

PHONETICS-PHONOLOGY INTERACTIONS IN PRE-SONORANT
VOICING

A THESIS
SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY (PhD)
IN THE FACULTY OF HUMANITIES

PATRYCJA STRYCHARCZUK

SCHOOL OF ARTS, LANGUAGES AND CULTURES
2012

CONTENTS

List of Figures	5
List of Tables	8
Abstract	11
Declaration	12
Copyright	13
Acknowledgements	14
1 Introduction	16
1.1 Pre-sonorant voicing. An overview	16
1.2 Theoretical issues	18
1.2.1 Do sonorants spread [voice]?	18
1.2.2 Fricatives vs. stops	19
1.2.3 Positional restrictions. The role of delaryngealisation	20
1.2.4 Phonology vs. phonetics	21
1.3 Current approach. Synopsis	22
2 The phonology-phonetics interface	25
2.1 Phonetic and phonological knowledge	25
2.2 Modularity	28
2.3 Arguments against the phonology-phonetics distinction	29
2.3.1 The duplication problem. Functional phonology	29
2.3.2 Gradience and overlap	31
2.4 A distinction without an interface?	32
2.5 The current approach	35
2.6 Empirical diagnostics for categoricity and gradience	37
2.7 Final remarks	39
3 West-Flemish	40
3.1 Pre-sonorant voicing and other voicing-related processes in West-Flemish	40
3.1.1 Issues. The current study	42
3.2 Materials and method	43
3.2.1 Speakers	43
3.2.2 Stimuli	43
3.2.3 Procedure	44
3.2.4 Acoustic analysis	44

3.3	Results	45
3.3.1	Stops	46
3.3.2	Fricatives	47
3.3.3	Summary. Categoricity and gradience	51
3.4	Discussion	53
3.5	Conclusion	60
4	Poznań Polish	62
4.1	Kraków-Poznań voicing and other voicing processes in Polish	62
4.1.1	Previous analyses	64
4.1.2	Issues concerning neutralisation in Polish	69
4.1.3	The current study	71
4.2	Stop#Sonorant sequences in Poznań and Warsaw Polish. Pilot study.	72
4.2.1	Materials and method	72
4.2.2	Results	74
4.2.3	Summary and methodological remarks	78
4.3	Experiment 2. Obstruent#Sonorant sequences in Western and Central dialects of Polish.	79
4.3.1	Materials and method	80
4.3.2	Results	81
4.3.3	Summary	89
4.4	Discussion	92
4.4.1	The role of delaryngealisation	92
4.4.2	A unified analysis of Polish voicing	95
4.5	Conclusion	96
5	Central Catalan	98
5.1	Voicing in Catalan	98
5.1.1	Issues related to Central Catalan voicing	100
5.1.2	The current study	103
5.2	Materials and method	104
5.3	Results	105
5.3.1	Realisation of voicing-related acoustic parameters	105
5.3.2	Individual variation	112
5.4	Discussion	118
5.4.1	Undergoer asymmetry	118
5.4.2	Trigger asymmetry	121
5.5	Conclusion	123
6	Quito Spanish	125
6.1	Quito /s/-voicing	125
6.1.1	Opacity	126
6.1.2	Gradience and categoricity	128
6.1.3	The current study	128
6.2	Materials and method	129
6.3	Results	131
6.3.1	Phonetic realisation	131
6.3.2	Statistical analysis	131
6.4	Discussion	143
6.4.1	/s/-voicing – categorical or gradient?	143

6.4.2 Consequences for formal models	144
6.4.3 Further alternatives	145
6.5 Conclusion	146
7 Explaining pre-sonorant voicing	147
7.1 Empirical properties of pre-sonorant voicing	147
7.1.1 West-Flemish	147
7.1.2 Poznań Polish	148
7.1.3 Central Catalan	148
7.1.4 Quito Spanish	149
7.1.5 Emerging generalisations	149
7.2 The role of delaryngealisation	151
7.3 Modulating factors	155
7.4 Constraints on pre-sonorant voicing in a diachronic and a synchronic perspective .	157
7.5 Arbitrariness and the role of phonology	160
7.6 The relationship between phonology and phonetics	162
8 Conclusions, main findings, directions for research	164
Bibliography	166
Appendix A Test items used in the experiment on Belgian Dutch	179
A.1 Test items	179
A.2 Control items	180
Appendix B Test items used in the experiment on Polish	181
B.1 Experiment 1	181
B.1.1 Test items	181
B.1.2 Control items	182
B.2 Experiment 2	183
B.2.1 Test items	183
B.2.2 Control items	184
Appendix C Test items used in the experiment on Catalan	185
C.1 Test items	185
C.2 Control items	186
Appendix D Test items used in the experiment on Quito Spanish	187

LIST OF FIGURES

3.1 Interaction between dialect (East-Flemish vs. West-Flemish) with the right-hand environment on voicing duration in fricatives.	48
3.2 Interaction between dialect (East-Flemish vs. West-Flemish) with the right-hand environment on voicing ratio in fricatives.	50
3.3 Interaction between dialect (East-Flemish vs. West-Flemish) with the right-hand environment on fricative duration.	51
3.4 Density plots of voicing duration in West-Flemish pre-sonorant fricatives and stops. The dashed line represents the density of obstruents in the baseline condition, i.e. obstruents followed by a voiceless or a voiced stop.	53
3.5 Density plots of voicing duration in East-Flemish pre-sonorant fricatives and stops. The dashed line represents the density of obstruents in the baseline condition, i.e. obstruents followed by a voiceless or a voiced stop.	53
3.6 Pre-sonorant /t/ (left) and /s/ (right) with a short voicing tail	54
3.7 Left panel: closure and burst durations of intervocalic coronal stops classified as either voiced or voiceless, when preceded by a high (top), or low (bottom) vowel. Figure reproduced from Ernestus (2000, 219). Right panel: closure and burst duration of pre-sonorant stops in the current data.	57
3.8 Left: Intensity difference (difference between intensity peak at high frequencies and intensity dip at low frequencies) as a function of voicing duration in pre-sonorant fricatives. Negative intensity difference means that the minimum low-frequency intensity is higher than the high-frequency maximum intensity; Pearson's product-moment correlation: $r=-0.65$, $p<0.001$	58
4.1 Interaction between the right-hand environment and the underlying voicing of the preceding stop on the voicing duration within the Warsaw and Poznań dialects. . .	75
4.2 Interaction between the right-hand environment and the underlying voicing of the preceding stop on the voicing ratio within the Warsaw and Poznań dialects. . . .	76
4.3 Density plots of voicing duration (left) and voicing ratio (right) in underlyingly voiced and underlyingly voiceless stops followed by a sonorant in the Poznań dialect	79
4.4 Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting burst duration.	85
4.5 Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting voicing duration in fricatives.	86
4.6 Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting voicing ratio in fricatives.	87
4.7 Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting intensity difference in fricatives.	90

4.8	Density plots of voicing duration (left) and voicing ratio (right) in underlyingly voiced and underlyingly voiceless stops followed by a sonorant in the Poznań dialect. Data from experiment 2	91
4.9	Density plots of voicing duration (left) and voicing ratio (right) in underlyingly voiced and underlyingly voiceless fricatives followed by a sonorant in the Poznań dialect. Data from experiment 2	91
5.1	Interaction between the undergoer's and trigger's manner of articulation in a model of voicing duration in Catalan word-final obstruents.	107
5.2	Interaction between the undergoer's and trigger's manner of articulation in a model of voicing ratio in Catalan word-final obstruents.	108
5.3	Interaction between the undergoer's and trigger's manner of articulation in a model of vowel duration preceding Catalan word-final obstruents.	110
5.4	Interaction between the undergoer's and trigger's manner of articulation in a model of f_1 (in Hz) of the preceding vowel measured at 10 ms before the offset (left), and 20 ms before the offset (right).	112
5.5	Plot of mean voicing duration (in ms) as a function of the undergoer's and trigger's manner of articulation per individual speaker.	114
5.6	Plot of mean voicing ratio as a function of the undergoer's and trigger's manner of articulation per individual speaker.	115
5.7	Random coefficients for undergoer's manner of articulation within speaker in a model of voicing duration in prevocalic obstruents (left), and voicing ratio in prevocalic obstruents (right).	116
5.8	Random coefficients for undergoer's manner of articulation within speaker in a model of voicing duration in pre-sonorant obstruents (left), and voicing ratio in pre-sonorant obstruents (right).	117
6.1	Fully voiced /s/ in <i>esmoquin</i> (left) and <i>muchos hombres</i> (right) pronounced by Speaker 1	131
6.2	Partially voiced /s/ in <i>budismo</i> (left) and <i>tres autores</i> (right) pronounced by Speaker 5	132
6.3	Partially voiced /s/ in <i>posdata</i> (left), and fully voiced <i>presbitero</i> (right) pronounced by Speaker 5	132
6.4	Limited voicing of /s/ in <i>microscopio</i> (left) and <i>aviso</i> (right) pronounced by Speaker 5	133
6.5	Conditional inference tree for voicing duration measured in ms based on a model with two predictors: condition and speech rate.	134
6.6	Conditional inference tree for voicing ratio based on a model with two predictors: condition and speech rate.	134
6.7	Conditional inference tree for fricative duration measured in ms based on a model with two predictors: condition and speech rate.	135
6.8	Conditional inference tree for intensity difference measured in dB based on a model with two predictors: condition and speech rate.	136
6.9	Coefficients of rate effects by speaker in a model with voicing duration (in ms) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.	137
6.10	Coefficients of rate effects by speaker in a model with voicing ratio as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.	139

6.11	Coefficients of rate effects by speaker in a model with fricative duration (in ms) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.	140
6.12	Coefficients of rate effects by speaker in a model with intensity differnce (in dB) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.	141
6.13	Coefficients of rate effects by speaker in a model with the duration of the phrase <i>otra vez</i> (in ms) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.	142

LIST OF TABLES

1.1 Pre-sonorant voicing in the four studied languages	18
3.1 Regression coefficients, with standard error, t and p values for a model predicting the duration of vocal fold vibration (in ms) in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.	46
3.2 Regression coefficients, with standard error, t and p values for a model predicting the voicing ratio in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.	46
3.3 Regression coefficients, with standard error, t and p values for a model predicting closure duration (in ms) in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.	47
3.4 Regression coefficients, with standard error, t and p values for a model predicting burst duration (in ms) in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.	47
3.5 Regression coefficients, with standard error, t and p values for a model predicting voicing duration (in ms) in word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant in West-Flemish.	48
3.6 Regression coefficients, with standard error, t and p values for a model predicting voicing ratio in word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant in West-Flemish.	49
3.7 Regression coefficients, with standard error, t and p values for a model predicting fricative duration (in ms) in word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant in West-Flemish.	50
3.8 Regression coefficients, with standard error, t and p values for a model predicting duration of the vowel (in ms) preceding word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant.	51
4.1 Regression coefficients, with standard error, t and p values for a model predicting the duration of vocal fold vibration during closure, measured in ms, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to an underlying /d/ followed by a sonorant in the Poznań dialect.	75
4.2 Regression coefficients, with standard error, t and p values for a model predicting ratio of voicing duration to closure duration in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to an underlying /d/ followed by a sonorant in the Poznań dialect.	76

4.3	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting f_0 of a preceding vowel, measured in Hz, in stops followed by sonorants as compared with stops followed by obstruents. The intercept corresponds to a stop followed by a sonorant.	77
4.4	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting closure duration measured in ms, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to a pre-sonorant stop. . .	77
4.5	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting burst duration measured in ms, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to an underlying /d/ followed by a sonorant in the Poznań dialect.	78
4.6	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the duration of voicing during closure. The intercept corresponds to the Central dialect.	82
4.7	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the voicing ratio. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.	82
4.8	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the vowel duration. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.	83
4.9	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting f_0 of a preceding vowel, measured in Hz, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to a female speaker. .	83
4.10	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting f_1 of a preceding vowel, measured in Hz, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to a female speaker. .	83
4.11	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the closure duration. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.	84
4.12	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting burst duration. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.	84
4.13	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the duration of voicing during the fricative. The intercept corresponds to an underlyingly voiced stop (/z/) followed by a sonorant produced by a male speaker in the Central dialect.	86
4.14	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the voicing ratio in fricatives. The intercept corresponds to an underlyingly voiced stop (/z/) followed by a sonorant produced by a male speaker in the Central dialect.	87
4.15	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the duration of a vowel preceding the fricative. The intercept corresponds to an underlyingly voiced stop (/z/) followed by a sonorant in the Central dialect. . . .	88
4.16	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting f_0 of a vowel preceding the fricative. The intercept corresponds to a female speaker. .	88
4.17	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting f_1 of a vowel preceding the fricative. The intercept corresponds to a female speaker. .	89
4.18	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting duration of the fricative, measured in ms. The intercept corresponds to an underlyingly voiced pre-sonorant fricative in the Central dialect.	89

4.19	Regression coefficients, with standard error, <i>t</i> and <i>p</i> values for a model predicting the difference between mean intensity of the fricative's low frequency portion (0-900Hz) and mean intensity of the fricative (measured in dB). The intercept corresponds to an underlyingly voiced pre-sonorant fricative in the Central dialect.	90
5.1	Regression coefficients, with standard error and <i>t</i> values for a model predicting the duration of vocal fold vibration (in ms) in word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.	106
5.2	Regression coefficients, with standard error and <i>t</i> values for a model predicting ratio of vocal fold vibration to obstruent duration in word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.	108
5.3	Regression coefficients, with standard error and <i>t</i> values for a model predicting obstruent duration (in ms) in word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.	109
5.4	Regression coefficients, with standard error and <i>t</i> values for a model predicting vowel duration (in ms) preceding word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.	110
5.5	Regression coefficients, with standard error and <i>t</i> values for a model predicting f_1 (in Hz) of the vowel preceding word-final obstruents at 10 ms before the vowel offset. The intercept corresponds to a word-final sibilant followed by a sonorant.	111
5.6	Regression coefficients, with standard error and <i>t</i> values for a model predicting the f_1 (in Hz) of the vowel preceding word-final obstruents at 20 ms before the vowel offset. The intercept corresponds to a word-final sibilant followed by a sonorant.	111
5.7	Regression coefficients, with standard error and <i>t</i> values for models predicting voicing duration and voicing ratio in prevocalic obstruents. The intercept corresponds to a sibilant.	116
5.8	Regression coefficients, with standard error and <i>t</i> values for models predicting voicing duration and voicing ratio in pre-sonorant obstruents. The intercept corresponds to a sibilant.	117
6.1	Example test and baseline items. The baseline items are indicated by shading.	129
6.2	Regression coefficients, with standard error, and <i>t</i> values for a model predicting the duration of voicing in the /s/-voicing environments, measured in ms. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. <i>entusiasmo</i>) at a fast rate.	138
6.3	Regression coefficients, with standard error, and <i>t</i> values for a model predicting the duration of voicing in the /s/-voicing environments, measured in ms. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. <i>entusiasmo</i>) at a fast rate.	138
6.4	Regression coefficients, with standard error, and <i>t</i> values for a model predicting fricative duration in the /s/-voicing environments, measured in ms. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. <i>entusiasmo</i>) a fast rate.	139
6.5	Regression coefficients, with standard error, and <i>t</i> values for a model predicting intensity difference between mean low-frequency intensity and total mean intensity in the /s/-voicing environments, measured in dB. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. <i>entusiasmo</i>) at a fast rate.	140
6.6	The life cycle of /s/-voicing	145

ABSTRACT

The present dissertation is a laboratory phonology exploration of obstruent voicing before a following sonorant sound. Pre-sonorant voicing shows a number of recurrent typological properties that are challenging from a phonological point of view. First, voicing applies in the environment of a segment that is not contrastively voiced. This raises the question whether sonorants may be redundantly specified for laryngeal features, and, if so, how such specifications may be assigned. Second, voicing is positionally restricted to environments where laryngeal neutralisation is also observed, which has led some phonologists to propose a conditioning link between the two processes. Third, pre-sonorant voicing displays an array of manner asymmetries, with fricatives being more prone to undergo the process than stops, even though there is no clear evidence for a preference for voiced fricatives over voiced stops in the languages of the world. The segmental and positional asymmetries involved in conditioning pre-sonorant voicing coupled with questions concerning the role of phonological contrast have been the subject of active debate in theoretical phonology. The present thesis attempts to illuminate this debate through systematic analysis of acoustic data from four languages: West-Flemish, Poznań Polish, Central Catalan and Quito Spanish. The primary research question driving the exploration of the results is whether the studied cases of pre-sonorant voicing show evidence of phonetic gradience or whether their phonetic behaviour is more consistent with a categorical interpretation. Building up on this, I consider which aspects of pre-sonorant voicing may be phonetically conditioned, and what level of phonological abstraction is required to accommodate the observed empirical facts.

Empirical diagnostics for categoricity and gradience point towards the existence of cases of categorical pre-sonorant voicing which call for a phonological analysis. However, there is little evidence to support the hypothesis that the formal representation of pre-sonorant voicing involves redundant laryngeal specifications for sonorants. Instead, I propose that obstruents subject to pre-sonorant voicing acquire a laryngeal specification not via feature spreading, but through a context-sensitive process of default feature assignment targeting delaryngealised sonorants. This process arises diachronically from listeners misperceiving passive voicing as the realisation of a voicing target. As previously proposed by Jansen (2004), delaryngealisation facilitates extended passive voicing by removing active devoicing gestures. Passive voicing may then be reanalysed as intentional. I extend this perceptual explanation to account for manner asymmetries in pre-sonorant voicing. The argument draws on perceptual asymmetries concerning acoustic cues to voicing in stops and fricatives, which are potentially more salient in passively voiced fricatives than in passively voiced stops.

While pre-sonorant voicing reflects some influences of functional pressures shaping its diachronic development, the resulting voicing patterns are not always synchronically functional. For instance, voicing in stop+sibilant clusters but not in singleton stops, as seen in Catalan, is unlikely to reflect the direct phonologisation of phonetic pressures. Rather, I propose a diachronic scenario involving rule telescoping: a pattern of intervocalic sibilant voicing is first reanalysed as prevocalic by rule generalisation, and the voicing target assigned to the prevocalic sibilant later triggers anticipatory assimilation. The role of abstract phonology, manifested here through rule generalisation, can be further observed in the distribution of pre-sonorant voicing, which in some cases is sensitive to the boundaries of abstract morphosyntactic domains, as instantiated by the cases of Catalan and Quito Spanish.

I propose that functional phonetic and abstract phonological pressures on pre-sonorant voicing are best reconciled in a diachronic perspective, where phonetic factors play a crucial role in the initial stages of sound change. These phonetic factors are understood as influences on language use, rather than being cognitively represented in the speakers' grammars. The primary role of phonology involves replicating the patterns present in the input without direct influence from phonetic factors. This idea involves a vision of phonology which is relatively abstract, but which nevertheless retains a transparent mapping to the phonetics. Pre-sonorant voicing data lend empirical support to such a conception of phonology by showing evidence of categorical phonetic behaviour which does not correspond to phonological categories defined top-down using the criterion on contrast. Based on these cases, I argue for a bottom-up model where abstract features emerge from the continuous phonetics, and where categoricity, rather than contrast, is the criterion that determines whether or not a distinction is represented in the phonology.

DECLARATION

I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT

- i The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s policy on Presentation of Theses

ACKNOWLEDGEMENTS

I have been incredibly lucky to have Ricardo Bermúdez-Otero and Yuni Kim as my supervisors. They have always been there for me, providing insight, guidance, and encouragement. Most importantly, they kept asking hard questions, always pushing me to take my ideas further. Ricardo has been a true mentor, and he has influenced the way I think about language in more ways than he knows. Yuni has been a constant source of inspiration, and she has shown me new ways of thinking about the relevance of my work. I am also grateful to the third member of my panel, Andrew Koontz-Garboden, for support and advice throughout the last three years.

A substantial portion of this dissertation has grown out of collaborative projects. I have worked on the Belgian Dutch study together with Ellen Simon. Many thanks to Ellen for agreeing to work on this project with me, and for all her contribution. Chapter 3 of the present thesis has evolved out of our joint paper ‘Explaining pre-sonorant voicing. The case of West-Flemish’. The arguments presented in the article and later developed in the thesis have also benefitted from the review process in *Natural Language and Linguistic Theory*. The study of Quito Spanish has been made possible thanks to collaboration with Marijn van ’t Veer, Martine Bruil and Kathrin Linke. Thanks to Martine for collecting all the data, and to Marijn and Kathrin for their work on the experiment and on the paper which became the foundation of Chapter 6.

I wish to thank everyone who helped me throughout my fieldwork on Polish, especially Jarek Weckwerth for providing access to recording facilities, as well as Mateusz Szymarek and Krystyna Strycharczuk for help with participant recruitment.

My fieldwork on Catalan was completed during a study trip to Universitat Autònoma de Barcelona in Autumn 2010. I would like to thank Eulàlia Bonet and Maria-Rosa Lloret for inviting me to Barcelona, and for their warm welcome and extensive help throughout the fieldwork. Special thanks to Eulàlia for providing me with access to the recording studio and work facilities at UAB, and for her help with setting up the experiment, and recruiting participants. Eulàlia and Maria-Rosa have also given me advice on the experimental design, as have Joan Mascaró, Laia Mayol, Clàudia Pons-Moll, Daniel Recasens and Xico Torres-Tamarit. I am thankful to Maria-Josep Solé for discussion of my study, and to Xico Torres-Tamarit for his input on the Catalan chapter. I would also like to thank Carles Salse Capdevila for technical support.

The collection of the data presented in this work would not have been possible without the speakers. Many thanks to everyone who volunteered to participate in the recordings.

It has been a wonderful experience to be a part of the linguistic community at the University of Manchester, not least thanks to my friends and colleagues: Jon Morris, Michael Ramsammy and Danielle Turton. A note of thanks is also due to my former phonology teachers at the University of Tromsø: Patrik Bye and Martin Krämer, who first encouraged my interest in the discipline. I was greatly inspired to take my research in a new direction through my participation in a mini-course on the phonetics-phonology interface at the University of Tromsø in 2009, taught by Abby Cohn and John Kingston, and organised by Bruce Morén-Duolljá.

I have received a lot of useful comments from the audiences at the 33rd GLOW Colloquium

ACKNOWLEDGEMENTS

Workshop on Positional Phenomena in Phonetics & Phonology, 18th and 19th Manchester Phonology Meeting, OCP8, and the 2011 LAGB Meeting. I am especially thankful to those who read and commented on earlier versions of this work, first and foremost Koen Sebregts who provided substantial feedback and inspiring discussion throughout the entire development of my thesis research. I would also like to thank Petya Rácz and Dan Silverman for their comments on the West-Flemish paper.

Thanks to everyone who has provided methodological ideas and technical assistance to facilitate data collection and analysis. I would like to thank Vincent van Heuven for his suggestions on intensity measurements, Peter Jurgec for help with Praat scripts, and Marcin Włodarczak for help with R code.

I am most grateful to Laura Arman, Jon Morris, Laura Partridge, Michael Ramsammy, Koen Sebregts and Ellen Simon for proofreading parts of this dissertation at very short notice.

Finally, I would like to thank my family and friends for all their love and continuous support.

The research presented in this thesis was funded by the Arts and Humanities Research Council grant no. AH/H029141/1.

CHAPTER 1

INTRODUCTION

A vegetable garden in the beginning looks so promising
and then after all little by little it grows nothing but
vegetables, nothing, nothing but vegetables.

Gertrude Stein
WARS I HAVE SEEN

This dissertation is a laboratory phonology investigation of pre-sonorant voicing in four languages: West-Flemish, Poznań Polish, Central Catalan and Quito Spanish. I use the term ‘pre-sonorant voicing’ to denote a process which conditions word-, or syllable- final obstruent voicing when a sonorant consonant or a vowel follows. Pre-sonorant voicing bears considerable resemblance to a typologically common and extensively researched process of voice assimilation to obstruents, but it is relatively less frequent, and it has not as yet received a systematic phonetic description. My thesis attempts to fill this gap, documenting the phonetic properties of pre-sonorant voicing in four languages, based on experimental acoustic data.

Apart from providing a detailed description of pre-sonorant voicing, this work is also driven by a theoretical aim of relating the obtained findings to the current issues in phonological theory. The focus point of the discussion is the question of what conditions pre-sonorant voicing. In an attempt to address this issue, I consider a number of synchronic and diachronic explanations, and discuss how their predictions pan out for the synchronic patterns of variation in pre-sonorant voicing brought forth by the current data.

1.1 Pre-sonorant voicing. An overview

I talk of ‘pre-sonorant voicing’ where a final obstruent becomes voiced before a following sonorant segment. ‘Final’ can mean word-, syllable-, or prefix- final, depending on the language, and it is used in the working definition to signal the fact that the process is positionally restricted, and that all the pre-sonorant voicing languages studied in this work also have sequences of voiceless obstruents followed by vowels, and (in a subset of cases) sequences of voiceless obstruents followed by sonorant consonants. I generally use ‘pre-sonorant’ as a cover term for environments where a sonorant consonants or a vowel follows. However, I distinguish voicing before vowels and sonorant consonants in languages where such a distinction is descriptively relevant.

The four languages studied in this thesis: West-Flemish, Poznań Polish, Central Catalan and Quito Spanish, were selected for the study based on literature reports of pre-sonorant voicing (see Table 1.1 for references). Other languages where pre-sonorant voicing has been reported, but which have not been included in this study include Catalan dialects other than Central (Jiménez & Lloret, 2008), Limburg Dutch (De Schutter & Taeldeman, 1986), Walloon (Francard & Morin, 1986), Breton (Ternes, 1970), Luxembourgish (François Conrad, p.c.), Slovak (Blaho, 2008; Bárkányi &

Kiss, 2012), dialects of Ukrainian (Zilyns'kyj, 1979), Tangale (Kidda, 1993), and Isthmus Zapotec (Herrera Z., 2000).

Even a very brief survey of pre-sonorant voicing languages delivers a number of interesting typological generalisations. First, as already mentioned, pre-sonorant voicing tends to be positionally restricted, and its application cannot be defined in purely segmental terms. For instance, in Catalan prevocalic voicing applies to word-final fricatives, but it does not apply word-medially before a vowel. Examples are in (1).

(1) Prevocalic voicing in Catalan

<i>vasos</i>	[βa.sus]	'glasses'
<i>vas antic</i>	[βa.zan.'tik]	'old glass'

Quito Spanish shows a similar pattern to Catalan when it comes to prevocalic voicing; voicing applies only to word final obstruents. Sonorant consonants, however, trigger voicing in preceding obstruents also word-medially. Examples from Quito Spanish are in (2).

(2) Prevocalic and pre-sonorant /s/-voicing in Quito Spanish (Bermúdez-Otero, 2011)

<i>gasita</i>	[ga.si.ta]	'gauze, dim.'
<i>gas acre</i>	[ga.za.kre]	'acrid gas'
<i>plasma</i>	[plaz.ma]	'plasma'
<i>gas noble</i>	[gaz.no.ble]	'noble gas'

Kraków-Poznań Polish shows the same voicing patterns before all sonorant sounds that Central Catalan and Quito Spanish have prevocalically. Word-finally pre-sonorant obstruents undergo voicing, but word-medially the voicing contrast is maintained before sonorants, including vowels and consonants. Examples from Polish are in (3).

(3) Pre-sonorant voicing in Polish

<i>brat rodzony</i>	[brad.rɔ̃.d̥zɔ̃.ni]	'blood brother'
<i>setna</i>	[sɛ.'tna]	'hundredth, FEM'
<i>sedna</i>	[sɛ.'dnɑ̃]	'crux, GEN. SG.'

West-Flemish shows a similar pattern to Kraków-Poznań Polish with pre-sonorant voicing limited to the external sandhi cases. However, this is mostly seen in fricative+vowel sequences (2). Word-medially, fricatives followed by sonorant consonants only occur in borrowed words such as *moslim* 'muslim', *islam*, and *kosmos*. The realisation of the fricative as voiced or voiceless in these words seems to be variable (cf. Simon (2010, 135)).

(4) Prevocalic voicing in West-Flemish

<i>dat mens is</i>	[dat.mɛn.zɪs]	'that person is'
<i>jassen</i>	[ja.sən]	'coats'

Apart from being positionally restricted, pre-sonorant voicing is frequently limited in its application to fricatives. An example is prevocalic voicing in Catalan, which applies to fricatives, as in (1), but it fails to apply in stops. Word-final voiceless stops in Catalan surface as voiceless before a following vowel, as illustrated in (5).

(5) No prevocalic voicing in Catalan stops (Jiménez & Lloret, 2008)

<i>sap aixo</i>	[sa.pə.'ʃɔ̃]	'I know that'
-----------------	--------------	---------------

A similar asymmetry between stops and fricatives as undergoers of voicing is found in West-Flemish, where reports of pre-sonorant voicing single out fricatives as the undergoers of voicing, while similar reports of stops are missing. Collins & Mees (1999, 214) note that voiced realisations are frequent in Dutch coda *fricatives* (emphasis added) followed by a vowel. Similarly, De Schutter

& Taeldeman (1986) report that in West-Flemish word-final fricatives are usually realised as voiced when the next word begins in a sonorant. Simon (2010) found pre-sonorant voicing in fricatives, but not in stops in the production of West-Flemish speakers. A summary of the the pre-sonorant voicing environments, triggers and undergoers in the four languages studied in this work is in Table 1.1.

Table 1.1: Pre-sonorant voicing in the four studied languages

Language	Domain for voicing	Triggers & Targets		References
West-Flemish	word-final	all sonorant sounds	fricatives	De Schutter & Taeldeman (1986) Weijnen (1991) Simon (2010)
Kraków-Poznań Polish	word-final	all sonorant sounds	all obstruents	Benni (1959) Ostaszewska & Tambor (2000)
Central Catalan	word-final syllable-final prefix-final	sonorant consonants vowels	all obstruents fricatives	Bonet & Lloret (1998) Wheeler (2005) Jiménez & Lloret (2008)
Quito Spanish	word-final syllable-final prefix-final	vowels sonorant consonants	sibilants	Robinson (1979) Lipski (1989)

1.2 Theoretical issues

As seen in the previous section, even a basic description of where pre-sonorant voicing applies can be somewhat complex. Theoretical complexities inevitably follow, as one tries to supplement a description with an explanation of why pre-sonorant voicing occurs when and where it does. In relation to this at least four issues can be identified:

1. Why are non-contrastively sonorants able to trigger voicing in a preceding obstruent? Can we admit the possibility that sonorants are laryngeally active despite the absence of voicing contrast, and how do we square that with the observation than in other languages non-contrastively voiced sonorants are not laryngeally active?
2. Why does prevocalic/pre-sonorant voicing only target fricatives, but not stops in Catalan and West-Flemish?
3. Where do the positional restrictions in pre-sonorant voicing come from?
4. Does pre-sonorant voicing apply at the level of phonology? Which aspects of pre-sonorant voicing require a phonological explanation, and which can/should be explained by extra-phonological factors?

1.2.1 Do sonorants spread [voice]?

The central question which transpires with relation to pre-sonorant voicing is whether or not the voicing consists in feature spreading from the sonorant. Allowing such possibility would entail equipping sonorants with a laryngeal feature, which I shall refer to as [voice] for the purpose of the current exposition, even though sonorants are laryngeally not contrastive. The existence of pre-sonorant voicing has been taken as *prima facie* evidence that lexical contrast alone does not determine featural specification. With relation to this issue Kiparsky (1995) draws a line between distinctive and redundant features. The former are lexically specified, while the latter may be assigned in the phonological derivation, typically postlexically. [voice] is not a distinctive feature for sonorants, as it carries no phonological contrast. However, allowing for [voice] assignment to sonorants in the course of a derivation predicts the existence of a system, where [voice] may

spread for sonorants. Kiparsky identifies Kraków-Poznań Polish as such a system, and contrasts it with Warsaw Polish, where [voice] assignment to sonorants follows the assimilation rule, and so pre-sonorant voicing does not apply. The prediction is cited in support of the argument that phonological features can be redundant. The same point is explicitly argued by Blaho (2008) based on pre-sonorant voicing in Slovak:

“Slovak provides evidence that *redundant feature specifications* can play a role in phonology: sonorants and vowels do not contrast for [voice] in this language, but the fact that they cause obstruents to become voiced shows that they are specified for this feature nevertheless.”(Blaho, 2008, p.235)

That sonorants should have redundant laryngeal features has been assumed by a number of formal analyses of pre-sonorant voicing, including Bethin (1984), Gussmann (1992) and Rubach (1996) for Polish, and Jiménez & Lloret (2008) and Bermúdez-Otero (2001) for Catalan.

However, it is not accepted by everyone that sonorants should be laryngeally active. Jansen (2004) states that only actively voiced sounds are able to trigger anticipatory voice assimilation, actively voiced sounds being the ones whose voicing is a result of a pre-planned gesture. Jansen argues that no such planning is present for sonorant voicing. Instead, sonorant voicing follows spontaneously from the low supraglottal pressure associated with open articulations¹.

Where does pre-sonorant voicing come from then? In the light of Jansen’s proposal, the actual source of voicing is the preceding vowel rather than the following sonorant. Sonorants, as previously explained, are inherently voiced. Once vocal folds have been set into motion, for instance during sonorant production, they naturally keep on vibrating for a certain amount of time (40 ms is the average estimate by Westbury & Keating (1986)) unless an actively voiceless target follows. If the following sound is *actively* devoiced, some devoicing gesture may be executed by the speaker that counteracts voicing, such as raising the larynx, tensing the vocal tract walls, glottalisation, or glottal abduction.

However, word-final obstruents in final-devoicing languages do not have their own voicing targets in Jansen’s (2004) model. Thus, in these obstruents uncounteracted passive voicing can continue longer. Prolonged passive voicing can make obstruents be perceived as categorically voiced by listeners. Such perceptions, in turn, might give rise to reinterpretation as a categorical pattern, which eventually stabilises as a result of language change. In that way, pre-sonorant voicing is synchronically a reflection of a perceptually motivated diachronic change, but the synchronic grammar need not necessarily directly refer to the factors that motivate the change in question.

1.2.2 Fricatives vs. stops

A recurrent peculiarity of pre-sonorant voicing is that it tends to target only a subclass of obstruents, typically fricatives or sibilants. Reports of pre-sonorant voicing in West-Flemish single out fricatives as the undergoers of voicing, while similar reports of stops are absent². Collins & Mees (1999, 214) note that voiced realisations are frequent in Dutch coda *fricatives* (emphasis added) followed by a vowel. Similarly, De Schutter & Taeldeman (1986) report that in West-Flemish word-final fricatives are usually realised as voiced when the next word begins in a sonorant. Finally,

¹Jansen (2004, 36) provides the following definition of spontaneous (passive) voicing: “Sounds or parts of sounds are said to be passively voiced if a closed equilibrium position of the vocal folds and normal subglottal pressure (according to Stevens (1998), 8000 dyne/cm² / 800 Pa is typical) are sufficient to initiate or maintain the physical conditions for vocal fold vibration. Sonorants are typical examples of passively voiced sounds: because their supralaryngeal articulations allow air to escape freely from the supraglottal vocal tract (either through the oral or nasal tract or both) the supraglottal pressure during these sounds remains approximately equal to atmospheric pressure.”

²There are, however, reports of stop voicing word-finally before a vowel, e.g. /po:t/ is / → [po:dis] in some Limburg dialects and in the northwest of East Flanders (De Schutter & Taeldeman, 1986)

Simon (2010) found pre-sonorant voicing in fricatives, but not in stops in the production of West-Flemish speakers. A similar situation is found in Ecuadorian Spanish, where prevocalic voicing is restricted to sibilants. Word-final stops are also exempted from pre-sonorant voicing in most Catalan dialects³.

Stops pattern with fricatives with respect to final devoicing and voice assimilation to obstruents, which makes the pre-sonorant asymmetry puzzling. Most authors attribute the disparate behaviour of the obstruent subclasses to phonetics. Wheeler (2005) proposes an articulatory explanation implemented in a specific constraint (**LAZYSIBILANT**) which voices sibilants before a vowel. However, the generalisation that intervocalic voicing is natural holds for both stops and fricatives, and Westbury & Keating (1986) show how voicing may be favoured in intervocalic stops. This being the case, should not a constraint like **LAZYSTOP** also be posited? An analysis that draws constraints from independently motivated observations concerning phonetic naturalness ought to admit **LAZYSTOP** alongside with **LAZYSIBILANT**. It might be that **LAZYSTOP** is implicitly assumed in Wheeler's (2005) analysis but is absent from the tableaux, as its low ranking suppresses its activity in the language. But what is the source of this putative low ranking? Is it a universal observation that fricatives are more prone to voicing than stops? Such generalisation does not square with the typological rarity of voiced fricatives, as reported by Ladefoged & Maddieson (1996)⁴. If, on the other hand, the ranking of **LAZYSIBILANT** and **LAZYSTOP** is free, and Catalan just happens to disfavour intervocalic stops, the question of where this sort of ranking may come from still remains.

Alternative proposals involve linking voicing to the degree of stricture (Jiménez & Lloret, 2008), or continuancy (Bermúdez-Otero, 2001). However, these accounts receive theoretical and typological, but not functional support, as the authors offer little discussion of why the degree of stricture or continuancy might affect voicing. In the absence of such arguments, the markedness restrictions used by Jiménez & Lloret (2008), and Bermúdez-Otero (2001) represent possible formalisations of the stop-fricative asymmetry, but they do not provide an explanation for why such an asymmetry should occur.

1.2.3 Positional restrictions. The role of delaryngealisation

One of the challenges to any analysis of pre-sonorant voicing is accounting for the environments where the process applies. Pre-sonorant voicing does not apply across the board in any of the languages studied in this thesis, unlike voice assimilation to obstruents which is relatively less positionally restricted. In Polish, for instance, regressive voice assimilation applies across word boundaries as well as within words, whereas pre-sonorant voicing is limited to external sandhi contexts. While the exact set of positions where pre-sonorant voicing is seen varies between languages (see Section 1.1 for an overview), the following generalisation holds for all the languages in the current sample: pre-sonorant voicing targets obstruents in the positions where voicing is non-contrastive. This includes the word-final position and in the case of Spanish and West-Flemish also word-medial codas.

The link between laryngeal neutralisation and pre-sonorant voicing has been singled out as crucial by virtually every analysis of pre-sonorant voicing to date. For Kraków-Poznań Polish Gussmann (1992) and Rubach (1996) all propose [voice]-spreading rules that target delaryngealised obstruents. Cyran (2012) departs from those earlier analyses in many aspects, but he also asserts that any analysis of Kraków-Poznań voicing must involve a delaryngealisation process. Similarly,

³ Jiménez & Lloret (2008) report a dialect continuum in Catalan, where all dialects except Central Valencian have sibilant voicing before vowels. In addition to sibilant voicing, Central Catalan has (variable) /f/ voicing, while Alicantino is reported to voice all obstruents.

⁴ For further discussion on the role of naturalness in stop vs. fricative voicing see Silverman (2006a, 164–65) and Chapter 3 of this thesis

word-final neutralisation is proposed by Bermúdez-Otero (2001) and Jiménez & Lloret (2008) for Catalan. For Quito Spanish Bermúdez-Otero (2011) argues that pre-sonorant voicing is conditioned by delaryngealisation at the Word Level. Colina (2009) rejects an ordered analysis, arguing for a monostratal model, but also argues that pre-sonorant voicing is conditioned by word-final obstruents and word-medial codas being unable to license their own laryngeal targets.

Delaryngealisation conditions pre-sonorant voicing also in the diachronic model proposed by Jansen (2004). As discussed in Section 1.2.1 word-final obstruents do not have their own laryngeal targets in Jansen's model. This property makes them more amenable to spontaneous voicing from the neighbouring sonorants. In that way intersonorant voicing is more likely to develop in positions where laryngeal contrast had been lost.

Delaryngealisation appears to be the common denominator of most analyses of pre-sonorant voicing, and a key to explaining where voicing is expected to apply. However, there is no agreement on how to conceptualise the link between delaryngealisation and voicing. In a diachronic model the relationship between the two consists in a succession of events in the course of linguistic change. The same kind of succession is replicated, as it were, in some synchronic models, including spreading analyses of Polish (Bethin, 1984; Gussmann, 1992; Rubach, 1996). Similarly, Bermúdez-Otero's multi-stratal model of Catalan allows for delaryngealisation to precede voicing. This kind of analysis, however, crucially requires some sort of a derivational framework, especially given the overapplication effects observed in some languages, notably Quito Spanish (a detailed discussion of the opacity facts is in Chapter 6). Those synchronic models that reject derivations, however, need alternative means of preserving the feeding relationship between delaryngealisation and pre-sonorant voicing. This brings us to the last point of contention concerning pre-sonorant voicing, namely that of the phonological or phonetic status of voicing.

1.2.4 Phonology vs. phonetics

A number of theoretical analyses of pre-sonorant voicing have implicitly assumed its phonological status and provided an account within a phonological model. However, the phonological status of the voicing process has occasionally been questioned. This issue has surfaced recently in the context of a debate on Quito /s/-voicing between Colina (2009) and Bermúdez-Otero (2011). Both authors acknowledge that a phonological analysis of prevocalic sibilant voicing in Quito Spanish requires minimally two steps of derivation due to overapplication effects. Overapplication consists in coda sibilant voicing before sonorants to word-final prevocalic sibilants which surface as onsets. Bermúdez-Otero (2011) argues that this overapplication cannot be explained as conditioned by correspondence to another surface form, as there is no appropriate base form with a word-final voiced variant. According to Colina (2009), however, the challenge is only apparent, assuming that a phonological output correspondence relationship can hold between two members of a paradigm that are phonetically quite distinct. Colina proposes that word-final pre-vocalic sibilants in Quito Spanish, as well as all pre-sonorant sibilants, undergo phonetic, but not phonological voicing. At the level of phonology, goes the argument, the /s/-voicing tokens are delaryngealised archiphonemes by correspondence to word-final prepausal sibilants.

A similar argument is made for Kraków-Poznań Polish by Cyran (2012). Cyran's proposal is couched within a non-derivational framework of Government Phonology, which precludes ordering phonological processes directly. In addition, Cyran rejects the idea that non-contrastively voiced sonorants could trigger spreading. Instead, he proposes that word-final obstruents undergo delaryngealisation at the level of phonology, but they undergo voicing before sonorants at the phonology-phonetics interface, and surface as phonetically voiced.

In response to Colina's (2009) proposal for Quito Spanish, Bermúdez-Otero (2011) points out that the assumption concerning the phonetic status of Quito /s/-voicing lends itself to an

empirical test. Spontaneous voicing operating on phonetically underspecified targets is expected to be variable and gradient. A phonological voicing process, on the other hand, is expected to be categorical. Phonological reasoning alone is insufficient to ascertain a categorical or gradient status of any speech phenomenon, as shown by laboratory research of the past two decades. For instance, an EPG study of American English palatal place assimilation by Zsiga (1995) shows that a process previously assumed to be phonological shows characteristics of phonetic gradience. Ellis & Hardcastle (2002), on the other hand, present evidence that external sandhi can also be categorical. Ellis & Hardcastle's articulatory data on English /n#k/ sandhi show in fact that even the same pattern may be gradient or categorical depending on the speaker. Further empirical investigations concerning categoricity and gradience include Barry (1992) Cohn (1993), Holst & Nolan (1995), Nolan et al. (1996), and Tucker & Warner (2010), *inter alia*. If there is one common conclusion to be drawn from these studies, it is that neither categoricity nor gradience can be safely assumed in the absence of empirical evidence. Consequently, the explanatory adequacy of any formal model critically hinges on whether or not its assumptions concerning categoricity and gradience are indeed met by the speech data.

Coming back to the case at hand, every phonological analysis of pre-sonorant voicing there is seems to have committed, explicitly or not, to treating the process as either gradient coarticulation (Colina (2009) for Quito Spanish, Cyran (2012) for Kraków-Poznań Polish), or categorical spreading (Bethin (1984), Gussmann (1992) and Rubach (1996) for Kraków-Poznań Polish, Bermúdez-Otero (2001) and Jiménez & Lloret (2008) for Catalan, Bermúdez-Otero (2011) for Quito Spanish). As a consequence of these commitments, all the above models make empirical predictions about the nature of pre-sonorant voicing, even though such predictions are rarely spelled out (Bermúdez-Otero (2011) is a notable exception), nor are they empirically supported.

