analysis, which revealed that non-natives did not prevoice fully-voiced stops as much as natives, and thus produced
them with much less voicing than natives. More specifically, advanced learners were found to prevoice more than intermediate learners but less than native speakers. (See sub-section 5.2.1. for further discussion).

5.1.2. Partial voicing neutralization

As Snoeren et al., (2006) observe, the process of assimilation is usually described in categorical terms; however, the degree of assimilation could be presented on a continuum, ranging from complete assimilation to a lack thereof, including partial assimilation. The results of the acoustic experiments demonstrated that even native speakers did not neutralize the voice distinction completely in C1 stops and fricatives in the assimilating environment. This finding supported earlier claims about the incomplete nature of acoustic neutralization of an underlying voice contrast (Burton & Robblee 1997; Charles-Luce 1993; Charles-Luce & Dinnsen 1984; Port & O’Dell 1985; Port & Crawford 1989; Slowiaczek & Dinnsen 1985; Warner et al., 2004). It should be noted that in these studies, generally a combination of different temporal properties, rather than one specific property, was reported to distinguish underlying voiced and voiceless obstruents; whereas, in this analysis the duration of voicing appeared to be the most reliable cue indicative of a voicing distinction. The apparent presence of incomplete neutralization only, or primarily, in the proportion of C1 voiced is consistent with the results for the larger effects of devoicing itself discussed above.

Additionally, in both experiments, V1 duration was in fact found to be longer before an assimilated C1 fricative than before a fully-voiceless fricative. This could be taken as evidence of incomplete voice neutralization as it affects the vowel duration. Recall that vowels usually shorten before voiceless obstruents; therefore, if
voicing neutralization was complete in the real and nonsense tokens, vowel duration was expected to be similar in assimilating and fully-voiceless environments (Charles-Luce 1985, 1993; Chen 1970). However, in the stop-fricative clusters in real words, the effect for V1 was in the opposite direction. Thus, differences in V1 duration in this data are likely to reflect the fact that minimal pairs did not exist for the real word contrasts, rather than supporting an incomplete neutralization effect.

Several studies that specifically addressed the issue of devoicing in Polish and Russian reported incomplete neutralization of obstruents. In their analysis of RVA across prepositional boundaries in Russian, Burton & Robblee (1997) found one specific instance of incomplete devoicing. In the /zt/ cluster there was more voicing of C1 than in the /st/ cluster, which suggested incomplete devoicing of a voiced C1 fricative in the assimilating environment. Likewise, Slowiaczek & Dinnsen (1985), who studied voice neutralization in Polish, claim that in their data voicing remained contrastive in word-final position, as manifested by longer vowel duration and, for labial stops, by more voicing into closure. That is, word-final obstruents that were supposed to devoice (as in grad gra[t] ‘hail’) were phonetically distinguishable from truly voiceless obstruents (as in grat gra[t] ‘old thing’).

However, the issue of incomplete neutralization does not have any bearing on the results and objectives of the present investigation because the research question was not whether L1 speakers devoice obstruents in the assimilating environment completely or not, but rather whether L2 learners produced RVA similarly to L1 speakers. Although native speakers were found to devoice C1 in the assimilating environment only partially, the amount of devoicing produced by natives was treated as the target pronunciation for non-natives. This is in accord with the results of other
studies on voicing which have shown that voicing into closure for devoiced obstruents does not have to be zero but can be around 30% of the closure duration (Darcy & Kugler 2007; Snoeren et al., 2006).

5.2. Implications for phonetic theory

In the following section, I will discuss the results of the acoustic experiments in light of research on L2 phonetic acquisition. In particular, I will address such issues as phonetic L1- L2 transfer, L2 approximation of L1 properties, and L2 phonetic learning.

5.2.1. Phonetic L1 – L2 transfer: evidence from the analysis of voiced stops and fricatives

In both experiments, L2 speakers were found to devoice C1 stops in all environments less than L1 speakers. However, they were also found to produce assimilated and fully-voiceless C1 fricatives in the real word experiment statistically similar to L1 speakers. As tempting as it could be to take this result as evidence of the successful acquisition of RVA by L2 speakers, this finding could be best accounted for by acoustically-motivated factors. First of all, had L2 speakers indeed mastered RVA, they would have applied it to the production of C1 stops in the assimilating environment as well. However, this was not the case. Another piece of evidence comes from acoustic studies of English fricatives, which reported that English native speakers normally do not completely voice even fully-voiced fricatives (Haggard 1978; Stevens at al., 1992), and, therefore, have less voicing to lose in the assimilating environment.
As discussed in Chapter One (cf. Section 1.4.4.), previous L2 research has shown that L2 learners identify certain non-native sounds with sounds in their native language even if these sounds are acoustically different (Boersma & Escudero 2004; Flege 1987; 1990, 2001; Flege & Eefting 1987). Therefore, it is possible that English L1-Russian L2 speakers simply transferred an L1 phonetic pattern into the L2, and produced even fully-voiced Russian fricatives with less voicing than natives. Consequently, they had less voicing to lose in C1 fricatives in the assimilating environment, as well. In other words, although L2 subjects appeared to produce assimilated fricatives as fully-voiceless, this could reflect their general lack of voicing during fricatives. The analysis of the C2 stop and fricative voicing proportion lends support to this claim. Although L2 speakers successfully produced voiceless C2, they were found to produce fully-voiced C2 with significantly less voicing than native speakers, similarly to C1 voicing production in the fully-voiced clusters. This finding could indicate that L2 learners have not yet acquired the pronunciation norms of target Russian fricatives, which require more voicing than English ones, and had to transfer L1 voicing properties into L2.

Research on L2 acquisition has also shown that since adult L2 learners already possess a phonetic system for their L1, they generally make more errors in L2 than early learners do. This is due to the fact that adult L2 learners utilize previously acquired structures to produce L2 sounds and therefore perceive and produce target structures in terms of L1 allophones and phonemes (Flege 1984, 1987, 1992, 2002; Flege et al., 1987; Major 2005; McKay et al., 2001). To take just one example, Flege et al., (1987) found that Chinese learners of English did not sustain closure voicing in word-final stops to the same extent as English native speakers. Therefore, their
mispronunciation of English stops was due to the transfer of L1 acoustic properties into L2.

Findings of L1 into L2 phonetic transfer have been reported in numerous studies by Flege and his colleagues (Flege 1991; Flege & Eefting 1987; Flege et al., 1995; Flege et al., 1999). Likewise in the present analysis, English L1 - Russian L2 learners arguably have exhibited a similar tactic. Recall that according to Flege’s Speech Learning Model (SLM), late bilinguals may expand their L1 phonetic inventory by adding phonetic categories for L2 sounds that are substantially different from any sound in L1 (cf. Section 1.4.4.1). However, new category formation is unlikely to occur if L1 and L2 sounds, although being acoustically different from each other, belong to the same phonetic category. Flege (1990) argued that late bilinguals fail to note acoustic differences between such related sounds. As a result, only ‘new’ but not ‘similar’ L2 sounds may be eventually mastered by late L2 learners. In this study, adult L2 learners of Russian failed to produce Russian obstruents in a native manner. In light of Flege’s SLM, this finding is attributed to the fact that Russian stops and fricatives and their English counterparts belong to the same overall categories. Although they are acoustically different, these acoustic differences are not used to contrast meaning in either English or Russian. As a consequence, English L1 – Russian L2 speakers might have failed to establish new phonetic categories for obstruents with voicing continuing throughout the consonant, as opposed to their native obstruents with only partial voicing, and instead they applied different realization rules to the same phonetic categories.

Another piece of evidence for an L1 phonetic effect on L2 speech comes from the analysis of C2 stop VOT. C2 VOT was not central to the predictions of the
experiment, and it was measured to be sure that all necessary information about the consonant clusters was tested. However, as explained in Footnote 1, the C2 VOT results provided evidence for Flege’s SLM by suggesting that at least some L2 speakers transferred L1 phonetic properties into L2. In the real-word experiment, both groups of non-native subjects produced fully-voiceless C2 stops with VOT values similar to the L1 values; thus, their fully-voiced stops were not nearly as prevoiced as L1 voiced stops. Interestingly, advanced speakers produced their fully-voiced stops with negative VOT values that were between native and intermediate speakers’ values. Since English speakers were shown to prevoice their voiced stops variably (Keating 1984:306), this finding could suggest that intermediate L2 learners produced Russian L2 voiced stops with the English L1 values, which is in accord with the findings reported in Flege & Hillenbrand (1984). In their investigation of the VOT production in English L2, it was found that some L2 learners produced English /p, t, k/ with short-lag values similar to VOT values for /p, t, k/ in their L1. Advanced learners, on the other hand, showed signs of at least partial learning of the target values. I will turn to the discussion of the advanced L2 learners’ production in the next sub-section.

5.2.2. Approximation of L2 phonetic properties: evidence from the analysis of voiceless and voiced stops

As discussed in the previous sub-section, non-native speakers were found to realize voiced stops with very little, if any, prevoicing. Advanced learners were shown to produce more voicing in voiced stops than intermediate learners, which could suggest that L2 speakers gradually approximated the target VOT values in voiced stops, with
advanced L2 speakers producing negative VOT values half-way between native and intermediate learners’ values.

Phonetic approximation has been thoroughly described in the work on L2 phonetic acquisition (Caramazza et al., 1973; Flege & Port 1981; Flege & Hammond 1982; Flege 1987; Flege 1990). It has been shown that even highly experienced L2 learners often do not produce an L2 sound authentically if it has a counterpart in L1 with which it can be identified. Such sounds are usually represented by the same IPA symbol (Flege & Hillenbrand 1984). For example, voiceless stops /t/ occurring in French and English could be perceived by L2 learners as different realizations of the same category because of their overall phonetic similarity (Flege & Hillenbrand 1984). It is also believed that with sufficient experience adult L2 learners will use merged categories and produce stops differently from L1 but will never perfectly match native speakers of the target language (Flege 1984). For example, in many languages (Dutch, French, Italian), voiceless stops /p, t, k/ are realized as unaspirated short-lag stops. Adult speakers of such languages produce English /p, t, k/ with longer VOT values than in their L1; however, these values are too short for English (Flege 1995).

Caramazza and colleagues (1973) reported that French Canadian subjects did not match English L1 speakers in producing English stops /p, t, k/. They were found to produce these stops with longer VOT values than in French stops produced by French monolinguals. Flege (1995) argues that the phonetic norms of the target language could be approximated in such a manner when L1 and L2 sounds are phonetically close to each other and are judged as being instances of the same category. A new category will thus not be established for an L2 sound, but rather
existing L1 and L2 categories will be collapsed into one. As a result, in adult L2 speech, a new L2 sound represents a merger of the phonetic properties of L1 and L2 sounds, and thus this could be taken as evidence that L2 speakers are capable of modifying previously established patterns of production.