1.3 Current approach. Synopsis

This study undertakes a systematic investigation of pre-sonorant voicing guided by the research questions opened up by the existing theoretical debate. My aim is to address the four issues outlined in the previous section from the perspective offered by phonetic data. The question of phonological vs. phonetic status of pre-sonorant voicing is taken as the starting point. Based on evidence for categoricity or gradience I discuss which aspects in the conditioning of pre-sonorant voicing belong to the phonological component of the grammar, and which seem to be phonetic. I also argue that a closer phonetic scrutiny of pre-sonorant voicing data coupled with a cross-linguistic comparison can inform various aspects of the ongoing theoretical debate. The argument is built up as follows.

Chapter 2 outlines the model of the phonology-phonetics interface that I shall use to approach the data. I discuss the rationale for assuming such a model in the context of conflicting views on the phonology-phonetics distinction, and I provide empirical diagnostics for distinguishing phonetic and phonological processes developed by earlier studies in phonetics and laboratory phonology.

Chapter 3 presents data on pre-sonorant voicing in West-Flemish. Based on the evidence of a bimodal distribution of voicing duration and voicing ratio I argue that fricatives in West-Flemish undergo optional but categorical voicing when they are followed by a sonorant consonant or a vowel in the next word. West-Flemish stops in comparison show no evidence of categorical pre-sonorant voicing. I then consider whether the West-Flemish voicing pattern can be adequately explained by assuming a synchronic analysis of [voice] spreading from a following sonorant. I argue that there is evidence against assuming [voice] specifications for sonorants, as sonorants trigger no phonetic coarticulation, or indeed phonological [voice] assimilation in the preceding stops. Instead, the synchronic pattern found in West-Flemish can be adequately explained in a diachronic perspective. I propose a scenario where the process originates as passive voicing, and

becomes perceived as intended to later become categorical without an intermediate step of [voice] assignment to sonorants. I also argue that a perception-driven language change model could explain the asymmetry we find in voicing between stops and fricatives, as perceptual cues to voicing in fricatives make them likely to be perceived as voiced than it is the case with passively voiced stops.

Chapter 4 discusses the issue of voicing in Poznań Polish. A careful phonetic investigation confirms the existence of a pre-sonorant voicing pattern in Poznań Polish, but the data also deliver two hitherto unreported generalisations. First, a subset of the data delivers evidence for non-neutralisation of the underlying voicing contrast in pre-sonorant voicing environments, as underlying voiced stops and fricatives tend to surface as fully voiced more frequently than their underlyingly voiceless counterparts. This finding challenges the generalisation that delaryngealisation is a prerequisite to pre-sonorant voicing, be it in a formal synchronic model, or in a chain of diachronic changes. However, the data do not constitute conclusive proof that pre-sonorant voicing may operate on laryngeally specified targets, as the variation can also be explained using a stochastic phonological model, where both delaryngealisation and pre-sonorant voicing operate optionally, but remain in a feeding relationship nonetheless. Therefore, although Poznań Polish could be a non-neutralising exception in the typology of pre-sonorant voicing languages, the exceptionality may in fact be only apparent. The Poznań data also provide further evidence against common laryngeal specifications for stops and fricatives, previously argued on the basis of the West-Flemish case. As pre-sonorant voicing and voice assimilation to obstruents apply with varying degrees of frequency in Poznań Polish, sonorants cannot be seen as but a subset of voiced sounds that all behave as a coherent group.

Chapter 5 illustrates pre-sonorant and prevocalic voicing in Central Catalan. Catalan displays an array of complex manner asymmetries with respect to pre-sonorant voicing, including a disparate behaviour of vowels and sonorants in triggering voicing in a preceding word-final obstruent. Some of these effects had previously been reported, such as the presence of prevocalic voicing in sibilants and stop+sibilant clusters in the absence of prevocalic stop voicing. Phonetic data presented in Chapter 5 confirm this generalisation. In addition, the data deliver new findings on variable application of pre-sonorant voicing depending on the undergoer's manner of articulation with a high degree of inter- and intra-speaker variation. The case of undergoer asymmetries with respect to pre-sonorant voicing is argued to pose a serious problem to functional synchronic grammars, as the presence of voicing in clusters but not in singleton stops is not functionally motivated. From a diachronic point of view the origin of such a pattern can be explained via rule generalisation followed by rule telescoping. The implication of this interpretation for synchronic grammars is that phonological learning must have an inductive component, and it may also involve a bias for formal simplicity. I further consider the role of markedness in the origin and learning of sound patterns, arguing that markedness effects may follow from induction rather than drive phonological learning. Finally, I discuss the asymmetry between vowels and sonorant consonants as triggers of voicing in Catalan. The asymmetry may indicate that pre-sonorant and prevocalic voicing developed as two separate processes, although a phonetic explanation for this does not seem to be readily available. At the same time, however, potential representational solutions to trigger asymmetries do not deliver either, as the feature specifications necessary to capture the observed variation are not supported by lexical or phonetic evidence.

The focus of Chapter 6 is the issue of categoricity vs. gradience in pre-sonorant and prevocalic voicing. The Quito data involve a potential case of phonological opacity, whose status, however, is sensitive to the /s/-voicing being categorical. The question of whether or not /s/-voicing is a gradient or a categorical process is addressed based on evidence from speech rate manipulations. The influence of speech rate on voicing duration and voicing ratio is taken as evidence in support of /s/-voicing being categorical for some speakers, whereas others show phonetic behaviour consistent with a gradient analysis. The existence of a categorical voicing pattern even in a subset of speakers

undermines the models which relegate voicing to the level of phonetics on theoretical grounds. In contrast, multi-level models of phonological computation receive support from the Quito data, as aspects of serial derivation appear necessary to provide a coherent phonological generalisation over the environments where /s/-voicing applies. Finally I consider how the Quito Spanish pattern may be accounted for in non-generative model by shifting some of the explanation for the observed distribution from phonological computation to diachronic influences and lexical storage.

Chapter 7 proceeds to reconcile pieces of evidence on the nature of pre-sonorant voicing provided by individual languages in a more general perspective. The main emerging empirical generalisations concerning pre-sonorant voicing involve a link between voicing and delaryngealisation, manner asymmetries with respect to voicing, and recurrent differences between pre-sonorant and pre-obstruent voicing. Based on these findings I argue that pre-sonorant voicing reflects a wide array of phonetic and phonological influences, which are best understood in a diachronic perspective as influences on sound change. Specifically, I propose that pre-sonorant voicing is diachronically conditioned by delaryngealisation, and that it develops through perceptual reinterpretation. This scenario provides a potential explanation for the manner asymmetries seen in pre-sonorant voicing, although a full appreciation of the manner effects also requires insights from phonological influences on the development of sound change. I further consider the role of the right-hand environment and the possibility that pre-sonorant voicing develops as a result of sonorants acquiring a laryngeal specification. I propose that a combination of articulatory, aerodynamic and perceptual factors may lead to the development of voicing in the absence of redundant laryngeal specifications for sonorants, and that the existence of such specifications is not supported by the patterns of variation found in the current data. Finally, I consider the implications of the present work to conceptions of abstract phonology and its relationship with phonetics.

Chapter 8 concludes the discussion, and comments on the falsifiability of my proposal, formulating some empirical predictions concerning speech perception and sound change. It also suggests directions for laboratory research on the role of trigger asymmetries in pre-sonorant voicing, and for computational research on phonological category learning.

CHAPTER 2

THE PHONOLOGY-PHONETICS INTERFACE

‘Oh, it doesn’t really matter. I’m just the Appointed Cynic. I merely react to other people’s ideas. What do *you* think?’

‘I’m just the Ideas Catcher. I catch ideas. I don’t have any of my own.’

Julian Barnes

ENGLAND ENGLAND

The present work is concerned with disentangling the phonological and phonetic aspects of pre-sonorant voicing, and central to this enterprise is the idea that phonology and phonetics are distinct, and distinguishable. This view, although axiomatic to some strands of linguistic research, is not universally shared. In this chapter I provide my rationale for assuming a phonology-phonetics distinction, in the context of the ongoing debate on the nature of the relationship between phonology and phonetics.

2.1 Phonetic and phonological knowledge

I understand phonology and phonetics to represent different aspects of the linguistic knowledge of a language user. I do not make any claims that phonetics and phonology are different modules in the sense of Fodor (1983), who argues for an architectural organisation of the brain with individual, highly specialised, genetically specified and largely independent units. Instead, I assume a phonology-phonetics distinction based on the observation that knowledge of language, or more specifically knowledge of the sounds of a language, involves different degrees of abstraction which can be operationalised by using two levels of representation: phonological (categorical and discrete) and phonetic (gradient and continuous). Phonetic knowledge allows language users to analyse the continuous acoustic speech signal in terms of abstract units, and to translate those units into specific articulations in speech production. Phonological knowledge, on the other hand, deals with the categories that language users abstract away from the phonetics in order to produce categorical phonetic behaviour in their own speech.

Evidence for phonetic knowledge goes back to experimental findings on language-specific phonetics (e.g. Cho, 1999). Languages differ systematically in their exact phonetic realisation of similar phonological categories, such as ‘front unrounded vowel’, or ‘aspirated labial stop’ (e.g. Pierrehumbert et al., 2001, pp.285–286, and references therein). This shows that the exact realisation of phonological categories in continuous phonetic space must be a part of a native speaker’s competence. Recent advances in sociophonetics show that phonetic knowledge is more than language-specific, as the kind of competence responsible for associating specific regions in the acoustic and articulatory space with specific phonological categories is also sensitive to social factors. For instance, in a study of spectral characteristics of /s/ realisations in Glaswegian speech

Stuart-Smith (2007) found a spectral difference between working-class and middle-class females. The working-class females had a lower centre of gravity in their /s/ realisations, and patterned together with Glaswegian male speakers. Similar findings of fine speaker control over phonetic detail as a marker of social class, age, gender, or ethnicity are reported by a growing body of sociophonetic studies (see Foulkes et al. (2010) for an overview).

The conception of phonetic knowledge that emerges from phonetic and sociophonetic research is rather powerful, as it allows language users to perceive, store and reproduce complex sound patterns down to very fine phonetic detail. This view of phonetics also calls for re-thinking of what, if any, is the role of phonology in encoding and shaping knowledge of sound patterns. Certainly the fact that sound patterns are to a large extent analysable in terms of abstract categorical interactions does not entail that those interactions are psychologically real for the language users. Yet, phonological influences on phonetics and on sound change, as well as psycholinguistic findings of category salience suggest that phonology plays a role in the organisation of language in the speaker's mind with abstract categories as an intermediate step in coupling sound with meaning.

In an insightful assessment of modular feedforward models (cf. 2.2), Pierrehumbert (2002) notes that the evidence for phonological categories is supported by early psycholinguistic research on speech errors, such as the study by Shattuck-Hufnagel (1979), whose findings are consistent with a model involving a phonological buffer between the lexicon and speech production (see also Goldrick & Blumstein (2006); Goldrick et al. (2011)). Pierrehumbert (2002) cites prosodic constraints on phonetic processes as further evidence for abstract phonology. Phrasal prosody has been shown to affect a wide-range of phonetic parameters, including f_0 , duration, aspiration and glottalisation. Since phrasal prosody can be assigned to novel utterances, it has to be computed online, which supports the idea of a modular architecture, where prosodic categories feed into phonetic realisation. Although Pierrehumbert (2002) does not ultimately subscribe to the view that grammar organisation is modular and feedforward, she proposes a hybrid model which to an extent relies on abstract phonological representations. Similarly, aspects of abstract phonological encoding are preserved in hybrid models of word recognition by Hawkins (2003), McLennan et al. (2003), and Goldinger (2007).

Salience of phonological categories receives striking support from psycholinguistic studies by Maye et al. (2002), and Maye et al. (2008) which show that category membership affects phonetic discrimination in infants. Maye et al. (2002) performed an experiment examining the perception of equidistant stimuli from a VOT continuum ([da]-[ta]) amongst infants of 6 and 8 months. The infants were familiarised with a continuum of [da]-[ta] realisations. The stimuli in the familiarisation stage were unimodally distributed for one group of infants, and bimodally distributed for the other group. The infants were then tested for their discrimination of the end-points of the continuum. In the test phase, only the infants from the bimodal condition showed discrimination of the end-point tokens. In a later study Maye et al. (2008) found that familiarising 8-month-old infants with bimodally distributed acoustic continua facilitated learning of difficult phonetic contrasts. These results suggest that category membership reflected in the statistical distribution of acoustic parameters plays a vital role in speech perception, and supports the idea of a hierarchical organisation of speech, where phonetic tokens map onto abstract units.

Phonological category membership also plays a role in sound change. Janda (2003) notes this in his discussion of Twaddell's (1938) account of High German umlaut. Twaddell (1938) proposes that Old High German did not indicate front rounded vowels, [ü(:)] and [ö(:)], as they were both allophones of back vowels, [u(:)] and [o(:)] in the context of a following high vowel or palatal glide. In Twaddell's account, it was only the reduction or loss of the conditioning environment that led to the creation of phonemes [ü(:)] and [ö(:)] in Middle High German. However, if [ü(:)] and [ö(:)] had indeed been only allophones in Old High German, the loss of conditioning environment should have triggered the loss of the allophone. The fact that [ü(:)] and [ö(:)] emerged as phonemes after

the loss of the following /i(:), j/ indicates that the fronted rounded vowels must have formed a separate category prior to the phonemic split.

The role of phonological categories is also of relevance to the way sound change typically develops over time. Janda (2003) notes that sound changes do not continue indefinitely, but they stabilise over time instead. Bermúdez-Otero & Trousdale (2012) take up this issue, arguing that stabilisation is associated with phonological category formation. Bermúdez-Otero & Trousdale analyse sound change as following a specific unidirectional life cycle, illustrated in (1).

- (1) The life cycle of phonological process Bermúdez-Otero & Trousdale (2012)



The diagram in (1) represents that sound change originates in the external factors, such as constraints on the vocal tract, or (mis)perception, as explored by Ohala (1981, 1983). The next step in the development of sound change happens when the externally motivated phonetic phenomenon is perceived and interpreted by listeners as a part of phonetic grammar, and is then implemented by the listeners when it is their turn to speak. This systematic implementation of gradient phonetic processes in a speech community has been observed for instance in the Northern Cities æ-tensing (Labov, 1981, 1994). The salience of phonetic categories plays a role in the next stage of sound change development, that is stabilisation (Hayes & Steriade, 2004; Bermúdez-Otero, 2007; Bermúdez-Otero & Trousdale, 2012). Bermúdez-Otero & Trousdale (2012) use the Ellis & Hardcastle (2002) data on English /n#k/ sandhi (cf. 2.3.2) to illustrate phonologisation and stabilisation. Ellis & Hardcastle (2002) present an EPG and EMA study of place assimilation in English, based on a comparison of realisations of underlying /n/ and underlying /ŋ/ before a word-initial /k/. The results show a number of different idiolectal strategies with respect to the realisation of /n#k/ sequences. Some speakers consistently executed a full coronal gesture in their realisation of /n/ before /k/. The next group of speakers showed gestural overlap with a residual coronal gesture blending with the dorsal one. The third group of speakers assimilated in a categorical fashion, with no trace of a coronal gesture. Finally, some speakers alternated between full assimilation with no residual coronal gesture and no assimilation at all with the coronal gesture fully realised. Bermúdez-Otero & Trousdale (2012) take the different idiolectal strategies revealed by the Ellis & Hardcastle (2002) study to correspond to three different stages of sound change development. The speakers who made a full coronal gesture in the /n#k/ sandhi context are the most conservative ones, with no phonetic coarticulation rule. The next group of speakers, showing gestural blending of a residual coronal and a dorsal gesture represent the next stage of the sound change development, as for them the coarticulation rule has been phonologised, i.e. it has entered the phonetic grammar. For the speakers who showed no evidence of a residual coronal gesture, Bermúdez-Otero & Trousdale argue that the coarticulation rule has entered the stabilisation stage of its life cycle, where the rule begins to affect phonological categories, in this case deleting the feature [Coronal]. Crucial support for the interpretation of assimilation in categorical terms comes from speakers who alternated between full assimilation with no residual coronal gesture and no assimilation at all. While either of those realisations could be interpreted as an endpoint in a continuum of gradient overlap, a categorical alternation between two extreme articulatory endpoints is not consistent with a gestural blending account. Instead, this type of variation suggests that a categorical reinterpretation has taken place, where the speakers optionally delete or retain the coronal place feature. This is not to say that only assimilation (but not coarticulation) is a part of the speaker's grammar. Rather, coarticulation and assimilation apply at different levels of grammar. Assimilation targets abstracted categories, deleting the [Coronal] feature in the case of most innovative speakers who showed no residual coronal gesture in their /n#k/ sandhi.

Coarticulation, on the other hand, takes place at the level of the phonetic grammar, which is still controlled by the speaker, but which does not affect the speaker's discrete representation of a word.

The role of categorical phonology also becomes apparent in the ways phonology interacts with morphology, as argued by Hyman (2001). Hyman views phonology as an intersection between phonetics and grammar. The phonetic side of phonology is conditioned by the diachronic development of phonology from phonetics via phonologisation (or stabilisation in Bermúdez-Otero & Trousdale's terms). Stabilisation of phonetically natural processes is, according to Hyman, the reason why phonological processes are largely natural. However, there is also an unnatural side to phonology, which can only develop through the interaction of grammatical pressures, or abstract reanalysis by means of e.g. rule telescoping, or rule inversion. Hyman (2001) discusses the case of phonetically unnatural progressive place assimilation in Noni (Bantu, Cameroon). Place assimilation tends to be regressive, a fact which Ohala (1990a) attributes to the fact that in a sequence of two stops the first one tends to be unreleased, which makes its place of articulation difficult to perceive. This perceptual difficulty might be responsible for a sound change where the place feature of the first consonant in a sequence is lost. While a similar explanation is available for nasal+stop sequences, which squares with the observation that nasals tend to assimilate its place to that of a following stop, place assimilation in Noni is a counterexample. As illustrated in (2), the first consonant of the suffix /-te/ in Noni assimilates its place to the preceding dorsal.

- (2) Realisation of the progressive suffix in Noni (Hyman, 2001, p.145)

cím	'dig'	cim-tè	'be digging'
dvum	'groan'	dvùm-tè	'be groaning'
cíŋ	'tremble'	ciíŋ-kè	'be trembling'
káŋ	'fry'	kaaŋ-kè	'be frying'

Hyman argues that the phonetically unnatural progressive place assimilation is a paradigmatic effect where preserving base features is preferred to preserving features of the affix. Similar phonological asymmetries between base and affixes are frequently observed, reinforcing the argument that while there is an intersection between some aspects of phonology and phonetics, the two are not one and the same thing, as phonetics also intersects with physical and physiological events external to language, while phonology intersects with higher levels of grammar, including morphology. While a handful of cases have been reported with morphological distinctions directly influencing phonetics have been reported (e.g. English *t/d*-deletion, Coetzee & Pater (2011)), their status is controversial. According to Bermúdez-Otero (2010), the currently available data do not warrant the conclusion that the phonological level is not involved in any of the purported cases of direct morphology-phonetics interaction.

2.2 Modularity

The empirical arguments for phonetic and phonological knowledge are related to the classical problem of dual nature of speech which deals with continuous and gradient events on the one hand, and symbolic categories on the other. Ladd (2011) traces this problem back to the 19th century and the origin of the phonemic principle, which states that speech can be represented by means of a string of contrastive phonemes. However, arguably, the abstractness of speech holds independently of the phonemic principle, as categorical and symbolic behaviour can be observed on a number of different levels, from phonological features to morphophonological categories. Linguists have long been debating the ways to model the abstract and the physical aspects of speech. The standard solution to the dual nature of speech adopted within the generative tradition is modularisation, i.e. the idea that the abstract and the physical aspect of speech are confined to different modules within the brain. This idea goes back to Chomsky & Halle (1968), and originally corresponded to

the Chomskyan competence-performance distinction, under which view phonetics was not seen as a part of grammar. Early generative hypotheses assumed a universal phonetics, which maps language specific featural specifications onto the physical domain. This hypothesis has not been upheld by advances in phonetic research, which reveal that phonetic realisations of similar phonological categories are language-specific. However, while these findings go against the idea of universal phonetics, or indeed the idea that phonetics is not grammar, a revised view of modular architecture has emerged based on more data-driven approaches to phonology and phonetics. This view is associated with works by Keating (1988, 1990, 1996), Pierrehumbert (1980) and Cohn (1990, 1993). The term *modular mapping* is from Cohn (2006), who provides the following succinct summary of the basic idea behind the approach:

“The phonology (...) is understood as the domain of discrete, qualitative patterns and the phonetics as the domain of the continuous, quantitative realization of those patterns. Intrinsic to this view is the idea that lexical entries and phonological patterns are represented in terms of distinctive features, taken to be abstract properties, albeit defined phonetically. These are then interpreted in a phonetic component, distinct from the phonological one. (Cohn, 2006, p.28)”

The revised view of modular mapping proposes that phonological representations are abstract, discrete and they consist of phonological features. These abstract categories are then translated into physical events by means of continuous language-specific functions, conceived of in terms of articulatory gestures, or acoustic targets. Related to this is the idea of phonetic underspecification (Keating, 1988), where a specific segment may or may not be associated with a specific feature. In phonetic terms this translates into the absence of a specific gesture/target which results in a variable, context-dependent, and phonetically gradient realisation of the given segment.

2.3 Arguments against the phonology-phonetics distinction

While the modular mapping approach has been quite influential with respect to both theoretical and empirical research, a number of criticisms have been raised against the idea that phonology and phonetics should be treated as distinct. The arguments against such distinction have mostly been based on the existence of many similarities between phonology and phonetics. Two approaches grow out of the idea that phonology and phonetics cannot be distinguished in any principled anyway: functional phonology, and Articulatory Phonology.

2.3.1 The duplication problem. Functional phonology

One of the arguments against the phonetics-phonology distinction is concerned with redundancy. In a modular approach there is a degree of duplication between phonetics and phonology, as very similar phenomena are considered as either phonetic, or phonological depending on whether they are gradient or categorical. An example of this, discussed by Flemming (2001), is vowel nasalisation. A vowel adjacent to a nasal consonant becomes nasalised in a number of languages. In English words like /dən/, this phenomenon is considered to be phonetic, following findings that nasalisation is gradient (Cohn, 1993). In some other languages, including Bengali and Nupe, the same phenomenon is said to be contrast-neutralising. While nasal and oral vowels are contrastive in most positions, the contrast is neutralised following a nasal consonant, which is considered to be an effect of phonological nasal place assimilation. Examples of this kind are abundant, and have led some linguists to abandon the distinction between phonology and phonetics. Central to this stance is the view that both phonology and phonetics are subject to the same influences which can be conceived of as a unified system. Flemming (2001) provides an explicit formalisation of such a

system within a constraint-based optimisation model, where phonological and phonetic phenomena result from an interaction of the same constraints. The constraints themselves refer to an array of structural, articulatory, and perceptual requirements, including factors like articulatory effort avoidance. Coarticulation and assimilation phenomena are thought of as conditioned by constraints which penalise articulatory effort. For instance, the previously discussed nasal assimilation is conditioned by a constraint MINIMISEEFFORT which militates against the kind of rapid velar closure that would be required to produce an oral vowel following a nasal consonant.

Flemming's account grows out of the functional approach to grammar, which assumes that the speaker has knowledge of phonetic influences on grammar, for instance effort reduction (Hayes & Steriade (2004), and individual contributions in Hayes et al. (2004)). Unfortunately, this particular assumption involves the very same problem that the entire approach had set out to avoid, namely redundancy. Articulatory, aerodynamic, and perceptual influences on language systems exist independently of grammar. For instance, as Flemming notes, postnasal vowel nasalisation is an 'almost inevitable consequence of moving the velum into and out of the lowered position required to produce a nasal consonant (Flemming, 2001, p.11)'. A cognitive constraint against rapid velic movements does not perform any job preventing a speaker from producing an oral vowel in the nasal context that is not already done by the shape of the vocal tract. Thus, positing such a constraint involves unnecessary duplication between properties of the outside world, and specific properties of language that have to be learnt by a user about his or her language.

A question that arises at this point is which kind of redundancy is more serious: the one that delegates apparently similar processes to distinct components of grammar, which is the case with a modular phonology-phonetics distinction, or the one which purports that learners need knowledge of constraints that independently affect factors like production and perception. One way to settle this question is to ask where both kinds of redundancies come from. Proponents of a phonetics-phonology distinction argue that the duplication between the two modules comes from phonologisation (Hyman, 1976, 2008), a diachronic change whereby an extrinsically motivated phonetic process becomes a part of the speakers' grammar. Under this view, similarities between phonology and phonetics are not an accident, but a consequence of sound change, as phonetic processes which originate from articulatory and perceptual constraints are perceived, interpreted, internalised and systematised within the speaker-hearer loop over time. The view of phonologisation fits in very closely with empirical evidence from the direction of sound change, as gradient phonetically motivated processes frequently evolve into categorical, systematic and idiosyncratic ones, where the phonetic effect is no longer a direct consequence of extrinsic phonetic factors. Hyman (2008) lists a number of examples, including tonal depression in Nguni languages, described by Traill (1990) as phonologised, as "there is no longer a transparent phonetic explanation for it, and (...) the phonetic effect has been exaggerated (Traill, 1990, p.166)".

A similar argument for why there should be a duplication between articulatory, aerodynamic, or perceptual influences on speech and speakers' grammars is difficult to adduce. Jakobson's argument in favour of grammatical markedness constraints is the apparent absence of sound changes leading to structures which violate implicational universals (Jakobson, 1929), such as final obstruent voicing (Kiparsky, 1995). There is, however, a number of conceptual and empirical issues associated with evidence based on absence of linguistic patterns. First, it is extremely difficult to distinguish between impossible linguistic patterns and accidental gaps, as pointed out for instance by Buckley (2009) in the context of Optimality Theory factorial typologies. Second, the absence of sound changes resulting in marked structures can also be explained by factors involved in phonologisation, as most marked structures frequently involve a combination of sound changes, which makes the patterns less likely to arise (Blevins, 2004). Third, some sound patterns violating implicational universals have in fact been attested, including final obstruent voicing in Lezgian (Yu, 2004).

One of the main reasons for positing speakers' knowledge of functional constraints within

functional phonology appears to be little more than a working assumption that all phonological and phonetic processes follow directly from a synchronic interaction of constraints on mental grammars. In order to accommodate functional optimisation strategies within this system, constraints on minimising effort and increasing perceptual distance have to be embedded as grammatical principles. However, the existence of such principles at the level of a speaker's grammar is not supported by empirical evidence. Sebregts & Scobbie (2011) point out that, according to functional approaches, articulatory effort is only performed by the speakers for listeners' sake. The minimisation of articulatory effort is challenged by Sebregts & Scobbie (2011), based on evidence from ultrasound tongue imaging, which reveal that speakers perform covert articulatory gestures that leave no acoustic trace. Such covert rhotic gestures have been found in speakers of Dutch (Scobbie & Sebregts, 2010), and Scottish English (Scobbie et al., 2008). Sebregts & Scobbie (2011) take this finding as evidence against both speaker selfishness and speaker altruism, as articulatory expenditure without acoustic pay off does not reflect either one of those principles in any obvious way. Instead, Sebregts & Scobbie (2011) argue that speakers' articulatory behaviour is more complex than follows from an 'ease-of-articulation-perceptual-distinctness-tug-of-war', and that although functional factors are at work in shaping speech phenomena, they are best understood as a subset of multifaceted and sometimes indirect influences on performance, communication, and, ultimately, grammar.

All in all, it appears that current criticisms of a modular phonetics-phonology distinction that draw on redundancy do not offer an alternative which would genuinely improve on the redundancy issue. On the contrary, the phonetics-phonology duplication which exists due to an independent force (sound change) is substituted with a duplication between extragrammatical (e.g. articulatory and aerodynamic) influences on speech, and speakers' implicit knowledge of those factors. This latter kind of duplication seems to follow from little more than a theoretical stipulation, and it is undermined by experimental evidence which shows that speakers do not necessarily avoid articulatory effort in the absence of an acoustic/communicative pay-off.

As a final remark on the issue of redundancy, it appears that avoidance of redundancy seems to have been elevated in the generative tradition from a working assumption to a theoretical principle. This argument is made by Cohn (2006) in her discussion of the role of Occam's Razor in the debate on modularity. Cohn points out that redundancy is ubiquitous in language, and that it is widely observed in phonetics, for instance in the duplication of phonetic cues to phonological categories and contrasts. Similarly, speakers' lexical representations have been shown to contain rich, and sometimes redundant, detail. Cohn concludes that "‘duplication’ is not a problem, but in fact an intrinsic characteristic of language (Cohn, 2006, 43–44)". Following this argument, however unparsimonious, a distinction between phonology and phonetics should not be easily dismissed on the grounds of economy alone. Instead, it seems more appropriate to ask where the apparent redundancy might come from, and whether or not incorporating it in a linguistic model might be beneficial to our understanding of speech and speech processes.

2.3.2 Gradience and overlap

Apart from functional approaches to grammar, proposals to abandon the distinction between phonology and phonetics have been put forward by researchers working within the framework of Articulatory Phonology (Browman & Goldstein, 1986, 1989, 1992). In a series of papers Browman and Goldstein propose a phonological model, whose representations are based directly on articulatory movements. One advantage of such a model is argued by the authors to lie in explicitly linking phonological categories (corresponding to phonological features in Autosegmental Phonology) to physical movements of articulators in speech production. In addition, Browman & Goldstein (1992) argue that gestural overlap explains straightforwardly a broad range of

phenomena previously analysed as autosegmental feature spreading phenomena as co-occurrence of several gestures. Browman & Goldstein (1992, p.30) further argue that gestural analyses have a broader scope than feature spreading accounts, and are successful in capturing ‘a wider range of behaviour’. This last argument relates to gestural analyses’ ability to accommodate gradient coarticulation. A number of studies published in the wake of Browman & Goldstein’s introduction of their model have revealed gradient properties of speech processes which had been previously conceived of as categorical assimilation, providing support for gestural overlap analyses over feature spreading accounts. One example is Zsiga’s (1995) study of /s/+/j/ coarticulation, which shows that realisation of /s/+/j/ sequences across an intervening word boundary involve gestural and durational properties different from those of a lexical /ʃ/, but consistent with an overlap of an alveolar and a palatal gesture tongue. Other studies showing effects of gestural overlap include Wright & Kerswill (1989), Barry (1992), and Romero (1999), *inter alia*.

However, some other authors present evidence of categorical assimilatory behaviour not predicted by gestural models. Examples of this kind of work include Holst & Nolan (1995), Nolan et al. (1996), Ladd & Scobbie (2003), and Ellis & Hardcastle (2002). Holst & Nolan (1995) present acoustic data on English /s/ to /ʃ/ assimilation which is shown to involve four distinct types of realisation: type A, where there is no assimilation, types B and C with partial assimilation, and type D, where the assimilation is categorical. While Articulatory Phonology is capable of modelling different degrees of assimilation, as different degrees of overlap, it crucially predicts that complete assimilation is a case of gestural blending, which produces a distinct kind of output with characteristics of both lexical /s/ and lexical /ʃ/, but not identical to either one of those. However, according to Holst & Nolan (1995), type D realisations are spectrally /ʃ/-like without any discernible /s/-characteristics, which is consistent with a feature spreading analysis, and not with a gestural blending account. While the results by Holst & Nolan (1995) could be disputed due to the lack of articulatory data (cf. 2.6), similar categorical effects have been found in an articulatory study of /n#k/ sandhi by Ellis & Hardcastle (2002), discussed in 2.1. Ellis & Harcastle found that some speakers realised underlying /n#k/ sequences in two categorically distinct ways: either with a full coronal gesture present, or with no coronal gesture at all. Importantly, these speakers did not produce intermediate realisations with a residual coronal gestures, which would suggest a continuum of articulatory realisations. The categorical opposition apparent in the speakers alternating between two articulations goes against the predictions made by gestural blending, and provides further support to the hypothesis that there are cases of categorical assimilation that involve a deletion of a gesture.

Accumulating experimental evidence suggests that gestural models do not necessarily cover a broader, but a different range of speech processes than feature spreading analyses. Gestural overlap accounts have been tremendously successful in modelling gradient coarticulation, and they have also inspired a great body of empirical work on categoricity and gradience. However, existing evidence does not support the hypothesis that a model in which articulatory gestures function as both cognitive units and instructions to articulators, is sufficient in accounting for the full range of coarticulation and assimilation phenomena. Instead, it appears that gestural accounts work well for the speech processes that a modular distinction treats as phonetic, but they are less successful with what might be considered phonological phenomena, thus reinforcing rather than obliterating the need to distinguish between phonology and phonetics.

2.4 A distinction without an interface?

The proposals to unify phonology and phonetics within a single model seem restricted to functional phonology and Articulatory Phonology, while most other approaches recognise that

phonetics and phonology differ in some way. However, the question of how to conceptualise the distinction between phonology and phonetics has proven highly divisive. A particularly contentious notion is that of a phonetics-phonology interface, which has been rejected by a number of authors, even those who do not equate phonetics and phonology, including Ohala (1990b), Pierrehumbert et al. (2000), and Scobbie (2005).

Ohala's criticism is based on the implications brought about by the notion of the interface. Ohala provides the following two working definitions of phonology-phonetics interface.

- (A) There are two domains in the universe of speech, one covered by phonology (the speakers' knowledge?), another covered by phonetics (the physical side of speech); there is a point at which these two meet, for example where phonological representations become implemented physically.
- (B) Phonology and phonetics are largely two autonomous disciplines whose subject matters may be similar (like those of chemistry and physics) but which study them in different ways; there is an area in between where the two can co-operate. (Ohala, 1990b, 153)

According to Ohala, the term 'interface' entails that phonology and phonetics are to be viewed as largely autonomous and independent of each other, and that such a view is untenable; the phonological component cannot exist without phonetics, and *vice versa*. Ohala proceeds to argue that the separation of phonology and phonetics into independent disciplines encourages the study of phonology as autonomous, a practice which is prone to 'circularity, reification, projection and myopia'. Indeed, studying phonology as an autonomous object is prone to ascribing properties to mental grammars that can follow independently from external factors. Ohala's criticism also applies to the practice of rescuing phonological theories by labelling some phenomena as phonetic on a strictly top-down basis. However, right as this criticism may be, it is really concerned with the practice of individual linguists, and not with the idea that phonology and phonetics might be interfaced. While the term 'interface' is perhaps not be the most fortunate one, it is ultimately just a label for the concept that phonological and phonetic representations translate systematically into one another in an integrated system, a notion that is difficult to dispense with once any kind of distinction is drawn between phonology and phonetics.

While Ohala (1990b) focuses on the criticism of the interface as separating phonetics and phonology into separate fields, Pierrehumbert et al. (2000) and Scobbie (2005) take up the issue of whether the demarcation line between phonology and phonetics can be reliably determined. Both give a negative answer, clearly stated in the following quotation from Pierrehumbert et al. (2000).

"Although MESM [Modified Extended Standard Modularization, cf. Pierrehumbert (1994)] asserts that the relationship of quantitative to qualitative knowledge is modular, this assertion is problematic because it forces us to draw the line somewhere between the two modules. Unfortunately, there is no place that the line can be cogently drawn. (Pierrehumbert et al., 2000, p.287)"

Pierrehumbert et al. (2000) go on to argue that aspects of categoricity and gradience feature in both phonology and phonetics. On the one hand, there is a growing body of experimental evidence which shows that some phonetic detail is stored with the representations of words/morphemes, which undermines the categorical view of phonology. On the other hand, phonetic knowledge is language specific and learned, which gives a 'phonological flavour' to the detailed phonetic knowledge of phonological categories. In addition, some aspects of categoricity are inherent to phonetic through the existence of quantal effects, as argued by Pierrehumbert et al. (2000) and Cohn (2006).

Cohn (2006, 2007) proposes that a modular distinction between phonology and phonetics might involve a simplification, due to the existence of indeterminate cases in between. Thus, similarly to Pierrehumbert et al. (2000), Cohn talks about phonology and phonetics as positioned alongside a continuum of granularity. Phonology is placed towards the categorical, and phonetics towards the gradient end of the continuum, but there is also an area of overlap in between. The overlap area, according to Cohn (2007) exists due to the inherently fluid nature of phonologisation. At the same time, however, Cohn recognises that the endpoints of the continuum are in some sense privileged. This proposal is backed up by a body of evidence, some of it discussed in Section 2.1 of the current chapter, which confirms that the core of phonology is concerned with abstract categorical interactions, while the physical aspect remains a chief property of phonetics. Cohn (2007) concludes that verifying whether the phonology-phonetics distinction is modular, or qualifying the strength of the categorical-phonology-gradient-phonetics correlation is down to empirical evidence.

As an alternative a modular distinction, Pierrehumbert et al. (2000) develop a view of phonology and phonetics as covering different aspects of a continuum, with the phonetic knowledge concentrated in the gradient end of the spectrum, and phonological knowledge concerned primarily with the other, categorical end. Related to this is a conception of phonology, whose discrete aspects are ‘embedded in a continuous description’, and which arise through cognitive processes responsible for discretisation of gradient variation. Some aspects of the discretisation have to do with acoustic non-linearities (i.e. quantal effects), while some others arise through the function of phonology as encoding lexical contrast. In addition, discretisation may arise through the neural network organisation, as observed in the phenomenon of perceptual magnet effect (Guenther & Gjaja, 1996).

The discussion by Pierrehumbert et al. is rooted in the exemplar theoretic view of phonology as a self-organising system, explored further by Scobbie (2005) from the point of view of phonetics-phonology distinction. Exemplar Theory (Nosofsky, 1986, 1988; Goldinger, 1997; Johnson, 1997) is based on a model of the mental lexicon which contains multiple traces, or exemplars of linguistic units, such as words (Wedel, 2004; Johnson, 2007), or speech sounds (Skousen, 1989; Pierrehumbert, 2001). These rich phonetic representations are paired with linguistic meaning through context, and the recurring sound-meaning pairs are strengthened. In this way, abstract categories can emerge from the phonetic detail. The emerging system is more discrete and shares many characteristics with phonology understood in the traditional generative sense. However, as pointed out by Scobbie (2005), the emergent exemplar-based phonology is not a separate module. Instead phonological categories and phonetic detail co-exist together in the lexicon, alongside with more in-between ‘fuzzy’ labels which share the properties of phonology and phonetics. Scobbie (2005), espousing the exemplar view of the lexicon, rejects the modular distinction in which phonology and phonetics are separate in favour of a phonetics-phonology overlap, where phonology and phonetics might not be the same, but they overlap in many areas.

Unfortunately, even though Exemplar Theory offers new ways of thinking between the relationship between phonetics and phonology, it involves an unresolved problem of the status of phonological categories and category labels. Boersma (2012a) argues that for combining phonetically-defined categories with abstract symbols used in phonological models there is a necessary intermediate step of assigning category labels. In the current models by Boersma this assignment has to be stipulated (‘category creation day’), but it is possible thanks to the model’s architecture, which distinguishes between lexical (underlying), phonological (surface) and phonetic representation. Exemplar-based models are also faced with the problem of assigning category labels, but the status of phonological label assignment is less clear if phonological form is not recognised as a distinct level of representation. Yet, as Boersma (2012a) points out, some exemplar theoretic models rely on category labels to model realistic linguistic behaviour (e.g. Wedel’s (2006) model of auditory dispersion). Boersma goes on to argue that apart from category labels, exemplar-

based models also need to incorporate relations between morphemes and underlying forms, as well as relations between underlying and surface forms to account for the full range of phonological phenomena, including sentence phonology. Thus, even though Exemplar Theory brings promise of reconceptualising the classical modular distinction between phonology and phonetics, it does not as yet provide a way of modelling categorical phenomena without an architecture where categorical effects belong to a separate level of phonology.

2.5 The current approach

The arguments against a modular phonology-phonetics distinction are extremely important, as are the empirical issues surrounding the notion of the interface. We are thus faced with the question how to conduct empirical research, and how to analyse phonetic data, acknowledging those issues on the one hand, yet not losing sight of the phenomena that motivated the phonetics-phonology distinction in the first place. What is important to keep in mind, is that even in the face of the arguments against strict modularity, the question of what exactly phonology is and what it does still persists. Regardless of one's views on the interface, or indeed on the lexicon, phonology must have some systemic properties for the language users to be able to communicate. Thus, some old phonological questions are as valid as ever under a rich memory view of the lexicon. What kind of units are abstracted by listeners from the phonetic signal? Are there hierarchical relationships between those units? How do the units interact? How broad are the generalisations made by language users about recurring linguistic patterns? How do abstract categories map onto articulatory gestures in production?

One of the difficulties of studying phonology stems from the fact that in our attempts to tackle phonological questions we rely largely on phonetic data. What we hear, perceive, and utter as language users, and what we record, measure and analyse as linguists, are the only actual data we have access to. Phonological categories that we can infer from these, the generalisations that we make, whatever we could call ‘phonological data’, are not objective tangible entities, but mostly ways of analysing the phonetics. Although not always made explicit, this is true for everything that counts as data in the phonological literature. Transcriptions are the ways the transcriber analyses the acoustic signal. Native speakers’ intuitions represent the ways native speakers analyse speech (to the extent that speakers are indeed able to introspect about speech). Phonological alternations involve linguists’ inferences that two entities are equivalent excepting some specific aspect. All actual language data are continuous. Discretisations are not data, but data analyses. Consequently, phonology has no object of study without phonetics. Therein lies the difficulty, but also the necessity to define phonology against phonetics. Because we can only catch a glimpse of the phonological organisation of speech through phonetics, we need some way of deciding which aspects of the phonetic data are about phonetics, and which are the ones that signify something about a higher-level phonological organisation. Hardly any phonological answers can even be attempted without at least a working definition of what counts as abstract and discrete, and what does not, which forces linguists to operationalise this distinction, even if it involves a simplification.

Simplification is inherent to any kind of modelling, and it does not automatically render a model useless. Newton’s three laws of motion, to take a non-linguistic example, are a simplification, but they provide a very close approximation of a wide range of physical phenomena. Similarly, a distinction between gradient phonetics and categorical phonology is a simplification, which may not reflect the complexity of the mental representation of speech, and which may fail to accurately represent the behaviour of a subset of more complex cases, but which nevertheless provides an extremely good approximation of how the majority of sound patterns work.

An architecture with two separate levels of phonetics and phonology seems indispensable

to the only currently existing explicit algorithms that link the emergence of categories from a phonetic continuum with abstract symbolic computations, as discussed in the previous section. It is primarily for this reason that a modular-like phonetics-phonology distinction is assumed in the current approach, despite the criticisms that such a distinction is not always workable. If explicit algorithms are created at some point which capture categorical behaviour with categories created bottom-up, but without positing a phonological level in its architecture, the current model could be seen as an approximation of how gradient and categorical behaviour are related to each other. Meanwhile, although not all phonetic patterns can be straightforwardly classified as either phonological or phonetic, even formulating where the difficulty lies can and should be considered progress. I would go so far as to argue that it is precisely the not-straightforward results that are of crucial importance, as they shed light on the grey areas, and show us what kind of difficulties arise in attempting to distinguish between phonetic and phonological processes. It might be that the current open question in research on the interface are due to empirical shortcomings. Indeed, for individual processes more conclusive data unravel when new methods are attempted, for instance as discussed in the case of /l/-darkening (Section 2.6). It might be that what we are after is not a single criterion, but a more complex probabilistic model of determining a phonetic or a phonological character of a given phenomenon. However, for the empirical research to progress in this direction, the phonetics-phonology distinction must not be a foregone objective, as there seems to be little merit in untangling a complex issue once ‘complex’ is accepted as a satisfying enough qualification.