As discussed above, previous studies have shown that learners of the L1 language that has short-lag VOT stops produce English stops with VOT values that are longer than appropriate for their L1 but shorter than required by English (Caramazza et al., 1973; Flege & Port 1981). Kewley-Port & Preston (1974, cited in Flege 1984) hypothesized that short-lag stops are generally easier to produce than long-lag stops. Flege (1984) argued that English L1- French L2 speakers in his study of the stop production succeeded better in producing the short-lag French stops than native speakers of French mastering the long-lag English stops. This observation can be extended to the present study, since Russian has short-lag stops as well. The results of the acoustic experiments suggest that it might indeed be easier for English native speakers to master short-lag VOT because in both experiments L2 subjects realized short-lag voiceless stops similarly to natives. This can be taken as an indication that both groups of L2 learners might have successfully approximated the positive VOT values in Russian.

The results of the analysis of the VOT values of fully-voiceless and fully-voiced C2 are thus in the accord with the existing studies on L2 phonetic acquisition, which provide evidence of the fact that L2 learners are capable of learning the phonetic norms of the target language by approximating appropriate acoustic L2 values and modifying the existing L1 sound categories.
5.3. Implications for phonological theory: learning of phonological patterns rather than categories

In the previous sub-sections, I discussed whether L2 phonetic learning has taken place in L2 learners, as manifested by partial C1 devoicing and the approximation of target VOT values. In this section, I will turn to the discussion of L1 and L2 production patterns.

5.3.1. Implications for phonological acquisition

It has long been recognized that certain language structures and patterns may be more difficult to acquire for adult language learners. As a response to this, a variety of phonological theories and approaches explaining language learners’ behavior have emerged over the years. In particular, while traditional Optimality Theory (Prince & Smolensky 1993; McCarthy & Prince 1993) demonstrates how a specific grammar can be achieved as a result of various constraint rankings and how grammars can be learned, a stochastic version of OT (Boersma & Hayes 2001) addresses the issue of variability and free ranking in production.

One of the research objectives of this dissertation was to offer an interpretation of the L2 production patterns which was done within the stochastic model of acquisition put forth by Boersma & Hayes (2001). More specifically, gradual L2 phonological acquisition of RVA observed in L2 speakers was analyzed through the application of the GLA (cf. Section 4.5.1.1.). Following the main tenets of the GLA, I demonstrated how a fixed target constraint ranking ensuring the emergence of RVA could be eventually obtained. However, the GLA model alone did not explain partial devoicing of an obstruent in the appropriate context and overall variability in L1 and L2 production.
In order to account for gradiency within a single token, following Boersma (1997, 1998) and Jun (1995, 2004), I decomposed the markedness constraint
\[ AGREE(VOICE)_{CG} \]
which allowed only voiceless obstruent clusters to surface in Russian L1 into a family of fine-grained phonetic constraints. Such an analysis details the re-ranking process within the same family of constraints which incorporate quantitative value. The re-rankings of the members of this constraint family within the family triggered by the available input resulted in gradient devoicing of C1 in clusters.

To sum up, the speech data obtained in this dissertation were analyzed using both traditional OT and its stochastic version. As previous and this research have shown, L2 production is usually characterized by a great deal of variability (cf. Section 4.5.2.1.). However, the present work has demonstrated that although traditional OT is capable of handling categorical fixed ranking of constraints, it does not succeed in explaining variability in the data. On the other hand, the stochastic version of OT (i.e., the GLA by Boersma & Hayes 2001) in conjunction with the concept of phonetically-detailed constraints (Boersma 1997, 1998; Jun 1995, 2004) has been shown to provide a satisfactory account of both L1 and L2 production patterns.

Therefore, one of the implications of the present analysis is that the phonological model that encodes the phonetic information directly into the constraints appears to account better for data exhibiting different types of variability\(^{25}\). More specifically, it has been demonstrated that different rankings of IDENT(VOICE) and

\(^{25}\) It is of course possible to generate an analysis of the data in this dissertation by separating phonetics from phonology and positing a separate phonetic component responsible for supplying language-specific phonetic detail (Keating 1984). Keeping in line with the spirit of the GLA, which is argued to be a better model for L2 learning in this dissertation, a phonetics-in-the-phonology analysis was only presented.
{AGREEC1aC2(VOICE)}_{CG}, and permutations of the constraints within the
{AGREEC1aC2(VOICE)}_{CG} family encompass a variety of grammars, which were
illustrated by different production patterns (cf. Section 4.5.4.).

Overall, this analysis can handle both language-specific details and cross-
linguistic, universal patterns. While language-specific assimilation patterns result
from language-specific constraint rankings, the universal patterns emerge from
universal rankings. Furthermore, the production patterns obtained in the phonetic
experiments provided the phonological analysis with empirical basis. Therefore, the
details of the proposed analysis are supported by real-life production data.