The set of theoretical assumptions taken as a starting point in this work are close to that of a modular approach along the lines of Keating (1988, 1990, 1996), Pierrehumbert (1980) and Cohn (1990, 1993). I do not follow the modular approach in assuming a modular organisation of the brain. I do, however, follow it in the following:

1. Phonetics and phonology are distinct and they are concerned with different types of processes.
2. Phonology is concerned with categorical and discrete patterns.
3. The existence of phonological categories is determined based on empirical evidence of categorical phonetic behaviour. Phonological categories are schematically represented by features. Features are taken to be language-specific and emergent, as proposed by Mielke (2008), and Boersma (2012a).
4. Phonetics connects abstract categories or features with their quantitative realisations, for instance by means of gestures, or targets.
5. Both, phonology and phonetics are language-specific and learnt (emergent). A child acquiring her L1 needs to abstract the phonological categories from the phonetics, and learn the language-specific mapping between features and their phonetic realisation in different contexts.
6. Phonological knowledge is defined as knowledge about categories (phonological features and phonological structures), and about how these categories interact in the language. Phonetic knowledge of a language user is defined as the knowledge of how phonological features and prosody are phonetically realised in production and perception.

The modular mapping approach grows out of the minimal memory models of the lexicon (term from Wedel (2007)), which assume that redundant phonetic detail is not a part of the lexical representation. However, the crucial insight of the modular mapping approach which I attempt to preserve in this work, that the phonology-phonetics distinction can be operationalised in terms of the opposition between categoricity and gradience, is not inherently tied to any particular view of the lexicon. The categoricity-gradience opposition is directly compatible with hybrid models

of speech production which involve both phonetically detailed lexical representations and abstract phonological categories, such as the model by Pierrehumbert (2002). Although the scope of specific phonological generalisations may differ between the rich and the minimal memory models of the lexicon, empirical findings concerning categoricity and gradience are of relevance under either view.

2.6 Empirical diagnostics for categoricity and gradience

From the empirical point of view, a great strength of research on the phonetics-phonology interface has been the accumulation of new phonetic data. The relevant studies include Barry (1992); Cohn (1993); Zsiga (1995); Holst & Nolan (1995); Nolan et al. (1996); Ellis & Hardcastle (2002), and recently Tucker & Warner (2010). In this section I discuss some of these works in greater detail in the context of specific methods used to diagnose categoricity and gradience based on phonetic data. I also acknowledge the empirical issues related to individual diagnostics.

The perhaps most frequent approach to differentiating between phonological and phonetic processes consists in comparing tokens derived by e.g. assimilation to tokens where a given phonological category is present underlyingly. The comparative method is used by Romero (1999), who analyses /s/ voicing assimilation in Castilian Spanish, based on EMA data supplemented with laryngeal transillumination. The study compares words and word sequences containing /s/+voiced obstruent clusters with words containing singleton voiced or voiceless obstruents. Romero observes that laryngeal gestures differ in their magnitude and synchronisation between these three speech conditions, and that the laryngeal gesture in /s/+voiced stop clusters occurs in between the two supralaryngeal gestures. Based on this observation, Romero (1999) argues that a gestural blending view of voice assimilation is more realistic than the one offered by traditional phonological accounts, presumably meaning feature spreading. The comparison test is also used in the previously discussed study by by Ellis & Hardcastle (2002), who compare the realisation of underlying /n#k/ and /ŋ/#/k/ sequences in order to determine whether there is a difference in how the two kinds of underlying nasals are articulated.

While highly successful for some kinds of data, the comparison test is not always applicable, for instance in situations where the occurrence of a particular category is positionally restricted. Such positional restrictions are involved for instance in the case of voice assimilation, which is the subject of the current work. Voice assimilation targets are typically found in the position where voice neutralisation otherwise occurs, e.g. word-final segments undergo regressive voice assimilation before an obstruent in the next word, and final devoicing before a sonorant or a pause. Consequently, voiced word-final obstruents are only found in the assimilation context, and cannot be compared against lexically voiced tokens in the same position.

Another way of diagnosing categoricity has to do with statistical distribution of phonetic parameters. Phonetically gradient phenomena typically give rise to continuous statistical distribution, whereas periods of discontinuity might be indicative of a categorical effect (Scobbie, 2005). This reasoning is frequently employed to diagnose whether the distribution of a phonetic parameter indicates the presence of more than one phonological category. The bimodality test has been used by Herd et al. (2010), who distinguish flapped and non-flapped variants of /t/ and /d/ in American English based on the discontinuity in the distribution of duration.

There are three potential problems associated with the bimodality criterion in diagnosing categorical phenomena. First, the bimodality test is easily confounded by experimental design, as pointed out by (Scobbie, 2005, 13), who notes that it is “very easy to find bimodal or multimodal distributions of values for phonetic parameters, where each mode is associated with some conditioning factor”. Scobbie (2005) goes on to discuss the example of vowel duration, which might be slightly decreased following aspiration, potentially giving rise to a bimodal distribution of

vowel duration in words such as English ‘bat’ and ‘cat’, even though post-aspiration vowel clipping can hardly be considered a phonological phenomenon in English. The second type of challenge to using the bimodality test come from the non-linear relationship between articulation and acoustics. Browman (1995) raises this issue in a response to Holst & Nolan (1995), pointing out that gradient shift in articulation might correspond to abrupt changes in the acoustic signal. Browman (1995) cites the findings of Perkell et al. (1979), which show an abrupt spectral change in a continuum of [s] to [ʃ] realisations. This kind of non-linearity is observed for a broad range of phonetic phenomena, and explored in a series of papers on Quantal Theory (Stevens et al., 1986; Stevens & Keyser, 1989; Keyser & Stevens, 2006; Stevens & Keyser, 2010). Based on the existence of quantal effects, where areas of acoustic stability corresponding to a fairly large range of articulations, Browman (1995) argues that categorical effects in the acoustic signal (which includes multimodality in the distribution of acoustic measurements) is not always indicative of genuine categoricity in speech. Finally, the existence of multiple modes in the distribution might also be concealed if the distance between the two peaks is not sufficiently large, as argued by Schilling et al. (2002), who show that combining two normal distributions may give rise to a continuous distribution if the difference between the two means is less than the sum of the standard deviations.

A yet another methodological approach to categoricity and gradience is taken by Cohn (1993), who differentiates between gradient and categorical in the case of nasalised vowels in French and English by means of a qualitative phonetic analysis. She finds that in English, the nasal flow on the vowel in the word *den* has the form of a cline, following the oral [d] and preceding the nasal [n], which is best explained as phonetic interpolation between two distinct articulatory states (velar closure and velar opening). In comparison, the nasal airflow in the French *daim* ‘deer’ has the form of two plateaus connected by a rapid transition, which is more typical of categorical nasalisation.

Studies of categoricity and gradience are also informed by speech rate effects. It has been observed that there is more coarticulation at faster speech rates (Byrd, 1994; Keating, 1996). As an increase in speech rate leaves less time for articulators to move away from each other, a considerable degree of overlap is expected to occur. Thus, more gradient coarticulation is typical of faster speech rates, whereas in slower speech coarticulation is limited. Categorical assimilation, on the other hand, is not directly influenced by speech rate. If assimilation is interpreted as the presence of a phonetic target, then it involves the presence of a pre-planned articulatory gesture rather than gestural overlap/interpolation. Such a gesture is likely to occur regardless of speech rate, even though rate might influence its duration and magnitude. Solé (1992, 1995) uses manipulations of speech rate as diagnostics for distinguishing between mechanical phonetic properties which follow automatically from the vocal tract constraints. While mechanical properties are not expected to be affected by variations in speech rate, the acoustic cues that are under speaker’s control undergo durational adjustments to maintain a steady ratio to the segment duration across different speech rates. Vowel nasalisation in Spanish and American English provide an illustration. Solé (1992, 1995) analysed the correlation of speaking rate duration of the nasal portion of a vowel followed by a nasal consonant. In Spanish the nasalised vowel portion retained a relatively stable duration, and did not correlate significantly with the speaking rate. In American English, on the other hand, the duration of nasalisation varied systematically with the speaking rate, while the ratio of nasal portion to the total vowel duration remained stable at 100%. Solé proposes that the short period of nasalisation observed in Spanish is automated and it follows from the velum lowering produced in the anticipation of the nasal consonant. However, this kind of overlap is insufficient to explain the extent of rate-independent nasalisation in American English. Thus, for American English, Solé argues that vowel nasalisation is speaker-controlled and actively targeted. The speech rate test is also applied by Cuartero Torres (2001) to voice assimilation in Catalan and English.

Similarly as in the case of comparison test, quantitative evaluation of categorical behaviour in the continuous data (Cohn, 1993), or speech rate effects in the extent of speech process application

(Solé, 1995, 2007) suit some types of data better than others. Cohn's approach works well for comparing nasal flow data for gradient nasalisation in English and categorical nasalisation in French, but the method relies on the existence of a similar process in the phonetics and phonology of two different languages. Solé's predictions for the effect of speech rate on voicing duration and voicing ratio are not readily extendible to e.g. place assimilation.

Finally, different empirical test for determining gradience and categoricity do not always converge, and conflicting findings are reported by different authors, depending on the method used, or even on the size of the sample. The case of /l/-darkening in English provides an interesting example. English /l/ is traditionally described as a case of allophony, with light /l/ occurring in syllable onsets, and dark /l/ in syllable codas. Sproat & Fujimura (1993) argue against a categorical clear-dark /l/ distinction in (American) English based on experimental data from four speakers recorded in a controlled experiment. Based on combined acoustic and X-ray data, Sproat & Fujimura (1993) report a correlation between the backness of the /l/ and the duration of the rhyme; darker /l/s were found in longer rhymes. Based on the duration conditioning of the /l/ darkness, Sproat & Fujimura (1993) argue for a gradient interpretation of /l/-darkening, and propose an articulatory account, where the execution of dorsal and coronal /l/ gestures varies with the position on the syllable, and the time window. However, a recent paper by Yuan & Liberman (2009) introduces new evidence for a categorical distinction between clear and dark /l/ in English. Yuan & Liberman (2009) analysed 21,706 tokens of /l/ taken from 2001 term of the SCOUS (Supreme Court of the United States) corpus, using forced alignment. Their results confirm that there is correlation between /l/-darkness and duration, but the correlation only holds for the rime, and not for the syllable-initial /l/. Yuan & Liberman (2009) argue that this result cannot be explained within Sproat and Fujimura's model, and that the distinction between pre-vocalic and postvocalic /l/ is best understood as categorical.

There is a number of important findings that follow from the accumulating body of empirical research into categoricity and gradience. First, categorical phonetic processes are relatively rare, certainly more so than it would follow from categorical transcription data of earlier days. However, they are not entirely nonexistent, as some categorical effects survive a careful experimental scrutiny based on continuous data. Classifying speech phenomena as gradient or categorical thus becomes an important descriptive task of experimental phonetics, even outside of the larger theoretical debate of how categoricity and gradience relate to cognitive organisation of speech. One has to bear in mind that specific methods for diagnosing categoricity in phonetic signal are sensitive to some empirical difficulties, as highlighted in the discussion above. Those difficulties, however, bring out the necessity to properly control for all the potential confounds, and to critically evaluate the results rather than render any of the specified tests invalid.

2.7 Final remarks

Studying phonetics can inform us about patterns of more abstract cognitive behaviour on the part of speakers, and it is with that aim that a phonetic study of pre-sonorant voicing is undertaken in this thesis. The idea is to explore what, if any, categorical effects emerge from continuous phonetic data on pre-sonorant voicing, and what these tell us about abstract phonological representations that speakers might have. The distinction between gradient phonetics and categorical phonology is adopted to provide a working approach to the data. While the ongoing debate and accumulating empirical evidence suggest that the relationship between phonology and phonetics might not be strictly dichotomous, it is worthwhile to establish in what aspects the dichotomy can and cannot account for the observed data. Thus, a relatively simple dichotomous model is taken as a starting point, with the view to expand its complexity as dictated by the data.

CHAPTER 3

WEST-FLEMISH

It may of course be argued that these ideas, too, are speculative, but on the whole speculations based on known phenomena are to be preferred to those which have their origin purely in fantasy.

Malcolm Dyson

This chapter presents the results of an acoustic study of voicing in obstruents followed by a sonorant across a word boundary in two dialects of Dutch: East- and West-Flemish. 6 West-Flemish and 6 East-Flemish speakers read two repetitions of word-final obstruents followed by word-initial sonorants and obstruents. In both varieties only gradient phonetic voicing was typically found in word-final stops when a sonorant followed in the next word. In addition, West-Flemish showed optional categorical voicing in word-final pre-sonorant fricatives, reflected in a bimodal distribution of voicing duration and voicing ratio. The voicing of fricatives is argued to be phonological, as it extends beyond the scope of automated coarticulation, and as the data pattern to form a distinct phonetic voicing target. However, the phonetic results do not support the hypothesis that West-Flemish sonorants are laryngeally specified, and thus able to spread voicing to neighbouring fricatives. Instead, I propose that fricative voicing is an optional positional realisation in West-Flemish. Although the process cannot be directly motivated by reference to the phonological specifications of the segments surrounding its target, it makes sense in terms of perceptual factors leading to diachronic reanalysis. The West-Flemish positional variation may occur when partially voiced fricatives are perceived and subsequently reanalysed as categorically voiced by listeners, as proposed by Jansen (2004). I further argue that fricatives are more likely than stops to be reinterpreted as voiced, as additional acoustic cues prevent voiced percepts in passively voiced stops.

3.1 Pre-sonorant voicing and other voicing-related processes in West-Flemish

West-Flemish is unique among dialects of Dutch in showing a process of fricative voicing word-finally when a sonorant consonant follows. As illustrated in (1), word-final fricatives can be voiced in Standard Dutch preceding vowel-initial words and in compounds, but not word-medially (De Schutter & Taeldeman, 1986).

- (1) Pre-sonorant voicing in Dutch
- | | | |
|---------------|---------------|------------------|
| /dat məns is/ | [dat.mən.zis] | 'that person is' |
| /ras+ɛχt/ | [ra.zɛχt] | 'pure bred' |
| /jasən/ | [ja.sən] | 'coats' |

The voicing of fricatives preceding a vowel has been reported to occur in all of the Southern Dutch dialects (De Schutter & Taeldeman, 1986). Additionally, the voicing of fricatives preceding sonorant consonants across word-boundaries has been reported for West-Flemish, as exemplified in (2).

- (2) Voicing of word-final fricatives preceding sonorant consonants in West-Flemish (De Schutter & Taeldeman, 1986)
- /zes jar/ [zɛz.jar] ‘six years’

Word-medially, fricatives followed by sonorant consonants only occur in borrowed words such as *moslim* ‘muslim’, *islam*, and *kosmos*. The realisation of the fricative as voiced or voiceless in these words seems to be variable (cf. Simon (2010, 135)).

An important aspect of pre-sonorant voicing is that it tends to target only a subclass of obstruents, typically fricatives or sibilants. Reports of pre-sonorant voicing in West-Flemish single out fricatives as the undergoers of voicing, while similar reports of stops are absent¹. Collins & Mees (1999, 214) note that voiced realisations are frequent in Dutch coda *fricatives* (emphasis added) followed by a vowel. Similarly, De Schutter & Taeldeman (1986) report that in West-Flemish word-final fricatives are usually realised as voiced when the next word begins in a sonorant. Finally, Simon (2010) found pre-sonorant voicing in fricatives, but not in stops in the production of West-Flemish speakers.

In addition to voicing before sonorants West-Flemish, just like other dialects of Dutch, shows two common laryngeal processes: final devoicing and regressive voice assimilation. Final devoicing affects obstruents in the final position, as exemplified in (3). A number of descriptions of Dutch voicing represent final devoicing as a case of laryngeal neutralisation, i.e. a complete loss of the underlying voicing contrast.

- (3) Final devoicing in Dutch (Zonneveld, 1983)

rond	[t]	‘round’	ronde	[d]	‘round, PL.’
hard	[t]	‘hard’	harde	[d]	‘hard, PL.’
huis	[s]	‘house’	huizen	[z]	‘houses’
hoef	[f]	‘hoof’	hoeven	[v]	‘hoofs’

Final devoicing is also reported by Zonneveld (1983) as applying prevocally at the end of a morpheme in compounds, as illustrated by the examples in (4).

- (4) Morpheme-final devoicing in Dutch compounds (Zonneveld, 1983)

rondrit	[t]	‘round trip’
hardlopen	[t]	‘race’
huisarts	[s]	‘family doctor’
hoefijzer	[f]	‘horseshoe’

Zonneveld’s data, however, are not confirmed by Booij (1995), who lists morpheme-final prevocalic fricatives (as in *huisarts*, and *hoefijzer*) as voiced. In addition, while there is a general consensus that word-final obstruents in Dutch are affected by devoicing, understood as a loss of vocal fold vibration, some experimental sources argue that the process is non-neutralising. Warner et al. (2004) and Ernestus & Baayen (2006) show that the underlying voicing contrast is reflected on the surface in small phonetic differences, including vowel duration and closure duration, which is consistent with an incomplete neutralisation analysis.

Another voicing process in Dutch that is of relevance to the current study is regressive voice assimilation. Word-final and morpheme-final obstruents undergo voicing when a voiced stop follows

¹There are, however, reports of stop voicing word-finally before a vowel, e.g. /po:t̪ is/ → [po:dis] in some Limburg dialects and in the northwest of East Flanders (De Schutter & Taeldeman, 1986)

in the next word (Zonneveld, 1983; Slis, 1985; Booij, 1995). Examples are in (5).

- (5) Regressive voice assimilation in Dutch (Jansen, 2004; Simon, 2010)
- | | | |
|---------------------|------|----------------|
| <i>weekdier</i> | [gd] | 'mollusc' |
| <i>visdief</i> | [zd] | 'common tern' |
| <i>laat boeken</i> | [db] | 'to book late' |
| <i>twaalf dozen</i> | [vd] | 'twelve boxes' |

Fricatives are reported not to trigger regressive voice assimilation, and they undergo devoicing following another obstruent (Booij, 1995; Jansen, 2004).

3.1.1 Issues. The current study

West-Flemish pre-sonorant voicing has been mentioned by a number of sources, but experimental evidence concerning its realisation is scarce. The only systematic study of the process to date appears to be by Simon (2010). Simon collected data on voicing before sonorants produced by West-Flemish and East-Flemish speakers as a part of a larger study on the realisation of voicing by speakers of Belgian Dutch, both in their native dialect and in their L2 English. Simon's data were predominantly based on spontaneous speech from recorded conversations between dyads of participants. The corpus was then transcribed, and the tokens in the relevant environments were coded as voiced or voiceless based on auditory classification performed by two transcribers. Pre-sonorant fricatives in West-Flemish were found to frequently surface as voiced (ca. 80-90% of the time depending on the underlying voicing value of the fricative), whereas in East-Flemish pre-sonorant fricative voicing was rare (it occurred in ca. 15-30% of cases depending on whether the fricative was underlyingly voiced or voiceless). Voicing was more common preceding vowel than sonorant consonants. While this result is important in corroborating previous reports of West-Flemish pre-sonorant fricative voicing, the fact that the study was based on perceptual classification limits potential insights into how the voicing is realised phonetically. Voiced perceptions are likely to be triggered by instances of partial voicing, as well as by cases of complete voicing. Consequently, given the perceptual results, West-Flemish pre-sonorant voicing is equally likely to be phonetically gradient, or categorical but optional.

Although drawing the distinction was not the aim of Simon's study, it is crucial for current work which aims to disentangle the phonetic and phonological influences involved in the evolution and synchronic representation of pre-sonorant voicing. In addition, differentiating between categoricity and gradience in pre-sonorant voicing is of theoretical consequence in the context of the debate surrounding Quito Spanish /s/-voicing between Colina (2009) and Bermúdez-Otero (2011). For theoretical reasons, which are discussed in detail in Chapter 6, Colina (2009) argues that Quito Spanish pre-sonorant voicing is best understood as a phonetic phenomenon. Relegating pre-sonorant voicing from the domain of phonology has the capacity to reduce formal accounts of the process in a number of aspects. However, such a reduction can only be fully justified if it is supported by empirical evidence of phonetic gradience.

The experiment presented in this chapter is a production study designed primarily with the question of categoricity vs. gradience in mind. In addition, the study explores two issues related to pre-sonorant voicing introduced in Chapter 1. The first problem has to do with laryngeal representations of sonorants. A number of formal analyses have assumed that sonorants can be laryngeally specified, despite being laryngeally non-contrastive, and that pre-sonorant voicing is an instance of [voice] spreading (examples of such analyses include Bethin (1984), Gussmann (1992), Rubach (1996) for Polish, as well as Bermúdez-Otero (2001), Wheeler (2005), and Jiménez & Lloret (2008); see Chapter 4, and 5 for details). The hypothesis involving a laryngeal specification for sonorants is evaluated here against phonetic evidence from West-Flemish. The second issue taken

up in this chapter is that of undergoer asymmetry reported for West-Flemish and Catalan, where pre-sonorant voicing only targets fricatives (in the case of Catalan this generalisation is true of prevocalic voicing). Thus, the data analysis and discussion presented in this chapter is centred on the following research questions:

1. Does a phonetic analysis of new data confirm the existence of pre-sonorant voicing in West-Flemish fricatives, as opposed to stops?
2. If so, do we find categorical (phonological) or gradient (phonetic) voicing?
3. Is there any phonetic evidence for sonorants having active voicing targets, or do we need to look for an alternative explanation?
4. What factors may be involved in the undergoer asymmetry with a preference for pre-sonorant voicing in fricatives?

3.2 Materials and method

A production experiment was carried out to study the phonetics of West-Flemish pre-sonorant voicing. Test stimuli were presented in writing to the participants, and the pronunciations were recorded.

3.2.1 Speakers

12 native speakers of Dutch participated: 6 from West Flanders (WF) and 6 from East Flanders (EF). The participants were all female, aged 19-47. Participation was voluntary, and the subjects were not paid.

3.2.2 Stimuli

The test items including stops were monosyllabic words ending in /aut/, or /aud/, followed by a trochaic word beginning in a sonorant segment (/m/, /n/, /r/, /l/, /w/, /j/, /ɛ/, /ɑ/), as exemplified in (6). The test items for fricatives were monosyllabic words ending in /i:s/, or /i:z/² followed by a trochaic word beginning in sonorant (including vowels), as shown in (6).

- (6) Sample test items Stops
 koud muntje
 ‘cold mint’ Fricatives
 Fries liefje
 ‘Frisian boyfriend/girlfriend’

The test items were embedded in a standard carrier sentence, as shown in (7).

- (7) The carrier sentence
 ‘*Was een koud muntje, we, geen koud nootje.*
 ‘It was a cold mint, not a cold nut.’

The dialectal marker *we* was used in the stimuli presented to the West-Flemish speakers. A corresponding marker *ze* was used in the version for the East-Flemish controls.

In order to establish the baseline for voicing measurements, tokens of word-final obstruents followed by obstruents in the next word were also included. Voiced obstruents are not expected

²While the vowel in the adjective of the target items was thus held constant before voiced and voiceless fricatives (/i:/), and before voiced and voiceless stops (/au/), no four adjectives could be found with the same vowel before underlying voiced and voiceless fricative and stops.

to surface word-finally in Dutch, except when a voiced stop follows in the next word. Although external voice sandhi is inherently variable (Slis, 1985), the sandhi context is the only available baseline for voicing in word final position, since Dutch obstruents are devoiced before a pause. Thus, the previously included test items were also put in the context where a voiced labial stop followed in the next word. Similarly, the context of following voiceless labial stops was used as the baseline for voiceless. Examples of the control items are in (8).

(8) Sample control items

koud boontje

'cold bean'

koud peertje

'cold pear'

Tokens of word-final obstruents followed by fricatives (/f/, /v/) in the next word were also included, but fricatives were consistently found to be the undergoers, rather than triggers of voice assimilation, which is in line with the established generalisations for voicing in Dutch (Booij, 1995). Thus, the results from the fricative context will not be counted or discussed here.

3.2.3 Procedure

The stimuli were randomised for each speaker, and presented, one at a time, on a computer screen. The experiment was self-timed, and the speakers were encouraged to produce the utterances at a natural speed. The recordings were made in a quiet room, using a Marantz Professional solid state recorder (PMD620), with a Sony condenser microphone (ECM-MS907) placed on a stand. The recordings were sampled at 44 kHz.

Altogether 4 (obstruents)*9 (contexts)*12 (speakers)*2 (repetitions)=864 utterances were recorded. 66 utterances were discarded due to reading errors, mispronunciations, or hesitations, leaving 798 utterances for analysis.

3.2.4 Acoustic analysis

Acoustic analysis was performed using Praat (Boersma & Weenink, 2010), on a 5 ms Gaussian window. The spectrograms were analysed visually and segmented manually. Based on the inserted boundaries, the following measurements were made.

(9) Acoustic measurements made for stops:

1. Duration of the preceding vowel.

2. Stop closure duration. The closure was taken to be the period of low acoustic energy between the preceding vowel and the following stop release. In case of stop+stop clusters where the first stop in the sequence was unreleased (29% of cases) stop closure duration measurements were not taken.

3. Duration of voicing during stop closure, based on the presence of the voicing bar on the spectrogram.

4. Duration of the burst, based on the presence of high frequency noise following the closure phase of the stop. The absence of burst was coded as 0.

5. f_0 at 20 ms and 10 ms before the onset of the stop, using the autocorrelation algorithm in Praat.

6. f_1 at 20 ms and 10 ms before the onset of the stop, using the Burg algorithm in Praat.

(10) Acoustic measurements made for fricatives:

- 1.Duration of the preceding vowel.
- 2.Duration of the frication noise.
- 3.Duration of voicing during frication, based on the presence of the voicing bar on the spectrogram.
4. f_0 at 20 ms and 10 ms before the onset of frication, using the autocorrelation algorithm in Praat.
5. f_1 at 20 ms and 10 ms before the onset of frication, using the Burg algorithm in Praat.
- 6.Maximum intensity at high frequencies (bandpass filtered from 500 to 10000 Hz).
- 7.Minimum intensity at low frequencies (bandpass filtered from 0 to 500 Hz).

The choice of acoustic variables was mostly based on previous findings on voice assimilation in Dutch by Slis (1985), Ernestus (2000) and Jansen (2004). All the duration, f_0 and f_1 measurements had been previously reported by these works to correlate with the voicing contrast and voicing effects in Dutch. The intensity measurements were motivated by Ladefoged & Maddieson (1996), who discuss the role of intensity as a voicing cue in fricatives.

During the analysis of the recordings, it became apparent that some West-Flemish speakers have a split in their production of the *ou* vowel. While all East-Flemish speakers produced it as a monophthong, the pronunciation of the West-Flemish speakers varied between monophthongal [u:] and diphthongal [au]. The variation was not categorically lexically conditioned (both variants were found in the production of *koud*, *zout* and *fout*), but the majority of monophthongal realisations were found with *koud*, which skewed the vocalic measurements. To avoid this problem, all of the vocalic measurements (vowel duration, f_0 and f_1) in the context of a stop were discarded from the subsequent statistical analysis.

3.3 Results

The statistical analysis was performed using R (R Development Core Team, 2005), version 2.11.1. The acoustic correlates of voicing were analysed with linear mixed-effects regression models using the lme4 package (Bates & Maechler, 2009). The analysis was run separately for stops and fricatives due to different variables measured for the two obstruent subclasses. For every phonetic variable a mixed-effects model was fitted with random intercepts for speaker and item. Predictors considered in the fixed part of the model included: dialect (West-Flemish vs. East-Flemish), the right-hand environment (sonorant consonant, vowel, voiced stop, or voiceless stop; abbreviated as ‘Environment’ tables) and the underlying voicing value of the word-final obstruent (voiced or voiceless). The contrast coding distinguished between obstruents followed by sonorant consonants and vowels following previous findings that prevocalic voicing is relatively more common (Simon, 2010). Fixed predictors were only retained in the model if they were found to significantly improve the model’s fit according to the log-likelihood test. Best fitting random-intercept only models were then compared to models with random slopes (i.e. models where the effect of predictors is allowed to vary within speaker/item) using the log-likelihood test. The best fitting models are presented throughout this section. The summaries of the fixed coefficients are presented in tables; each table includes the β -coefficient, standard error, and t -value for the intercept and for the fixed predictors. p -values were calculated based on Markov Chain Monte Carlo sampling, using the pvals function in the languageR package (Baayen, 2011). The p -values are included in the models’ summaries. Significant interactions are represented graphically. For vectors with more than two levels the models were re-run with different baselines to obtain the relevant significance values. Those values are quoted throughout the text, but they are not included in the tables.

3.3.1 Stops

Stop voicing was analysed using linear mixed-effects regression models fitted to the following variables: duration of vocal fold vibration, ratio of voicing to closure duration, duration of closure duration, and duration of burst.

Duration of vocal fold vibration The most parsimonious model of vocal fold vibration in stops had just one main effect, that of the right-hand environment. Voiced stops were found to trigger significantly more voicing in a preceding word-final stop than sonorants ($\beta=28.14$, $SE=2.61$, $t=10.77$, $p=0.001$). The voicing duration was on average shorter before voiceless stops than before sonorants ($\beta=-4.69$, $SE=2.31$, $t=-2.03$, $p=0.040$). There was no significant difference in vocal fold vibration between the prevocalic and pre-sonorant environment ($\beta=4.94$, $SE=3.09$, $t=1.60$, $p=0.111$). The model did not improve significantly upon including the effect of dialect ($\chi^2=2.11$, $df=1$, $p=0.146$), or the effect of underlying voicing ($\chi^2=1.46$, $df=1$, $p=0.227$). This result shows that East- and West-Flemish speakers did not differ significantly with respect to vocal fold duration in word-final stops, and that the duration of vocal fold duration was not sensitive to the underlying voicing value of the stop. The model's summary is in Table 3.1.

Table 3.1: Regression coefficients, with standard error, t and p values for a model predicting the duration of vocal fold vibration (in ms) in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.

Term	Level	β	SE	t	p
(Intercept)		19.76	1.59	12.40	0.001
Environment	voiced stop	28.14	2.61	10.77	0.001
Environment	voiceless stop	-4.69	2.31	-2.03	0.040
Environment	vowel	4.94	3.09	1.60	0.111

Voicing ratio Ratio of vocal fold vibration to closure duration (voicing ratio) was also modelled using linear mixed-effects regression. Similarly as in the case of voicing duration, the most parsimonious model of voicing ratio had only one fixed main effect, that of the right-hand environment. Pre-sonorant stops had a significantly lower voicing ratio than stops followed by voiced stops ($\beta=0.63$, $SE=0.09$, $t=7.30$, $p=0.001$), but they did not differ significantly in terms of voicing ratio from stops followed by voiceless stops ($\beta=-0.02$, $SE=0.07$, $t=-0.22$, $p=0.782$), or from stops followed by vowels ($\beta=0.14$, $SE=0.08$, $t=1.85$, $p=0.065$). The fit of the model did not improve significantly, according to the log-likelihood test, upon adding a main effect of dialect ($\chi^2=1.38$, $df=1$, $p=0.240$), or an effect of the underlying voicing of the stop ($\chi^2=0$, $df=1$, $p=1$).

Table 3.2: Regression coefficients, with standard error, t and p values for a model predicting the voicing ratio in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.

Term	Level	β	SE	t	p
(Intercept)		0.32	0.04	8.42	0.001
Environment	voiced stop	0.63	0.09	7.30	0.001
Environment	voiceless stop	-0.02	0.07	-0.22	0.782
Environment	vowel	0.14	0.08	1.85	0.065

Closure duration The final model of closure duration chosen based on the selection criteria defined above had a significant main effect of the right-hand environment. The model is summarised

in Table 3.3. Closure duration was relatively greatest in pre-sonorant stops, but the difference was only significant in comparison to stops followed by voiced stops ($\beta=-12.27$, $SE=4.67$, $t=-2.63$, $p=0.012$), and to stops followed by voiceless stops ($\beta=-12.94$, $SE=4.05$, $t=-3.19$, $p=0.002$). Adding the effect of dialect, or underlying voicing did not significantly improve the fit of the model (dialect: $\chi^2=0.30$, $df=1$, $p=0.582$; underlying voicing: $\chi^2=0$, $df=1$, $p=1$).

Table 3.3: Regression coefficients, with standard error, t and p values for a model predicting closure duration (in ms) in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.

Term	Level	β	SE	t	p
(Intercept)		65.13	3.36	19.41	0.001
Environment	voiced stop	-12.27	4.67	-2.63	0.012
Environment	voiceless stop	-12.94	4.05	-3.19	0.002
Environment	vowel	-5.45	4.70	-1.16	0.247

Burst duration Similarly as in the case of other variables, modelling burst duration yielded a model with a single main effect, that of the right-hand environment. The model's summary is in Table 3.4. Burst duration was on average longest preceding sonorants than in any other environment, though the difference was only significant at the level of voiced stops. It has to be borne in mind, however, that a number of stops followed by another stop were unreleased. This was the case with 29% of stop+stop clusters. The burst duration was coded as 0 in such cases, which may have skewed the results of the model.

Table 3.4: Regression coefficients, with standard error, t and p values for a model predicting burst duration (in ms) in word-final stops in East- and West-Flemish. The intercept corresponds to a stop followed by a sonorant.

Term	Level	β	SE	t	p
(Intercept)		30.01	4.10	7.33	0.001
Environment	voiced stop	-20.23	8.92	-2.27	0.008
Environment	voiceless stop	-7.86	7.60	-1.03	0.238
Environment	vowel	-9.44	7.31	-1.29	0.198

3.3.2 Fricatives

Voicing in fricatives was analysed by mixed of linear mixed-effects regression using a series of models fitted to the following variables: duration of voicing during frication, ratio of voicing duration to duration of frication, duration of frication, duration of the preceding vowel, as well as f_0 and f_1 at 10 and 20 ms before the vowel offset.

Voicing duration The model of voicing duration in fricatives was chosen according to the previously established model selection criteria had a main effect of dialect and the right-hand environment in its fixed part, as well as an interaction between these two factors. Adding the underlying voicing of the fricative as an additional predictor did not significantly improve the fit of the model ($\chi^2=0.74$, $df=1$, $p=0.390$). A summary of the selected model is in Table 3.5. The significant main effects involved increased voicing duration in pre-sonorant fricatives compared to fricatives followed by voiceless stops ($\beta=-45.71$, $SE=5.20$, $t=-8.90$, $p=0.001$), as well as greater voicing duration in West-Flemish compared to East-Flemish ($\beta=-46.31$, $SE=5.06$,

$t=-9.16$, $p=0.001$). There was no significant difference between voicing duration in pre-sonorant and prevocalic fricatives ($\beta=8.35$, $SE=5.72$, $t=1.46$, $p=0.128$). As illustrated in Figure 3.1, dialect interacted with the right-hand environment. West-Flemish had greater voicing duration than East-Flemish within the pre-sonorant and pre-vocalic environment. The inter-dialect contrast did not differ significantly between those two environments ($\beta=-0.05$, $SE=7.24$, $t=-0.01$, $p=0.944$). While there was a large difference between voicing duration in East-Flemish and West-Flemish fricatives before a following sonorant consonant or a vowel, the two dialects patterned together with respect to voicing duration in fricatives followed by voiced and voiceless stops. The difference between the two dialects was significantly greater within the pre-sonorant environment than within the context of a following voiced stop ($\beta=53.12$, $SE=5.79$, $t=9.17$, $p=0.001$), or a following voiceless stop ($\beta=43.97$, $SE=5.90$, $t=7.46$, $p=0.001$).

Table 3.5: Regression coefficients, with standard error, t and p values for a model predicting voicing duration (in ms) in word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant in West-Flemish.

Term	Level	β	SE	t	p
(Intercept)		58.55	3.76	15.58	0.001
Environment	voiced stop	0.08	5.20	0.02	0.994
Environment	voiceless stop	-45.71	5.14	-8.90	0.001
Environment	vowel	8.35	5.72	1.46	0.128
Dialect	East-Flemish	-46.31	5.06	-9.16	0.001
Env.: dialect	voiced stop: EF	53.12	5.79	9.17	0.001
Env.: dialect	voiceless stop: EF	43.97	5.90	7.46	0.001
Env.: dialect	vowel: EF	-0.05	7.24	-0.01	0.944

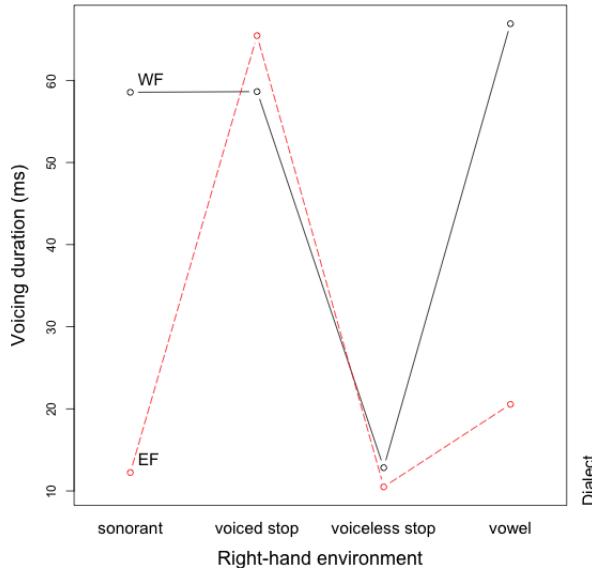


Figure 3.1: Interaction between dialect (East-Flemish vs. West-Flemish) with the right-hand environment on voicing duration in fricatives.

Voicing ratio The best-fitting model of voicing ratio had a main effect of dialect and the right-hand environment, as well as an interaction between these two predictors in its fixed part. In addition, a random slope for environment within speaker was included in the final model, as its inclusion resulted in a significantly better-fitting model ($\chi^2=22.68$, $df=9$, $p=0.007$). Including the

underlying voicing of the fricative as a predictor, on the other hand, did not significantly improve the model's fit ($\chi^2=0$, $df=1$, $p=1$). The summary of the fixed effects is in Table 3.6. The summary does not include p -values, as the current version of the lme4 package does not support MCMC sampling for models with random slopes. However, it can be inferred from the summary that the size of the effects and their associated t -values is consistent with the result of voicing duration modelling. Pre-sonorant fricatives showed a higher voicing ratio than fricatives followed by voiceless stops ($\beta=-0.60$, $SE=0.05$, $t=-11.22$). The mean voicing ratio was higher in West-Flemish than in East-Flemish ($\beta=-0.65$, $SE=0.06$, $t=-10.83$). Also, as in the model of voicing duration, the effect of the right-hand environment on voicing ratio interacted with dialect. The interaction is illustrated in Figure 3.2. The voicing ratio was very high in West-Flemish in the pre-sonorant and prevocalic contexts, patterning with the voicing ratio in fricatives followed by voiced stops. In comparison, East-Flemish speakers had a limited voicing ratio within these two environments. The two dialects patterned together with respect to voicing ratio in fricatives preceding voiced and voiceless stops.

The effects described above did not apply to the same extent for all the speakers, as indicated by the improvement in model's fit associated with allowing the effect of the right-hand environment to vary within speaker. An analysis of the random coefficients revealed that two speakers were exceptional with respect to the trends otherwise found in their dialect. One of the East-Flemish speakers (EF4) had a high coefficient of voicing ratio in prevocalic fricatives. One of the West-Flemish speakers (WF1) showed a relatively lower voicing ratio in fricatives followed by sonorant consonants compared to fricatives followed by voiced stops than other West-Flemish speakers, but she had a similar voicing ratio to other speakers in fricatives followed by a vowel.

Table 3.6: Regression coefficients, with standard error, t and p values for a model predicting voicing ratio in word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant in West-Flemish.

Term	Level	β	SE	t
(Intercept)		0.76	0.04	18.03
Environment	voiced stop	0.07	0.07	0.99
Environment	voiceless stop	-0.60	0.05	-11.22
Environment	vowel	0.18	0.11	1.61
Dialect	East-Flemish	-0.65	0.06	-10.83
Env.: dialect	voiced stop: EF	0.62	0.10	6.41
Env.: dialect	voiceless stop: EF	0.61	0.08	8.00
Env.: dialect	vowel: EF	0.01	0.16	0.08

Fricative duration The model of fricative duration involved main effects of the right-hand environment and dialect, as well as an interaction between the two in its main part, and a random slope for the right-hand environment within speaker. The inclusion of the random slope was motivated by the associated improvement in model's fit ($\chi^2=20.70$, $df=9$, $p=0.014$) compared to a model with random intercepts only. No improvement in fit was found upon including the fricative's underlying voicing value as a predictor ($\chi^2=0$, $df=1$, $p=1$). The summary of the fixed part of the model is in Table 3.7, and the interaction between the right-hand environment and dialect is in Figure 3.3. Fricatives were on average longer in East-Flemish than in West-Flemish, as indicated by a large main effect of dialect ($\beta=37.74$, $SE=8.26$, $t=4.57$). This difference between the two dialects was larger within the pre-sonorant fricatives than within fricatives followed by voiced stops ($\beta=-27.46$, $SE=7.16$, $t=-3.84$), or fricatives followed by voiceless stops ($\beta=-24.55$, $SE=7.17$, $t=-3.43$). The interaction was not significant at the level of a following vowel compared with the intercept ($\beta=-4.85$, $SE=10.74$, $t=-0.45$). The main effect of the right-hand environment does not

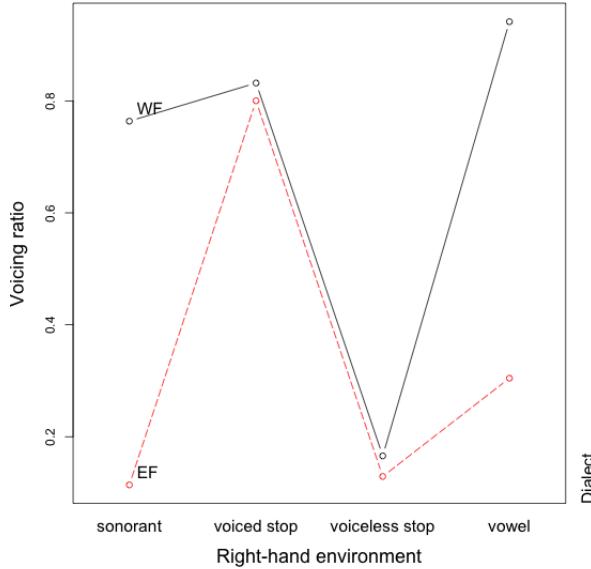


Figure 3.2: Interaction between dialect (East-Flemish vs. West-Flemish) with the right-hand environment on voicing ratio in fricatives.

lend itself to any conclusive generalisations. The effect was associated with very high standard error values at every level, and it was found to vary significantly across individual speakers.

Table 3.7: Regression coefficients, with standard error, t and p values for a model predicting fricative duration (in ms) in word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant in West-Flemish.

Term	Level	β	SE	t
(Intercept)		78.62	6.09	12.90
Environment	voiced stop	-5.94	6.88	-0.86
Environment	voiceless stop	2.34	6.81	0.34
Environment	vowel	-8.31	8.61	-0.97
Dialect	East-Flemish	37.74	8.26	4.57
Env.: dialect	voiced stop: EF	-27.46	7.16	-3.84
Env.: dialect	voiceless stop: EF	-24.55	7.17	-3.43
Env.: dialect	vowel: EF	-4.85	10.74	-0.45

Vowel duration In the case of vowel duration the model selection criteria yielded a model with only one fixed effect, that of the right-hand environment, and no random slopes. The fit of the model did not improve with the inclusion of dialect or the fricative's underlying voicing as fixed predictors (dialect: $\chi^2=0$, $df=1$, $p=1$; underlying voicing: $\chi^2=0.52$, $df=1$, $p=0.47$), or with the inclusion of the right-hand environment as a random effect within speaker ($\chi^2=0$, $df=1$, $p=1$). The summary of the model is in Table 3.8. Vowels were on average longer preceding fricatives in the context of a following sonorant than in the context of a following voiceless stop ($\beta=-6.15$, $SE=2.35$, $t=2.62$, $p=0.016$). Vowels preceding prevocalic fricatives were longer still ($\beta=8.92$, $SE=2.87$, $t=3.11$, $p=0.004$ compared to vowels in the context of a pre-sonorant fricative).

f₀ and f₁ No significant main effects of dialect, right-hand environment or the underlying voicing of the fricative were found in the process of mixed-effects modelling of f₀ and f₁ at 10 ms before the vowel offset. At 20 ms before the vowel offset there was an effect of f₀ and f₁ lowering preceding

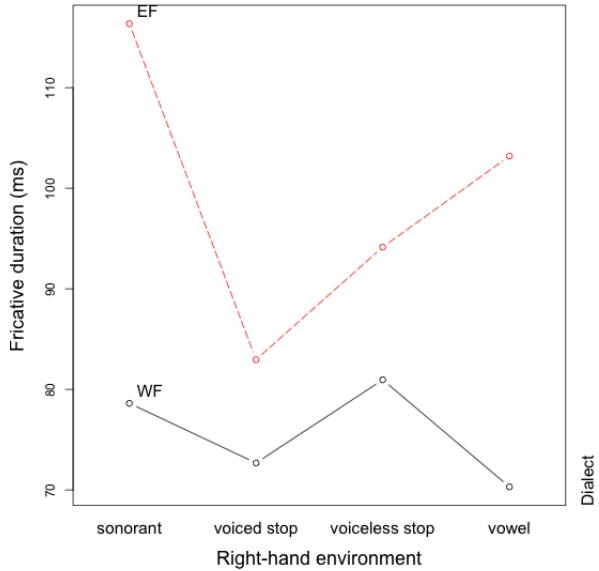


Figure 3.3: Interaction between dialect (East-Flemish vs. West-Flemish) with the right-hand environment on fricative duration.