5.3.2. Gradual acquisition of L2 phonological patterns
The ultimate goal of L2 phonological acquisition, beyond acquisition of new sound
categories, is acquiring new phonological patterns. In addition to demonstrating that
L2 learning of sound categories has taken place, the acoustic experiments also
provided evidence of L2 acquisition, albeit gradual, of phonological rules. This was
demonstrated by the differences between the production patterns of advanced and
intermediate learners, with advanced subjects performing somewhat better than
intermediate learners, although not on all measures. Recall that all groups of L2
subjects succeeded in producing voiceless obstruents in the fully-voiceless
environment (cf. Sections 2.3.2.2. & 3.3.2.2.). Thus, one would expect L2 learners to
be able to produce devoiced C1 in the appropriate contexts as well, as required by
Russian phonology. However, this was not the case. Recall that there were at least
four production patterns revealed by the histograms (cf. Section 4.5.2.1.), which
differed in the distribution of fully-devoiced, fully-voiced and partially-devoiced C1
in the assimilation environment. The histograms presented evidence that at least some
learners might be aware of the Russian phonological pattern and had made some progress in learning to produce it. Yet, other L2 learners were shown to have a bias toward fully-voiced and/or partially-devoiced obstruents. This finding and the fact that on the whole advanced learners fared better than intermediate subjects in their production of RVA (cf. Sections 2.3.2.2. & 3.3.2.2.) serve as evidence that not all intermediate L2 learners have acquired the target phonological pattern. Overall, the results of the voicing proportion analysis, which demonstrated that the amount of voicing in C1 in the assimilating environment was dependent on the speakers’ proficiency level, provided evidence for gradual phonological acquisition.

5.4. Implications for L2 teaching

5.4.1. Russian L2 learners’ production

Research on L2 phonological and phonetic acquisition has important implications for the L2 classroom in that it provides detailed information about overall language learners’ production, and the specific difficulties and challenges facing them. Knowledge of the difficulties adult language learners encounter while learning a foreign language and factors affecting their L2 production can be employed by language instructors to develop appropriate teaching materials and techniques that will target problematic areas in students’ production. For example, the findings of the acoustic studies described in this work and the conclusions drawn from them not only discovered particular problems in L2 Russian speech (e.g., non-native speakers do not devoice an obstruent in the assimilating environment to the same extent as native speakers do and thus fail to produce RVA) and pointed out their underlying causes (e.g., L1-L2 phonetic and phonological influences), but also demonstrated that indeed
L2 learners’ production improved as a result of practice and increased exposure to the target language. In addition, this work offers insights into the gradual nature of L2 phonological acquisition and awareness of the acquisition of these phonological patterns, and could inform language teachers about the stages through which L2 learners pass. This research project has also confirmed the findings of previous research, namely that in early stages of acquisition adult learners transfer L1 properties into L2 and produce L2 sounds with L1 values. This claim is supported by the finding that less proficient L2 learners produced Russian L2 voiced stops with virtually no prevoicing; that is, they employed the English L1 values in their production of Russian stops. Thus, the current results demonstrate an interplay between L1-L2 phonetic influence on sound categories, and incomplete acquisition of L2 phonological patterns. Both influences combine to create L2 learners’ specific type of non-native productions.

5.4.2. Does formal L2 instruction matter?

Whether formal instruction facilitates L2 phonetic/phonological acquisition has been open to debate. Overall, there is little evidence that formal classroom instruction of an L2 affects the degree of foreign accent in long-term bilinguals (Flege et al., 1995; Piske et al., 2001). Yet, given the production patterns and pronunciation accuracy in intermediate vs. advanced L2 learners and the subjects’ relatively early stage of learning, it appears that knowledge regarding gradual rather than categorical acquisition of the RVA pattern can be used in developing appropriate training practices and teaching materials that would emphasize, and thus facilitate, its acquisition. It is unclear, however, whether the fact that advanced L2 learners having fared better than intermediate learners is due to the amount of L2 classroom
instruction, or rather overall exposure to the target language in and outside the classroom. Advanced participants in this study have not only had more classroom (and otherwise) training in Russian, but have also studied or lived in Russia or some other predominantly Russian-speaking country. Advanced L2 speakers have thus been immersed in the target language environment, experiencing authentic language and communicating in it on the daily basis. As a result, they might have learned to perceive fine acoustic details of the RVA in the native speech and have started devoicing fully-voiced obstruents in the environment of a fully-voiceless one. On the whole, by the time of the experiments, intermediate learners had had a relatively small number of instruction hours and very little overall exposure to Russian. However, since they were found to devoice voiced obstruents to at least some degree, this could be taken as evidence that however minimal their exposure to the target language was, it was sufficient to trigger some assimilation in their speech.

Among those studies that did find a correlation between formal instruction and the pronunciation accuracy rate was the study of the acquisition of Spanish trills by English L1 speakers by Johnson (2008). More specifically, it was found that 3rd and 4th year students who had, in addition, studied abroad improved in comparison with lower level learners. However, his work was concerned with the acquisition of a discrete segment, which is absent from the English phonological inventory, and thus is new and unfamiliar to English native speakers. English L1- Spanish L2 speakers had to introduce a new phone to their inventory and learn how to articulate it. The focus of this project, on the other hand, is on the acquisition of a phonological pattern, namely RVA.
Although RVA does not apply in English obstruent clusters per se, the concept of voice assimilation is not unfamiliar to English speakers, given progressive voicing assimilation, as in *dogs* – do[gz]. Therefore, rather than acquiring a new phone altogether, L2 Russian speakers have to attune to and re-learn the pattern they already know by switching the direction of assimilation from progressive to regressive. In order to trigger the L2 speakers’ re-learning of the phonological pattern, their attention should be drawn to this fact. Therefore, it seems that providing explanations of the RVA pattern in an explicit and straightforward manner might be beneficial, at least in early stages of acquisition. Piske et al. (2001) cite the results of three studies (Bongaerts et al., 1997; Missaglia 1999; Moyer 1999) which showed that if L2 classroom teaching involves special training in the perception and production of L2 sounds, it may have a larger effect on L2 pronunciation accuracy. As language learning proceeds, L2 learners equipped with knowledge of RVA become progressively more and more sensitive to it both in perception and their own production. In later stages of acquisition, practice and exposure to the target language in authentic contexts will facilitate the recognition and production of RVA.