Table 3.8: Regression coefficients, with standard error, t and p values for a model predicting duration of the vowel (in ms) preceding word-final fricatives in East- and West-Flemish. The intercept corresponds to a fricative followed by a sonorant.

Term	Level	β	SE	t	p
(Intercept)		74.92	2.01	37.37	0.001
Environment	voiced stop	1.07	2.31	0.46	0.682
Environment	voiceless stop	-6.15	2.35	-2.62	0.016
Environment	vowel	8.92	2.87	3.11	0.004

underlyingly voiced fricatives. However, considering that a similar effect was not observed closer to the vowel offset suggests an influence of the preceding sound. In all the underlyingly voiceless fricatives there was a rhotic preceding the vowel (*Fries*), whereas all the underlyingly voiced fricatives were preceded by an obstruent (*vies*). This was necessitated by lexical restrictions, but is likely to have confounded the f_0 and f_1 results.

3.3.3 Summary. Categoricity and gradience

Phonetic variables associated with voicing in prevocalic and pre-sonorant stops in Belgian Dutch pattern in accordance with earlier descriptions. No significant dialectal variation was found with respect to stop voicing. East-Flemish and West-Flemish speakers had an extensive portion of vocal fold vibration, both in absolute terms and relatively to closure duration, in stops followed by another voiced stop. In comparison, stops followed by voiceless stops, sonorants and vowels had a very short period of voicing. Pre-sonorant stops had a significantly longer voicing duration than stops followed by voiceless stops, but they did not differ significantly from prevocalic stops. Modelling closure and burst duration did not reveal any significant dialectal difference. None of the analysed variables showed a significant effect of the underlying voicing of the stop.

Results from pre-sonorant and prevocalic fricatives are also largely in line with previous reports on voicing in Belgian Dutch. Significant interactions between dialect and the right-hand environment in models of voicing duration, voicing ratio, and fricative duration. Models of voicing

duration and voicing ratio showed a considerable dialectal difference within pre-sonorant and prevocalic fricatives. Fricatives within these two environments patterned with fricatives followed by voiced stops in West-Flemish. In East-Flemish, on the other hand, pre-sonorant and prevocalic fricatives were close in their mean values of voicing duration and ratio to fricatives followed by a voiceless stop. Prevocalic fricatives appeared to have more voicing on average than pre-sonorant fricatives, but neither the main effect of vowel, nor the relevant interactions were significant, presumably due to individual variation. This hypothesis is supported by the result from the model of voicing ratio, which showed a significant effect of environment within speaker. A closer analysis of the random variation confirmed that some East-Flemish speakers, one in particular, had high levels of prevocalic fricative voicing, whereas for other speakers prevocalic voicing was limited. A degree of individual variation was also found with respect to the effect of environment on fricative duration. This is interesting, considering that no other duration measurement (e.g. voicing and vowel duration) showed such an effect. This would suggest that speakers make active use of fricative duration to realise a given voicing target. What is also noteworthy is that, despite all this variation, a fairly stable pattern emerges for the population with respect to ratio of voicing to fricative duration.

Before discussing the significance of these findings, it is necessary to ask to what extent pre-sonorant and prevocalic fricative voicing in West-Flemish can be seen as a categorical process. Partial support for the categoricity hypothesis comes from the fact that pre-sonorant and prevocalic fricatives in West-Flemish patterned with fricatives followed by voiced stops, showing very high values of voicing duration and voicing ratio. In addition to this, independent support for categoricity in West-Flemish voicing comes from bimodality. Figure 3.4 presents a density plot, reflecting the distribution of obstruents with specific voicing durations during friction/closure. Pre-sonorant and pre-vocalic obstruents had been pooled together, which is justified considering the lack of significant differences uncovered between these two groups. From now on, I shall use the term ‘pre-sonorant’ in the context of West-Flemish as encompassing the prevocalic environment as well. In the case of pre-sonorant fricatives there is a very clear bimodal distribution, which approximates the bimodality found before voiced and voiceless stops (represented by dashed lines)³. In comparison, a different situation is found in stops. While pre-sonorant stops are not entirely normally distributed (there are a few tokens with extended voicing during closure), a clear bimodality is missing. The bimodality test confirms that two categories emerge from the continuous data in the case of pre-sonorant fricatives, which is consistent with an interpretation that a categorical voicing process applies there in an optional fashion.

On top of the two voicing categories, as observed in WF pre-sonorant fricatives, evidence of gradient voicing was also found in the data. For instance, as shown in the right panel of Figure 3.4, some phonetic voicing was present in most pre-sonorant stops in West-Flemish. The East-Flemish produced hardly any pre-sonorant obstruents with categorical voicing. Instead, as shown in Figure 3.5, EF speakers realised their pre-sonorant obstruents either as categorically devoiced (with no phonetic voicing whatsoever), or with gradient voicing up to 40 ms in duration⁴.

The remaining question is which part of the stop is affected by the gradient voicing. What transpired from the visual spectrographic analysis was that partial voicing was invariably found in the initial part of the closure. Figure 3.6 shows two typical tokens of a postvocalic voicing tail, as observed in the data.

³The first peak in the distribution is associated with somewhat shorter voicing in fricatives followed by voiceless stops than in fricatives followed by sonorants. A possible explanation lies in coarticulation which might be present when a voiceless stop follows, but is absent in the presence of a following sonorant.

⁴Figure 3.5 shows a dip in the distribution of voicing duration in voiceless obstruents produced by EF speakers. Interesting though it is, I do not attempt to provide a full account for this bimodality, and I refrain from discussing whether it should in any way be reflected in the feature inventory of East Flemish. For the purpose of the current investigation it is only crucial to observe that EF speakers did not typically produce categorical voicing before a sonorant.

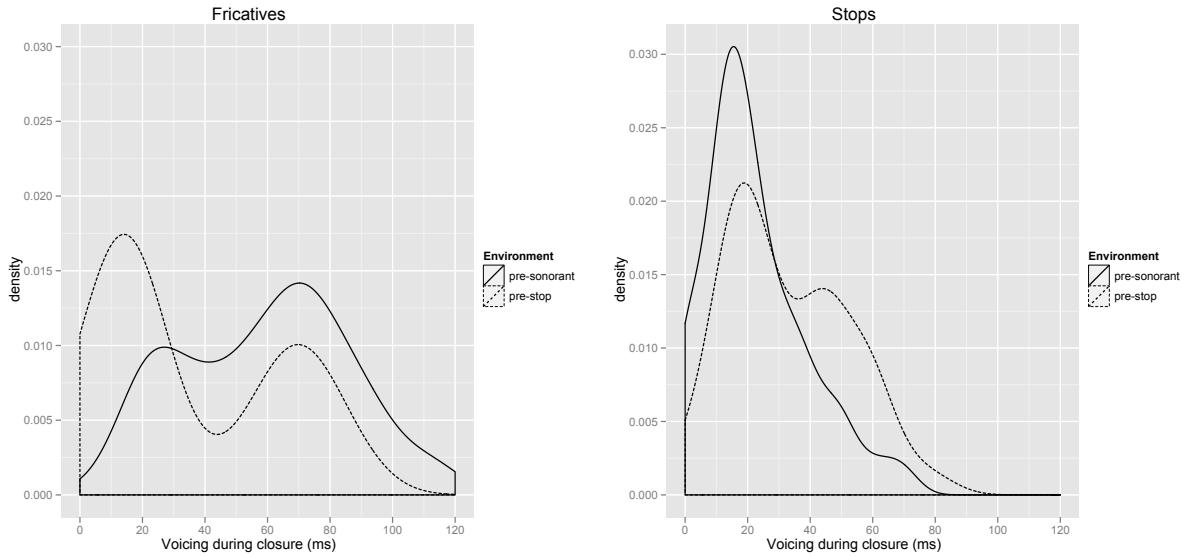


Figure 3.4: Density plots of voicing duration in West-Flemish pre-sonorant fricatives and stops. The dashed line represents the density of obstruents in the baseline condition, i.e. obstruents followed by a voiceless or a voiced stop.

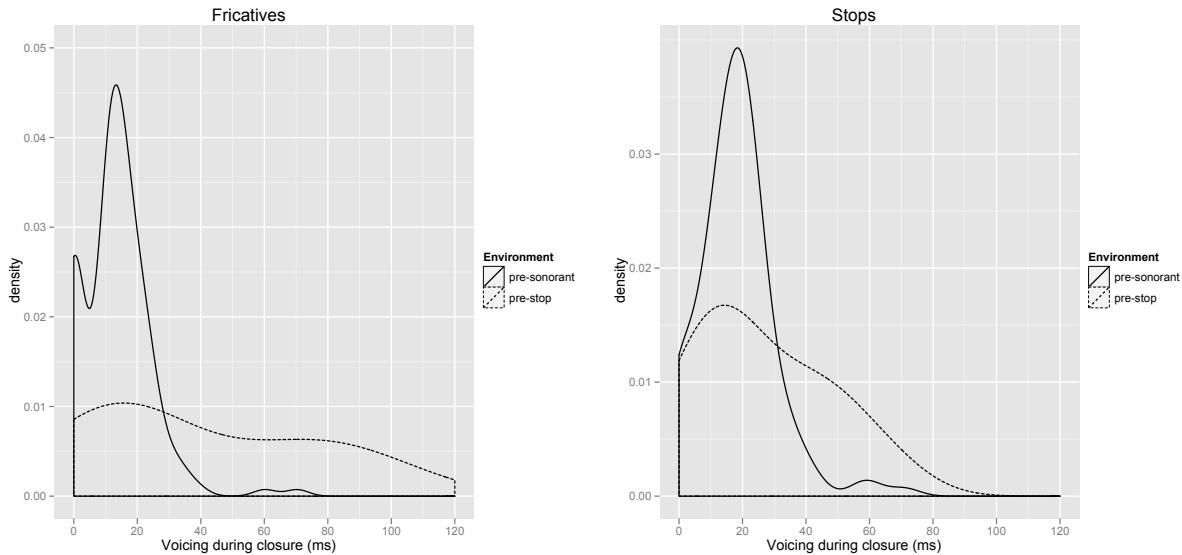


Figure 3.5: Density plots of voicing duration in East-Flemish pre-sonorant fricatives and stops. The dashed line represents the density of obstruents in the baseline condition, i.e. obstruents followed by a voiceless or a voiced stop.

3.4 Discussion

The phonetic results presented in the previous section furnish important information about the nature and the source of pre-sonorant voicing. Crucially, the results do not confirm that pre-sonorant voicing in West-Flemish word-final obstruents derives from the following sonorant, be it by means of gradient coarticulation, or categorical spreading.

The predictions made by the coarticulation analysis are inconsistent with the finding that voicing in the West-Flemish fricatives is optional, but categorical. The core idea behind the coarticulatory explanation is that word-final /s/ surfaces as a delaryngealised archiphoneme [S], which undergoes passive voicing in the phonetics (Colina (2009) for Ecuadorian Spanish). This analysis cannot be extended to West-Flemish: while some pre-sonorant fricative tokens surfaced

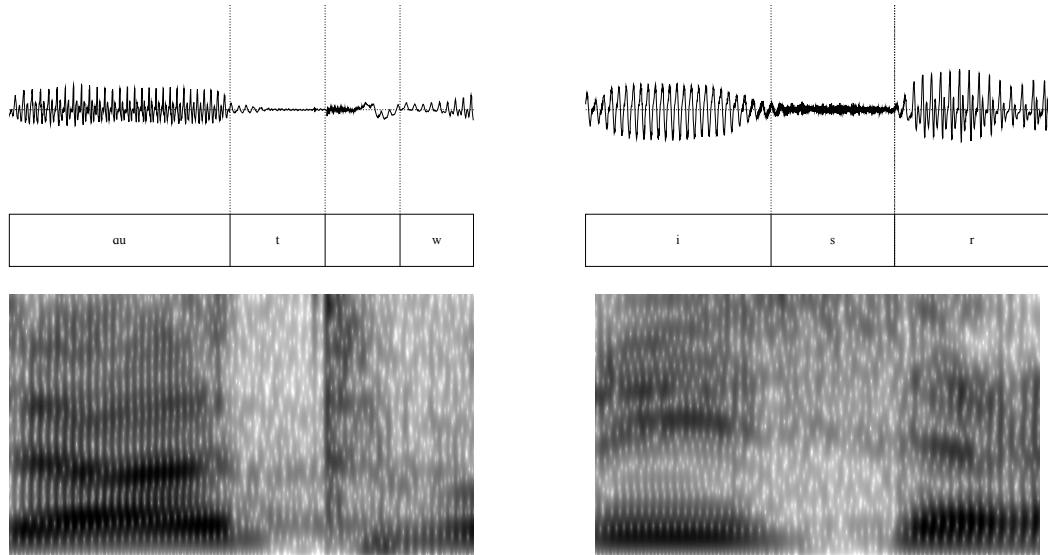


Figure 3.6: Pre-sonorant /t/ (left) and /s/ (right) with a short voicing tail

as gradiently voiced, in the majority of cases the fricative was fully voiced, with the duration of glottal pulsing exceeding the effect of passive voicing. This categorical application refutes the crucial prediction of a passive voicing analysis, and necessitates a phonological account of the West-Flemish pre-sonorant fricative voicing.

When it comes to categorical spreading analyses, these also make predictions which are not borne out by the current results. Specifically, the outcome of the experiment does not confirm that West-Flemish sonorants are redundantly laryngeally specified⁵, as proposed by Bethin (1984), Gussmann (1992) and Rubach (1996) for Polish, and by Bermúdez-Otero (2001), Wheeler (2005) and Jiménez & Lloret (2008) for Catalan. Three sources of evidence converge on the lack of laryngeal targets in West-Flemish sonorants.

First, sonorants do not trigger increased phonetic voicing in West-Flemish pre-sonorant stops, which would be expected if they were indeed laryngeally specified, and thus had an actively voiced target. This type of coarticulatory effect is observed by Jansen (2004) for the voiced fricative [z] in English⁶. Jansen (2004) found significantly more voicing during closure in word-final stops followed by [z] than in stops followed by [r] (e.g. the [g] voicing was longer in *Limburg zombie* than in *Limburg relish*). In stops followed by the actively devoiced [s] (*Limburg satin*), the voicing was shorter still. These findings provide support to the hypothesis that sounds with a voicing target, are phonetically active with respect to voicing. Following this logic, if West-Flemish sonorants were indeed phonologically voiced, they should trigger more voicing in the preceding obstruent than voiceless stops. However, this prediction is not confirmed by the West-Flemish sonorants. The model of voicing duration in stops showed a significant effect of the right-hand environment on voicing duration, with sonorants triggering more phonetic voicing than voiceless stops in preceding word-final stops ($\beta=-4.69$, $SE=2.31$, $t=-2.03$, $p=0.040$). However, this effect was no longer significant when the model was refitted to the subset of the data representing West-Flemish stops only ($\beta=-2.87$, $SE=3.01$, $t=-20.95$, $p=0.418$). A significant effect was found in a model of the subset from East-Flemish speakers ($\beta=-6.31$, $SE=2.17$, $t=-2.90$, $p=0.018$), which

⁵This generalisation holds whether the feature is generic [(+)voice], or a more sonorant-specific feature, e.g. [Sonorant Voice] (cf. Rice (1993)). I leave aside the issue of whether the feature in question is privative, binary or ternary at the phonological level.

⁶In the absence of phonological voice assimilation English provides a good test case of for the effect of voicing targets on phonetic voice coarticulation.

suggests that whatever effect of sonorants on preceding stop voicing we see in the data from the whole population is driven by the East-Flemish speakers, although the size of the dialect effects or the statistical power were not large enough to yield a significant effect of dialect in the model of voicing duration. In the case of West-Flemish speakers it seems that sonorants trigger as little coarticulatory phonetic voicing as voiceless stops. This result suggests that the voicing in sonorants is not phonetically active.

Second, other than not triggering phonetic voice assimilation, sonorants pattern differently from phonologically voiced stops. Unlike stop voicing, sonorant voicing is not contrastive. Pre-sonorant voicing also has a different distribution from regressive voice assimilation before stops, as [(+)voice] can spread from voiced stops, but not from sonorants, to the preceding word-final stops.

Third, the coarticulatory voicing hypothesis is corroborated by the observation that partial obstruent voicing in pre-sonorant position is invariably perseverative, as illustrated in Figure 3.6. The presence of a voicing tail readily follows from the aerodynamics of voicing: once set into motion (as during the articulation of a sonorant segment), the vocal folds naturally continue to vibrate until their movement is ceased by a rise in the supraglottal pressure, unless an extra articulatory gesture is performed to counteract the voicing. This is a typical coarticulatory effect which does not necessitate positing phonological assimilation⁷. This interpretation is also consistent with the observation made by Daniloff & Hammarberg (1973) that perseverative coarticulation is more typical of articulatory inertia, than of speech planning.

Challenging the proposal that pre-sonorant voicing is sourced by the following sonorant brings about the question of where the pre-sonorant voicing might come from. Following Jansen (2004), I propose that the actual source is the vowel preceding a word-final fricative. I further argue that the categorical application of pre-sonorant voicing, as observed in West-Flemish fricatives, emerges through diachronic perceptual reinterpretation.

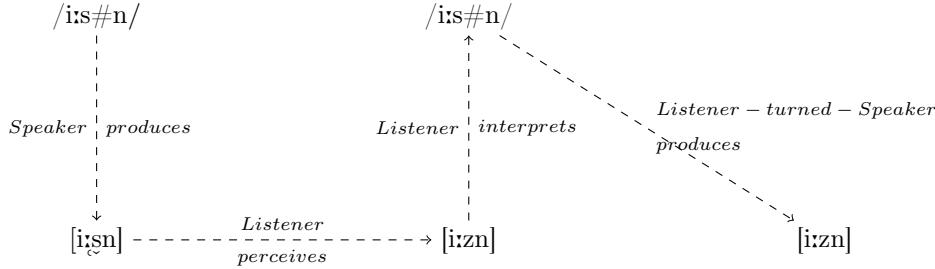
Following Jansen, I take it that vocal fold inertia gives rise to an automatic coarticulatory effect of partial voicing in word-final obstruents. The vocal folds are set into vibration for the articulation of a preceding vowel, and some of the vibration continues into the word final obstruent. This voicing is not counteracted in any way, due to the final obstruents' lack of voicing targets⁸. Jansen also notes that passively voiced obstruents, being able to sustain voicing for longer than actively devoiced ones, can be perceived as categorically voiced by listeners. In consequence, listeners may start to reinterpret these obstruents as categorically voiced, and realise voiced targets in their own production. In that way a categorical voicing pattern, such as the one we find in West-Flemish word-final fricatives, may arise.

An important aspect of the model outlined above is that it considers the roles of both speakers and listeners in the evolution of phonological patterns. Speakers initiate the change by producing coarticulation. However, the propagation of the change relies crucially on the listeners. The idea, going back to Ohala (1981), is that listeners participate in sound change by reproducing the linguistic input *as perceived*. If listeners perceive a coarticulatory process as a categorical change, they will produce the innovative pattern in their own production. The diagram in (11), modelled after Ohala (1981), illustrates the process with respect to West-Flemish fricatives. Speakers produce a word-final pre-sonorant /s/ with partial voicing which continues from the preceding vowel. A listener perceives the partial voicing as categorical, and maps the underlying /s/ onto the surface [z] in her own production.

(11) Reinterpretation of voicing in West-Flemish fricatives

⁷In comparison, stops followed by a voiced stop typically had continuous voicing extending throughout the entire cluster

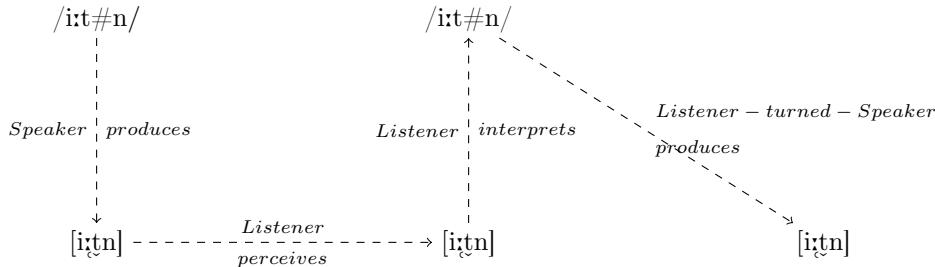
⁸The prediction that word-final obstruents in final-devoicing languages do not have their own voicing targets is confirmed by the current data, as no effect was found of the underlying voicing on the phonetic realisation of voicing in word-final stops or fricatives.



Although the diagram is meant to represent the change in time, its rightmost part also reflects what happens in the synchronic phonology and in acquisition. Children acquiring the West-Flemish system will perceive the variable voicing of pre-sonorant fricatives, and reproduce the same variation. However, they will also reconstruct the underlying voicing of the fricative (/s/ or /z/) through their knowledge of alternations involving suffixation when the underlying voicing surfaces phonetically. In this way the learners effectively arrive at mastering the pre-sonorant voicing process, without going through an intermediate step of assigning [(+)voice] to sonorants.

A perception-driven model has an additional advantage in allowing us to analyse the observed voicing asymmetry between stops and fricatives as resulting not from increased stop resistance to voicing, but from a voiceless perception of partially voiced stops. The diagram in (12) schematises the following scenario: even though coarticulatory voicing is present in stops, just like in fricatives, it does not result in a voiced perception by the listeners.

(12) No reinterpretation of voicing in West-Flemish stops



A motivation for why the perception of voicing might differ between subclasses of obstruents can be found in how voicing is cued in stops and fricatives. First, there are acoustic cues to stop voicing which are absent from fricatives. As noted by Slis & Cohen (1969), when it comes to obstruent voicing, it is more relevant to talk about the presence of voicing in the case of fricatives, and the onset of voicing in the case of stops. Onset of voicing is associated with the VOT parameter (Voice Onset Time), found by Lisker & Abramson (1964) to correlate significantly with the realisation of stop voicing in 11 languages, including Dutch. According to Lisker & Abramson (1964), voicing starts earlier relative to the release for voiced stops than for voiceless stops. The way this difference manifests itself in Dutch, especially in intervocalic position, is by means of the opposition between a negative and a positive VOT. In voiced sounds the VOT tends to be negative, i.e. the continuous voicing of the sound following the stop starts before the release. In the case of voiceless sounds, the VOT is positive, i.e. voicing only starts after the onset of the burst. A commutation test administered by Slis & Cohen (1969) confirms that the presence of a voice lag (a portion of voicelessness following the release of a plosive) leads to voiceless perceptions by listeners. Let us now consider the VOT values in West-Flemish pre-sonorant stops. According to the definition of VOT (distance between the onset of the release and the onset of *continuous* voicing of the following sound), stops with a short initial period of voicing have a positive VOT, because there is a lag between the initial closure voicing and the continuous voicing of the next segment. 119 out of 127

presonorant stops (93.7%) realised by the West-Flemish speakers in the experiment had a positive VOT. Thus, a cue for voicelessness was present in the vast majority of pre-sonorant stops.

Secondly, in addition to positive VOT, voicelessness was cued in most pre-sonorant stops by a long burst of 28 ms on average. In the absence of my own data on burst duration in voiced and voiceless stops in a position of contrast⁹, I shall compare this figure to the findings of Ernestus (2000) on Dutch¹⁰. Ernestus analysed the duration of closure and burst in single intervocalic stops in casual Dutch conversations, and evaluated the measurements against a categorical classification of the stop (voiced vs. voiceless) by 3 trained phoneticians who were also native speakers of Dutch. As illustrated by the scatterplots in the left panel of Figure 3.7, voiced percepts of coronal stops were found with shorter bursts, both absolute and relative to the closure. The right panel shows the ratio of burst to closure in the current data¹¹. It transpires that West-Flemish pre-sonorant stops tend to pattern like Standard Dutch intervocalic voiceless stops. What this means is that although some coarticulatory voicing is present in the closure, the voicing is not enhanced by a shorter burst. In consequence, the duration of the burst is longer in partially voiced pre-sonorant stops than in intervocalic voiced stops, potentially preventing voiced percepts.

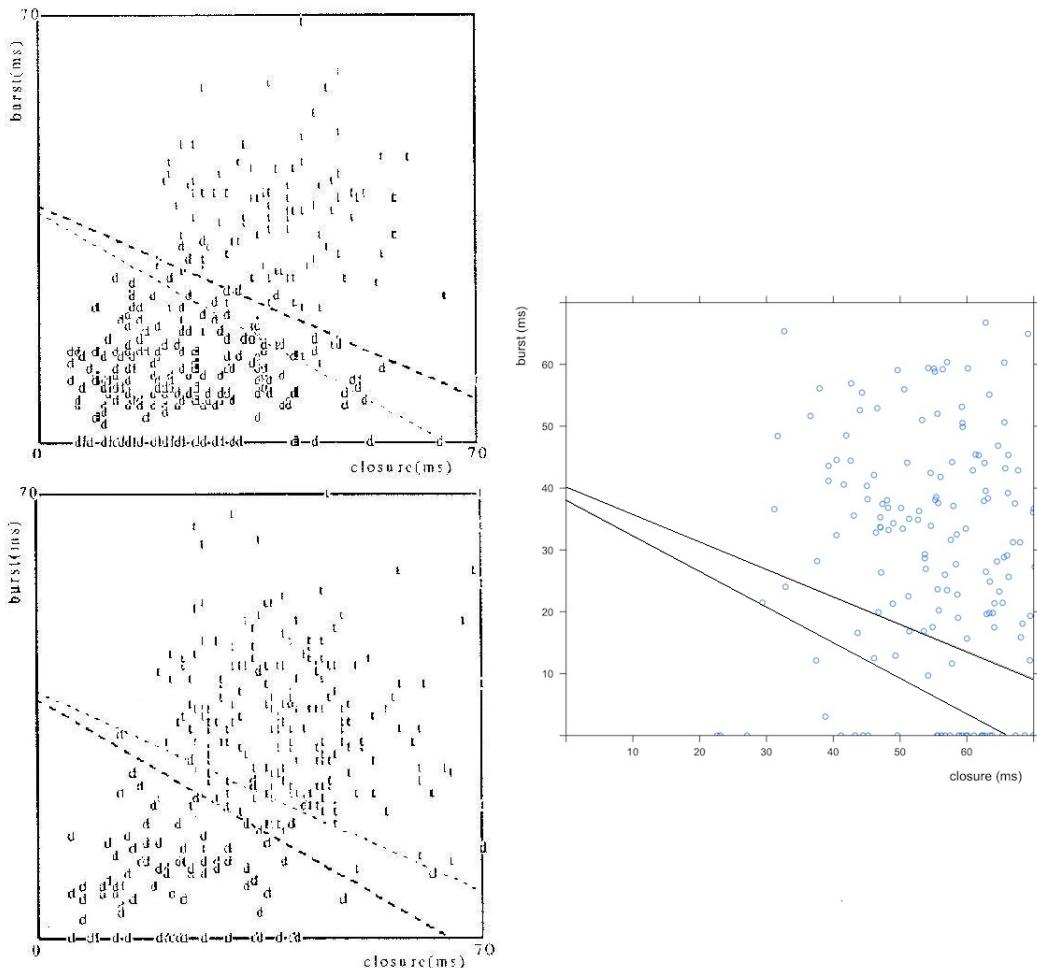


Figure 3.7: Left panel: closure and burst durations of intervocalic coronal stops classified as either voiced or voiceless, when preceded by a high (top), or low (bottom) vowel. Figure reproduced from Ernestus (2000, 219). Right panel: closure and burst duration of pre-sonorant stops in the current data.

⁹The participants in the experiment neutralised the underlying voicing contrast in the word-final position. Consequently, word-final stops studied in the current experiment do not provide information on how the voicing is cued.

¹⁰I would like to thank Mirjam Ernestus for granting me permission to reproduce the left panel of Figure 3.7.

¹¹Some observations are missing as the scale has been adjusted to correspond to Ernestus (2000). Also the durations are noticeably longer in the current data, which might be an effect of the reading task.

In fricatives, unlike in stops, the presence of voicing is more salient than the onset of voicing, as noted by Slis & Cohen (1969). To support this view, they present the results of Forrez (1966), who found that adding voice to the friction noise leads to [z] perceptions. Stevens et al. (1992) report a similar result, having analysed the acoustic correlates of voicing in the production and perception of voiced fricatives in American English. They found increased duration of glottal pulsing during the friction noise in voiced fricatives. Further, they examined the effect of voicing during friction in a forced choice perception experiment. In the classification of voiced and voiceless intervocalic fricatives listeners were found to rely on the duration of the friction noise where no glottal pulsing was present. 60 ms (or more) of voiceless friction noise triggered voiceless fricative percepts. Importantly, this result considers the duration of voicing relative to the duration of the friction noise. Coarticulatory voicing is expected to either not affect the duration of a word-final fricative, or even to shorten the friction noise¹². Either way, the presence of coarticulatory voicing both enhances the duration of a voiced part of the friction, and shortens the portion where the fricative is voiceless. Both these cues facilitate voiced percepts.

Voicing in fricatives is also cued by the relative intensity of high-frequency noise, as high energy at low frequencies masks the intensity of high-frequency friction (Ladefoged & Maddieson, 1996). In order to assess the scale of this effect in West-Flemish, the maximum intensity at high frequency (band-pass filtered from 500 to 10000 Hz) was measured, and it was further compared with the minimum intensity at low frequency (0 to 500 Hz). The maximum high-frequency intensity in fully voiced pre-sonorant fricatives was on average 6.72 dB lower than the minimum low-frequency intensity. With such intensity differences, the masking effect is likely to occur, preventing voiceless percepts. Moreover, the intensity difference correlated significantly with the duration of voicing during friction. As illustrated in Figure 3.8, the longer the voicing tail, the lower the high-frequency intensity as compared with the low-frequency one. In that way, the capacity for masking high-frequency noise increases steadily with the duration of voicing.

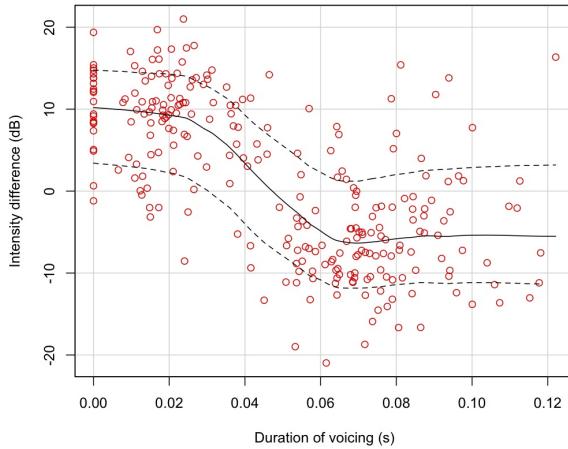


Figure 3.8: Left: Intensity difference (difference between intensity peak at high frequencies and intensity dip at low frequencies) as a function of voicing duration in pre-sonorant fricatives. Negative intensity difference means that the minimum low-frequency intensity is higher than the high-frequency maximum intensity; Pearson's product-moment correlation: $r=-0.65$, $p<0.001$

In addition, the peak intensity at high frequencies was positively correlated with the duration of the voiceless portion of a fricative (Pearson's product-moment correlation: $r=0.66$, $p<0.01$). It appears that in the fricatives with a shorter voiceless portion, the high-frequency intensity does not

¹²Jansen (2004) notes that friction noise is mechanically linked to voicing, and so it might shorten as an effect of passive voicing.

reach quite as high values as when the voiceless frication continues for longer. From a perceptual point of view, the longer the voiceless frication, the more robust the intensity cue for voiceless fricatives. This observation squares with the finding by Stevens et al. (1992), who argue that what matters in the perception of voice in fricatives is the length of not only the duration of voicing, but also the duration of the voiceless portion of the frication.

The significance of that argument is crucial, when one considers what happens in West-Flemish pre-pausal fricatives. If a word-final fricative has a voicing tail from the preceding vowel, and if voicing can be reinterpreted as categorical, then in principle word-final fricatives could undergo voicing before a voiceless stop, or before a pause. Voicing before voiceless stops is unlikely to occur in a language like Dutch, where there is regressive voice assimilation, and where obstruent clusters with conflicting voicing are unattested. The final voicing prediction, however, is more troublesome, especially if one wants to avoid stipulating independent constraints against word-final voicing.

Yet the prediction no longer holds if the duration of coarticulatory voicing is considered relative to the duration of the fricative. In discourse pre-pausal fricatives typically occur at the end of an utterance, sometimes at the end of a phrase. These are also the positions of increased final lengthening (Klatt, 1975). Cooper & Danly (1981) studied the effect of final lengthening on obstruents and vowels in different positions within a word in American English. Within fricatives they found the greatest amount of lengthening in the word-final position at the end of an utterance. A similar result was replicated for word-final fricatives in Dutch by Hofhuis et al. (1995), who found that word-final fricatives are twice as long at the Utterance boundary, as at the end of a word within discourse. In the light of the results by Stevens et al. (1992) and Hofhuis et al. (1995) the following scenario seems plausible. Coarticulatory voicing from the preceding vowel affects pre-sonorant and prepausal fricatives alike. This type of voicing is more likely to be reinterpreted as categorical in shorter fricatives. In effect, listeners might perceive gradient voicing as categorical in pre-sonorant position, but not prepausally, where the fricative is longer. Once this kind of gap in perception occurs, it is reflected in the production, so that pre-sonorant voicing is found in a language alongside with final devoicing before a pause.

The lack of prepausal voicing is an additional argument against laryngeal specifications in West-Flemish sonorants. As we have previously seen, partial obstruent voicing from sonorants is perseverative, rather than anticipatory. Thus, a parsimonious and phonetically accurate account should assume the same direction of categorical voicing. However, if voicing did indeed consist in rightward feature spreading, it would be expected that fricatives should surface as voiced following vowels in word-final position. Feature spreading, being a categorical process, is not sensitive to the relative duration of the target, and cannot differentiate between word-final fricatives before sonorants and fricatives followed by a pause. The most straightforward way to employ durational conditioning, and reconcile pre-sonorant voicing, final devoicing and the left-to-right directionality is through the perceptual filter of a language user in a diachronic perspective. Listeners perceive voicing in the (relatively shorter) pre-sonorant fricatives, and reanalyse it as categorical voicing in pre-sonorant position. Pre-pausal fricatives, being longer, are perceived as voiceless, and the reanalysis yields devoicing before a pause.

To sum up, the main idea of the perceptual re-interpretation model is that pre-sonorant voicing is a direct reflection of how listeners and learners perceive coarticulatory voicing in word-final obstruents. Initial coarticulatory voicing in stops does not affect VOT, or burst duration, both of which provide cues to voicelessness. While this observation does not warrant an immediate conclusion that partial closure voicing is not perceived by listeners, the cues to voicing are most likely offset by the co-occurring cues to voicelessness. A similar effect is absent from fricatives. Instead, there is the opposite trend, where adding voiced portions to the frication noise is likely to bring about voiced percepts. Even coarticulatory voicing is associated with high intensity at low frequencies which might mask higher frequency energy, similarly to what happens in actively

voiced fricatives. The acoustic differences provide phonetic grounds for the different perception of voicing in stops and fricatives. Those perceptions, in the diachronic model, give rise to the different patterns observed in obstruent subclasses. Perceptual factors also prevent fricatives from voicing before a pause, as the longer prepausal fricatives favour voiceless perceptions, which lead to voiceless (or more accurately: gradually voiced), and not categorically voiced productions.

One might wonder whether the occurrence of pre-sonorant voicing, and the stop-fricative asymmetry could not be explained by aerodynamic factors. The issue of positional application aside, intersonorant voicing is cross-linguistically a fairly common pattern. Westbury & Keating (1986) offer some explanation for this typological tendency, based on their aerodynamic model of voicing. Westbury and Keating show that voicing may be harder to terminate than to maintain in the intervocalic position¹³. In a phonological model where phonetic naturalness can trigger phonological change, the naturalness of intersonorant voicing could perhaps play a role in the initiation of the change I hypothesise to have occurred in West-Flemish.

However, whereas aerodynamic facts may explain why intersonorant obstruents are voiced, they cannot account for another trend present in the data, i.e. the asymmetry in the voicing of West-Flemish stops and fricatives. From a certain point of view, voicing could be seen as more natural in fricatives than in stops, as voicing is most difficult to maintain with the full closure in the vocal tract. However, this phonetic reasoning has been considered in Ohala (1983)¹⁴, and rejected on typological grounds. Ohala notices that languages are far more likely to have a voicing contrast in stops but not in fricatives than vice versa. As a potential functional explanation for this asymmetry, he considers the inherent aerodynamic conflict between voicing (low intraoral pressure) and frication (high intraoral pressure). This results in voiced fricatives being aerodynamically more demanding than voiced stops, which could explain why they are typologically dispreferred. Silverman (2006a, 165) offers some further typological discussion on the naturalness of stop vs. fricative voicing. He comments on the uncommonness of a diachronic change where a stop which has undergone intervocalic voicing is further spirantised. This observation could possibly be taken as tangential evidence for the unnaturalness of intervocalic fricative voicing *vis-à-vis* intervocalic stop voicing.

The upshot of Silverman's (2006) discussion of the relative naturalness of stop and of fricative voicing, which I fully endorse, is that we simply do not know whether it is easier to voice an intervocalic stop or an intervocalic fricative. In the absence of a quantifiable measure of effort involved in the production of stop and fricative voicing, and in the face of conflicting typological evidence, the hypothesis that fricative voicing evolves due to relative ease is unfalsifiable.

3.5 Conclusion

The data from West-Flemish show that pre-sonorant voicing can be phonological. Phonetic evidence reveals that the voicing of West-Flemish fricatives in pre-sonorant position is optional, but categorical, and must therefore take place in the phonology. However, both phonetic and phonological evidence speak against positing laryngeal specifications for sonorants. Based on this evidence, I have proposed that the categorical voicing in word-final fricatives in West-Flemish does not come from the neighbouring sonorant via feature spreading. Instead, I have argued that the West-Flemish pattern can evolve through language change based on categorical re-analysis of postvocalic coarticulatory voicing. I have also argued that acoustically voicing is more amenable

¹³Westbury & Keating (1986) focus on the intervocalic position in their discussion, but their conclusions seem readily extendable to the broader intersonorant context, since the aerodynamics of voicing is similar in sonorant consonants and in vowels.

¹⁴"If the problem with stops and voicing is that the accumulation of air in the oral cavity eventually quenches voicing, then this constraint should be less evident with fricatives since they have continuous venting of oral air pressure. So much for a priori prediction, since this turns out not to be true."(Ohala, 1983, 201)

to reinterpretation in fricatives than in stops. This proposal finds support in what we know about the perception of voicing (e.g. Stevens et al. 1992), and it makes predictions which are borne out closely by the production data obtained in the study. In addition, the perceptual hypothesis also has the advantage of employing explicit and falsifiable empirical premises which can be tested by future research.

CHAPTER 4

POZNAŃ POLISH

(...) [T]he award-winning Kemper Arena collapsed, with no loss of life, just twenty-four hours after hosting the American Institute of Architects Convention. The laws of physics are not to blame - merely the fact that sometimes what the models leave out is more important than what they include.

Tim Hartford

THE UNDERCOVER ECONOMIST

This chapter presents acoustic data on the realisation of word-final obstruents before a following sonorant sound in the Poznań dialect of Polish. Results from 16 native speakers of Poznań Polish confirm the existence of a voicing process in this dialect that affects word-final obstruents when they occur before a sonorant in the next word (Benni, 1959; Dejna, 1973; Ostaszewska & Tambor, 2000). However, in addition to what previous research has reported, the results allow for two crucial generalisations about presonorant voicing in Poznań Polish to be formulated. First, voicing before sonorants has phonetic characteristics of an optional but categorical process, but it applies less frequently than voicing before obstruents. In addition to this, the fact that underlying voiced obstruents in a word-final presonorant context are more prone to surface with full voicing than voiceless obstruents in the same position supports the view that voicing is non-neutralising in Poznań Polish. The co-occurrence of non-neutralisation with pre-sonorant voicing makes Poznań Polish exceptional in the typology of pre-sonorant voicing languages, and this, in turn, has potential implications for the claim that pre-sonorant voicing is conditioned by delaryngealisation. However, the implication is not genuinely falsified by the Poznań data, as the empirical result is also consistent with a scenario where delaryngealisation and pre-sonorant voicing apply sequentially, but optionally. Following an initial discussion of delaryngealisation, I consider the current data in the context of the existing formal analyses of Kraków-Poznań voicing, including Bethin (1984); Gussmann (1992), Rubach (1996), and Cyran (2012). I argue that the theoretical agenda of these approaches to provide a unified analyses of all the voicing facts in Polish is misguided, and that an empirically adequate model of Poznań Polish voicing needs to distinguish between pre-sonorant voicing and voice assimilation to obstruents.

4.1 Kraków-Poznań voicing and other voicing processes in Polish

Kraków-Poznań voicing is an oft-cited feature of south-western dialects of Polish that involves the voicing of word-final obstruents if a sonorant consonant or a vowel follows in the next word. Reports of the process go back to (Benni, 1959), who gives the examples listed in (1).

- (1) Examples of environments for Kraków-Poznań voicing (Benni, 1959)¹

		SW dialects	NE dialects
	<i>brat rodzony</i>	'blood brother'	[brad rɔd̪zɔni] [brat rɔd̪zɔni]
	<i>okrzyk radości</i>	'cry of joy'	[ɔkʃig radɔetci] [ɔkʃik radɔetci]
	<i>las rąbać</i>	'to fell the forest'	[las rɔmbatɛ] [las rɔmbatɛ]

Reports of Kraków-Poznań voicing are also found in more recent descriptive grammars of Polish. For instance, Ostaszewska & Tambor (2000) report that word-final obstruents undergo voicing before a sonorant in South-Western dialects of Polish: they illustrate this with the examples listed in (2) below.

- (2) Examples of environments for Kraków-Poznań voicing (Ostaszewska & Tambor, 2000)

		SW dialects	NE dialects
	<i>kosz malin</i>	'basket of raspberries'	[kɔʒ malin] [kɔʃ malin]
	<i>talerz malin</i>	'plate of raspberries'	[taleʒ malin] [talef malin]
	<i>kosz jabłek</i>	'basket of apples'	[kɔʒ jabwɛk] [kɔʃ jabwɛk]
	<i>kot rudy</i>	'ginger cat'	[kɔd rudi] [kɔt rudi]

Importantly, voicing before sonorants is limited to the external sandhi context. Voicing is contrastive in pre-sonorant obstruents word-medially: this is evident from a number of minimal pairs such as those listed in (3).

- (3) Word-medial voicing contrast before sonorants

<i>wiosła</i>	'oars'	[vjɔswa]
<i>wiozła</i>	'carried, FEM, PAST'	[vjɔzwa]
<i>setna</i>	'hundredth, FEM'	[setna]
<i>sedna</i>	'crux, GEN. SG.'	[sedna]

Other voicing-related processes in Polish that are relevant to the present discussion include word-final devoicing and voice assimilation to obstruents, both of which are reported to apply in all varieties of Polish. Word-final devoicing causes word-final obstruents to have a voiceless realisation (Benni, 1959; Wierzchowska, 1980). The devoicing is observed phrase finally (cf. (4) below), and in the case of North-Western dialects of Polish, also before a pause as shown in (1) and (2).

- (4) Final devoicing in Polish (Jassem & Richter, 1989)

<i>rów</i>	'ditch'	[ruf]	<i>rowy</i>	'ditches'	[rɔvi]
<i>ruf</i>	'stern, GEN. PL.'	[ruf]	<i>rufy</i>	'sterns'	[rufi]
<i>snob</i>	'snob'	[snɔp]	<i>snoby</i>	'snobs'	[snɔbi]
<i>snop</i>	'sheath'	[snɔp]	<i>snopy</i>	'sheaths'	[snɔpi]

Voice assimilation between obstruents is regressive and applies across word-boundaries, as illustrated by the alternations in (5).

- (5) Voice assimilation across word boundaries

<i>kosz gruszek</i>	'basket of pears'	[kɔʒ gruʃɛk]	<i>kosze</i>	'baskets'	[kɔʃe]
<i>talerz płaski</i>	'flat plate'	[taleʃ pwaski]	<i>talerze</i>	'plates'	[taleʒe]
<i>kot Basi</i>	'Basia's cat'	[kɔd bac̪i]	<i>koty</i>	'cats'	[kɔti]
<i>kod pocztowy</i>	'postal code'	[kɔt pɔtʃtɔvi]	<i>kody</i>	'codes'	[kɔdi]

¹The data presented in this section are based on the existing sources as indicated. IPA transcriptions have been added, and they reflect the voicing processes described by the sources.

Voice assimilation is described as spreading from right to left, regardless of the underlying voice specification and manner of the target and the trigger.

Word-internally consonant clusters always agree in voicing, and the agreement is supported by morphological alternations and loanwords adaptation, as illustrated in (6).

(6) Voice agreement word-internally

a. In morphological paradigms

<i>tawka</i>	‘bench’	[wafka]		<i>tawek</i>	‘benches, GEN.’	[wavek]
<i>laska</i>	‘cane’	[laska]		<i>lasek</i>	‘canes, GEN.’	[lasék]

b. In loanwords (Gussmann, 2007)

<i>futbol</i>	‘football’	[fudbol]
<i>Makbet</i>	‘Macbeth’	[magbet]

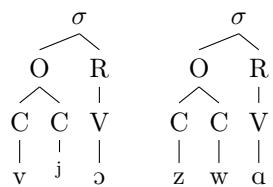
4.1.1 Previous analyses

Previous analyses of Kraków-Polish voicing have focused chiefly on explaining how the difference between the South-Western and North-Eastern dialects is to be understood theoretically, as well as on where the positional restrictions on pre-sonorant voicing come from. All these analyses attempt to reconcile pre-sonorant voicing with other Polish voicing phenomena using a minimal set of theoretical tools. The individual analyses are outlined below.

Bethin (1984)

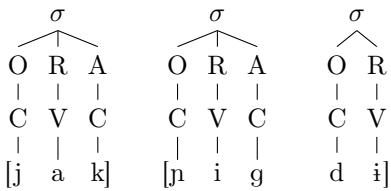
Bethin (1984) proposes a non-linear syllable-based analysis of Polish voicing. The syllable structure involves an onset and a rhyme, the latter consisting of a nucleus and a coda. The only possible codas are sonorant consonants. Obstruents which cannot be syllabified into an onset position (for instance word-final obstruents) occupy a slot in a syllable appendix. Syllabification applies prior to voicing, and it is generally guided by onset maximisation and the sonority principle, which predicts that word medial obstruent+sonorant clusters form a complex onset. An example is the word *wiozła* ‘carried, FEM, PAST’, whose syllabification is given in (7).

(7) Syllabification of word-medial obstruent+sonorant clusters (Bethin, 1984, p.25)



Furthermore, syllabification is also sensitive to word boundaries: thus, word-final obstruents in obstruent#sonorant clusters typically surface in coda position. This kind of syllabification is illustrated in (8) for *jak nigdy* ‘as never’. This example shows that Bethin’s syllabification rules predict that word-medial sequences of two obstruents should be separated by a syllable boundary.

(8) Syllabification of obstruent#sonorant clusters (Bethin, 1984, p. 31)



A = appendix

Having laid out the basic syllabification rules, Bethin (1984) formulates two dialect-specific voice-assimilation rules, which differ in the specification of the trigger and the environment. The North-Eastern rule of voicing assimilation, illustrated in (9), is triggered only by contrastively voiced obstruents, and it is not sensitive to syllable structure. The rule is said to apply word-medially, and also across word boundaries in fast or casual speech.

- (9) Voicing assimilation rule (Bethin, 1984, p.18, 28)

$$[-\text{sonorant}] \rightarrow [\alpha \text{ voice}] / \text{_____} \left[\begin{array}{c} \alpha \text{ voice} \\ -\text{sonorant} \end{array} \right]$$

For Kraków-Poznań Polish, Bethin proposes that voice assimilation is triggered by all voiced sounds, and that it targets coda obstruents only. The relevant rule is given in (10).

- (10) Kraków-Poznań appendix voicing assimilation rule (Bethin, 1984, p.28)

$$\begin{array}{c} \text{A} \\ | \\ \text{C} \quad \rightarrow [\alpha \text{ voice}] / \text{_____} [\alpha \text{ voice}] \\ [-\text{sonorant}] \end{array}$$

The rule stated in (10) voices all coda obstruents when they occur before a voiced sound. This includes both word-final obstruents and word-medial obstruents when they are followed by another obstruent. However, word-medial obstruents in pre-sonorant position do not meet the structural requirements of the voicing rule since they surface in onsets (cf. (7)). Therefore, pre-sonorant voicing applies across word-boundaries, but not within a word.

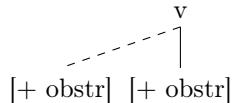
Gussmann (1992)

Like Bethin (1984), Gussmann (1992) argues for a syllable-conditioned view of voice assimilation. Gussmann provides a detailed discussion of syllabification rules in Polish and their application to complex consonant clusters. For the purpose of the current discussion, however, it will suffice to mention the case of two-way clusters. Gussmann's syllabification rules predict that word-medial pre-sonorant obstruents surface in onsets, but word-medial obstruents followed by another obstruent surface in codas, as do all word-final obstruents. Furthermore, Gussmann proposes that laryngeal features are only licensed in onsets: thus, only syllable-initial consonants may bear a contrastive [voice] specification. In codas, all obstruents are argued to undergo delaryngealisation. In this way, voice assimilation is seen as spreading from a laryngeally specified segment to an unspecified one. The asymmetry between voice licensing conditions in codas and onsets therefore predicts that delaryngealised coda obstruents will undergo $[\pm \text{voice}]$ spreading from the following onset.

Unlike Bethin, Gussmann distinguishes between word-internal voice assimilation and assimilation

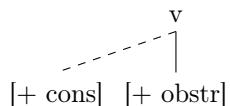
across word boundaries. An argument for the distinction is based on variability: while word-internal voice assimilation is argued to be obligatory, assimilation across word boundaries is subject to variation conditioned by speech rate and style (Gussmann, 1992, p.53)². To account for this difference, Gussmann proposes two separate ordered rules. The rule for word-internal voice assimilation (stated in (11)) is obligatory and spreads voicing from an obstruent to a preceding obstruent.

- (11) Obligatory rule of word-internal voice assimilation (Gussmann, 1992, p.44)



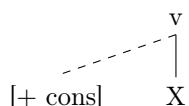
In addition to the obligatory word-internal assimilation rule, Gussmann proposes an optional rule, which applies across word boundaries³. In Warsaw Polish this optional rule involves spreading of [\pm voice] from voiced obstruents to preceding consonants (12).

- (12) Optional voice assimilation across word boundaries in Warsaw Polish (Gussmann, 1992, p. 53)



In Poznań Polish optional [\pm voice] spreading across word boundaries, illustrated in (13), can be conditioned by any segment. This is consistent with the assumption that “[i]n Polish, as elsewhere, sonorants are made [+voice] in the course of the phonological derivation” (Gussmann, 1992, p.42). Thus, taking into account (i) the specific view of [voice] specification that Gussmann advocates, (ii) the claim mentioned above about [voice]-licensing in onsets, and (iii) the rule given in (13) below, the Kraków-Poznań voicing data described in Section 4.1 can be accounted for.

- (13) Optional voice assimilation across word boundaries in Kraków-Poznań Polish (Gussmann, 1992, p. 54)



Rubach 1996

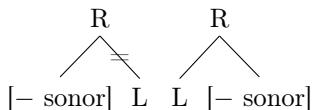
Rubach (1996) criticises the syllable-based analyses discussed above on empirical grounds. The main point of criticism is the syllabification itself. Specifically, Rubach shows that Gussmann and Bethin's syllabification rules predict that some word-medial pre-sonorant obstruents should surface in coda position. This prediction follows from the role that empty slots play in the syllabification

²According to Gussmann (1992), further motivation for distinguishing between voice assimilation word-internally and across word boundaries comes from the disparate behaviour of obstruent+sonorant#obstruent clusters compared with obstruent#sonorant+obstruent cluster. Some issues concerning this argument are discussed in Strycharczuk (2012).

³The specification of the rule in (11) as applying to consonants rather than obstruents reflects Gussmann's formulation. The reasons for this formulation are orthogonal to the current discussion; for details on why sonorants might be seen as targets of voice assimilation the reader is referred to Gussmann (1992).

algorithms, the fine details of which shall not concern us here. However, the crucial flaw of this model is that Kraków-Poznań voicing is falsely predicted to apply word-medially in certain cases, such as *wiosna* [v̥jɔsna] ‘spring’ (Rubach, 1996, pp. 107–108). As an alternative to the syllable-based proposals, Rubach analyses Polish voicing as in terms of Laryngeal nodes adjacency, along the lines of Cho (1990). A crucial aspect of the analysis is the link between delaryngealised and voicing: [\pm voice] spread targets only delaryngealised obstruents, and delaryngealisation may apply to consonants either in prosodic-word-final position or in pre-obstruent position. Delaryngealisation of obstruents followed by another obstruent is conditioned by a rule of Obstruent Delinking, illustrated in (14).

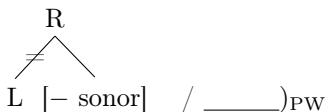
- (14) Obstruent Delinking (Rubach, 1996, p. 78)



R = Root node; L = Laryngeal node

As a consequence of the Obstruent Delinking rule, the underlying [voice] specification of any obstruent surfaces only in a pre-sonorant position. Prosodic-word-final delaryngealisation applies as shown in (15).

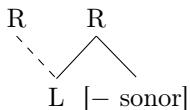
- (15) Final Devoicing (Rubach, 1996, p. 77)



R = Root node; L = Laryngeal node; PW = phonological word

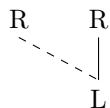
Once delinked from their laryngeal node, obstruents become potential targets for spreading, and so the delinking rules determine which obstruents are susceptible to voice assimilation. The spreading triggers, in turn, are defined by specific spreading rules. For Warsaw Polish Rubach proposes a rule that spreads [\pm voice] between obstruents. The rule is illustrated in (16).

- (16) Warsaw Spread (Rubach, 1996, p. 78)



For Kraków-Poznań Polish Rubach proposes a dialect-specific rule which involves spreading from any laryngeally specified segment, in line with the analyses proposed by Bethin (1984) and Gussmann (1992). This, again, rests on the assumption that sonorants acquire [voice] specification in the course of the derivation. Following the voice assignment, sonorant spread [voice] in Kraków-Poznań Polish, as conditioned by the rule in (17).

- (17) Cracow Spread (Rubach, 1996, p. 82)



What is interesting is that unlike Gussmann (1992), Rubach asserts that “the rules of voice assimilation in Polish are not only fully productive, but also entirely exceptionless”(Rubach, 1996, p.70). In support of this claim, Rubach cites the argument of language transfer, as seen in L2 pronunciations of English sequences like ‘these people’ [sp] or ‘this boy’ [zb]. The occurrence of voice assimilation in such pronunciations is taken by Rubach to indicate that voice assimilation is an obligatory process in Polish.

Cyran (2012)

In a recent paper, Cyran (2012) addresses the issue of Kraków-Poznań voicing working in a non-derivation, government phonological framework. One of the crucial assumptions Cyran makes, following Harris & Lindsey (1993), is that phonological representations are phonetically interpretable. One of the consequences of this approach is a departure from default-filling rules, which entails no phonological [voice] specification for sonorants and vowels, and in turn, no [voice] spreading from sonorants either. As an alternative to an approach involving [voice] specification of sonorants and vowels, Cyran proposes an analysis inspired by laryngeal realism (Iverson & Salmons, 1995, 2006; Jessen & Ringen, 2002; Honeybone, 2005; Beckman et al., 2011). In a nutshell, the proponents of laryngeal realism advocate that phonological representations should use laryngeal features which directly reflect the phonetic nature of a voicing contrast in a language. Thus, aspirating languages, which contrast short lag VOT with a long-lag VOT, have been proposed to involve a feature like [spread glottis], where tense obstruents are marked, and lenis obstruents are unmarked. Phonetic evidence in favour of this type of representation has been provided for instance by Jessen & Ringen (2002) for German, and by Beckman et al. (2011) for Swedish. Cyran (2012) proposes that Kraków-Poznań Polish is also a [spread glottis]-like system, though in the absence of direct phonetic motivation, Cyran dubs his approach ‘laryngeal relativism’. The representations assumed are as follows. Warsaw Polish contrasts neutral obstruents (C^0) with obstruents that contain the element |L|, or Low tone (C^L). C^L obstruents are always interpreted phonetically as voiced. Kraków-Poznań Polish, on the other hand, contrasts neutral obstruents (C^0) with obstruents containing the element |H|, or High tone (C^H). C^H are always pronounced with a short-lag VOT, whereas neutral obstruents in the H-system of Kraków-Polish voicing may in fact be prevoiced. Hence, the laryngeal relativism representations do not have a universal phonetic interpretation. They are said, however, to have a regular phonetic interpretation within a particular system (H-system, or L-system). Within an H-system, like that of Kraków-Poznań Polish, all voicing is spontaneous, and it depends on the manner of articulation, as well as on the prosodic and segmental environment. Voiced obstruents, sonorants and vowels have a common phonological representation (i.e. they are laryngeally neutral), but they may undergo spontaneous voicing in a variety of situations. Spontaneous voicing is always present for vowels, and it is usually present for sonorants, except when a sonorant is flanked by two voiceless obstruents, or a voiceless obstruent and a word-boundary. Neutral obstruents are only voiced before a sonorant, or a vowel, although in case of obstruent clusters, an entire cluster may receive spontaneous voicing in a pre-vocalic/pre-sonorant position.

Cyran considers spontaneous voicing to be an interface phenomenon, which relegates most of the Kraków-Poznań Polish voicing facts from the domain of phonological computation. However, Cyran’s analysis crucially requires one phonological operation, namely that of word-final devoicing which removes the element |H| from word-final underlyingly voiceless obstruents. C^H obstruents

are always pronounced as voiceless, and so they could not serve as input to spontaneous voicing at the interface. Thus, Cyran proposes that word-final devoicing is still a core property of Polish laryngeal phonology (in Kraków-Poznań Polish as well as Warsaw Polish). Devoicing only applies word-finally, which also results in the lack of pre-sonorant voicing in the word-medial position: underlyingly voiced obstruents are C^H in the output of phonology, and hence they are not amenable to spontaneous voicing.

Points of discussion and agreement

Although the summary offered in Section 4.1.1 is sketchy when it comes to the finer points of some of the analyses, it highlights the issues that much of the theoretical debate has focused on. The earlier papers mostly disagree on what conditions [voice] licensing in Polish: whether it is syllable structure, as proposed by Bethin (1984) and Gussmann (1992), or laryngeal node adjacency, as argued by Rubach (1996). In addition to this, Gussmann makes a distinction between word-internal voice assimilation and assimilation across word-boundaries because of the latter rule's variable application, whereas Rubach explicitly denies that any such variation exists. Both Gussmann and Rubach propose that voice assimilation is fed by delaryngealisation in order to derive the positional restrictions that limit Kraków-Poznań voicing to the word-final position. Bethin (1984), Gussmann (1992) and Rubach (1996) all share the assumption that sonorants and vowels are laryngeally specified in Polish. Laryngeal specifications for sonorants and vowels are used in all the analyses to account for the fact that voicing occurs before a non-contrastively voiced segment. The difference between South-Western and North-Eastern dialects is then also derived by a relatively small adjustment to the voicing rule, as the Warsaw rule only has obstruents as voice assimilation triggers, whereas in Kraków and Poznań voicing is said to be triggered by any voiced sound. Thus, in South-Western dialects, pre-sonorant voicing is not stipulated by means of an additional pre-sonorant voicing rule; once sonorants/vowels acquire their laryngeal features, they are said to behave just like obstruents do. While Cyran (2012) departs from the previous analyses in abandoning [voice] specifications for sonorants and vowels, he also assumes that sonorants, vowels and voiced obstruents have a common (neutral) laryngeal representation within the Kraków-Poznań voicing dialect. In addition to this, Cyran follows the earlier analyses in adopting the view that Kraków-Poznań voicing is conditioned by final devoicing.

4.1.2 Issues concerning neutralisation in Polish

Previous findings

As discussed at length in the previous section, all of the theoretical analyses of Kraków-Poznań voicing assume or predict laryngeal neutralisation at the Prosodic Word boundary in all dialects of Polish. However, this assumption warrants a closer scrutiny given previous empirical findings. In an experiment with five speakers of Polish, Slowiaczek & Dinnsen (1985) found that the underlying voicing contrast was typically preserved in the phonetics, as a significant main effect of the underlying voicing was found on the duration of the preceding vowel, and on the duration of vocal fold vibration within labial stops. Based on these findings, Slowiaczek & Dinnsen argue against the existence of laryngeal neutralisation in the word-final position in Polish, linking the case of Polish to the reports of near-neutralisation in other languages, including Catalan (Dinnsen & Charles-Luce, 1984) and German (Port & O'Dell, 1985). What is interesting is that the findings of Slowiaczek & Dinnsen (1985) were not replicated in a (1989) study by Jassem & Richter, who found no consistent phonetic trends that would suggest incomplete neutralisation in word-final obstruents. The underlying voicing specification did not significantly affect any of the four voicing-related phonetic cues analysed in the study, including duration of the preceding vowel, duration of

voicing during closure/frication, duration of closure/frication, and duration of stop release.

Jassem & Richter argue that the discrepancies between the results of the two studies might be due to methodological differences. Participants in the Slowiaczek and Dinnsen's study were native speakers living in an English-speaking country at the time of the experiment. The stimuli presented to the participants included metalinguistic evaluations, e.g. *Marysia wymawia _____ poprawnie/okropnie* ‘Mary pronounces _____ correctly/terribly.’, which might have triggered hypercorrection. Hypercorrection may also have been induced by the use of written stimuli, as the underlying voicing contrast in Polish is reflected in the spelling. Finally, the tokens of word-final devoicing analysed in the study were not phrase-final, but appeared before a vowel in some cases. In their discussion of this methodological point, Jassem & Richter (1989) state overtly that the environment for laryngeal neutralisation in Polish is the phrase, not the word. They further hypothesise that while word-final devoicing effects do occur, their status might be different from phrase-final devoicing, and dialect-specific.

In their own study, Jassem & Richter (1989) used monolingual native speakers of South-Western dialects. The pronunciations were elicited, rather than presented in writing, and all the word-final obstruents analysed in the study they were also phrase-final. Thus, due to these methodological differences between the studies by Slowiaczek & Dinnsen (1985) and Jassem & Richter (1989), it is impossible to conclude whether word-final devoicing involves a laryngeal neutralisation, and to what extent there might be a dialect effect. What is more, according to Jassem & Richter, there is also a an effect of prosodic boundary, which follows from their assertion that the domain of final devoicing in Polish is “the *phrase*, not the *word*” (Jassem & Richter, 1989, p. 318, emphasis original). Ultimately, in the absence of experimental evidence concerning laryngeal neutralisation at the Prosodic Word boundary in South-Western dialects of Polish, caution is certainly warranted in asserting that such neutralisation exists.

Near-neutralisation and non-neutralisation

The discrepancy between the results reported by Jassem & Richter (1989) and Slowiaczek & Dinnsen (1985) echoes the situation with German, where some authors report incomplete neutralisation of the voicing contrast word-finally (Port et al., 1981; Port & O'Dell, 1985; Port & Crawford, 1989). Others, however, found that the contrast was completely neutralised (Fourakis & Iverson, 1984). Similarly, incomplete neutralisation is reported by Ernestus & Baayen (2006) and Warner et al. (2004) for Dutch, while complete neutralisation was found by Jongman et al. (1992) and Baumann (1995). To an extent, the degree of incomplete neutralisation seems to depend on the experimental paradigm. Baumann (1995) argues that incomplete neutralisation effects follow largely from the use of read speech, which might involve hypercorrective pronunciation. That hypercorrection might indeed play a role is supported by the results reported by Port & Crawford (1989), who analysed the degree of incomplete neutralisation under different speaking styles. Port & Crawford elicited minimal pairs from five speakers in different stylistic conditions, including reading words embedded in different contexts, repeating words pronounced by an assistant, and dictating words to an experimenter. The degree of neutralisation was found to be influenced by style (with relatively greatest contrast when the speakers were dictating), but the near-neutralisation effects were not limited to conditions where increased hypercorrection was expected. Similarly, it is likely that orthography might trigger a degree of incomplete neutralisation, given that more neutralisation is typically found in studies that do not use written stimuli. However, incomplete neutralisation has also been found in languages like Catalan (Dinnsen & Charles-Luce, 1984; Charles-Luce, 1993), where the underlying voicing contrast is not represented orthographically.

Apart from the non-convergence of experimental evidence another troubling aspect of incomplete neutralisation is the effect size. Virtually every finding of incomplete neutralisation concerns very

small differences, typically in vowel duration. Slowiaczek & Dinnsen (1985) found that the near-neutralisation in Polish involved primarily a difference in vowel duration of about 10-15 ms. For German (Port & O'Dell, 1985) report effects on vowel and burst duration of around 15 ms, as well as a ca. 5ms effect on closure voicing. Charles-Luce's (1993) of study of Catalan also showed a 15 ms effect on vowel duration restricted to the cases where neutralisation would create semantic ambiguity. Warner et al. (2004) report an effect as small as 3.5 ms in vowel duration, and they also found a 9 ms effect on burst duration for phonemically long vowels in Dutch. Braver (2011a) reports a 3.5 ms effect and a 9 ms effect, both in vowel duration before /d/- and /t/-flaps in American English, based on two production studies.

While all of the near-neutralisation effects listed above were significant in production, one might wonder to what extent effects as small as these may play a role in perception. Slowiaczek & Szymanska (1989) tested the perception of the Polish minimal pairs obtained in the (1985) study by Slowiaczek and Dinnsen. Polish listeners were able to identify the underlying voicing of the word-final obstruent 61% of the time, which was significantly better than chance. A similar success rate (59% correct identification) was achieved by non-Polish speaking English listeners who also participated in the perception study. Using a similar procedure that involved playing minimal pairs obtained in a production study Port & O'Dell (1985) found 61% correct identification of the underlying voicing value in German. For minimal pairs recorded in the study on Dutch final devoicing Warner et al. (2004) found a better than chance identification when natural speech was used. The identification rate was around 60%, depending on the speaker producing the stimuli. Based on results from manipulated speech Warner et al. also report that listeners made use in perception of cues that are not systematically used in the production of incompletely neutralised voicing, such as closure duration. All in all, it would appear that the accuracy of perceptual identification is typically better than chance, but it is not entirely reliable with success rates not much higher than 60%. However, just as in the case of production-based studies on near-neutralisation, the perception-based studies of the same phenomenon are not without methodological problems. Braver (2011b) points out that that perception studies using real lexical items might be confounded by frequency effects, or perceptual biases. In a perception study of /t/- and /d/-flaps in American English in which some attempt was made to control for these potential confounds Braver (2011b) found no reliable identification of the underlying voicing value. In addition, Braver tested discrimination of the test items, which he reports to be limited to an AXB task. In an AB task listeners were unable to reliably discriminate /t/- and /d/-flaps. Thus, Braver's results introduce an additional facet to the issue of near-neutralisation, which involves a small but systematic, yet apparently imperceptible contrast.

Near-neutralisation emerges from the above survey as a very complex phenomenon indeed, but one that displays some cross-linguistically recurrent properties. It involves contrasts that are significant but invariably very small, typically in vowel duration. In addition, the contrasts are not reliably perceptible, even if identification is better than chance. I shall use these properties to identify cases of near-neutralisation in the current data, in accordance with research aims introduced below.

4.1.3 The current study

The current study is a phonetic exploration of pre-sonorant voicing in the Poznań dialect of Polish. As this is the first experimental investigation of Kraków-Polish voicing, one of my aims is to verify the earlier impressionistic records that such a process does indeed apply. This evaluation is achieved by means of a comparison between pre-sonorant and pre-obstruent voicing, and a comparison between Poznań speakers and speakers of other dialects. In addition, this study sets out to investigate whether there is an effect of the underlying voicing value on Poznań

voicing. The existence of an effect is consistent with Jassem and Richter's (1985) claim that laryngeal neutralisation applies in Polish only at the Phrase level, whereas the lack of an effect is predicted by the theoretical analysis of Kraków-Poznań voicing, all of which assume word-final delaryngealisation. Due to the issues surrounding near-neutralisation (cf. Section 4.1.2), I shall not merely class the effect of underlying voicing as significant or non-significant, but I will also consider the size of potential effects.

The investigation is based on data from two controlled experiments. The first experiment was a pilot study intended as an initial exploration of the phonetics of Poznań-Kraków voicing. The realisation of pre-sonorant voicing was tested based on word-final pre-sonorant stops (O#S) sequences embedded in meaningful Polish sentences. Following analysis of these results and the identification of further relevant research questions, a follow-up study was conducted. This experiment was designed to examine the realisation of word-final pre-sonorant obstruents (stops and fricatives) in a strictly controlled prosodic context (O#S sequences embedded in the same carrier phrase).

4.2 Stop#Sonorant sequences in Poznań and Warsaw Polish. Pilot study.

The pilot study was set up as a preliminary investigation into the voicing of word-final stops followed by a sonorant in the next word. The aim of the experiment was to establish whether the underlying voicing of the stop is recoverable in O#S clusters. 12 speakers participated in the experiment. 6 speakers were born in Poznań and were resident there at the time the experiment was conducted. They were all females, aged 24-26. The remaining 6 speakers were originally from Warsaw or the surrounding area, and they were resident in Warsaw at the time of the experiment. All were females aged between 20 and 24. All the participants were naive as to the purpose of the experiment, and were not paid for their participation.

4.2.1 Materials and method

The test items were 14 disyllabic words ending in a coronal stop (/t/ or /d/) followed by a sonorant (/m/, /n/, /r/, /l/, /w/, /j/, /ɔ/) in the next word. All the stops were preceded by a high-back vowel /u/. The test items were embedded in meaningful Polish sentences, as illustrated in (18).

- (18) Sample stimulus sentence
Kryzys gospodarczy powoduje nawrót lęku w społeczeństwie.
 ‘The economic crisis triggers a return of anxiety in the society.’

In addition, 20 tokens of word final coronal stops (/t/ or /d/) followed by an obstruent (/p/ or /b/, /f/, /v/, /s/, /z/) in the next word were included as controls. Like the test items, the control items were embedded in meaningful Polish sentences. The sentences were presented to the participants on cards, one at a time. The stimuli were randomised by re-shuffling for each participant.

The recordings were made in a sound-treated room, using a Behringer-B1 condenser microphone. The speakers were positioned 30 cm away from the microphone and instructed to read the sentences at a comfortable rate. They were encouraged to correct themselves if they made a mistake. Audio data were sampled at 44.1 kHz. Segmentation and acoustic analysis were carried out in Praat (Boersma & Weenink, 2010) on a 5 ms Gaussian window (spectrogram bandwidth 260 Hz). Boundaries were inserted manually based on visual inspection of the spectrograms. Altogether 2(repetitions)*(14(test stimuli)+20(control stimuli))*12(subjects)= 816 utterances were recorded.

14 utterances were excluded due to deletions, mispronunciations, or segmentation difficulties, leaving 802 test utterances for analysis.

The following acoustic measurements were taken. All of the measurements for stops and vowels had been previously recorded in studies on the voicing contrast in Polish, including Keating (1980), Slowiaczek & Dinnsen (1985), Jassem & Richter (1989), as well as in studies on laryngeal neutralisation in other languages, including Fourakis & Iverson (1984), Port & O'Dell (1985), Charles-Luce (1985), and Barry (1988). A wide array of measurements was considered for completeness, although some of them, particularly vowel duration, f_0 and f_1 measurements might have been vulnerable to confounds from phrasal stress, which was not controlled for.

1. Duration of glottal pulsing during closure (voicing duration). The presence of vocal fold vibration has been shown to be a primary acoustic correlate of the voicing contrast in true voice languages, including Polish (Keating, 1980). Voiced segments in Polish are typically realised with longer glottal pulsing, both in absolute terms and relative to closure duration, than their voiceless counterparts. Duration of glottal pulsing was measured manually based on the presence of the voicing bar on the spectrogram and periodicity in the waveform. Absence of voicing was coded as 0.
2. Stop closure duration. The voicing contrast has been shown to influence the duration of stop closure in some languages, where word-medial and word-final phonologically voiced stops have a shorter closure phase than phonologically voiceless stops (Chen, 1970; Kluender et al., 1988). Although this effect does not seem to have been observed for Polish, closure duration was recorded to help contextualise the measurements of the duration of glottal pulsing in terms of duration of occlusion. Closure was measured manually, based on the presence of low acoustic energy between the preceding vowel and the following stop release.
3. Vowel duration. Lengthening of the preceding vowel has been observed before phonologically voiced stops for a number of languages (Peterson & Lehiste, 1960; Chen, 1970; Kluender et al., 1988). Slowiaczek & Dinnsen (1985) report this effect for Polish, but their finding was not replicated by Jassem & Richter (1989). Keating (1980) finds that duration of the preceding vowel does not correlate with the voicing contrast. Vowel duration was measured manually. The measurement marker for the beginning of the vowel was placed at the beginning of the formant structure for vowels preceded by obstruent. For vowels preceded by sonorants, the initial boundary was placed at the onset of the formant steady state. The boundary between the vowel and the following stop was placed at the onset of low acoustic energy at higher frequencies.
4. Duration of the burst. Phonologically voiced stops have been found to have a weaker and shorter burst than phonologically voiceless stops (Fischer-Jørgensen, 1954; Slis & Cohen, 1969), although the effect in Polish is said to be fairly weak (Keating, 1980). The burst was identified based on the presence of high frequency noise following the closure.
5. f_0 and f_1 preceding the closure (20 ms and 10 ms before the onset of the closure) and following the release (10 ms and 20 ms following the release). House & Fairbanks (1953) and Kingston & Diehl (1994), *inter alia*, report that f_0 and f_1 are relatively lower following voiced consonants. The effect is most prominent following the obstruent, and fades over time. Jansen (2004) reports similar low-frequency lowering effects for vowels preceding lenis obstruents in Dutch, although the effect is less prominent than in the case of following vowels.

All of the duration measurements, including voicing, stop closure, burst, and vowel duration, were taken in milliseconds. f_0 and f_1 were measured in Hertz.

4.2.2 Results

Statistical analysis was performed in R (R Development Core Team, 2005), version 2.13.1, using the lme4 package (Bates & Maechler, 2009). The analytical method used was mixed-effects linear regression used stepwise to select a model that best fit the data in each case. The choice of the final model was based in all cases on whether or not the individual predictors/interactions were significant, and whether or not the inclusion of a predictor/interaction resulted in a model improvement according to the log-likelihood test. Non-significant predictors that did not result in a model improvement were not retained. The best models obtained for each variable are presented in tables throughout this section. For each model the intercept is included, together with the effect size, and the standard error for all the individual predictors, as well as the *t*- and *p*-value. The *p*-values were calculated using Markov Chain Monte Carlo (MCMC) sampling in the pvals function (languageR package; Baayen (2011)). For those predictors that were factors with more than two levels, the variables were re-levelled in order to obtain significance values as relevant. These values are quoted in the discussion of the individual models as appropriate, but they are not included in the model summaries presented in tables.

In order to explore the phonetic realisation of voicing in different environments, a series of mixed-effects models were fitted predicting the realisation of phonetic voicing before sonorants and before obstruents. The predictors considered in the models included the right-hand environment (i.e. whether the following segment was a sonorant, voiced obstruent, voiceless obstruent. This predictor is abbreviated as ‘Environment’ in the model summaries;), dialect (Poznań vs. Warsaw) and the underlying voicing value of the first stop in the sequence (abbreviated as ‘C1’; /d/ vs. /t/). Speaker and item were treated as random effects in all the models.

Duration of vocal fold vibration Table 4.1 below presents the summary of the first model, which was examined the effect of the duration of vocal-fold vibration during the stop closure. Significantly more voicing during closure was found in the Poznań dialect as compared with Warsaw ($p=0.002$), and in underlying /d/s as compared with underlying /t/s ($p=0.001$). Pre-sonorant stops showed significantly more voicing than stops followed by voiceless obstruents ($p=0.001$), as well as significantly less voicing than stops followed by voiced obstruents ($p=0.001$). In addition to the main effects, significant interactions were also found. There were significant two-way interactions between dialect and the right-hand environment, as well as between dialect and underlying voicing. The two dialects differed significantly more in the voicing duration of stops when these occurred before a sonorant than when they occurred before either a voiced or voiceless obstruent (both significant at $p=0.001$). The Poznań dialect also showed more of a /d/-/t/ contrast than the Warsaw dialect ($p=0.001$). A three-way interaction between dialect, right-hand environment and the underlying voicing of the stop was also significant. Figure 4.1, which illustrates the interaction, shows that underlying /d/s and underlying /t/s differed to a similar (small) extent in all environments in Warsaw dialect. In the Poznań dialect, by contrast, there was a relatively larger /d/-/t/ contrast before sonorants than before voiced or voiceless obstruents.

Ratio of voicing to closure The duration of voicing fold vibration relative to the duration of closure (henceforth, voicing ratio) was also modelled using mixed-effects linear regression. The results were similar as for voicing duration. There were significant main effects of dialect, right-hand environment and underlying voicing. The voicing ratio was higher in the Poznań dialect than in the Warsaw dialect ($p=0.001$). The ratio was significantly higher in stops followed by sonorants than in stops followed by voiceless obstruents ($p=0.001$), but lower than in stops followed by voiced obstruents ($p=0.006$). Underlying /d/s had a higher voicing ratio than underlying /t/s. There were significant two-way interactions between dialect and the right-hand environment, between dialect

Table 4.1: Regression coefficients, with standard error, t and p values for a model predicting the duration of vocal fold vibration during closure, measured in ms, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to an underlying /d/ followed by a sonorant in the Poznań dialect.

Term	Level	β	SE	t	p
(Intercept)		29.98	2.21	13.57	0.001
Dialect	Warsaw	-10.58	2.93	-3.61	0.002
Environment	voiced obstruent	14.98	1.93	7.77	0.001
Environment	voiceless obstruent	-9.54	2.15	-4.44	0.001
C1	/t/	-7.07	1.84	-3.85	0.001
Dialect: environment	Warsaw:voiced obstruent	8.69	2.19	3.97	0.001
Dialect: environment	Warsaw:voiceless obstruent	11.30	2.43	4.65	0.001
Dialect:C1	Warsaw:t	8.98	2.10	4.28	0.001
Environment: C1	voiced obstruent:/t/	3.67	2.80	1.31	0.190
Environment: C1	voiceless obstruent:/t/	4.82	2.94	1.64	0.090
Dialect:environment:C1	Warsaw:voiced obstruent:/t/	-5.56	3.21	-1.73	0.100
Dialect:environment:C1	Warsaw:voiceless obstruent:/t/	-8.81	3.34	-2.64	0.010

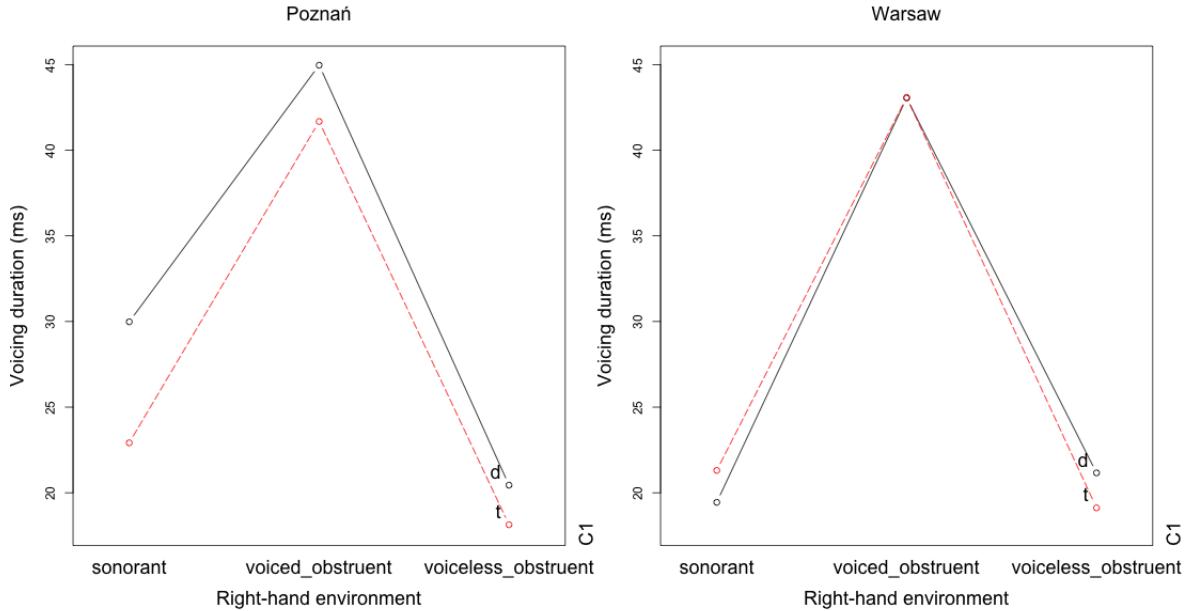


Figure 4.1: Interaction between the right-hand environment and the underlying voicing of the preceding stop on the voicing duration within the Warsaw and Poznań dialects.

and underlying voicing, as well as between dialect and the right-hand environment. Pre-sonorant stops differed more between the two dialects than stops followed either by a voiced ($p=0.001$) or voiceless ($p=0.001$) obstruent. Pre-sonorant stops had a higher voicing ratio in Poznań than in Warsaw. The Poznań dialect also had a more robust /t/-/d/ contrast than the Warsaw dialect in terms of voicing ratio ($p=0.001$). Underlying /t/s and /d/s contrasted more before sonorants than before voiced obstruents ($p=0.004$), or before voiceless obstruents ($p=0.070$). Finally, the model of voicing ratio yielded a significant three-way interaction between dialect, underlying voice specification of the stop, and the right-hand environment. The regression coefficients are listed in Table 4.2, whereas the interaction is plotted in Figure 4.2. The interaction shows that there is a /t/-/d/ distinction in pre-sonorant position in the Poznań dialect, but less so in other environments in this dialect. In Poznań underlying /d/s had a higher voicing ratio before sonorants than before voiceless obstruents, but the ratio was lower than when a voiced obstruent followed. Underlying

/t/s followed by a sonorant in the Poznań dialect had a similar voicing ratio to underlying /t/s followed by a voiceless obstruent. The Warsaw dialect did not show an increased pre-sonorant voicing ratio, compared to the voicing ratio preceding a voiceless obstruent. Additionally, the Warsaw dialect did not show much /t/-/d/-distinction in any of the environments.

Table 4.2: Regression coefficients, with standard error, t and p values for a model predicting ratio of voicing duration to closure duration in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to an underlying /d/ followed by a sonorant in the Poznań dialect.

Term	Level	β	SE	t	p
(Intercept)		0.72	0.05	15.31	0.001
Dialect	Warsaw	-0.27	0.06	-4.43	0.001
Environment	voiced obstruent	0.27	0.04	5.97	0.001
Environment	voiceless obstruent	-0.13	0.05	-2.58	0.006
C1	/t/	-0.20	0.04	-4.61	0.001
Dialect: environment	Warsaw:voiced obstruent	0.30	0.05	5.88	0.001
Dialect: environment	Warsaw:voiceless obstruent	0.24	0.06	4.20	0.001
Dialect:C1	Warsaw:t	0.21	0.05	4.29	0.001
Environment: C1	voiced obstruent:/t/	0.21	0.07	3.26	0.004
Environment: C1	voiceless obstruent:/t/	0.12	0.07	1.75	0.070
Dialect:environment:C1	Warsaw:voiced obstruent:/t/	-0.25	0.07	-3.38	0.004
Dialect:environment:C1	Warsaw:voiceless obstruent:/t/	-0.16	0.08	-2.08	0.022

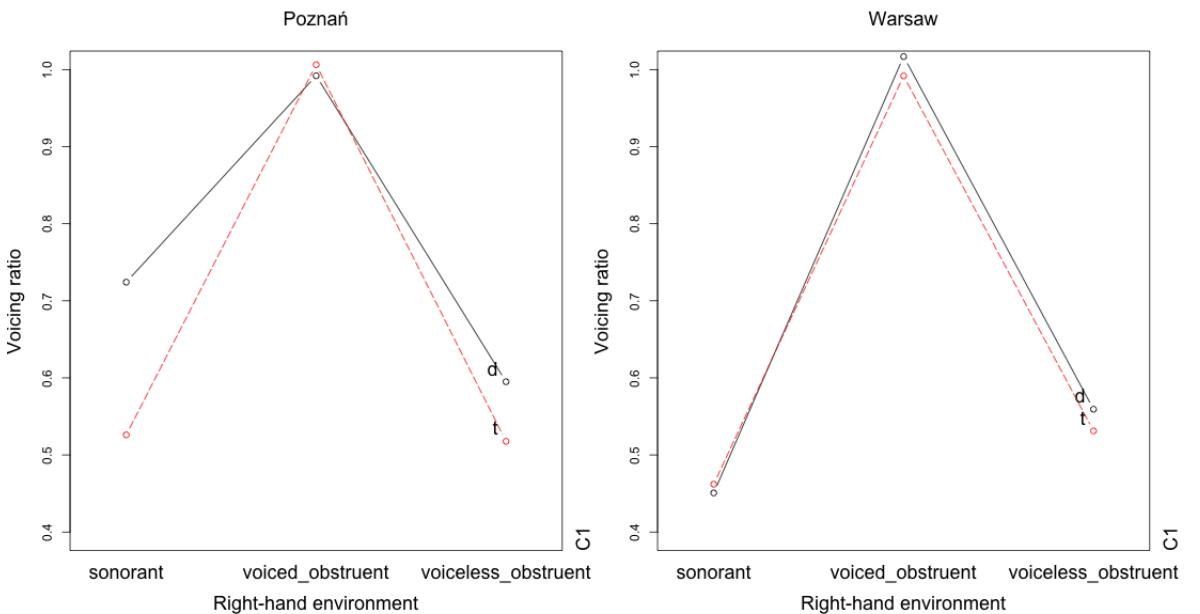


Figure 4.2: Interaction between the right-hand environment and the underlying voicing of the preceding stop on the voicing ratio within the Warsaw and Poznań dialects.

Vowel duration Underlying voicing, dialect and the right-hand environment were considered as predictors in modelling the vowel duration preceding word final stops followed by a sonorant in the next word. None of these predictors were found to have a significant effect on vowel duration.

f₀ and f₁ of the preceding vowel Mixed-effect linear regression model were fitted to the data predicting the f₀ and f₁ of the preceding vowel at 10 and 20 ms before the offset of the vowel.