5.4.3. Suggestions for Russian L2 instruction

As pointed out by Cook (1996:46), in today’s classroom, most language instructors use “integrated pronunciation teaching’ in which L2 pronunciation is taught as ‘an incidental to other aspects of language.” Such teaching usually involves incidental correction of mispronounced phonemes, and pronunciation work is included in other activities, such as reading texts or acting out dialogs. One clear implication that follows from the present work is that the goal of explicit instruction of RVA should be twofold. Since the successful acquisition of RVA encompasses both the mastery of
acoustic properties of Russian sounds and the knowledge of *when* to produce RVA, classroom instruction should focus on both aspects.

Given the production patterns of RVA in Russian exhibited by the L2 learners in the two experiments (cf. Section 5.4.2.1.), it is not unreasonable to suggest a number of techniques related to learners’ particular stages of development. Recall that at least some L2 Russian learners produced (nearly) fully-voiced C1 in the assimilating environment, thus demonstrating their lack of implementation of RVA whereas others maintained a significant amount of voicing. Impressionistically, the lack of RVA in L2 Russian speech is one of the factors contributing to the overall perceived foreign accent. Therefore, in the early stages of classroom instruction, the RVA pattern should be explicitly introduced and explained, as perhaps, a rule that is required by the phonology of the Russian language.

This can be done by presenting minimal pairs contrasting the clusters containing fully-voiced/voiceless obstruents and clusters containing an assimilated obstruent (e.g., *obzharit* - *o/bʒ*/arit ‘to stir-fry’ vs. *obsharit* - *o/pʃ*/arit ‘to rummage’) in various tasks. More specifically, task types capitalizing on the [voice] contrast in minimal pairs can be devised. One of the activities can involve both listening comprehension and identification of the intended obstruent in terms of voicing specification. For example, after listening to the list of minimal pairs with voiced and assimilated clusters, language learners can write them down in Russian orthography. The students can be assigned to read out the minimal pairs pointing out voiced vs. assimilated clusters.
Since RVA is not reflected in orthography (e.g., obšharit - o/pʃ/ərit ‘to rummage’), the discrepancy between the spelling of a word and its pronunciation should be accentuated, as well. Recall that in some cases RVA applies at the prefix-stem boundary (e.g., ob – sharit), and the morpheme-final obstruent devoices if followed by a voiceless obstruent (cf. Section 1.2.1.). Therefore, students could be familiarized with the (rather limited) list of Russian prefix morphemes ending in a voiced obstruent (e.g., ob-, pod-, etc.) which do not have voiceless counterparts such as *op- and *pot-. Thus, they should learn to spell these prefixes with a voiced prefix-final obstruent but to pronounce them with a devoiced counterpart in the appropriate context. This strategy might prove to be useful in recognizing such morphemes in writing and reading tasks.

Furthermore, English – Russian bilinguals were found to transfer voicing characteristics of L1 voiced fricatives into L2 (cf. Section 2.3.2.2.). As a result, fully-voiced Russian L2 fricatives were produced with much less voicing than Russian L1 counterparts. In order for the learners to pronounce L2 fricatives in a native-like manner, they should be trained to fully voice these segments. This can be achieved by using a simple pronunciation exercise which can be performed in a computer lab. The students listen to a list of lexical items containing fully-voiced fricatives in various contexts read out by a Russian native speaker. They are recorded reproducing these items and the recording is played back to them. This exercise can provide the students with the opportunity to compare their own production to that of a native speaker.

Although one should not expect L2 learners to skip the phonological L1 transfer into L2 stage altogether, heightening their awareness of RVA could possibly facilitate the process of its acquisition. Likewise, explicit training in pronunciation of
discrete phonemes and learning a new phonetic contrast of Russian (voiced vs. voiceless unaspirated) as opposed to English (voiced vs. voiceless aspirated and voiceless unaspirated) could be beneficial to the overall pronunciation.

5.5. Conclusion and further research

5.5.1. Overall conclusion

One of the features that sets the speech of an adult L2 learner apart from the speech of an L1 speaker is foreign accent, which is manifested by a distorted production of sounds and sequences of sounds. L2 speech is usually characterized by the interference from the native language’s phonetics and phonology. That is, non-native production is a result of the interaction of L1 and L2 phonological and phonetic realization rules. The issues of the extent to which adult learners rely on their L1 while acquiring sounds of L2 and the specifics of the interaction of the two sound systems were analyzed in this dissertation.

As expected, it was found that the overall L2 pronunciation improved with increased proficiency, but even at the advanced level, on the whole L2 production was not native-like. This was manifested by the production of Russian L2 sounds with English L1 acoustic properties as a result of L1 phonetic and phonological transfer. However, at least some more proficient learners approximated L2 values and produced them in a near-native manner.

Generally, native speakers of a language are successful at detecting foreign accents and identifying the language background of an L2 speaker. However, when asked to describe the features that make somebody’s speech sound non-native, L1 speakers are rarely able to pinpoint the exact characteristics that contribute to the
accentedness of L2 speech. When native speakers are able to provide a response, they usually point to segment substitution in L2 production. For example, adult Russian L1 – English L2 speakers are notorious for substituting English interdental fricatives with Russian bilabial fricatives, as in /ð/-/z/ and /θ/-/f/. In light of Flege’s SLM (1995), this could be interpreted as that Russian L1 speakers treat English interdental fricatives as instances of Russian bilabial fricatives and thus do not establish separate categories for English sounds. However, as evidenced by the present work, the relationship between L1 and L2 is more complex than that, and the source of foreign accent could be a mis-production at the level of sequences of sounds. Nevertheless, increased exposure to the target language and consistent practice do result in a lesser degree of accentedness and improved overall pronunciation.

5.5.2. Further research questions

This dissertation has provided information about the production and acquisition of regressive voicing assimilation in obstruent clusters in Russian. The major findings were that the L2 acquisition of RVA was gradual in nature, with learners improving somewhat over the course of L2 learning, and that some L1 and L2 speakers devoiced C1 in the appropriate contexts gradiently. These findings lead to more research questions.

First, since the acquisition of RVA was found to be gradual, one question that arises is: what are the exact stages of acquisition that L2 learners go through? The re-ranking stages assumed in the analysis in Section 4.5.4. were hypothesized based on the re-ranking posited by the GLA. However, a longitudinal study closely monitoring the progress of L2 speakers learning RVA might be able to test this assumption and provide a definite answer. Also, given the fact that some research has found no effect
of formal instruction on the pronunciation accuracy (cf. Section 5.4.3.), future research has to establish whether explicit classroom instruction in RVA, as well as explicit training in producing fully-voiced vs. fully-voiceless obstruents prove to be beneficial to language learners.

Second, did English L1 – Russian L2 bilinguals consider Russian obstruents to be instances of corresponding English ones, or did they establish a merged L1 – L2 category? The variability in the learners’ production of L2 obstruents was explained by the postulates advanced by the SLM (Flege 1995). More specifically, it was assumed that adult L2 learners did not create a new phonetic category for L2 sounds, but treated English L1 and Russian L2 stops and fricatives as belonging to the same respective categories. Flege (1991) also argues that proficient L2 speakers create a single merged L1-L2 category. As a result, language learners produce an L2 sound with ‘compromise’ values, intermediate between their L1 and L2. More importantly, their production of L1 sounds is also affected by this merger, and their values shift toward L2 values. Given that in this dissertation English L1- Russian L2 learners were tested on their production of L2 obstruents only, it would be beneficial to compare their production of L1 English stops and fricatives to the production of monolingual English speakers’ in order to detect any changes. If changes are detected, it will prove that English - Russian bilinguals have merged their L1 and L2 categories into one.

Third, what is the nature of the native grammar? It has been found that there are at least three distinct patterns in native production which correspond to different constraint rankings. Does this finding mean that there is the native grammar, and the observed patterns which employ the same basic set of constraint are merely the result of speaker-specific permutations of these constraints and/or the result of individual
differences in speakers? Or is there, in fact, a number of distinct native
(sub)grammars? Also, as the results of the experiments have suggested, some
speakers switch from one ranking/grammar to another for real vs. nonsense words.

Thus, the next, fourth, question is: What factors influence speakers’ choice of
rankings? The switch from the categorically assimilating pattern in real words to a
different pattern in non-words could be due to the task effect. That is, native and non-
native speakers alike were confused when presented with made-up words and were
led by the orthography rather than phonology. This issue leads to the next question.

Fifth, did the task type affect subjects’ production? One of the criticisms of
this research could be that the data was obtained in a list-reading task. It was possible
that both in the real and nonsense words, the subjects were influenced by the
orthography rather than by their knowledge of Russian phonology. Therefore,
different tasks that would eliminate this possibility are necessary. The data can be
collected in a natural setting through oral communication with the subjects or by
presenting them with a fill-in the blank type of task, or the combination of the two.
For example, the subjects would be presented with a nonsense form containing an
assimilated obstruent in a cluster and be required to generate a related form containing
the obstruent in question in isolation, thus demonstrating its underlying voicing, and
vice versa.

Furthermore, the variability in the amount of devoicing can be attributed to the
fact that both high and low frequency words were included in the real-word
experiment. The subjects, including the natives, might treat high and low frequency
words in a different manner, and treat nonsense items as instances of low frequency or
unfamiliar existing words. It is not unconceivable that at least native speakers could
devoice C1 in high frequency words to a greater extent due to their familiarity with such words, having heard them more often than having seen them in a written form. Therefore, the final questions are: Given the paucity of the real words containing obstruent clusters requiring RVA, how can one eliminate the possible effect of high vs. low frequency words? Are some obstruent combinations more frequent than others? Do native and non-native speakers treat them differently?

Finally, since a part of this work has been couched in terms of the SLM which is concerned with the acquisition of sound categories, while the main focus of this research project has been on the acquisition of L2 phonological patterns which is accounted for within the GLA framework, the question of the relationship between the SLM and GLA should be addressed. Is each of these models responsible for the kind of learning relevant to its respective domain, or should an integrated theory which introduces the concept of phonetic-level constraints into phonology be used instead? As this work has shown, phonetically detailed constraints are necessary to account for at least some speakers’ production. Some phonological models do not make a formal distinction between phonology and phonetics. For example, Flemming (1995, 2001) has proposed incorporating the existence of phonetic categories into the phonological constraint system, which derive the similar properties of phonetics and phonology from the same constraints. However, this approach raises the question of how the GLA can cover the acquisition of the constraints that are responsible for phonological patterns in production grammar, but not the constraints that create categories like the SLM.