In the case of f_0 , there was a significant effect of the right-hand environment. f_0 at 10 ms before the onset was significantly lower in vowels followed by stops+voiceless obstruent clusters than in vowels followed by stop+sonorant, or stop+voiced obstruent clusters ($p=0.001$ for both). A similar effect of f_0 lowering in vowels followed by stop+voiceless obstruent sequences was observed at 20 ms before the offset of the vowel.

There was no significant effect of dialect, or the underlying voicing of the stop, and neither of these effects were retained in the final model. A summary of the coefficients for the two models of f_0 is in Table 4.3

Table 4.3: Regression coefficients, with standard error, t and p values for a model predicting f_0 of a preceding vowel, measured in Hz, in stops followed by sonorants as compared with stops followed by obstruents. The intercept corresponds to a stop followed by a sonorant.

f_0 at 10 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		249.16	11.09	22.47	0.001
Environment	voiced obstruent	1.38	12.74	0.11	0.844
Environment	voiceless obstruent	-28.62	13.51	-2.12	0.002

f_0 at 20 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		253.66	11.41	22.24	0.001
Environment	voiced obstruent	1.38	13.25	0.10	0.886
Environment	voiceless obstruent	-28.61	14.05	-2.04	0.004

No significant effects and no interactions were found in the modelling of f_1 at 10 and 20 ms before the offset. The measurements of f_0 and f_1 after the release of the stop were not modelled, unlike all the other voice-related predictors, since those measurements cannot be systematically compared across different manner of articulation in the right-hand environment. For instance, for stops followed by voiceless obstruents, fundamental frequency is not expected following the release.

Closure and burst duration A model predicting the duration of closure of the first stop in the cluster showed a significant main effect of the right-hand environment, but no significant effect of dialect, or the underlying voicing of the stop. Closure duration was also significantly shorter before voiceless obstruents than before sonorants ($p=0.001$). The summary of the model is given in Table 4.4.

Table 4.4: Regression coefficients, with standard error, t and p values for a model predicting closure duration measured in ms, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to a pre-sonorant stop.

Term	Level	β	SE	t	p
(Intercept)		45.40	1.86	24.43	0.001
Environment	voiced obstruent	-1.72	1.37	-1.26	0.230
Environment	voiceless obstruent	-8.46	1.45	-5.82	0.001

A model predicting the duration of burst showed a significant main effect of the right-hand environment. The bursts were longer before sonorants than before voiceless obstruents ($p=0.001$), but there was no significant difference in burst duration between stops followed by sonorants and stops followed by voiced obstruents. In addition, dialect interacted with the underlying voicing. In the Poznań dialect underlying /d/s had shorter bursts than underlying /t/s, whereas in the Warsaw dialect the opposite effect was observed. In both cases the differences were smaller than

5 ms, and there was no significant main effect of dialect or underlying voicing. A summary of all the coefficients is in Table 4.5.

Table 4.5: Regression coefficients, with standard error, t and p values for a model predicting burst duration measured in ms, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to an underlying /d/ followed by a sonorant in the Poznań dialect.

Term	Level	β	SE	t	p
(Intercept)		25.06	1.80	13.92	0.001
Environment	voiced obstruent	-9.25	1.54	-5.99	0.001
Environment	voiceless obstruent	-2.41	1.63	-1.48	0.122
Dialect	Warsaw	4.17	2.04	2.04	0.080
C1	/t/	2.35	1.52	1.54	0.124
Dialect:C1	Warsaw:/t/	-3.50	1.56	-2.25	0.025

4.2.3 Summary and methodological remarks

The results confirm that there is a difference between the Poznań and Warsaw dialects with respect to phonetic voicing in pre-sonorant stops. Pre-sonorant stops in the Poznań dialect had significantly more voicing during closure compared to pre-sonorant stops produced by the Warsaw speakers, both in absolute terms and relative to closure duration. In the Poznań dialect, pre-sonorant stops also showed significantly more voicing than stops followed by voiceless obstruents. However, even though increased voicing was observed in Poznań pre-sonorant stops compared to the Warsaw dialect, and compared to stops in other environments, the results do not support the standard accounts of Poznań voicing in two regards.

First, Poznań pre-sonorant voicing in the current data is non-neutralising, as indicated by the effect of underlying voicing in the models of voicing duration and voicing ratio: underlying /d/ is more prone to voicing than underlying /t/. Non-neutralisation was dialect-specific and observed only for the Poznań speakers. This is of methodological importance as one could ask to what extent the surface phonetic contrast between underlying /t/ and /d/ might have been triggered by the use of orthographic stimuli (cf. the discussion on the neutralising status of final devoicing in Section 4.1.1). While it is difficult to assess the effect of a reading task without comparison of spontaneous speech, the dialectal difference suggests that the lack of word-final presonorant devoicing obtains because of something more than an aberrant task effect. Had the voicing contrast been induced by reading, one would expect to find non-neutralisation in the Poznań and the Warsaw dialects alike.

The non-neutralisation finding also warrants some further comment about how voicing duration and voicing ratio are distributed in underlyingly voiced pre-sonorant stops. From the models of voicing duration and voicing ratio (Table 4.1 and 4.2; Figure 4.1, 4.2) it follows that, while underlyingly voiced pre-sonorant stops in the Poznań dialect show significantly more voicing than stops followed by voiceless obstruents, they also show significantly less voicing than stops followed by voiced obstruents. As illustrated in Figure 4.2, stops followed by voiced obstruents typically surface as fully voiced (with the voicing ratio of 1) in the Poznań dialect, whereas the effect was smaller by ca. 0.27 in /d/s followed by a sonorant. This kind of effect could result from two different scenarios. One possibility is that pre-sonorant underlyingly voiced stops typically surface with partial voicing, i.e. the voicing duration is greater than in the case of stops followed by voiceless obstruents, but it does not amount to 100% of voicing during closure. The other possibility is that pre-sonorant /d/ optionally surfaces as fully voiced. The distribution of voicing duration and voicing ratio, illustrated in Figure 4.3, is more consistent with the second scenario. Both distributions are clearly bimodal in the case of underlying /d/, which means that underlying

/d/ tends to fall into one of two distinguishable categories. The first category is associated with limited partial voicing of about 20 ms duration, which yields stops with a voicing ratio of about 0.4. The second group is made of tokens with considerably longer voicing (ca. 40 ms), i.e. fully voiced stops. Most underlying /d/s surface as fully voiced.

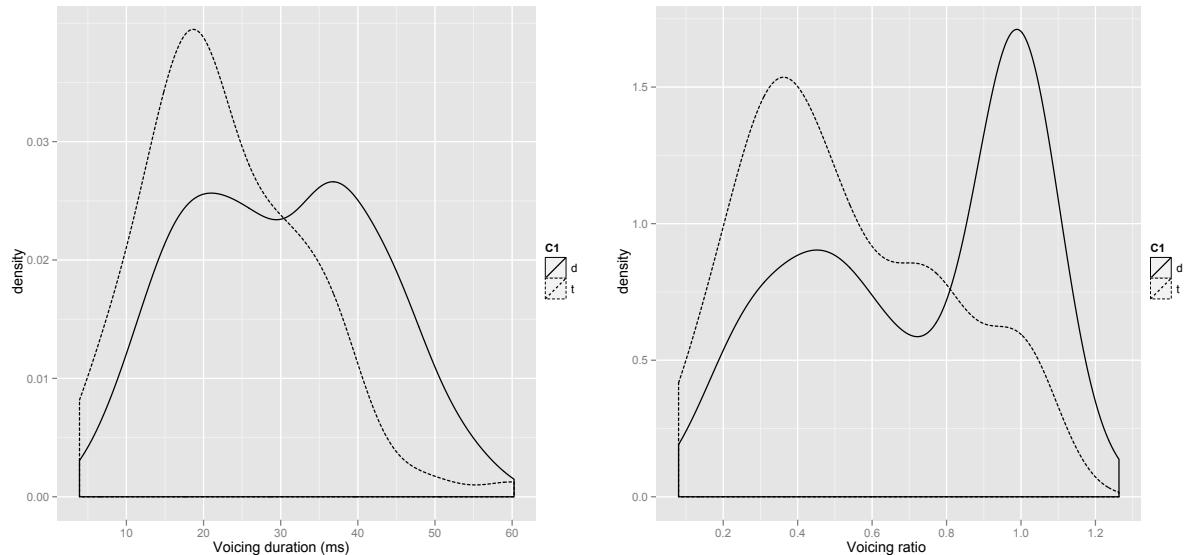


Figure 4.3: Density plots of voicing duration (left) and voicing ratio (right) in underlyingly voiced and underlyingly voiceless stops followed by a sonorant in the Poznań dialect

In sum, it appears that voicing of word-final pre-sonorant stops in the Poznań dialect is chiefly determined by the underlying voice specification of the stop. On the one hand, underlyingly voiceless stops typically surface with very little voicing, becoming phonetically indistinguishable from stops preceding voiceless obstruents. Underlyingly voiced stops, on the other hand, tend to surface with 100% of voicing during closure. In some cases, however, these stops undergo devoicing and surface with a very limited amount of vocal fold vibration. Consequently, the mean effect size is not entirely representative of the /t/-/d/ contrast observed in Poznań pre-sonorant stops. In some cases the contrast is neutralised, but when the contrast is maintained, it tends to be very robust.

Experiment 1 also delivered a surprising finding concerning f_0 , which showed significant lowering in vowels preceding stop+voiceless obstruent clusters, compared with vowels followed by stop+voiced obstruent, or stop+sonorant clusters. This is contrary to the prediction stated in Section 4.2.1 that f_0 lowering is expected preceding voiced sounds. This unexpected effect direction might suggest that the effect is incidental and conditioned by a confound in the experimental design. One potential source of such a confound is phrasal stress, which was not strictly controlled for, due to the use of different meaningful sentences as carrier phrases. Unfortunately, this might also have affected other measurements, especially vowel duration. As the potential confounding effects are hard to assess, the experiment was re-run with a design that controlled for phrasal stress.

4.3 Experiment 2. Obstruent#Sonorant sequences in Western and Central dialects of Polish.

The results of the pilot study confirm that there is a cross-dialectal difference between Poznań and Warsaw speech, which is related to the voicing of word-final pre-sonorant stops. However, the conclusions that we can draw from these results are somewhat tentative, given that phrasal

stress was not controlled for. In order to resolve this issue, another experiment was designed that strictly controlled for all the factors that might influence duration-measurements of relevance to voicing. The control was achieved by the use of invariant carrier phrase, as opposed to a set of different sentences used in the pilot experiment. In addition, data on word-final fricatives were collected as well as data on word-final stops. 16 native speakers of Polish participated in the experiment. 8 speakers came from Poznań and were living there at the time of the experiment. They were 4 females and 4 males, aged 21-53. The control group consisted of 8 speakers of central Polish dialects: 6 were female aged 28-53, and 2 were male aged 28 and 33. The purpose of the experiment was not explained to the speakers until after the recording. The participants were not paid.

4.3.1 Materials and method

The test items were words ending in coronal stops (/t/, /d/), or coronal fricatives (/s/, /z/) followed by a sonorant in the next word (/m/, /n/, /r/, /l/, /w/, /j/, /a/). All the items were genitive NPs where the NP head is a disyllabic noun ending in /u/+coronal stop (/t/ or /d/), and the complement is the genitive form of a trochaic proper name. The test items were embedded in a fixed carrier phrase, as exemplified in (19).

- (19) *To był dowód Marka, a nie dowód Nadii.*
'It was Marek's ID, not Nadia's ID.'

Each test item was included twice, but not in the same clause. The order of the clauses was reversed in the repetition in order to control for the effect of focus and associated stress. Thus, a repetition of the items exemplified in (19) would be as shown in (20).

- (20) *To był dowód Nadii, a nie dowód Marka.*
'It was Nadia's ID, not Marek's ID.'

In addition to the test tokens, sequences with following obstruents were also included. Labial obstruents (/p/, /b/, /f/, /v/) were used for easier segmentation. 2 repetitions of each baseline token were recorded. These baseline items followed the same template as the test items, and they were similarly varied with respect to focus.

The phonetic analysis followed the same procedure as in the pilot study. Spectrograms were labelled manually in Praat, and acoustic measurements were taken based on the inserted boundaries. The following measurements were made for pre-sonorant stops, following the same procedure as in the pilot study.

1. Duration of glottal pulsing into stop closure;
2. Stop closure duration;
3. Duration of the preceding vowel;
4. Duration of the burst;
5. f_0 at 10 ms and 20 ms before the end of the preceding vowel, as well as 10 and 20 ms following the onset of the following sound;
6. f_1 at 10 ms and 20 ms before the end of the preceding vowel, as well as 10 and 20 ms following the onset of the following sound.

For fricatives the following measurements were made.

1. Duration of glottal pulsing during the fricative. As with stops, increased glottal pulsing is expected to mark the surface voicing of a fricative.
2. Duration of the frication noise. Fricative duration has been shown to be an exponent of voicing in a number of languages, including Dutch (Slis & Cohen, 1969) and English (Crystal & House, 1988). Voiced fricatives tend to surface as shorter than voiceless fricatives. Stevens et al. (1992) have also found a duration effect on the perception of voice in fricatives, where shorter frication noise brings about the perception of voicing in listeners.
3. Duration of the preceding vowel.
4. f_0 at 10 ms and 20 ms before the end of the preceding vowel, as well as 10 and 20 ms following the onset of the following sound;
5. f_1 at 10 ms and 20 ms before the end of the preceding vowel, as well as 10 and 20 ms following the onset of the following sound;
6. Mean intensity of the fricative. The intensity measurement was normalised by being subtracted from the mean intensity of the low-frequency portion of the fricative (bandpass-filtered from 0 to 900Hz). The measurement was taken following Gradoville (2010), who found that it correlates with perceptions of gradient fricative voicing in Argentinian Spanish.

All duration measurements reported here are in milliseconds (ms). Frequency measurements are in Hertz (Hz), and intensity is measured in decibels (dB). For both stops and fricatives voicing ratio was calculated based on the duration of glottal pulsing into closure or into frication.

4.3.2 Results

As in the pilot study, statistical analysis was based on mixed-effects linear regression. In all the models the effects of speaker and item were treated as random. Model comparison was carried out by means of analysis whether individual predictors are significant or not, and whether or not models with additional predictors achieved a significantly better fit according to the log-likelihood test. Predictors that did not improve the fit of the model were discarded. Only significant predictors are presented in the model summaries throughout this section, unless otherwise indicated. Significance values were obtained based on MCMC sampling. For factors with more than two levels, the models were refitted to obtain significance values. These values are not included in the tables, but they are quoted in the discussion wherever relevant.

Realisation of voicing in stops

A series of mixed-effects linear models were fitted to the data, predicting the realisation of various voice-related acoustic parameters in stops. The predictors considered in the models included the right-hand environment (abbreviated as ‘Environment’ in the model summaries; sonorant, voiced obstruent, voiceless obstruent), dialect (Poznań vs. Central), the underlying voicing value of the first stop in the sequence (abbreviated as ‘C1’; /d/ vs. /t/), and the sex of the speaker (female vs. male). Only significant predictors are presented in the model summaries. The effects of speaker and item were treated as random in all the models.

Voicing duration The model predicting voicing duration in stops had a significant main effect of dialect and the right-hand environment, but no significant effect of underlying voicing. The model summary is shown in Table 4.6. The Poznań dialect had a higher voicing duration than the Central dialect ($p=0.001$). Voicing duration in stops followed by sonorants was greater than in stops followed by voiceless obstruents ($p=0.002$), but smaller than in stops followed by voiced obstruents

($p=0.001$). In addition, dialect interacted with the right-hand environment. The difference between voicing duration in the two dialects was greater within the pre-sonorant environment than when a voiced ($p=0.001$), or a voiceless ($p=0.004$) stop followed. In the Poznań dialect the voicing duration was greater before sonorants than before voiceless obstruents, but there was little difference between voicing duration before sonorants and voiced obstruents. In the Central dialect pre-sonorant stops were close to the stops followed by voiceless obstruents when it comes to voicing duration, with stops followed by voiced obstruents showing the largest voicing duration of all. No significant effect of the speaker sex was found.

Table 4.6: Regression coefficients, with standard error, t and p values for a model predicting the duration of voicing during closure. The intercept corresponds to the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		21.91	2.21	9.90	0.001
Dialect	Poznań	20.32	3.03	6.70	0.001
Environment	voiced obstruent	26.40	2.39	11.03	0.001
Environment	voiceless obstruent	-5.33	2.48	-2.15	0.032
Dialect: environment	Poznań: voiced obstruent	-27.11	2.99	-9.08	0.001
Dialect: environment	Poznań: voiceless obstruent	-8.34	2.99	-2.79	0.004

Voicing ratio The best-fitting model of voicing ratio had a significant main effect of dialect and the right-hand environment, but no significant main effect of underlying voicing. The model summary is presented in Table 4.7. The Poznań dialect had a higher voicing ratio than the Central dialect ($p=0.001$). Pre-sonorant stops had a significantly lower voicing ratio than stops followed by voiced obstruents ($p=0.001$), but they did not differ significantly from stops followed by voiceless obstruents. Dialect interacted with both underlying voicing and the right-hand environment. There was a significantly greater difference between the two dialects in pre-sonorant stops than in stops followed by voiced or voiceless obstruents ($p=0.001$ in both cases). Pre-sonorant stops in the Poznań dialect had a greater voicing ratio than pre-sonorant stops in the Central dialect. Poznań speakers also made more of a distinction in voicing ratio between underlying /t/ and underlying /d/ than Central dialect speakers ($p=0.004$). The right-hand environment did not interact with the underlying voicing of the stop. No significant effect of the speaker sex was found.

Table 4.7: Regression coefficients, with standard error, t and p values for a model predicting the voicing ratio. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		0.34	0.043	7.77	0.001
Dialect	Poznań	0.47	0.06	7.74	0.001
Environment	voiced obstruent	0.54	0.03	15.71	0.001
Environment	voiceless obstruent	-0.04	0.03	-1.24	0.232
C1	/t/	0.01	0.03	0.52	0.662
Dialect: environment	Poznań: voiced obstruent	-0.39	0.05	-8.21	0.001
Dialect: environment	Poznań: voiceless obstruent	-0.15	0.05	-3.24	0.002
Dialect:C1	Poznań:/t/	-0.10	0.04	-2.75	0.004

Vowel duration The model predicting the duration of the preceding vowel had three main effects and no interactions. There was a main effect of dialect, with increased vowel duration in Poznań, compared to the Central dialect ($p=0.001$). Vowels followed by stop+sonorant clusters were

significantly longer than vowels followed by stop+voiced obstruent ($p=0.001$), or stop+voiceless obstruent clusters ($p=0.052$). A significant main effect of underlying voicing was also observed: vowels preceding underlyingly voiceless stops were longer than vowels preceding underlyingly voiced stops ($p=0.001$). There was no significant effect of the speaker sex.

Table 4.8: Regression coefficients, with standard error, t and p values for a model predicting the vowel duration. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		62.31	2.97	20.99	0.001
Dialect	Poznań	-12.56	4.07	-3.09	0.001
Environment	voiced obstruent	-6.30	1.50	-4.20	0.001
Environment	voiceless obstruent	-2.92	1.56	-1.87	0.052
C1	/t/	4.61	1.47	3.13	0.001

f_0 and f_1 of the preceding vowel Underlying voicing of the stop, the right-hand environment and dialect were considered in the modelling of f_0 and f_1 at 10 and 20 ms before the offset of the preceding vowel. No significant effects were found in any of the models, except for the sex of the speaker. Male speakers had significantly lower f_0 and f_1 than female speakers.

Table 4.9: Regression coefficients, with standard error, t and p values for a model predicting f_0 of a preceding vowel, measured in Hz, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to a female speaker.

f_0 at 10 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		197.30	3.95	49.94	0.001
Sex	male	-71.16	6.37	-11.16	0.001
f_0 at 20 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		198.81	3.87	51.41	0.001
Sex	male	-71.48	6.26	-11.41	0.001

Table 4.10: Regression coefficients, with standard error, t and p values for a model predicting f_1 of a preceding vowel, measured in Hz, in stops followed by sonorants, as compared with stops followed by obstruents. The intercept corresponds to a female speaker.

f_1 at 10 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		343.09	7.37	46.54	0.001
Sex	male	-24.76	11.92	-2.08	0.012
f_1 at 20 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		356.93	7.43	48.01	0.001
Sex	male	-30.81	11.97	-2.57	0.008

f_0 and f_1 of the following segment were not considered in the modelling since these measurements cannot be systematically compared across manners of articulation, especially voiceless obstruents.

Closure and burst duration In the model predicting closure duration dialect and the right-hand environment emerged as significant predictors, and a significant interaction between the two

was also observed. The model summary is in Table 4.11. Closure duration was significantly shorter in the Poznań dialect than in the Central dialects ($p=0.01$) in all three environments, but the inter-dialectal difference was significantly greater in the pre-sonorant environment in comparison to the pre-voiceless obstruent environment ($p=0.012$). There was also a main effect of the right-hand environment: closure duration in pre-sonorant stops was longer than in pre-obstruent stops (this effect was significant at the 0.001 level both for voiced and for voiceless obstruents). There was no significant effect of the underlying voicing, or the sex of the speaker.

Table 4.11: Regression coefficients, with standard error, t and p values for a model predicting the closure duration. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		69.76	3.40	20.55	0.001
Dialect	Poznań	-11.81	4.38	-2.69	0.010
Environment	voiced obstruent	-13.82	3.43	-4.03	0.001
Environment	voiceless obstruent	-13.35	3.78	-3.54	0.001
Dialect: environment	Poznań: voiced obstruent	1.49	3.02	0.49	0.592
Dialect: environment	Poznań: voiceless obstruent	7.81	3.03	2.58	0.012

The model predicting burst duration (Table 4.12) had a significant main effect of dialect and the right-hand environment. The Poznań speakers showed on average a shorter burst duration than the Central dialect speakers ($p=0.004$). Bursts were longer preceding sonorants than preceding voiced or voiceless obstruents (voiced obstruents: $p=0.012$; voiceless obstruents: $p=0.046$). In addition to these main effects, dialect interacted with underlying voicing and the right-hand environment. In the Central dialect /d/s surfaced with relatively longer bursts than /t/s, whereas in Poznań the reverse was found. The difference in burst duration between the two dialects was greater in the pre-sonorant environment than preceding a voiced obstruent ($p=0.001$) or a voiceless obstruent ($p=0.001$). Finally, there was a significant three-way interaction between dialect, underlying voicing and the right-hand environment. The interaction is illustrated in Figure 4.4. In the Poznań dialect underlying /d/s generally had longer bursts than underlying /t/s, except in the pre-sonorant environment, where the reverse was found. In the Central dialect there was the same general tendency for /d/ to surface with a longer burst than /t/s, except when a voiceless obstruent followed.

Table 4.12: Regression coefficients, with standard error, t and p values for a model predicting burst duration. The intercept corresponds to an underlying voiced stop (/d/) followed by a sonorant in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		26.19	2.35	11.15	0.001
Dialect	Poznań	-10.51	3.04	-3.46	0.004
Environment	voiced obstruent	-7.16	2.75	-2.60	0.012
Environment	voiceless obstruent	-6.17	3.03	-2.04	0.046
C1	/t/	-1.79	2.01	-0.89	0.034
Dialect:C1	Poznań:/t/	4.71	2.11	2.23	0.026
Dialect: environment	Poznań: voiced obstruent	8.99	2.98	3.02	0.001
Dialect: environment	Poznań: voiceless obstruent	13.22	3.03	4.37	0.001
Environment: C1	voiced obstruent:/t/	-0.66	3.98	-0.17	0.886
Environment: C1	voiceless obstruent:/t/	3.48	4.30	0.81	0.398
Dialect:environment:C1	Poznań:voiced obstruent:/t/	-5.91	4.31	-1.37	0.148
Dialect:environment:C1	Poznań:voiceless obstruent:/t/	-9.19	4.31	-2.13	0.033

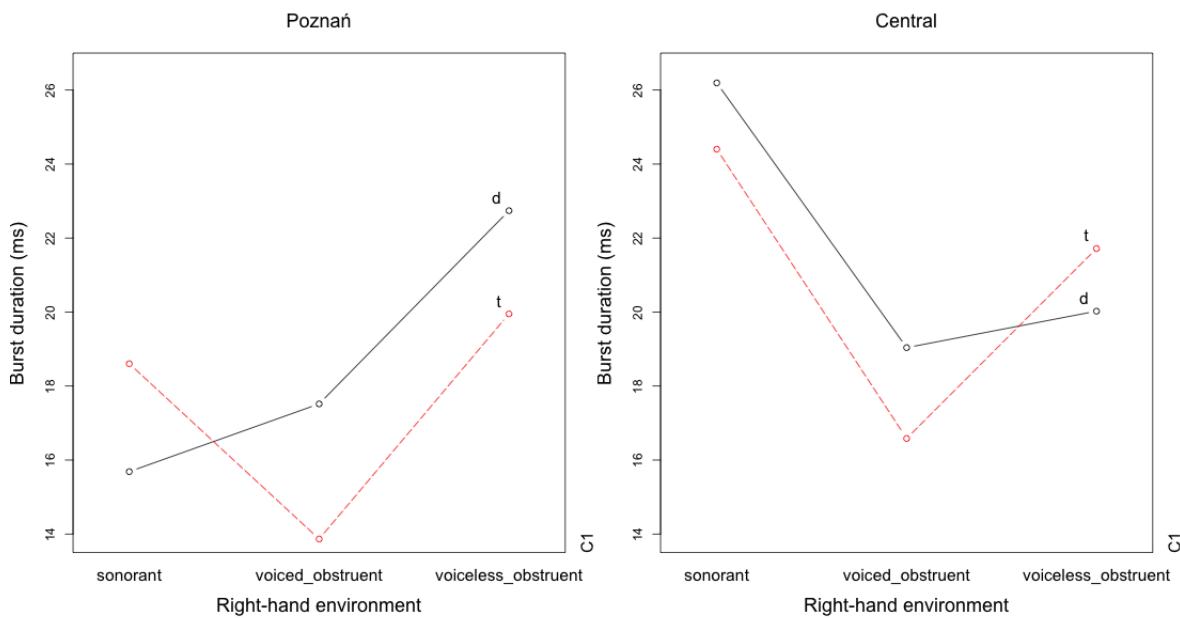


Figure 4.4: Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting burst duration.

Realisation of voicing in fricatives

A series of mixed-effects linear models were fitted to the data, predicting the realisation of various voice-related acoustic parameters in fricatives. The predictors considered in the models included the right-hand environment (abbreviated as ‘Environment’ in the model summaries; sonorant, voiced obstruent, voiceless obstruent), dialect (Poznań vs. Central), the underlying voicing value of the first fricative in the sequence (abbreviated as ‘C1’; /s/ vs. /z/), and the sex of the speaker (Sex; female vs. male). The effect of speaker and item were treated as random in all the models.

Voicing duration The model predicting the duration of voicing during frication had three significant main effects. On average greater voicing duration was found in the Poznań dialect than in the Central dialect ($p=0.001$); and likewise, duration of voicing was longer preceding voiced obstruents than preceding sonorants ($p=0.001$). There was also a significant main effect of the speaker sex, with greater voicing duration in male than in female speakers ($p=0.010$). No significant main effect of the underlying voicing was found. Dialect interacted with the underlying voicing and the right-hand environment. The /s/-/z/ contrast was significantly greater with regard to voicing duration in the Poznań dialect than in the Central dialect ($p=0.001$). Voicing duration in the two dialects also differed more in the pre-sonorant environment than before voiced ($p=0.004$), or voiceless obstruents ($p=0.004$). In addition there was a significant three-way interaction between dialect, right-hand environment and the underlying voicing. The interaction is plotted in Figure 4.5. The figure shows that increased voicing duration was found in fricatives followed by a sonorant in the Poznań dialect, but only for underlying /z/s. Underlying /s/s showed a similar voicing duration before sonorants and before voiceless stops. In the Central dialect the voicing duration was similar before sonorants and before voiceless obstruents for underlying /z/ and /s/ alike. The model summary is given in Table 4.13.

Voicing ratio The model predicting voicing ratio produced similar results to the model of voicing duration. There were significant main effects of dialect, right-hand environment and speaker sex.

Table 4.13: Regression coefficients, with standard error, t and p values for a model predicting the duration of voicing during the fricative. The intercept corresponds to an underlyingly voiced stop (/z/) followed by a sonorant produced by a male speaker in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		14.22	3.37	4.212	0.001
Dialect	Poznań	18.83	4.33	4.35	0.001
Environment	voiced obstruent	43.10	5.99	7.19	0.001
Environment	voiceless obstruent	-6.95	4.13	-1.68	0.094
C1	/s/	-1.60	3.06	-0.53	0.560
Sex	male	11.27	3.93	2.86	0.010
Dialect:C1	Poznań:/s/	-13.55	3.54	-3.83	0.001
Dialect: environment	Poznań: voiced obstruent	-23.98	6.84	-3.50	0.004
Dialect: environment	Poznań: voiceless obstruent	-13.37	4.57	-2.93	0.004
Environment: C1	voiced obstruent:/s/	-6.96	7.17	-0.97	0.302
Environment: C1	voiceless obstruent:/s/	4.11	7.28	0.57	0.552
Dialect:environment:C1	Poznań:voiced obstruent:/s/	20.40	8.20	2.49	0.016
Dialect:environment:C1	Poznań:voiceless obstruent:/s/	14.31	8.24	1.74	0.090

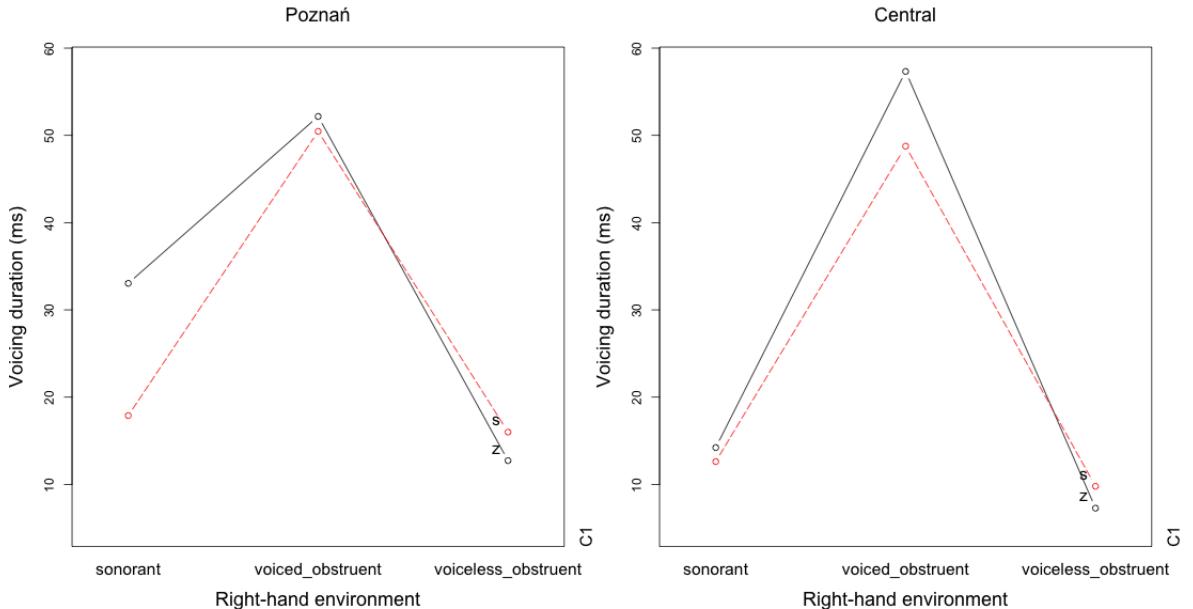


Figure 4.5: Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting voicing duration in fricatives.

The Poznań dialect had on the whole a higher voicing ratio than the Central dialect ($p=0.001$). Significantly more voicing was found before voiced obstruents than before sonorants ($p=0.001$), but the difference between pre-sonorant fricatives and fricatives followed by voiceless obstruents was not significant. Male speakers had higher voicing ratios than female speakers ($p=0.004$). Dialect interacted with class and underlying voicing, as the Poznań dialect had significantly higher voicing ratio than the Central dialect in the context of a following sonorant ($p=0.001$), but not in any other environment. Underlying /z/s had a significantly higher voicing ratio than underlying /s/s in the Poznań dialect ($p =0.001$), but no significant difference in voicing ratio depending on the underlying voicing of the fricative was observed for the Central dialect. There was also a significant three-way interaction between dialect, right-hand environment and underlying voicing, illustrated in Figure 4.6. There was increased voicing ratio in pre-sonorant /z/s in the Poznań dialect compared to underlying /s/s, or fricatives followed by voiceless obstruents. The ratio for pre-sonorant /z/s in

Poznań was slightly above 0.4, compared to ca. 0.7 for fricatives followed by voiced obstruents. In the Central dialect the voicing ratio was high (ca. 0.7) in fricatives followed by voiced obstruents, but fricatives followed by sonorants or voiceless obstruents had low voicing ratios (around 0.1). The summary of the model is shown in Table 4.14.

Table 4.14: Regression coefficients, with standard error, t and p values for a model predicting the voicing ratio in fricatives. The intercept corresponds to an underlyingly voiced stop (/z/) followed by a sonorant produced by a male speaker in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		0.14	0.05	2.89	0.002
Dialect	Poznań	0.27	0.07	3.91	0.001
Environment	voiced obstruent	0.56	0.07	8.41	0.001
Environment	voiceless obstruent	-0.08	0.04	-1.73	0.078
C1	/s/	-0.005	0.03	-0.15	0.884
Sex	male	0.17	0.06	2.74	0.004
Dialect:C1	Poznań:/s/	-0.18	0.05	-3.70	0.002
Dialect: environment	Poznań: voiced obstruent	-0.23	0.09	-2.48	0.016
Dialect: environment	Poznań: voiceless obstruent	-0.21	0.06	-3.39	0.001
Environment: C1	voiced obstruent:/s/	-0.07	0.08	-0.91	0.354
Environment: C1	voiceless obstruent:/s/	0.04	0.08	0.53	0.606
Dialect:environment:C1	Poznań:voiced obstruent:/s/	0.24	0.11	2.20	0.022
Dialect:environment:C1	Poznań:voiceless obstruent:/s/	0.23	0.11	2.10	0.052

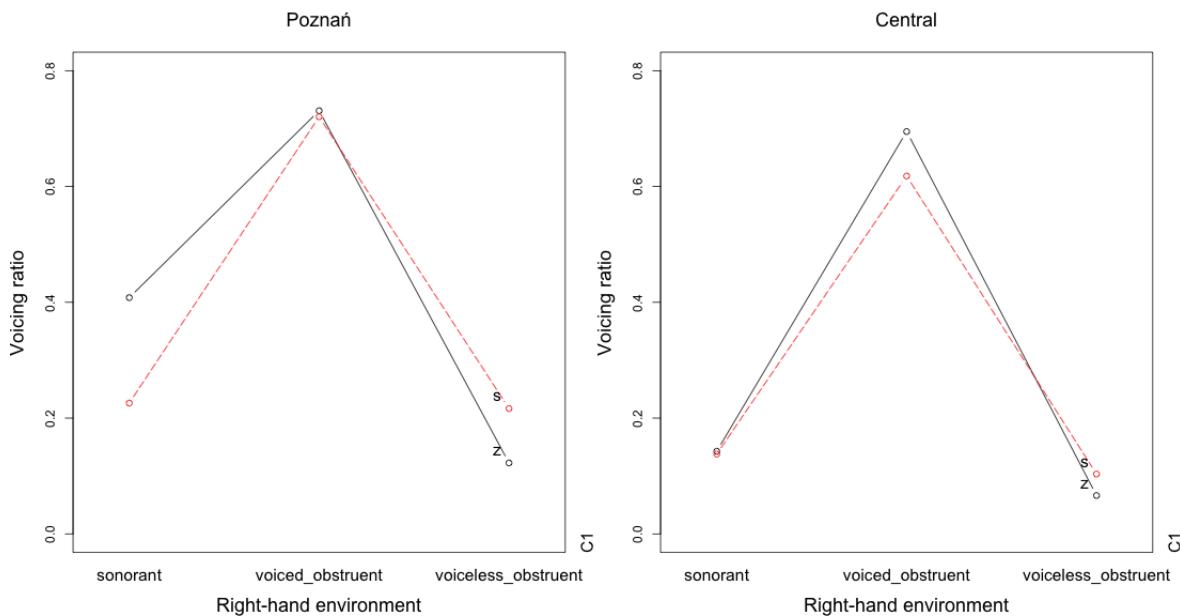


Figure 4.6: Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting voicing ratio in fricatives.

Vowel duration The model predicting the duration of the preceding vowel had two significant main effects, but no significant interactions. The model summary is presented in Table 4.15. Vowels followed by fricative+sonorant clusters were significantly longer than vowels followed by fricative+voiceless obstruent clusters ($p=0.046$), but they did not differ significantly from vowels followed by fricative+voiced obstruent clusters. The difference in preceding vowel duration between fricatives followed by voiced and voiceless obstruents was significant at $p=0.001$. There was also

a significant main effect of the underlying voicing of the fricative. Vowels preceding underlyingly voiced fricatives were significantly longer than vowels preceding underlyingly voiceless fricatives, with an effect size of ca. 7ms ($p=0.001$).

Table 4.15: Regression coefficients, with standard error, t and p values for a model predicting the duration of a vowel preceding the fricative. The intercept corresponds to an underlyingly voiced stop (/z/) followed by a sonorant in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		71.08	2.56	27.76	0.001
Environment	voiced obstruent	1.25	2.30	0.54	0.564
Environment	voiceless obstruent	-4.80	2.45	-1.96	0.046
C1	/s/	-7.28	1.77	-4.12	0.001

f₀ and f₁ of the preceding vowel The models of f₀ of the preceding vowel had only one significant main effect, namely that of speaker sex. Male speakers had a significantly lower f₀ than female speakers, both at 10 and 20 ms before the offset of the vowel. In both of the models the effect was significant at $p=0.001$. The models are summarised in Table 4.16.

Table 4.16: Regression coefficients, with standard error, t and p values for a model predicting f₀ of a vowel preceding the fricative. The intercept corresponds to a female speaker.

f ₀ at 10 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		196.29	3.49	56.31	0.001
Sex	male	-73.34	5.70	-12.88	0.001
f ₀ at 20 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		199.06	3.56	55.97	0.001
Sex	male	-73.47	5.81	-12.64	0.001

The models of f₁ also showed a significant lowering effect in male speakers which is significant at $p=0.001$, but only at 10 ms before the offset. At 20 ms before the offset, male and female speakers did not differ significantly. Similarly, the effect of the right-hand environment was only significant in the case of f₁ measured at 10 ms before the offset, with f₁ lowering in vowels followed by fricative+voiced obstruent clusters, compared with vowels followed by fricative+sonorant clusters ($p=0.048$). At 20 ms before the offset this effect was not significant ($p=0.174$). The two models of f₁ are summarised in Table 4.17. The non-significant predictors were retained in the model in the case of f₁ measured 20 ms before the offset for the purpose of comparison with the other f₁ model.

Fricative duration The model predicting the duration of the fricative had two significant main effects: one for underlying voicing and another for the right-hand environment. The model summary is shown in Table 4.18. Pre-sonorant fricatives were significantly longer than fricatives followed by voiced obstruents ($p=0.001$). There were no significant effects of underlying voicing, dialect or speaker sex.

Low-frequency intensity - mean intensity difference A mixed-effects linear model was fitted to the data from fricatives, predicting the difference between mean intensity of the fricatives' low frequency portion (0-900Hz) and the mean intensity of the fricative, measured in dB. The mean intensity of the fricative was higher than the mean intensity of the filtered portion, hence

Table 4.17: Regression coefficients, with standard error, t and p values for a model predicting f_1 of a vowel preceding the fricative. The intercept corresponds to a female speaker.

f_1 at 10 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		307.27	13.12	23.41	0.001
Sex	male	67.70	17.81	3.80	0.001
Environment	voiced obstruent	-45.98	22.55	-2.04	0.048
Environment	voiceless obstruent	6.90	23.32	0.30	0.722

f_1 at 20 ms before the offset					
Term	Level	β	SE	t	p
(Intercept)		321.66	9.30	34.61	0.001
Sex	male	12.58	13.78	0.91	0.380
Environment	voiced obstruent	-18.48	13.44	-1.37	0.174
Environment	voiceless obstruent	11.07	13.61	0.81	0.468

Table 4.18: Regression coefficients, with standard error, t and p values for a model predicting duration of the fricative, measured in ms. The intercept corresponds to an underlyingly voiced pre-sonorant fricative in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		94.11	3.74	25.18	0.001
Environment	voiced obstruent	-18.09	4.39	-4.12	0.001
Environment	voiceless obstruent	-6.91	4.70	-1.47	0.114

all the values are negative. The following discussion will refer to absolute values, so a difference of -3 dB is considered smaller than a difference of -7 dB. Relatively smaller differences occur where there is high intensity of the fricative's low frequency portion, i.e. relatively smaller differences are associated with the presence of voicing.

The model of intensity difference is summarised in Table 4.19. The model had a significant main effect of the right-hand environment, with a smaller intensity difference before sonorants than before voiced obstruents ($p=0.001$). The intensity difference before sonorants was not significantly greater than before voiceless obstruents. Dialect interacted with the right-hand environment, with a greater difference between the two dialects preceding voiceless obstruents than in other environments. There was also an interaction between dialect and underlying voicing, as Poznań speakers had a greater /s/-/z/ contrast than Central speakers. A three-way interaction was found between dialect, right-hand environment and underlying voicing, as illustrated in Figure 4.7. There was again an effect on Poznań pre-sonorant fricatives that was also sensitive to the underlying voicing; underlying /z/s in Poznań had a smaller intensity difference before sonorants than /s/s in the same position, or /z/s followed by a voiceless obstruent. However, the intensity difference was smaller still in fricatives followed by voiced obstruents. In the Central dialect pre-sonorant fricatives patterned more closely with fricatives followed by voiceless obstruents, and no /s/-/z/ contrast was maintained before a sonorant in that dialect.

4.3.3 Summary

The results concerning voicing in stops delivered by experiment 2 are different than in experiment 1. There was again evidence for increased duration of vocal fold vibration in the Poznań dialect before sonorants, but unlike in experiment 1 this effect did not interact with underlying voicing. This means that increased voicing was observed in underlying /t/s and /d/s alike. The distribution of voicing duration and voicing ratio depending on the underlying voicing of the stop

Table 4.19: Regression coefficients, with standard error, t and p values for a model predicting the difference between mean intensity of the fricative's low frequency portion (0-900Hz) and mean intensity of the fricative (measured in dB). The intercept corresponds to an underlyingly voiced pre-sonorant fricative in the Central dialect.