Moreover, some research provides evidence against the claim that phonological constraints are directly phonetic. For example, Smith (to appear)
analyzes the relationship between the constraints ONSET and *ONSET/X which have the same functional basis but distinct formal properties. Smith argues that if phonetic information is encoded directly into phonological constraints, then the shared functional basis of such constraints would imply that they are not formally distinct. The model that places the phonetics into the phonology would predict that ONSET is actually *ONSET/∅, which bans null onsets, the highest ranked constraint of the *ONSET/X family. However, the data from several languages including Sestu Campidanian Sardinian, Guugu Yimidhirr, and Pitta-Pitta, suggest that ONSET cannot be redefined as *ONSET/∅ because they evaluate distinct phonological structures, and ONSET has no fixed universal ranking with respect to the *ONSET/X family.

To conclude, as Flemming (2001:36) points out, “the separation of phonetics and phonology more often figures as an assumption rather than a topic of research, so the literature contain few explicit arguments in favour of such a model.” Therefore, future research on phonetic and phonological phenomena is needed in order to establish more definitively the extent of the overlap of the two components and its consequences for the models of phonetic/phonological acquisition.

5.5.3. Conclusion

The results of this investigation have demonstrated that L2 speech is characterized by both L1-L2 phonetic and phonological transfer and L2 phonetic approximation, and suggested that the learning of phonological patterns is possible in adult learners. Also, it has been shown how gradual learning and gradient devoicing can be accounted for by the Gradual Learning Algorithm in conjunction with phonetically detailed constraints. Overall, this research has demonstrated a surprising range of individual
variability, both in native and non-native phonology, and this raises questions about how speakers move among various phonological patterns.
APPENDIX A
TEST REAL WORDS FOR EXPERIMENT ONE

Stop – fricative cluster

Я говорю слово Обшить
Ya govori slovo obshyt
I am saying the word to saw around

Я говорю слово Обжить
Ya govori slovo obzhyt
I am saying the word to settle down

Я говорю слово Запшить
Ya govori slovo zapshykat
I am saying the word to spray with

Я говорю слово Обшарить
Ya govori slovo obsharit
I am saying the word to rummage

Я говорю слово Обжарить
Ya govori slovo obzharyt
I am saying the word to saw around

Я говорю слово Лапша
Ya govori slovo lapsha
I am saying the word noodles

Я говорю слово Обсудить
Ya govori slovo obsudit
I am saying the word to discuss

Я говорю слово Обзор
Ya govori slovo obzor
I am saying the word a view

Я говорю слово Рапсодия
Ya govori slovo rapsodya
I am saying the word a rhapsody

Я говорю слово Абсент
Ya govori slovo absent
I am saying the word absinth
Я говорю слово Лобзик
Ya govoriu slovo lobsik
I am saying the word a fret-saw

Я говорю слово Гипсовый
Ya govoriu slovo gipsovyi
I am saying the word cast

Я говорю слово Подсунуть
Ya govoriu slovo podsunut
I am saying the word to slide under

Я говорю слово Подзуживать
Ya govoriu slovo podzyzhivat
I am saying the word to provoke

Я говорю слово Отсудить
Ya govoriu slovo otsudit
I am saying the word to sue

Я говорю слово Подсечь
Ya govoriu slovo podsech
I am saying the word to cut

Я говорю слово Подземка
Ya govoriu slovo podzemka
I am saying the word subway

Я говорю слово Отсек
Ya govoriu slovo otsek
I am saying the word a unit

Я говорю слово Подшить
Ya govoriu slovo podshyt
I am saying the word to trim by sawing

Я говорю слово Поджечь
Ya govoriu slovo podzhech
I am saying the word to set on fire

Я говорю слово Отшить
Ya govoriu slovo otshyt
I am saying the word to dump somebody

Я говорю слово Подшовный
Ya govoriu slovo podshovnyi
I am saying the word under a seem
Я говорю слово Поджать
Ya govoriu slovo podzhat
I am saying the word to purse, pucker

Я говорю слово Отшагать
Ya govoriu slovo otshagat
I am saying the word to stride

Fricative – stop cluster

Я говорю слово Глазка́
Ya govoriu slovo glazka
I am saying the word an eye (diminutive)

Я говорю слово Лузга́
Ya govoriu slovo luzga
I am saying the word sunflower shells

Я говорю слово Волоска́
Ya govoriu slovo voloska
I am saying the word a piece of hair

Я говорю слово Сказка
Ya govoriu slovo skazka
I am saying the word a fairy-tale

Я говорю слово Розги
Ya govoriu slovo rozgi
I am saying the word birch rods

Я говорю слово Каска
Ya govoriu slovo kaska
I am saying the word a helmet

Я говорю слово Ложка
Ya govoriu slovo lozhka
I am saying the word a spoon

Я говорю слово Обожги
Ya govoriu slovo obozhgi
I am saying the word to burn

Я говорю слово Су́шка
Ya govoriu slovo sushka
I am saying the word a small bagel
Я говорю слово Везти
Ya govoriu slovo vezti
I am saying the word to drive

Я говорю слово Громоздить
Ya govoriu slovo gromozdit
I am saying the word to pile up

Я говорю слово Вести
Ya govoriu slovo vesti
I am saying the word to lead
APPENDIX B

TEST NONSENSE WORDS FOR EXPERIMENT TWO
GIVEN IN THE FRAME SENTENCE ‘I AM SAYING THE WORD ___ NOW’