Term	Level	β	SE	t	p
(Intercept)		-5.83	0.89	-6.52	0.001
Environment	voiced obstruent	2.90	0.97	2.99	0.004
Environment	voiceless obstruent	-0.49	0.67	-0.73	0.432
Dialect	Poznań	-0.23	1.23	-0.19	0.81
C1	/s/	0.424	0.49	0.86	0.38
Dialect: environment	Poznań: voiced obstruent	0.70	1.06	0.66	0.526
Dialect: environment	Poznań: voiceless obstruent	-1.76	0.71	-2.48	0.020
Environment: C1	voiced obstruent:/s/	0.13	1.16	0.12	0.856
Environment: C1	voiceless obstruent:/s/	0.51	1.18	0.44	0.644
Dialect:C1	Poznań:/s/	-2.35	0.55	-4.27	0.001
Dialect:environment:C1	Poznań:voiced obstruent:/s/	1.71	1.27	1.35	0.188
Dialect:environment:C1	Poznań:voiceless obstruent:/s/	2.73	1.28	2.14	0.034

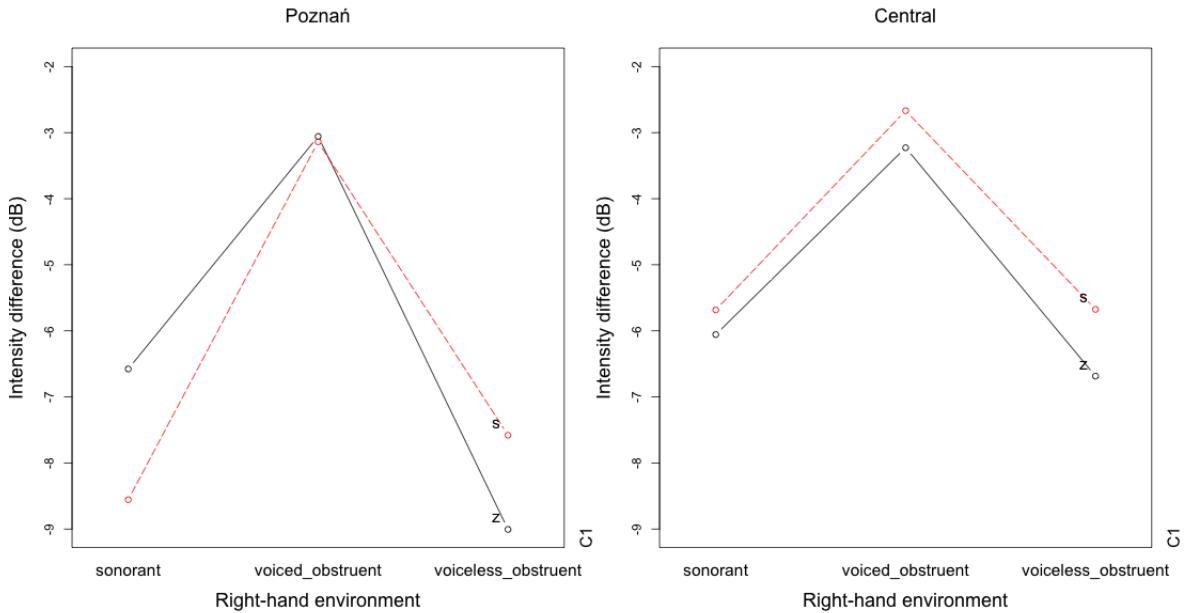


Figure 4.7: Interaction between right-hand environment and underlying voicing within the two dialects in a model predicting intensity difference in fricatives.

is shown in Figure 4.8. Two clear modes can be distinguished in the distribution of voicing ratio for both underlying /d/s and underlying /t/s. The first peak in both distributions is at 0.5-0.6 voicing to closure ratio, whereas the second peak is associated with full voicing during closure.

Dialect did not interact with underlying voicing in the models predicting voicing, closure or burst duration. There was a significant main effect of dialect in a number of models, with the Poznań dialect showing overall greater voicing duration and voicing ratio than the Central dialect, as well as shorter closure and shorter burst. All of these predictors are consistent with the generalisation that there is relatively more phonetic voicing in the Poznań dialect. For some of the models dialect interacted with the right-hand environment. The models of voicing duration and voicing ratio in particular show the presence of increased stop voicing in the pre-sonorant position in Poznań. At the same time, pre-sonorant stops in Poznań had a significantly smaller voicing ratio than stops followed by voiceless obstruents.

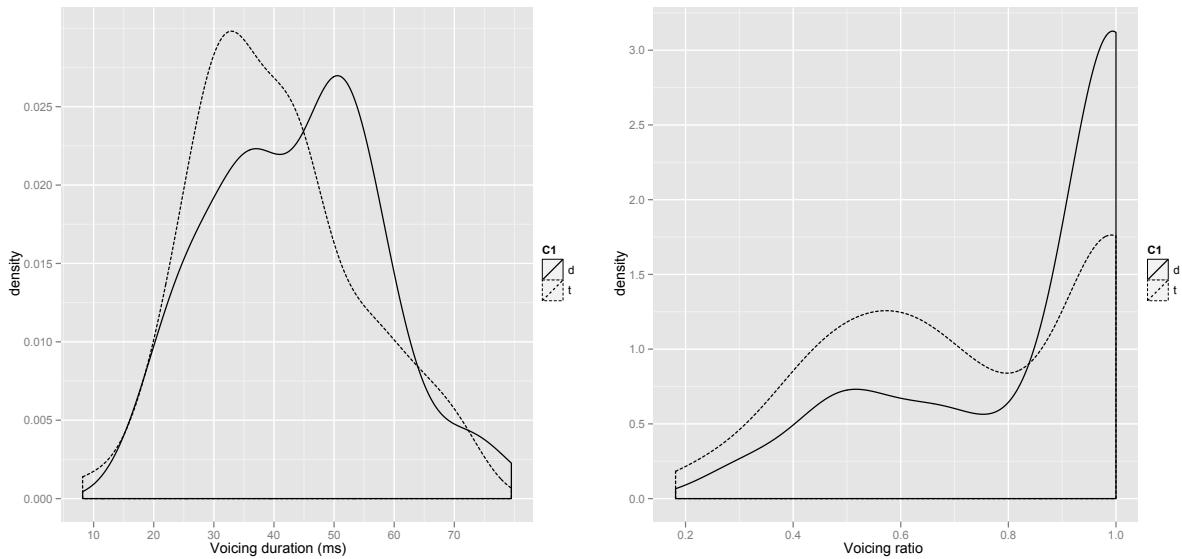


Figure 4.8: Density plots of voicing duration (left) and voicing ratio (right) in underlyingly voiced and underlyingly voiceless stops followed by a sonorant in the Poznań dialect. Data from experiment 2

The data from fricatives, on the other hand, show non-neutralising pre-sonorant voicing. Models of voicing duration and voicing ratio both showed increased voicing in the pre-sonorant position in the Poznań dialect. In the model of voicing ratio there was a significant three-way interaction between dialect, right-hand environment and the underlying voicing; voicing ratio was higher for pre-sonorant /z/s compared with pre-sonorant /s/s in Poznań. The voicing distributions are illustrated in Figure 4.9. Clear bimodality can be observed for underlyingly voiced, as well underlyingly voiceless fricatives. Underlyingly voiced fricatives typically surfaced as fully voiced, while underlyingly voiceless fricatives typically surfaced as partially voiced.

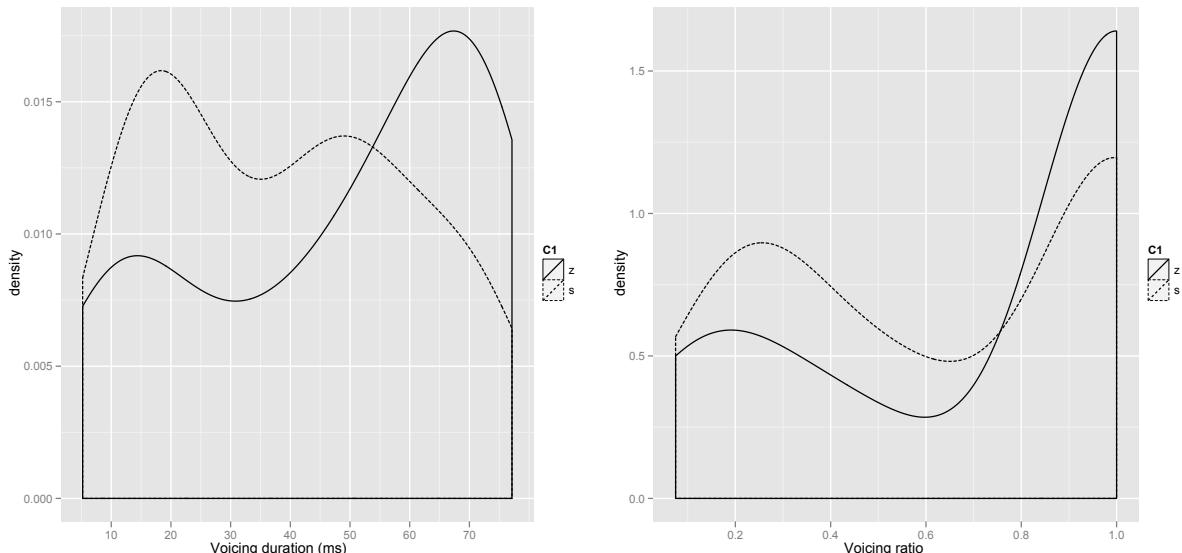


Figure 4.9: Density plots of voicing duration (left) and voicing ratio (right) in underlyingly voiced and underlyingly voiceless fricatives followed by a sonorant in the Poznań dialect. Data from experiment 2

A three-way interaction between dialect, underlying voicing and the right-hand environment

was also found in a model predicting the difference between mean intensity of the fricative's low portion and the mean intensity of the fricative. Relatively larger differences were found in fricatives followed by voiced obstruents than in fricatives followed by voiceless obstruents. In the Central dialect pre-sonorant fricatives patterned with fricatives followed by voiceless stops and there was no effect of underlying voicing. Within the Poznań dialect pre-sonorant underlying /z/s showed a smaller difference, compared with pre-sonorant underlying /s/s.

The effect of f_0 lowering preceding voiceless stops, previously observed in experiment 1, was not replicated in experiment 2. However, experiment 2 showed two non-dialect-specific near-neutralisation effects. Both concerned vowel duration: vowels were 4.61 ms longer preceding underlying /t/s compared to underlying /d/s, and 7.28 ms shorter preceding underlying /s/s compared to underlying /z/s.

4.4 Discussion

The combined results of the two experiments confirm that there is a dialectal difference between Poznań Polish and the Warsaw/Central dialect with respect to the realisation of voicing in word-final obstruents followed by a sonorant in the next word. Phonetically, the difference consists mostly in increased duration of voicing, both in absolute terms and relative to the obstruent's duration, as indicated by the existence of significant interactions between dialect and the right-hand environment in models predicting voicing duration and voicing ratio in stops and fricatives. However, even though the results confirm the existence of increased pre-sonorant voicing in the Poznań dialect, the data deliver two important pieces of evidence that have not formed a part of the standard description of Poznań voicing. First, pre-sonorant voicing is sensitive to the underlying voicing of the obstruent. Second, pre-sonorant voicing in Poznań is highly variable, unlike voice assimilation to obstruents.

4.4.1 The role of delaryngealisation

Some of the variation in pre-sonorant voicing has been found to be conditioned by the underlying voicing of the obstruent. The difference was very robust in the data from experiment 1, where underlyingly voiced stops varied between voiced and voiceless realisations, whereas underlyingly voiceless stops typically showed only partial voicing with a voicing ratio in the range of around 0.4. In comparison, considerably more voicing was found for underlyingly voiceless stops in the data from experiment 2, where a clear bimodal distribution of voicing ratio was observed for both underlying /t/ and underlying /d/ (cf. Figure 4.8): this included a peak for both categories at the ratio of 1. Consequently, underlyingly voiced and voiceless stops were neutralised before sonorants in the experiment 2 data; even though there was some difference with respect to how often underlyingly voiced and voiceless stops surfaced as fully voiced, the underlying voicing was not a significant predictor of voicing ratio within stops in the Poznań dialect. Data from pre-sonorant fricatives, on the other hand, showed a clear non-neutralisation pattern, with significant effects for underlying voicing in pre-sonorant fricatives pronounced by Poznań speakers emerging from models of voicing duration, voicing ratio and low-mean frequency intensity.

The results concerning the non-neutralisation of the underlying voicing contrast in Poznań Polish are accompanied by findings of near-neutralisation in both dialects. While no near-neutralisation effects were found in the pilot study, some effects emerge in the data from experiment 2. Near-neutralisation was characterised by small-scale differences in the duration of a preceding vowel depending on the underlying voicing of the following obstruent. Interestingly, the direction of the effect was not stable across the different samples. Cross-linguistically the presence of voicing is associated with an increased duration of the preceding vowel, and this kind of effect was also

observed in the data from word-final pre-sonorant fricatives (experiment 2). However, for word-final pre-sonorant stops in experiment 2 the opposite effect was found, with increased vowel duration preceding underlyingly voiceless stops. As the potential prosodic confounds were controlled for in experiment 2, the unexpected direction of the vowel lengthening preceding stops is largely unexplained, and it remains to be seen whether or not it can be replicated in future studies. Meanwhile, in view of the general controversy surrounding the significance of near-neutralisation, I shall refrain from making theoretical arguments based on any near-neutralisation effects. Instead, I focus here on discussing the non-neutralisation patterns in the current data.

The non-neutralisation patterns found in the current data were restricted to word-final pre-sonorant position, which is also the environment for pre-sonorant voicing. The co-occurrence of non-neutralisation and pre-sonorant voicing in the same position makes Poznań Polish unique in the typology of pre-sonorant voicing languages. Crucially, it seems to go against the predictions of Jansen (2004) concerning how pre-sonorant voicing may develop diachronically (cf. 3.4). In Jansen's theory obstruents may undergo intersonorant voicing only following the loss of their own laryngeal targets. A similar intuition is expressed in synchronic generative analyses of Kraków-Poznań voicing (Gussmann, 1992; Rubach, 1996; Cyran, 2012) summarised in Section 4.1.1 of this chapter. The shared proposal of these analyses is that word-final delaryngealisation explains the positional restrictions on pre-sonorant voicing which applies only in word-final obstruents. Word-final obstruents are said to undergo delaryngealisation, which makes them potential targets for pre-sonorant voicing, whereas word-medial obstruents are able to license their own laryngeal node, which prevents spreading. An important aspect of this proposal is that it needs to distinguish between voiceless and delaryngealised obstruents in order to capture the difference in the way word-medial and word-final voiceless obstruents behave.

There is, however, a caveat when it comes to the significance of the non-neutralisation findings. While the current data do not support the view that the voicing contrast is neutralised word-finally in Poznań Polish, they are not consistent with a gradient view of final devoicing either, as argued for instance by Dinnsen & Charles-Luce (1984) for Catalan. Based on findings of incomplete neutralisation, Dinnsen and Charles-Luce propose that final devoicing is not a phonological process, but that it consists in a gradual loss of voicing cues in word-final obstruents. In the case of Poznań Polish, however, a strictly phonetic analysis of final devoicing does not predict the bimodality observed in the distribution of voicing ratio for underlyingly voiced obstruents followed by a sonorant in the next word. This bimodality is more consistent with an interpretation whereby final devoicing in Poznań Polish involves a category change, but it is nevertheless optional.

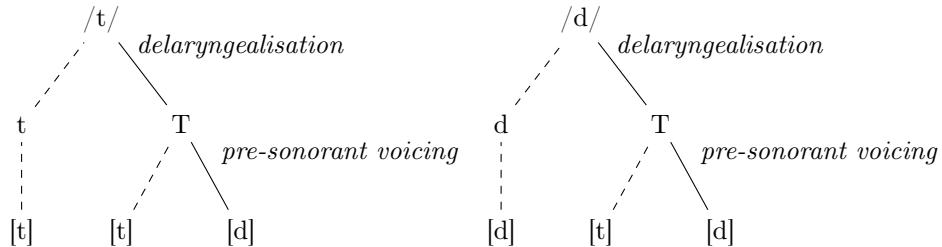
The finding that the distribution of voicing ratio values patterns in a bimodal fashion in underlying voiceless pre-sonorant obstruents also supports the view that presonorant voicing is categorical but optional in Poznań Polish. While surface voicing in underlyingly voiced obstruents can be explained by an optional suspension of final devoicing, the presence of extended glottal pulsing in the phonetic output of underlyingly voiceless obstruents suggests the activity of a voicing process. As two surface categories are discernible with respect to voicing ratio, it again appears that pre-sonorant voicing applies optionally, but nevertheless categorically.

The kind of variation observed in Poznań data, i.e. optional but categorical voicing coupled with the recoverability of underlying voicing in some tokens could be explained in two ways. The first possibility is within a model like the one proposed by Ernestus & Baayen (2006). Ernestus & Baayen argue for a view in which the lexical representation is informed by analogy between inflectionally related words. In this kind of framework, the realisation of word-final pre-sonorant obstruents would be influenced by a generalisation that voicing may occur optionally in this position. However, at the same time, the realisation would also be influenced by analogy between lexically related forms, including those members of the paradigm where voicing is contrastive. Stochastic activation of such forms coupled with a ‘voicing rule’ could yield a mixture of near-

neutralisation, non-neutralisation and pre-sonorant voicing.

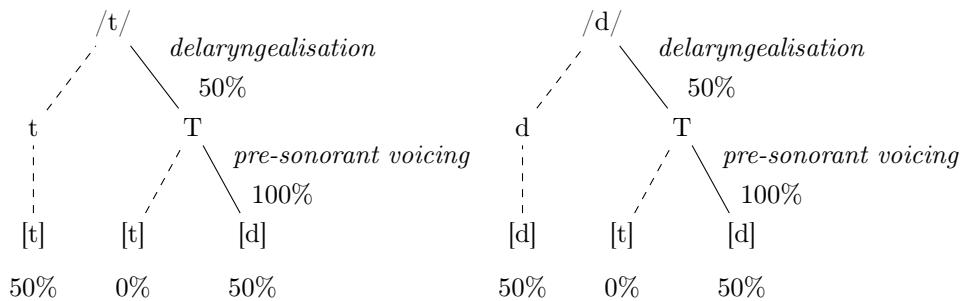
The second option is a stochastic feed-forward model in which delaryngealisation and pre-sonorant voicing are indeed in a feeding relationship, but they are both optional. Let us consider the model in (21), which illustrates a feeding relationship between delaryngealisation and pre-sonorant voicing. The ordering could be interpreted either as diachronic succession, or a synchronic derivation. Solid lines indicate results of the application of a given process, whereas dashed lines indicate what happens when a process fails to apply.

(21)



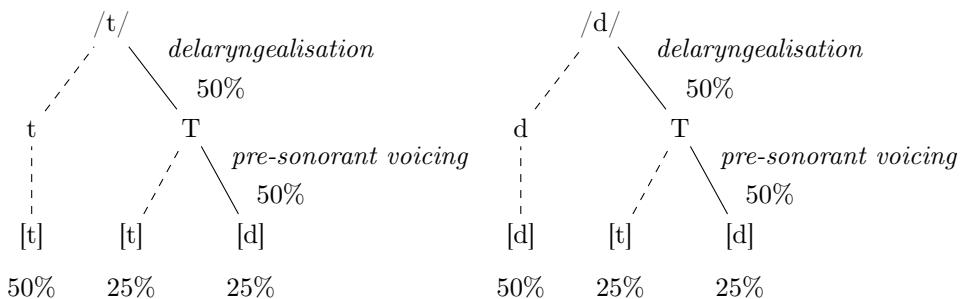
If both processes applied 100% of the time, the outcome would always be a [d] output regardless of the input, which is the prediction of previous analyses, but it is not borne out by all of the current data. Let us now consider what happens if only delaryngealisation applies optionally. An example is schematised in (22) with delaryngealisation applying 50% of the time.

(22)



For an underlying /t/ the output is a mixture of [t] and [d] forms. For an underlying /d/, however, the model can never produce a [t] output, as lack of delaryngealisation and categorical pre-sonorant voicing of delaryngealised stops both produce voiced outputs. Thus, in order to capture the fact that underlying /d/ does sometimes surface as voiceless before sonorants in Poznań Polish, both delaryngealisation and pre-sonorant voicing must be optional. For instance, if both processes applied 50% of the time, this would return a 75% of [t] realisations for an underlying /t/, and 75% [d] realisations for an underlying /d/, as shown in (23).

(23)



An interesting prediction that follows from the model is that the output will match the underlying voicing of the input more often than it follows from the frequency with which delaryngealisation

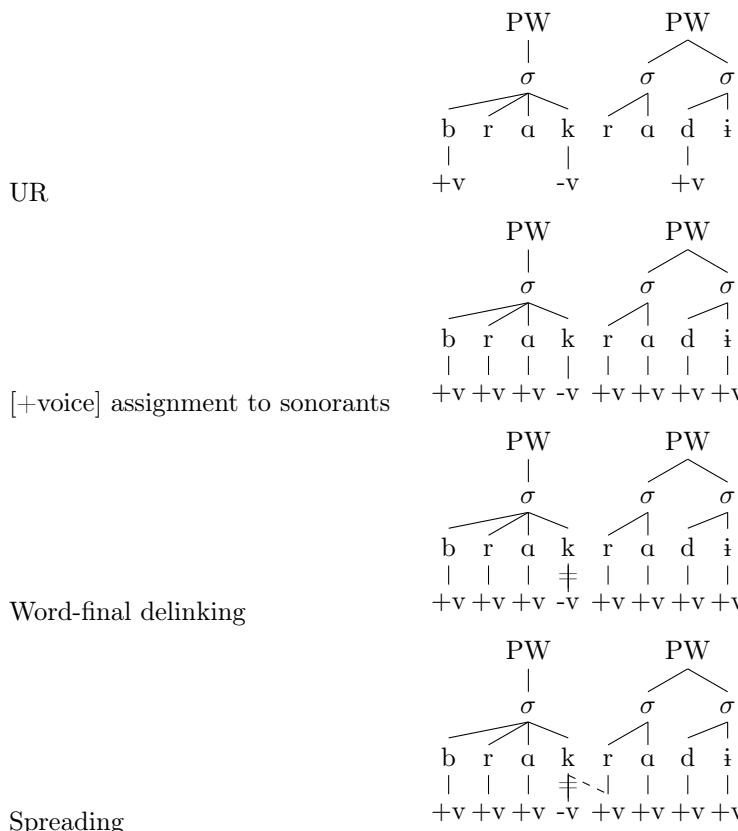
fails to apply. Even if delaryngealisation happened as often as 90% of the time, a 50-50 split in the frequency of pre-sonorant voicing, would produce a 55% of outputs that agree in their laryngeal specification with the input. This in turn could lead to a significant non-neutralising result even though the triggering factor, i.e. failure of neutralisation, is relatively infrequent.

Empirically an analogy-driven model and a stochastic rule-based model seem roughly equivalent for the current data. Certainly, any analysis that presupposes a dependence between delaryngealisation and pre-sonorant voicing would need to include optionality at the very least. At the same time, just because the stochastic possibility is there, the Poznań data do not constitute a definite counterexample to the generalisation that pre-sonorant voicing is conditioned by delaryngealisation, diachronically or synchronically. Thus, Poznań Polish remains an interesting case in the typology of pre-sonorant voicing, but its significance for the discussion on the role of delaryngealisation remains unclear at present.

4.4.2 A unified analysis of Polish voicing

While the results concerning the role of delaryngealisation are inconclusive, the current data are of immediate consequence to unified models of Polish voicing. As discussed in Section 4.1.1, most analyses avoid stipulating a separate rule for pre-sonorant voicing, and argue for a more theoretically elegant solution, where all phonetically voiced sounds (sonorant consonants, voiced obstruents, vowels) share a common laryngeal representation. This, in turn, makes them trigger similar laryngeal operations. For instance, Bethin (1984), Gussmann (1992) and Rubach (1996) propose that pre-sonorant voicing involves laryngeal feature spreading from a redundantly specified sonorant. Rubach's (1996) is shown in (24) for illustration.

- (24) Spreading analysis of Poznań-Kraków voicing by Rubach (1996)



The idea here is that sonorants undergo redundant [voice] assignment in the course of the derivation, which makes them indistinguishable from voiced obstruents as far as laryngeal activity

is concerned. Thus, sonorants are predicted to spread [voice], just as voiced obstruents do, to laryngeally unspecified targets, such as word-final obstruents which had undergone laryngeal feature delinking by that point in the derivation. The only dialect-specific adjustment has to do with the formulation of the spreading rule: while voice assimilation across word-boundaries is limited to voiced obstruents in North-Eastern dialects of Polish, the South-Western dialects permit spreading from all voiced sounds.

Any such unified analysis of all external sandhi voicing processes in Polish is challenged by the variation found in the current data. Variable application is quite common in external sandhi, as confirmed by instrumental studies of a wide range of processes, including voice assimilation (Slis, 1985; Jansen, 2004, *inter alia*). Therefore, it is not unexpected that we should see pre-sonorant voicing applying optionally, as demonstrated by the bimodal distributions of voicing ratio in Figures 4.3, 4.8 and 4.9. However, the degree of variation in the application of pre-sonorant voicing is different from the degree of variation observed in the pre-voiced obstruent environment. Evidence for this comes from the effect of the right-hand environment on phonetic voicing in the Poznań dialect. The effect was significant in the models of voicing ratio for all three subsets of data: i.e. the stops from experiment 1, and the stops and the fricatives from experiment 2. Voicing ratio was consistently found to be significantly greater in obstruents followed by a voiced obstruent, compared with obstruents followed by a sonorant. Although the experimental design was not suited to quantify the variation in regressive voice assimilation between obstruents, the existence of a significant difference in voicing ratio between sonorants and voiced obstruents as voicing triggers supports the generalisation that voicing applies more systematically before voiced obstruents than before sonorants. By linking voice assimilation to obstruents and pre-sonorant voicing to a more general principle of [voice]-spreading, all unified analyses predict that there should be no difference concerning the frequency with which the two processes apply. The way the Kraków-Poznań voicing autosegmental spreading rules are formulated, they cannot see the manner of articulation of the trigger, and thus the grammar has no way of differentiating between [voice] spreading from sonorants and [voice] spreading from obstruents.

The problem faced by analyses based on default-filling and spreading is also relevant for Cyran's (2012) analysis of Kraków-Poznań voicing. Even though Cyran rejects the idea that sonorants are redundantly specified for [voice], he also assumes that word-initial sonorant consonants, vowels and voiced obstruents share a common laryngeal specification characterised by the lack of the element |H|. Thus, all these sounds are expected to show consistent phonological behaviour with respect to the conditioning of laryngeal processes. The problem is less immediate than in the case of default-filling rules, as Cyran proposes that the actual voicing in word-final obstruent is an interface phenomenon. Under this view, the output of phonology does not state which obstruents are voiced, it only produces neutral obstruents amenable to spontaneous voicing at the interface. However, while the non-phonological status of voicing in Cyran's analysis could allow for the variation issue to be dealt with in the phonetics, the theory also needs to specify what spontaneous voicing is, when and where it applies, how it is conditioned and what its phonetic characteristics are. Without such an explication, the analysis fails to make meaningful predictions, and it can therefore hardly be considered complete.

4.5 Conclusion

The data from Poznań Polish problematise the generalisation that delaryngealisation conditions pre-sonorant voicing, as previously argued on the basis of the West-Flemish data. Word-final delaryngealisation appears to be optional in Poznań Polish. Consequently, one needs to consider the possibility that pre-sonorant voicing may operate on laryngeally specified targets, although the

empirical predictions of this proposal are matched by a model in which pre-sonorant voicing applies optionally to just the delaryngealised subset of word-final obstruents. In addition, the Poznań data are of consequence for the proposal that sonorants are specified for [voice], just like obstruents, and likewise to the claim that pre-sonorant voicing and pre-obstruent voicing constitute a uniform phenomenon. As pre-obstruent and pre-sonorant voicing apply with different frequency in the same prosodic environment, obstruents and sonorants must be distinguished as triggers at some level of analysis.

CHAPTER 5

CENTRAL CATALAN

'It isn't logical, I know, but through the centuries this became the established usage,' he said, as if begging the young Frenchwoman to take pity on a language cursed by history.

'I am glad that you admit it. It isn't logical. But a language *must* be logical,' Brigitte said.

Milan Kundera

IMMORTALITY

This chapter presents data on pre-sonorant and prevocalic voicing in Central Catalan. The current descriptions of Catalan voicing (Bonet & Lloret, 1998; Wheeler, 2005) report a number of manner-conditioned effects in Catalan voicing. Prevocalic voicing is said to affect word-final and prefix-final sibilants and stop+sibilant clusters to the exclusion of singleton stops. Voicing before sonorant consonants, on the other hand, applies to all word-final and prefix-final obstruents. The reports concerning prevocalic voicing are largely confirmed by the acoustic data collected in this study from six speakers of Central Catalan. Obstruents followed by sonorant consonants, on the other hand, show more limited voicing than hitherto reported, as well as a considerable degree of inter- and intra-speaker variation. I consider the consequences of these acoustic findings in the context of the existing formal analyses of Catalan voicing (Bermúdez-Otero, 2001; Wheeler, 2005; Jiménez & Lloret, 2008), and from the point of view of sound change, paying special attention to the predictions of a perception-driven reanalysis model developed in Chapter 3 to explain manner of articulation asymmetries in West-Flemish. I propose that the extension of intervocalic sibilant voicing to stop+sibilant clusters resulted from rule generalisation followed by rule telescoping. As a consequence of this step, the resulting pattern evades a successful functional explanation in synchronic terms. Formal synchronic grammars are able to capture the relevant patterns, but in doing so they require a mechanism for inductive learning. I then turn to discussing the issue of the asymmetry between prevocalic and pre-sonorant voicing, and the problems it poses to an evolutionary model, and to formal representational approaches to pre-sonorant voicing.

5.1 Voicing in Catalan

Based on the existing descriptions Catalan presents the most complex case of manner effects on pre-sonorant voicing out of all the languages considered so far. Wheeler (2005) reports that all word-final obstruents undergo voicing when the next word begins with a sonorant consonant. The relevant examples are in (1).

- (1) Word-final obstruent voicing before sonorant consonants (Wheeler, 2005)

<i>sap riure</i>	[br]	'knows how to laugh'
<i>poc whisky</i>	[gw]	'not much whiskey'
<i>cas notable</i>	[zn]	'remarkable case'
<i>mateix líquid</i>	[ʒl]	'same liquid'

Voicing before sonorant consonants also affects word-medial codas, as shown in (2).

- (2) Word-medial obstruent voicing before sonorant consonants (Wheeler, 2005)

<i>hipnosi</i>	[bn]	'hypnosis'
<i>ovni</i>	[vn]	'UFO'
<i>esnob</i>	[zn]	'snob'

Prevocalic voicing, however, is sensitive to the manner of articulation of the undergoer. According to Wheeler (2005) prevocalic voicing is limited to sibilants, i.e. it does not apply to non-sibilant fricatives, or stops. Examples are in (3).

- (3) Word-final obstruent voicing before a vowel (Wheeler, 2005)

a. Non-sibilants

<i>sap ajudar</i>	[p]	'knows how to help'
<i>llarg any</i>	[k]	'long year'
<i>xef únic</i>	[f]	'sole chef'

b. Sibilants

<i>cas extrem</i>	[z]	'extreme case'
<i>mateix element</i>	[ʒ]	'same element'
<i>mig any</i>	[dʒ]	'half a year'

While Wheeler (2005) groups labiodental fricatives with stops as resisting prevocalic voicing, Bonet & Lloret (1998) treat them as undergoing prevocalic voicing. However, both sources acknowledge that their generalisations do not take into account that prevocalic voicing is variable for labiodental fricatives. Such variation is noted by Recasens (1993), and is described as being conditioned by the placement of phrasal stress.

All consulted sources mentioned above agree that Central Catalan shows pre-sonorant and prevocalic voicing in consonant clusters ending in sibilants, as shown in (4).

- (4) Prevocalic voicing in consonant clusters with a final sibilant (Bonet & Lloret, 1998; Wheeler, 1979, 2005)

<i>cops amagats</i>	[bz]	'hidden blows'
<i>amics íntims</i>	[gz]	'close friends'
<i>discs antics</i>	[zgz]	'old discs'

The prevocalic voicing pattern of fricatives described above occurs not only at the end of a word, but also prefix-finally, as illustrated in (5).

- (5) Prefix-final voicing(Bonet & Lloret, 1998; Bermúdez-Otero, 2001)

<i>des-estimar</i>	[z]	'to rule out'
<i>trans-humància</i>	[z]	'transhumance'
<i>sots-arrendar</i>	[dʒ]	'to sublet'

Other voice-related processes in Catalan include final devoicing and voice assimilation to obstruents. As illustrated in (6), only voiceless obstruents are found word-finally before a pause, but voicing is contrastive word-medially.

- (6) Voice neutralisation and contrast (Wheeler, 2005)

Final devoicing			Voicing contrast		
<i>sap</i>	[sap]	'I know'	<i>sabem</i>	[sa.'βem]	'we know'
<i>tip</i>	['tip]	'fed up, MASC.'	<i>tipa</i>	['ti.pə]	'fed up, FEM.'
<i>cas</i>	['kas]	'case'	<i>casos</i>	['ka.zus]	'cases'
<i>braç</i>	['bras]	'arm'	<i>braços</i>	['bra.sus]	'arms'
<i>serf</i>	['serf]	'serf, MASC.'	<i>serfa</i>	['ser.fə] ~ ['ser.βə]	'serf, FEM.'
<i>buf</i>	['buf]	'puff'	<i>bufar</i>	[bu.'fa]	'to puff'
<i>mig</i>	['mitʃ]	'half, MASC.'	<i>mitja</i>	['mid.dʒə]	'half, FEM.'
<i>despatx</i>	[dəs.'patʃ]	'office'	<i>despatxos</i>	[dəs.'pa.tʃus]	'offices'

In obstruent clusters regressive voice assimilation applies with the [voice] value determined by the rightmost segment in the cluster, as shown in (7).

- (7) Voice assimilation in Catalan (Cuartero Torres, 2001; Wheeler, 2005)

a.	<i>nap gran</i>	[bg]	'big turnip'
	<i>gos danès</i>	[zð]	'Danish dog'
	<i>set terrible</i>	[t:]	'terrible thirst'
	<i>cas terrible</i>	[st]	'terrible case'
b.	<i>futbol</i>	[db] ~ [b:]	'football'
	<i>afgà</i>	[vy]	'Afghan'
	<i>dubte</i>	[pt]	'doubt'
	<i>tastar</i>	[st]	'taste'

5.1.1 Issues related to Central Catalan voicing

Similar to other cases of pre-sonorant voicing analysed so far, Central Catalan shows three puzzling characteristics that need to be accounted for. First, there is the issue of where the pre-sonorant and prevocalic voicing come from. The second question has to do with positional restrictions which limit the voicing process to word-final and prefix-final positions. Third, Catalan presents the problem of undergoer manner asymmetry, as only (sibilant) fricatives, but not stops, are reported to undergo prevocalic voicing. However, in Catalan the effect of manner of articulation on voicing is more complex than that, since there is also an asymmetry between sonorant consonants and vowels as triggers of voicing. This interaction of manner of articulation effects on voicing makes Catalan unique in the typology of pre-sonorant voicing languages considered so far. It also presents a considerable analytical challenge to any account of Catalan voicing.

Previous analyses

Bermúdez-Otero (2001) Bermúdez-Otero (2001) proposes that vowels and sonorants in Central Catalan are redundantly specified for [+voice], i.e. they share a uniform laryngeal specification with voiced obstruents. Being equipped with a laryngeal feature, vowels and sonorants are able to spread it to a neighbouring segment. Building up on this assumption, Bermúdez-Otero's analysis attempts to explain the restrictions on pre-sonorant and prevocalic voicing. Pre-sonorant voicing is analysed as conditioned by delaryngealisation at the Word Level, which affects all prefix-final and word-final obstruents. Coda obstruents then undergo postlexical voice assimilation to a following onset consonant.

The delaryngealisation-cum-voicing generalisation stated above predicts no voicing in word-final prevocalic obstruents, as they undergo resyllabification at the Phrase level, and surface as onsets. However, this prediction only holds for stops, while prevocalic fricatives surface as voiced. To account for this, Bermúdez-Otero proposes that prevocalic sibilant voicing is a

different process altogether from pre-sonorant voicing. While pre-sonorant voicing is conditioned by delaryngealisation, prevocalic voicing is proposed to result from alignment requirements on continuancy and voicing. The idea is that fricative+vowel sequences share a [+continuant] feature due to OCP requirements. In sequences like this a constraint *CONTVOILAG conditions leftward [+voice] spreading from the vowel to prevent the onset of voicing from lagging behind the onset of continuancy. The constraint cannot be violated by sequences of voiceless stops followed by a sonorant, as stops are [-continuant]. Voicing in word-final prevocalic stops is then blocked by lower ranked markedness constraints that penalise the [+voice] feature in obstruents. As a result of the constraint interaction, word-final prevocalic fricatives surface as voiced, but stops are voiceless in the same position.

Wheeler (2005) Wheeler (2005) proposes a parallel Optimality Theory analysis of Catalan voicing. Despite assuming a different number of levels and relying on a different constraint set, Wheeler's analysis shares some fundamental ideas with Bermúdez-Otero (2001) when it comes to explaining the source of voicing and positional restrictions. Sonorants and vowels are once again treated as laryngeally specified for [+voice], and thus subject to constraint evaluation of [voice] agreement with the flanking obstruents. Pre-sonorant voicing is analysed as coda delaryngealisation-cum-spreading; word-final obstruents followed by sonorant consonants lose their own laryngeal specifications to then undergo spreading so they agree in their [voice] value with the following sonorant. The voicing of word-final prevocalic obstruents which surface as onsets, however, is ascribed to a different mechanism. In order to capture sibilant voicing in this environment Wheeler (2005) introduces a constraint LAZYSIBILANTS which requires word-final fricatives to be voiced before vowels. Stops are exempt from this requirement by the constraint formulation, and they surface as voiceless word-finally before a following vowel.

Jiménez & Lloret (2008) Jiménez & Lloret (2008) propose a solution to the stop-fricative asymmetry based on a link between voicing and the degree of stricture. Jiménez & Lloret model word-final pre-sonorant voicing as conditioned by coda delaryngealisation-cum-spreading from the following redundantly specified sonorant. The spreading is formalised as resulting from an AGREE[±voice] constraint which requires adjacent segments to share the same voicing value. The same AGREE constraint conditions the spreading of [+voice] from a word-initial vowel to a preceding fricative. A similar spreading process is blocked from applying to stops by the activity of another high-ranked constraint No-LINK-VC_{≥Ocl}. No-LINK-VC_{≥Ocl} belongs to a family of constraints which prohibit the sharing of a feature [+voice] between a vowel and a following obstruent, depending on the obstruent's degree of constriction (Jiménez, 1999). No-LINK-VC_{≥Ocl} prohibits voicing of postvocalic stops, but permits voicing in segments of relatively lesser constriction, such as fricatives.

Issues and predictions

All of the above proposals resolve the issue of why fricatives are more susceptible to prevocalic voicing than stops by positing various kinds of markedness restrictions on the interaction of manner and voicing. However, while these restrictions achieve the required effect from the formal point of view, a functional motivation for assuming them is problematic.

Let us first consider a link between voicing and continuancy, as proposed by Bermúdez-Otero (2001) and Jiménez & Lloret (2008). At the heart of Bermúdez-Otero's proposal is the idea that voicing is somehow preferred in continuants above non-continuants when another [+continuant] sound follows. In support of this generalisation Bermúdez-Otero cites cases of prevocalic fricative voicing, such as the case of the Kentish dialect of Middle English where Old English [f, s, θ] underwent voicing to [v, z, ð] (e.g. OE *fæder* > ME *vader*).

Jiménez & Lloret, who link voicing to degree of stricture, look for evidence in the typology of Catalan dialects. They report that Catalan dialects form a continuum with respect to which obstruents undergo prevocalic voicing. On the one end of the spectrum the Alicante dialect voices all word-final prevocalic obstruents, including stops. Central Catalan voices all obstruents in the same position, excluding stops (as well as variably the labial fricative). Two Valencian dialects (which they term A and B) are relatively more restrictive than Central Catalan, with voicing applying prevocally in sibilant fricatives, but not in stops, labial fricatives, or affricates. The typology is completed by Central Valencian, where prevocalic obstruent voicing does not occur at all. Jiménez & Lloret argue that the typology reflects a hierarchy of susceptibility to prevocalic voicing, as evidenced by the absence of dialects where, for instance, voicing applies prevocally in affricates, but not in stops. The authors observe that the likelihood of voicing correlates with the degree of stricture, and postulate a hierarchy of No-LINK-VC constraints which disallow the voicing of a postvocalic segment depending on the degree of stricture.

The markedness constraints and markedness relationships proposed by Jiménez & Lloret (2008) and Bermúdez-Otero (2001) may receive some support from typological evidence, but their specific functional grounding remains unclear. Constraints on alignment between [+continuant] and [+voice] (Bermúdez-Otero, 2001), or restrictions on feature sharing imposed by degree of stricture (Jiménez & Lloret, 2008) cannot be taken as an explanation of what conditions prevocalic fricative voicing unless they can motivate the preference within an explicit articulatory, aerodynamic or acoustic model. Thus, the proposed markedness restrictions may be successfully used as a formal tool for capturing the observed variation, but the markedness-based solution is only tentative unless the relevant constraints and hierarchies can be grounded in external factors that underlie the effects of manner on susceptibility to voicing.

Wheeler (2005) argues that constraints need not necessarily be functionally motivated. He notes that sibilant liaison is a common process in Romance, as seen in Portuguese, Catalan, Occitan and French, and that it reflects a series of sound changes where singleton intervocalic obstruents underwent lenition, which was followed by a simplification of geminates. These two processes applying in succession rendered voicing opaque, and synchronically unnatural. Although Wheeler (2005) makes it clear that he sees Catalan voicing as a product of historical lenition rather than as a synchronically active lenition process, he nevertheless formalises his proposal by means of a constraint from the *LAZY family, which makes direct reference to avoidance of articulatory effort as a cognitive pressure on synchronic grammar. Expressing a generalisation which Wheeler admits to be synchronically arbitrary by means of an apparently functional constraint is questionable, and the resulting synchronic grammar lends itself to the same criticism levelled at the proposals by Bermúdez-Otero (2001) and Jiménez & Lloret (2008); all the analyses, although descriptively adequate, rely on stipulating a markedness relationship.

Setting aside the question of what motivates prevocalic voicing in word-final fricatives, the analyses listed above make a couple of common predictions concerning voicing before sonorant consonants. By treating sonorants as a subset of [+voice] sounds, and analysing pre-sonorant voicing as [+voice] spreading the analyses predict that obstruent voicing before sonorants and obstruents should be a uniform process. Since the product of categorical spreading is in both cases an obstruent specified as [+voice], and since the spreading is conditioned by common restrictions on [voice] agreement, voicing is expected to apply with the same frequency preceding sonorants and voiced obstruents. Prevocalic voicing, on the other hand, may apply with a different frequency from pre-sonorant voicing, since the two processes are conditioned by different factors. Although the consulted sources state the voicing generalisations as categorical, variation is not unexpected, just as in any sandhi process. Finally, all the analyses predict that the underlying voicing value of the obstruent that undergoes pre-sonorant voicing should be neutralised, since the voicing is fed by delaryngealisation.

Relevance to the current model

In my previous analysis of West-Flemish (Chapter 3) and Poznań Polish (Chapter 4) I have largely sidestepped the issue of the trigger's manner of articulation in the analysis of pre-sonorant voicing languages considered so far, and I have treated sonorants as a uniform group comprising sonorant consonants and vowels, based on the lack of significant statistical differences between vowels and sonorant consonants as triggers of voicing. However, the descriptions by Bonet & Lloret (1998) and Wheeler (2005) make it clear that this kind of approach would not be appropriate in the case of Central Catalan, given that prevocalic voicing is reported to only apply to sibilants, whereas all obstruents undergo voicing before a sonorant consonant. What is more, the way the voicing undergoer's manner of articulation interacts with the manner of articulation of the trigger poses a problem to the previously proposed model of pre-sonorant voicing as conditioned by perceptual re-analysis (cf. 3.4). The basic idea of the proposal was that pre-sonorant voicing is driven by perceptual reanalysis of passive voicing in delaryngealised obstruents as involving an active voicing target. I have argued in Chapter 3 that different acoustic cues to voicing in stops and fricatives may explain why pre-sonorant voicing is more likely in fricatives than in stops. Trigger's manner of articulation effects, however, are not inherently predicted by this model. The perception-driven model relies on delaryngealisation as the main factor triggering the occurrence of extended passive voicing, and delaryngealisation occurs before sonorant consonants and vowels alike. A trigger asymmetry is not counter-predicted, but accommodating it within the model requires a further consideration of the phonetic motivation behind trigger asymmetries and the stage of the sound change at which it becomes relevant.

Another surprising feature of Catalan voicing from the point of view of perceptual reanalysis is prevocalic voicing of stop+sibilant clusters. In my discussion of West-Flemish I have argued, following Jansen (2004), that passive voicing in word-final obstruents is perseverative, and that it consists of a continuation of previously initiated vocal fold vibration uncounteracted by the presence of a voicing target. This kind of voicing is arguably more likely to be perceivable in the case of sibilants than stops. However, since perseverative voicing is more likely to affect the initial part of the segment/cluster, clusters of stops followed sibilants would be expected to behave more like prevocalic stops than like prevocalic sibilants.