Stop – fricative cluster

Я говорю слово Кобсá
Ya govoriu slovo kobsa

Я говорю слово Копсá
Ya govoriu slovo kopsa

Я говорю слово Кобзá
Ya govoriu slovo kobza

Я говорю слово Вабсá
Ya govoriu slovo vabsa

Я говорю слово Вапсá
Ya govoriu slovo vapsa

Я говорю слово Вабзá
Ya govoriu slovo vabza

Я говорю слово Лабсá
Ya govoriu slovo labsa

Я говорю слово Лапсá
Ya govoriu slovo lapsa

Я говорю слово Лабзá
Ya govoriu slovo labza

Я говорю слово Хобсá
Ya govoriu slovo hobsa

Я говорю слово Хопсá
Ya govoriu slovo hopsa

Я говорю слово Хобзá
Ya govoriu slovo hobza

Я говорю слово Тобсá
Ya govoriu slovo tobsa
Я говорю слово Топса́
Ya govoriu slovo topsa

Я говорю слово Тобзá
Ya govoriu slovo tobza

Я говорю слово Зобша́
Ya govoriu slovo zobsha

Я говорю слово Зопша́
Ya govoriu slovo zopsha

Я говорю слово Зобжá
Ya govoriu slovo zobzha

Я говорю слово Пабша́
Ya govoriu slovo pabsha

Я говорю слово Папша́
Ya govoriu slovo papsha

Я говорю слово Пабжá
Ya govoriu slovo pabzha

Я говорю слово Кобша́
Ya govoriu slovo kobsha

Я говорю слово Копша́
Ya govoriu slovo kopsha

Я говорю слово Кобжá
Ya govoriu slovo kobzha

Я говорю слово Вабша́
Ya govoriu slovo vabsha

Я говорю слово Вапша́
Ya govoriu slovo vapsha

Я говорю слово Вабжá
Ya govoriu slovo vabzha

Я говорю слово Собша́
Ya govoriu slovo sobsha

Я говорю слово Сопша́
Ya govoriu slovo sopsha
Я говорю слово Собжá
Ya govoriu slovo sobzha

Я говорю слово Падсá
Ya govoriu slovo padsa

Я говорю слово Патсá
Ya govoriu slovo patsa

Я говорю слово Падза
Ya govoriu slovo padza

Я говорю слово Кадса
Ya govoriu slovo kadsa

Я говорю слово Катса
Ya govoriu slovo katsa

Я говорю слово Кадза
Ya govoriu slovo kadza

Я говорю слово Бадса
Ya govoriu slovo badsa

Я говорю слово Батса
Ya govoriu slovo batsa

Я говорю слово Бадза
Ya govoriu slovo badza

Я говорю слово Жадса
Ya govoriu slovo zhadsa

Я говорю слово Жатса
Ya govoriu slovo zhatsa

Я говорю слово Жадза
Ya govoriu slovo zhadza

Я говорю слово Кадша
Ya govoriu slovo kadsha

Я говорю слово Катша
Ya govoriu slovo katsha
Я говорю слово Каджа
Ya govoriu slovo kadzha

Я говорю слово Ладша
Ya govoriu slovo lodsha

Я говорю слово Латша
Ya govoriu slovo latsha

Я говорю слово Ладжа
Ya govoriu slovo ladzha

Я говорю слово Бадша
Ya govoriu slovo badsha

Я говорю слово Батша
Ya govoriu slovo batsha

Я говорю слово Баджа
Ya govoriu slovo badzha

Я говорю слово Падша
Ya govoriu slovo padsha

Я говорю слово Патша
Ya govoriu slovo patsha

Я говорю слово Паджа
Ya govoriu slovo padzha

Fricative – stop cluster

Я говорю слово Казта
Ya govoriu slovo kazta

Я говорю слово Каста
Ya govoriu slovo kasta

Я говорю слово Казда
Ya govoriu slovo kazda

Я говорю слово Грозта
Ya govoriu slovo grozta
Я говорю слово Гроста
Ya govori slovo grosta

Я говорю слово Грозда
Ya govori slovo grozda

Я говорю слово Вазта
Ya govori slovo vazta

Я говорю слово Васта
Ya govori slovo vasta

Я говорю слово Вазда
Ya govori slovo vazda

Я говорю слово Базта
Ya govori slovo bazta

Я говорю слово Баста
Ya govori slovo basta

Я говорю слово Базда
Ya govori slovo bazda

Я говорю слово Зазка
Ya govori slovo zazka

Я говорю слово Заска
Ya govori slovo zaska

Я говорю слово Зазга
Ya govori slovo zazga

Я говорю слово Базка
Ya govori slovo bazka

Я говорю слово Баска
Ya govori slovo baska

Я говорю слово Базга
Ya govori slovo bazga
Я говорю слово Хазка
Ya govoriu slovo hazka

Я говорю слово Хаска
Ya govoriu slovo haska

Я говорю слово Хазга
Ya govoriu slovo hazga

Я говорю слово Пазка
Ya govoriu slovo pazka

Я говорю слово Паска
Ya govoriu slovo paska

Я говорю слово Пазга
Ya govoriu slovo pazga

Я говорю слово Тажка
Ya govoriu slovo tazhka

Я говорю слово Ташка
Ya govoriu slovo tashka

Я говорю слово Тажга
Ya govoriu slovo tazhga
REFERENCES


International Conference on Linguistics, 45-50.


algorithm and some basic complexity results. Technical Report CU-CS-678-93, Department of Computer Science, University of Colorado, Boulder.