The behaviour of clusters is equally problematic for proposals which explain the stop-fricative asymmetry based on aerodynamic conditioning. Recasens (2002) argues that the observed asymmetry between stops and fricatives is consistent with the generalisation that voicing is harder to maintain in sounds with relatively more constriction, as obstruction to airflow conditions a rise in the intraoral pressure and a cessation of voicing¹. In the light of this explanation, however, voicing is not expected in prevocalic stop+sibilant clusters since just like stops they involve a closure phase where the obstruction in the airflow quenches voicing. One might argue that the closure phase is perhaps longer in singleton stops than in stop+sibilants clusters, and that the longer the constriction, the less likely the voicing. This explanation, however, is short lived considering that extended duration counteracts voicing in any obstruent as supraglottal pressure rises over time. Since clusters are longer than singleton stops they are relatively less likely to undergo voicing. The opposite effect that we find in Central Catalan is a puzzle which calls for a closer investigation.

5.1.2 The current study

This study probes whether any additional insight into the nature of pre-sonorant and prevocalic voicing in Central Catalan can be gained through an analysis of acoustic data. One of the aims of the investigation is descriptive. Despite the relatively high number of phonetic investigations into Catalan voicing contrast and sandhi voicing (Dinnsen & Charles-Luce, 1984; Charles-Luce, 1993;

¹Some problems concerning this proposal are discussed in 5.4.1 and 7.4.

Cuartero Torres, 2001, *inter alia*), pre-sonorant and prevocalic voicing does not seem to have as yet been studied instrumentally. Therefore, the existing descriptions based on introspection and auditory transcriptions await an experimental confirmation.

My second aim is to test a number of phonetic predictions which follow from the existing formal analyses of Catalan voicing (cf. Section 5.1.1) by addressing the following research questions: 1) Does voicing apply to the same extent preceding sonorant consonants, vowels and voiced obstruents?; 2) How does the effect of the right-hand environment interact with the manner of articulation of the undergoer?; 3) Is the realisation of voicing affected by the underlying voicing value of the obstruent?

5.2 Materials and method

The data presented in this chapter were collected in a production study. Six native speakers of Central Catalan from the Barcelona region read four repetitions of test items embedded in a carrier phrase. There were five female speakers and one male. The age range was 18-32. The speakers were not aware of the purpose of the experiment. They were not paid for participation.

The test items were words ending in word-final obstruents or obstruent clusters (/p/, /b/, /s/, /z/, /bz/) followed by a vowel or a sonorant consonant in the next word. An equal number of underlyingly voiced and voiceless tokens were used in the case of singleton stops and fricatives. Stop+sibilant clusters all had an underlyingly voiced obstruent followed by a morphological marker -s. Place of articulation of the undergoer was not strictly controlled for, since labial stops and coronal fricatives were used. This was necessary in order to avoid using word-final coronal stops, which undergo place assimilation before following laterals and nasals (Wheeler, 2005, p. 167). Labial fricatives were also avoided as prevocalic voicing in their case had been reported as highly variable (Recasens, 1993; Bonet & Lloret, 1998). For every type of undergoer three vocalic contexts were included ([i], [ə], [u]), and they were varied systematically. The sonorant contexts included a following /m/, /n/, /r/, /l/, and /ʎ/. The test items were embedded in the same carrier phrase. An example is in (8).

- (8) Example stimulus sentence
Digui ‘sap anglès’ un altre cop.
‘Say ‘knows English’ one more time.’

In addition, 20 tokens of word-final obstruents followed by an obstruent (/t/, /d/, /s/, /z/) in the next word were included as controls. For ease of segmentation following fricatives were used when the word final obstruent was a plosive, whereas word-final sibilants were placed in the context of a following stop. Similarly to the test items, the control items were embedded in a carrier phrase.

The recordings were made in a sound-treated room on a Pioneer PDR-609 CD recorder, using a Sennheiser MKH20 P-48 microphone. The speakers were positioned 30 cm away from the microphone and instructed to read the sentences at a comfortable rate. They were encouraged to correct themselves if they made a mistake. The test items were randomised for each speaker and presented on a computer screen, one at a time. The experiment was self-timed. The speakers read four repetitions of the materials.

The recordings were sampled at 44.1 kHz. Segmentation and acoustic analysis were carried out in Praat (Boersma & Weenink, 2010) on a 5 ms Gaussian window (spectrogram bandwidth 260 Hz). Boundaries were inserted manually based on visual analysis of the spectrograms. Altogether (40(test stimuli)+10(control stimuli))*6(speakers) *4(repetitions)= 1200 utterances were recorded. 74 utterances were excluded due to deletions, mispronunciations, segmentation difficulties, instances of prevocalic glottalisation, or place assimilation (e.g. /s#r/→[r:]). This left

1126 test utterances for analysis.

The following acoustic measurements were taken.

1. Duration of glottal pulsing during closure (voicing duration). Previous research shows that the presence of vocal fold vibration is the primary cue to voicing in Catalan (Cuartero Torres, 2001). In addition, Cuartero Torres reports extended glottal pulsing in segments followed by a voiced obstruent. Duration of glottal pulsing was measured manually based on the presence of the voicing bar on the spectrogram and periodicity in the waveform. Absence of voicing was coded as 0.
2. Duration of the segment. For fricatives duration of friction was measured, whereas for stops the duration measurement was taken to correspond to the closure phase. Frication was identified based on the presence of high-frequency noise in the spectrogram. Stop closure was identified based on the presence of low acoustic energy between the preceding vowel and the following stop release. Duration measurements were taken primarily with the aim of contextualising voicing duration. However, in some languages segmental duration has been shown to be a correlate of voicing (Chen, 1970; Kluender et al., 1988).
3. Vowel duration. Charles-Luce (1993) reports that in some cases vowel lengthening is observed before underlyingly voiced obstruents in word-final position. Similar effects have been reported for a number of other languages, notably by studies of incomplete neutralisation (e.g. Port & O'Dell (1985) for German, Warner et al. (2004) for Dutch). Vowel duration was measured manually. The beginning of the vowel was placed at the beginning of the formant structure for vowels preceded by obstruent. For vowels preceded by sonorants, the initial boundary was placed at the onset of the formant steady state. The boundary between the vowel and the following stop was placed at the onset of low acoustic energy at higher frequencies.
4. f_0 and f_1 preceding the closure (20 ms and 10 ms before the onset of the closure). f_0 and f_1 lowering has been observed preceding voiced stops in Dutch by Jansen (2004). f_0 and f_1 measurements in the following segment were not taken, since some of the following segments involved voiceless obstruents, where the measurements would be impossible or unreliable. f_0 and f_1 were measured using the autocorrelation and Burg algorithms in Praat, respectively.

All of the duration measurements, including voicing, stop closure, burst, and vowel duration, were taken in milliseconds. f_0 and f_1 were measured in Hertz.

5.3 Results

5.3.1 Realisation of voicing-related acoustic parameters

The data were analysed using linear mixed-effects regression. The analysis was run in R (R Development Core Team, 2005), using the lme4 package (Bates & Maechler, 2009). A series of models were fitted to predict the realisation of the following acoustic measurements associated with the voicing contrast: duration of vocal fold vibration, ratio of the duration of vocal fold vibration to obstruent duration (henceforth voicing ratio), obstruent duration, duration of the preceding vowel, f_0 and f_1 at 10 and 20 ms before the onset of the obstruent. For all the models the effects of speaker and item were treated as random. The fixed effects considered in the modelling included: obstruent's manner of articulation (whether it was a sibilant, a stop or a cluster), manner of articulation of the following segment (whether it was a sonorant consonant, a vowel, a voiced/voiceless obstruent), and the underlying voicing of the obstruent. The obstruent's

manner of articulation is abbreviated as ‘undergoer’, and the manner of articulation of the following segment is abbreviated as ‘trigger’ in tables and figures throughout this section. In the process of modelling, individual predictors were considered by being added to the model one by one. Predictors were retained in the fixed part of the model, or as a random slope within speaker/item only if they significantly improved the model’s fit on the log-likelihood test. The best fitting models for every variable are presented in this section. For every model a table of coefficients is included, summarising the β values, as well as the standard error, and the t -values. No p -values were obtained for any of the models presented in this section due to the presence of random slopes, as MCMC sampling has not been implemented for such models in the current version of the lme4 package. Significant interactions are represented graphically.

Voicing duration

The first model was fitted to the data predicting the duration of vocal fold vibration in the word-final obstruent. The best fitting model had an interaction between the manner of articulation of the trigger and of the undergoer in the fixed part. The random part of the model included random intercepts for speaker and item, as well as a random slope for undergoer’s manner of articulation within speaker. The model was not significantly improved by the inclusion of the undergoer’s underlying voicing as a predictor. The model’s summary is in Table 5.1.

Table 5.1: Regression coefficients, with standard error and t values for a model predicting the duration of vocal fold vibration (in ms) in word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.

Term	Level	β	SE	t
(Intercept)		46.25	5.78	8.01
Undergoer	cluster	2.10	8.28	0.25
Undergoer	stop	-2.37	7.76	-0.31
Trigger	voiced obstruent	20.70	6.97	2.97
Trigger	voiceless obstruent	-23.92	6.93	-3.45
Trigger	vowel	23.10	5.23	4.41
Undergoer: trigger	cluster: voiced obstruent	31.30	11.94	2.62
Undergoer: trigger	stop: voiced obstruent	-1.78	9.70	-0.18
Undergoer: trigger	cluster: voiceless obstruent	0.21	11.92	0.02
Undergoer: trigger	stop: voiceless obstruent	6.37	9.62	0.66
Undergoer: trigger	cluster: vowel	42.91	8.90	4.82
Undergoer: trigger	stop: vowel	-42.04	7.44	-5.65

The main effect of the undergoer’s manner of articulation was overall very small, and the associated small t -values (smaller than $|0.5|$) suggest that it was also not significant. There were, however, relatively large differences in voicing duration associated with the trigger’s manner of articulation. Voicing duration was considerably greater in obstruents followed by a voiced obstruent, as compared with pre-sonorants stops ($\beta=20.70$, $SE=6.97$, $t=2.97$). At the same time, presonorant obstruents were associated with increased vocal fold duration compared to obstruents followed by voiceless obstruents ($\beta=-23.92$, $SE=6.93$, $t=-3.45$). Vowels triggered on average more voicing in preceding obstruents than sonorants did ($\beta=23.10$, $SE=5.23$, $t=4.41$). This last effect was also found to interact with the undergoer’s manner of articulation. The interaction is illustrated in Figure 5.1. While within sibilants and clusters there was more voicing before a vowel than before a sonorant, the opposite effect was found in stops, where there was virtually no prevocalic voicing; prevocalic stops patterned with stops followed by voiceless obstruents in terms of voicing duration. Sibilants, stops and clusters had a similar amount of voicing duration before sonorants and before voiceless stops. Before voiced obstruents clusters showed increased duration of voicing compared

with both singleton sibilants and stops. Prevocalic clusters also showed greater voicing duration than prevocalic sibilants.

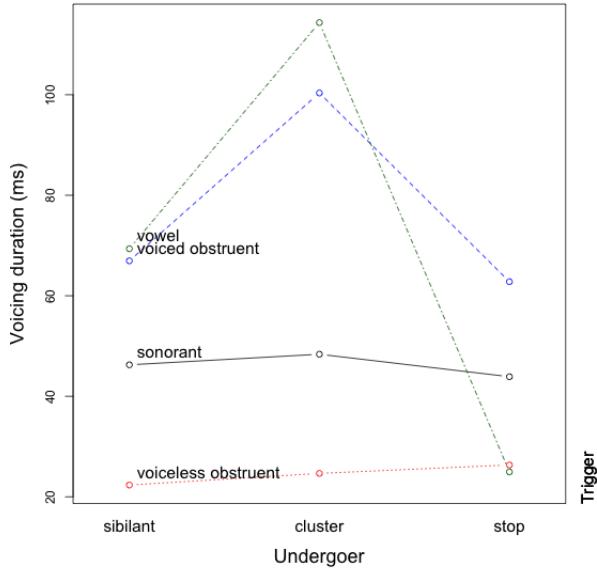


Figure 5.1: Interaction between the undergoer’s and trigger’s manner of articulation in a model of voicing duration in Catalan word-final obstruents.

The fit of the model improved significantly with the inclusion of a random effect of undergoer within speaker, which indicates that there were individual differences with respect to how much voicing was found on average in sibilants, stops and clusters. I will return to the issue of individual variation in Section 5.3.2.

Voicing ratio

A linear mixed-effects regression model was fitted to the data predicting the ratio of voicing duration to obstruent duration. The best fitting model of voicing ratio had an interaction of undergoer’s and trigger’s manner of articulation in its fixed part. The model did not improve significantly with the inclusion of underlying voicing as a predictor. The model did improve when its random part was expanded by adding random slopes for trigger’s and undergoer’s manner of articulation. The model’s summary is in Table 5.2.

From the effects’ summary it follows that voicing ratio was generally lower in obstruent clusters than in sibilants ($\beta=-0.23$, $SE=0.07$, $t=-3.45$), but there was very little difference between voicing ratio in sibilants and in stops ($\beta=0.03$, $SE=0.10$, $t=0.32$). The main effect of trigger’s manner on voicing ratio was similar to its effect on voicing duration. Voicing ratio was lower in obstruents followed by sonorants than in obstruents followed by a voiced obstruent ($\beta=0.36$, $SE=0.08$, $t=4.32$), but higher than in obstruents followed by a voiceless obstruent ($\beta=-0.28$, $SE=0.09$, $t=-3.19$). Overall the voicing ratio was higher in prevocalic than in presonorant obstruents ($\beta=0.34$, $SE=0.06$, $t=5.58$), except when the obstruent was a stop. An interaction between trigger’s and undergoer’s manner of articulation is plotted in Figure 5.2. Prevocalic stops patterned with stops followed by a voiceless obstruent in terms of voicing ratio. A relatively low voicing ratio was also observed in pre-sonorant clusters.

The model of voicing ratio with random slopes for undergoer’s and trigger’s manner of articulation fitted the data significantly better than a model with random intercepts (for speaker and item) only ($\chi^2=144.6$, $df=20$, $p<0.001$). Significant improvements were also found in a stepwise

Table 5.2: Regression coefficients, with standard error and t values for a model predicting ratio of vocal fold vibration to obstruent duration in word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.

Term	Level	β	SE	t
(Intercept)		0.56	0.06	8.87
Undergoer	cluster	-0.23	0.07	-3.45
Undergoer	stop	0.03	0.10	0.32
Trigger	voiced obstruent	0.36	0.08	4.32
Trigger	voiceless obstruent	-0.28	0.09	-3.19
Trigger	vowel	0.34	0.06	5.58
Undergoer: trigger	cluster: voiced obstruent	0.07	0.13	0.53
Undergoer: trigger	stop: voiced obstruent	-0.17	0.10	-1.64
Undergoer: trigger	cluster: voiceless obstruent	0.12	0.13	0.92
Undergoer: trigger	stop: voiceless obstruent	0.08	0.10	0.78
Undergoer: trigger	cluster: vowel	0.11	0.09	1.22
Undergoer: trigger	stop: vowel	-0.57	0.08	-7.21

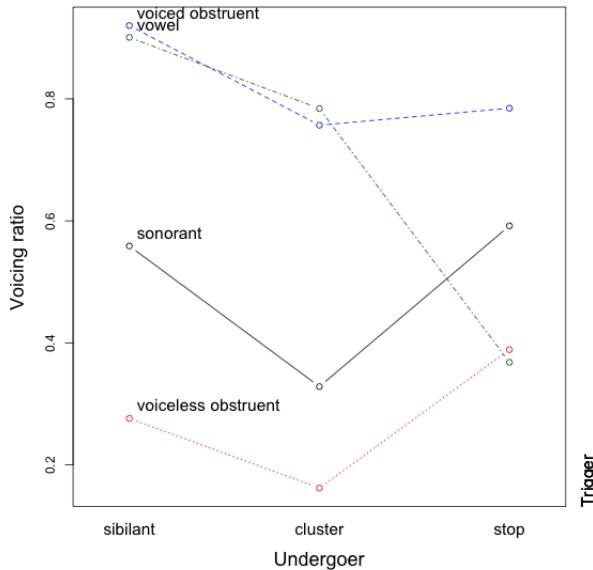


Figure 5.2: Interaction between the undergoer's and trigger's manner of articulation in a model of voicing ratio in Catalan word-final obstruents.

model selection between models with only one random effect within speaker. This, again, is indicative of a considerable amount of individual variation, and therefore the generalisations that emerge from the model for the population might not accurately reflect the behaviour of individual speakers.

Obstruent duration

A model of obstruent duration achieved the best fit of the data with two main fixed effects: of undergoer's and trigger's manner of articulation. The fit of the model did not significantly improve when an interaction between these two factors was added to the model ($\chi^2=10.34$, $df=6$, $p=0.11$), nor when a main effect of obstruent's underlying voicing was added ($\chi^2=0$, $df=1$, $p=1$). The summary of the fixed effects is in Table 5.3. The random part of the best fitting model included random intercepts for speaker and item, as well as random slopes for undergoer's and trigger's

manner of articulation within speaker.

Table 5.3: Regression coefficients, with standard error and t values for a model predicting obstruent duration (in ms) in word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.

Term	Level	β	SE	t
(Intercept)		86.16	6.35	13.58
Undergoer	cluster	68.88	5.46	12.61
Undergoer	stop	-5.38	6.70	-0.80
Trigger	voiced obstruent	-8.56	6.26	-1.37
Trigger	voiceless obstruent	-3.97	5.39	-0.74
Trigger	vowel	-7.07	3.14	-2.25

Clusters had a considerably higher duration than sibilants ($\beta=68.88$, $SE=5.46$, $t=0.32$), whereas there was very little difference in the average duration of frication and stop closure ($\beta=-5.38$, $SE=6.70$, $t=-0.80$). Presonorant obstruents were on average longer than prevocalic ones ($\beta=-7.07$, $SE=3.14$, $t=-2.25$), as well as obstruents followed by another obstruent, although the small t - values indicate a non-significant effect.

Vowel duration

The model of preceding vowel duration was found to provide the best fit of the data with the following structure: an interaction between between undergoer's and trigger's manner of articulation in the fixed part, and random effects for speaker, item as well as undergoer's and trigger's manner of articulation within speaker. The summary of the fixed part of the model is in Table 5.4. Vowels preceding sibilants were considerably longer than vowels preceding stops or stop+sibilant clusters (stops: $\beta=-29.22$, $SE=3.05$, $t=-9.57$; clusters: $\beta=-30.53$, $SE=3.02$, $t=-10.12$). Vowels preceding prevocalic obstruents were relatively longest ($\beta=12.18$, $SE=2.47$, $t=4.93$ compared to the pre-sonorant context), followed by vowels preceding obstruent+voiced obstruent clusters ($\beta=6.41$, $SE=3.42$, $t=1.87$ compared to the pre-sonorant context). Vowels followed by an obstruent in the context of a voiceless obstruent were relatively shortest ($\beta=-11.28$, $SE=4.38$, $t=-2.57$ compared to the pre-sonorant context). The comparison of trigger effects between voiced and voiceless obstruents confirms the existence of vowel lengthening in the context of a following voiceless sound. The difference in length between vowels followed by obstruent in the context of a voiced and voiceless obstruent (established by re-running the model with voiced obstruent as a baseline) was 10.27 ms ($SE=3.83$, $t=-2.68$). This effect of trigger's manner of articulation (vowel>voiced obstruent>sonorant>voiceless obstruent) was observed within stops, sibilants and clusters, but the size of the effect of trigger differed somewhat depending on the undergoer. The interaction between undergoer and trigger is plotted in Figure 5.3.

f_0 of the preceding vowel

Modelling of f_0 of the vowel preceding a word-final obstruent did not yield any significant main effects.

f_1 of the preceding vowel

The best-fitting model of the preceding f_1 had an interaction between manner's and trigger's manner of articulation in its fixed part, random intercepts for speaker and item, as well as a random slope for undergoer's manner of articulation within speaker. f_1 modelling showed an effect of f_1 lowering in the context where voicing is expected. This generalisation follows from the main

Table 5.4: Regression coefficients, with standard error and t values for a model predicting vowel duration (in ms) preceding word-final obstruents. The intercept corresponds to a word-final sibilant followed by a sonorant.

Term	Level	β	SE	t
(Intercept)		115.29	4.38	26.35
Undergoer	cluster	-30.53	3.02	-10.12
Undergoer	stop	-29.22	3.05	-9.57
Trigger	voiced obstruent	6.41	3.42	1.87
Trigger	voiceless obstruent	-11.28	4.38	-2.57
Trigger	vowel	12.18	2.47	4.93
Undergoer: trigger	cluster: voiced obstruent	-4.82	5.08	-0.95
Undergoer: trigger	stop: voiced obstruent	-3.51	4.12	-0.85
Undergoer: trigger	cluster: voiceless obstruent	7.59	5.07	1.50
Undergoer: trigger	stop: voiceless obstruent	8.39	4.07	2.06
Undergoer: trigger	cluster: vowel	-7.99	3.82	-2.09
Undergoer: trigger	stop: vowel	-11.97	3.18	-3.76

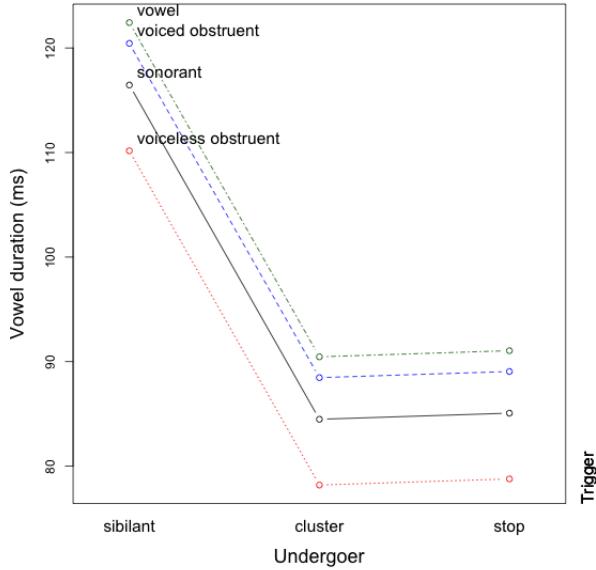


Figure 5.3: Interaction between the undergoer’s and trigger’s manner of articulation in a model of vowel duration preceding Catalan word-final obstruents.

effect of trigger’s manner of articulation. In a model of f_1 measured at 10 ms before the vowel offset there was a main effect of trigger’s manner, with f_1 lowering in the context of an obstruent followed by a voiced obstruent. f_1 was lower by 70.96 Hz in the context of a voiced obstruent (across another intervening obstruent) than in the context of a voiceless obstruent ($\beta=70.96$, $SE=18.18$, $t=3.90$). The context of a following sonorant consonant had a similar effect on f_1 to the context of a voiced obstruent. The main effect of vowel involved f_1 lowering compared to sonorants ($\beta=-37.91$, $SE=10.84$, $t=-3.50$), and compared to voiced obstruents ($\beta=-32.70$, $SE=15.79$, $t=-2.07$). This effect, however, was mostly conditioned by vowels followed by prevocalic sibilants. An interaction between trigger’s and undergoer’s manner of articulation on f_1 at 10 ms before the vowel offset is plotted in the left panel of Figure 5.4. The effect of trigger’s manner of articulation varied considerably within sibilants, with f_1 lowering in prevocalic sibilants, and relatively high f_1 before sibilants followed by voiceless obstruents. Within stops and clusters the effect of trigger’s manner was more limited, although clusters show some f_1 lowering in the preceding vowel when followed

by voiced obstruents.

Table 5.5: Regression coefficients, with standard error and t values for a model predicting f_1 (in Hz) of the vowel preceding word-final obstruents at 10 ms before the vowel offset. The intercept corresponds to a word-final sibilant followed by a sonorant.

Term	Level	β	SE	t
(Intercept)		616.15	23.78	25.91
Undergoer	cluster	24.27	23.21	1.05
Undergoer	stop	32.32	24.65	1.31
Trigger	voiced obstruent	-5.21	14.30	-0.37
Trigger	voiceless obstruent	65.74	14.08	4.67
Trigger	vowel	-37.91	10.84	-3.50
Undergoer: trigger	cluster: voiced obstruent	-17.19	24.20	-0.71
Undergoer: trigger	stop: voiced obstruent	3.95	19.45	0.20
Undergoer: trigger	cluster: voiceless obstruent	-56.07	24.07	-2.33
Undergoer: trigger	stop: voiceless obstruent	-61.72	19.06	-3.24
Undergoer: trigger	cluster: vowel	40.25	18.38	2.19
Undergoer: trigger	stop: vowel	52.55	15.28	3.44

A model of f_1 measured at 20 ms before the vowel offset showed very similar results to the model of f_1 measured at 10 ms before the offset. Both models achieved the best fit with the same effects: an interaction between manner's and undergoer's manner of articulation as a fixed effect, as well as a random effect of speaker, item and undergoer's manner of articulation within speaker. The model of f_1 at 20 ms before the offset is summarised in Table 5.6, and the interaction between undergoer's and trigger's manner of articulation is shown in the right panel of Figure 5.4.

Table 5.6: Regression coefficients, with standard error and t values for a model predicting the f_1 (in Hz) of the vowel preceding word-final obstruents at 20 ms before the vowel offset. The intercept corresponds to a word-final sibilant followed by a sonorant.

Term	Level	β	SE	t
(Intercept)		668.60	20.55	32.54
Undergoer	cluster	30.85	21.40	1.44
Undergoer	stop	45.40	21.32	2.13
Trigger	voiced obstruent	-5.99	12.37	-0.48
Trigger	voiceless obstruent	52.61	12.16	4.33
Trigger	vowel	-28.56	9.39	-3.04
Undergoer: trigger	cluster: voiced obstruent	-3.98	20.90	-0.19
Undergoer: trigger	stop: voiced obstruent	2.91	16.76	0.17
Undergoer: trigger	cluster: voiceless obstruent	-43.26	20.77	-2.08
Undergoer: trigger	stop: voiceless obstruent	-54.07	16.38	-3.30
Undergoer: trigger	cluster: vowel	35.99	15.91	2.26
Undergoer: trigger	stop: vowel	36.64	13.22	2.77

Summary and residual questions

The results of mixed-effects modelling largely confirm the literature reports concerning the realisation of voicing in word-final prevocalic obstruents. Prevocalic sibilants and clusters were typically realised with an extended portion of vocal fold vibration, and they patterned with sibilants and clusters followed by voiced obstruents in terms of voicing duration and voicing ratio. The effect of voicing in prevocalic sibilants was also manifested by f_1 lowering of the preceding vowel, although in clusters this effect was relatively smaller. Unlike sibilants and stop+sibilant clusters, singleton stops did not undergo voicing in the prevocalic position. Word-final prevocalic stops patterned

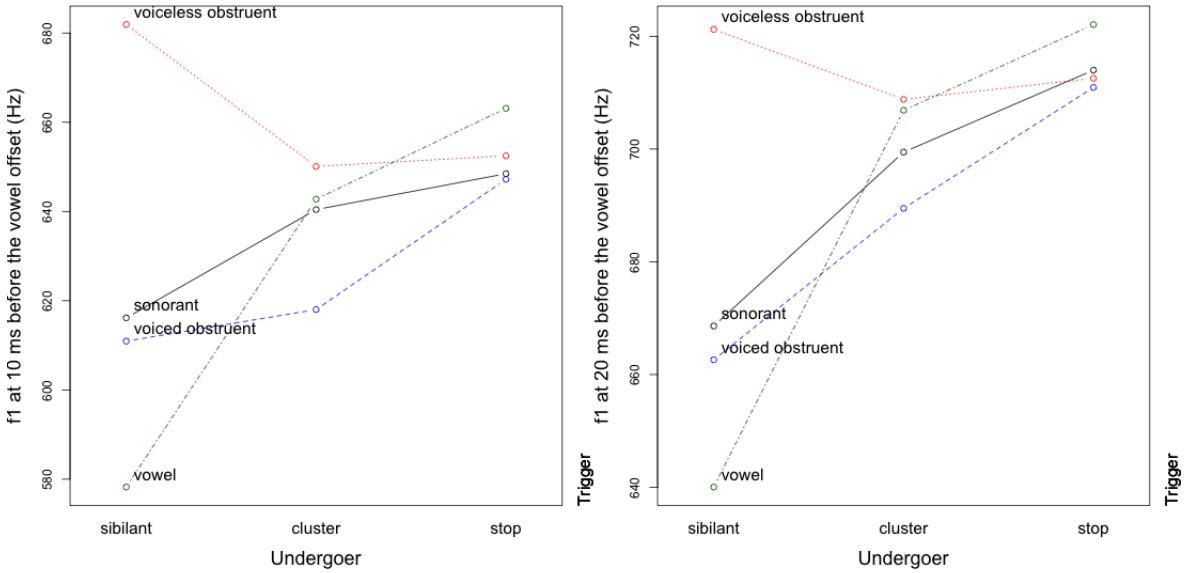


Figure 5.4: Interaction between the undergoer’s and trigger’s manner of articulation in a model of f_1 (in Hz) of the preceding vowel measured at 10 ms before the offset (left), and 20 ms before the offset (right).

with stops followed by a voiceless obstruent with respect to most acoustic predictors, including voicing duration, voicing ratio, closure duration, and f_1 .

Voicing before sonorant consonants was more limited than would be expected based on the literature. Word-final obstruents followed by a sonorant consonant in the next word did show increased voicing duration and voicing ratio, as well as f_1 lowering compared to obstruents followed by a voiceless obstruent in the next word. At the same time, however, average voicing duration and ratio were considerably lower in pre-sonorant obstruents than in obstruents followed by a voiced obstruent, with a difference in means of 20.7 ms (for voicing duration) and 0.36 (for voicing ratio), and high associated t -values.

The effect size in the model of voicing duration and ratio points to prevocalic voicing being categorical in sibilants and clusters, but absent from stops. The results concerning pre-sonorant voicing, however, are less clear. The mean effect of voicing in pre-sonorant obstruents could arise from a variety of factors. First, it could be conditioned by categorical intra-speaker variation, where all speakers voice their pre-sonorant obstruents in an optional but categorical fashion, as previously found for West-Flemish and Poznań Polish. A similar situation could arise from a gradient type of voicing within speaker, where obstruents are partially voiced in pre-sonorant position. Finally, the effect could also follow from intra-speaker variation of either the categorical-but-optional, or the gradient kind. The results of mixed-effects modelling indicate the existence of inter-speaker variation in the data. All of the models presented above were found to fit the data significantly better, once the effect of trigger’s manner of articulation and/or the undergoer’s manner of articulation were allowed to vary within speaker. The degree of inter-speaker variation is considered below in Section 5.3.2.

5.3.2 Individual variation

The results of the regression modelling confirm that individual speakers differed with respect to the effect of undergoer’s and trigger’s manner of articulation, but the differences are too difficult to gauge based on the models’ results. The reason for this is that the manner of articulation of

the trigger interacted with the manner of articulation of the undergoer (as seen for instance in the stop-fricative asymmetry with respect to prevocalic voicing), but it was not possible to include interactions between these two factors as random effects within speaker, as such models would not converge. Consequently, studying the coefficients from the random part of the models allows us to analyse the effect of trigger and/or undergoer on a speaker-by-speaker basis, but not how the two factors interacted within individual participants. In the face of this problem, an informal graphical exploration of the data was carried out, followed by a regression analysis of data subsets. Only voicing duration and ratio were analysed for individual speakers, based on the fact that the duration of vocal fold vibration (absolute and relative to obstruent duration) is the primary phonetic correlate of voicing in Catalan, as reported previously (Cuartero Torres, 2001), and as confirmed for the current data by the results of the regression analysis (Section 5.3.1).

Let us first consider the pattern of means based on raw data. Figure 5.5 shows mean values of voicing duration as a function of the undergoer's and trigger's manner of articulation in individual speakers. All speakers with the exception of B show similar patterns in the relative voicing of prevocalic obstruents with prevocalic sibilants and clusters showing an equal amount or more voicing than sibilants and clusters followed by voiced obstruents. Speaker B was exceptional in showing considerably less voicing than other speakers in all the environments, and by showing relatively less voicing in prevocalic clusters than in clusters followed by voiced obstruents. All the speakers were consistent in having very little voicing in prevocalic stops. In comparison to prevocalic voicing, voicing duration before a sonorant consonant involved a lot more inter-speaker variation. Most speakers showed a limited amount of voicing duration in their pre-sonorant obstruents, although the relative duration of pre-sonorant voicing depending on the undergoer's manner of articulation varied between speakers. Speaker A showed a considerable degree of pre-sonorant voicing in sibilants, equal to the speaker's mean for voicing duration in prevocalic sibilants. Clusters and stops, on the other hand, had considerably less voicing before a sonorant than sibilants in the same position. Speaker G had the greatest mean voicing duration before a sonorant of all speakers, with the greatest voicing duration in pre-sonorant clusters.

Figure 5.6 provides an illustration of the individual strategies in the realisation of voicing ratio. Again, the individual strategies for realising prevocalic voicing are similar in individual speakers, with the exception of Speaker B, who had a much lower ratio in prevocalic clusters than in prevocalic sonorants. Voicing ratio in presonorant obstruents was subject to greater individual variation. For all the speakers, with the exception of G, clusters had a lower voicing ratio before a sonorant than sibilants or stops in the same environment. Speaker G was exceptional not only in having a relatively higher voicing ratio in pre-sonorant clusters than in pre-sonorant stops, but also in having relatively highest voicing ratio out of all the speakers. There were also inter-speaker differences with respect to the relative voicing ratio in pre-sonorant sibilants and stops. For speaker A pre-sonorant sibilants had a very high mean voicing ratio, patterning with prevocalic sibilants, whereas in stops the voicing ratio was limited. For speaker F there was the opposite trend: pre-sonorant stops had the average voicing ratio of 0.096, whereas in sibilants the ratio did not reach 0.5.

The graphical exploration of individual strategies in the realisation of voicing duration and voicing ratio indicates that prevocalic voicing is categorical for sibilants and clusters, with high mean values for every speaker (with the exception of prevocalic clusters realised by speaker B). Most of the inter-speaker variation shown by the regression modelling seems to be contained to obstruents followed by sonorant consonants. As for the nature of variation in the pre-sonorant obstruents across speakers, it involved both the overall degree of voicing (ranging from fairly limited voicing in speaker B, to a considerable degree of voicing in speaker G), as well as the effect of undergoer's manner of articulation, as individual speakers showed different relative levels of voicing in sibilants, clusters and stops.

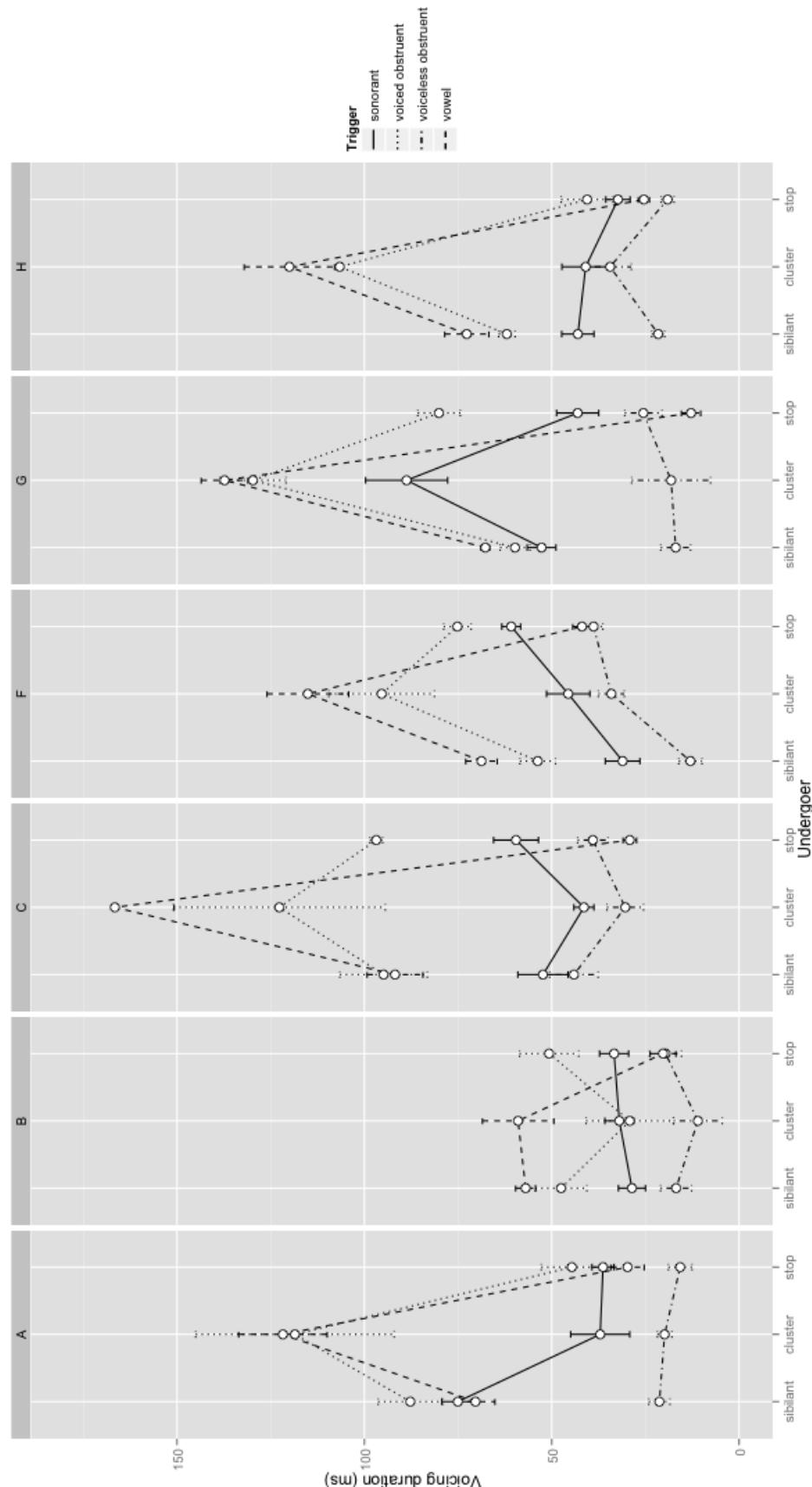


Figure 5.5: Plot of mean voicing duration (in ms) as a function of the undergoer's and trigger's manner of articulation per individual speaker.

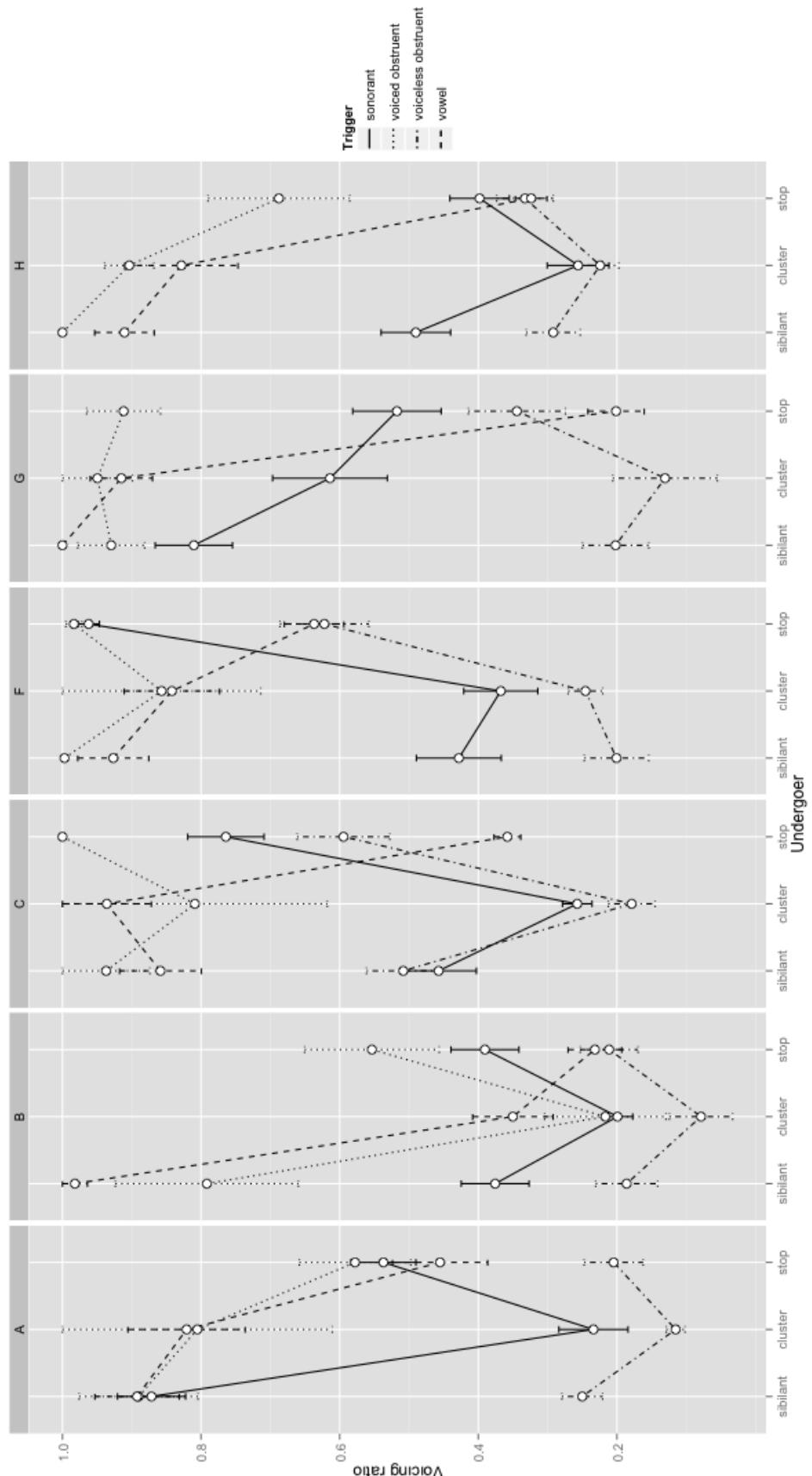


Figure 5.6: Plot of mean voicing ratio as a function of the undergoer's and trigger's manner of articulation per individual speaker.

In order to confirm these informal generalisations based on the apparent patterning of the data, linear mixed regression models of voicing duration and ratio were fitted to the data from the prevocalic ($n=333$) and pre-sonorant ($n=541$) environments. In that way the effect of undergoer's manner of articulation in prevocalic and pre-sonorant voicing could be analysed within speaker. All the models had the same fixed structure, which included only one main effect, that of undergoer's manner of articulation. The random part of the model included a random intercept for speaker and item in all cases. A model with random intercepts would then be compared to a model with a random effect of undergoer within speaker to test whether the inclusion of the random slope improved the model's fit.

The models of voicing duration and ratio before a vowel improved significantly upon including a random effect of undergoer's manner of articulation within speaker (model of voicing duration: $\chi^2=30.91$, $df=5$, $p <0.001$; model of voicing ratio: $\chi^2=34.37$, $df=5$, $p <0.001$). The random coefficients for undergoer within speaker are plotted in Figure 5.7. The fixed coefficients of both models are summarised in Table 5.7.

Table 5.7: Regression coefficients, with standard error and t values for models predicting voicing duration and voicing ratio in prevocalic obstruents. The intercept corresponds to a sibilant.

Voicing duration				
Term	Level	β	SE	t
(Intercept)		70.36	5.77	12.19
Undergoer	cluster	44.18	12.66	3.49
Undergoer	stop	-45.59	8.10	-5.63
Voicing ratio				
Term	Level	β	SE	t
(Intercept)		0.91	0.03	28.70
Undergoer	cluster	-0.13	0.10	-1.32
Undergoer	stop	-0.55	0.09	-6.39

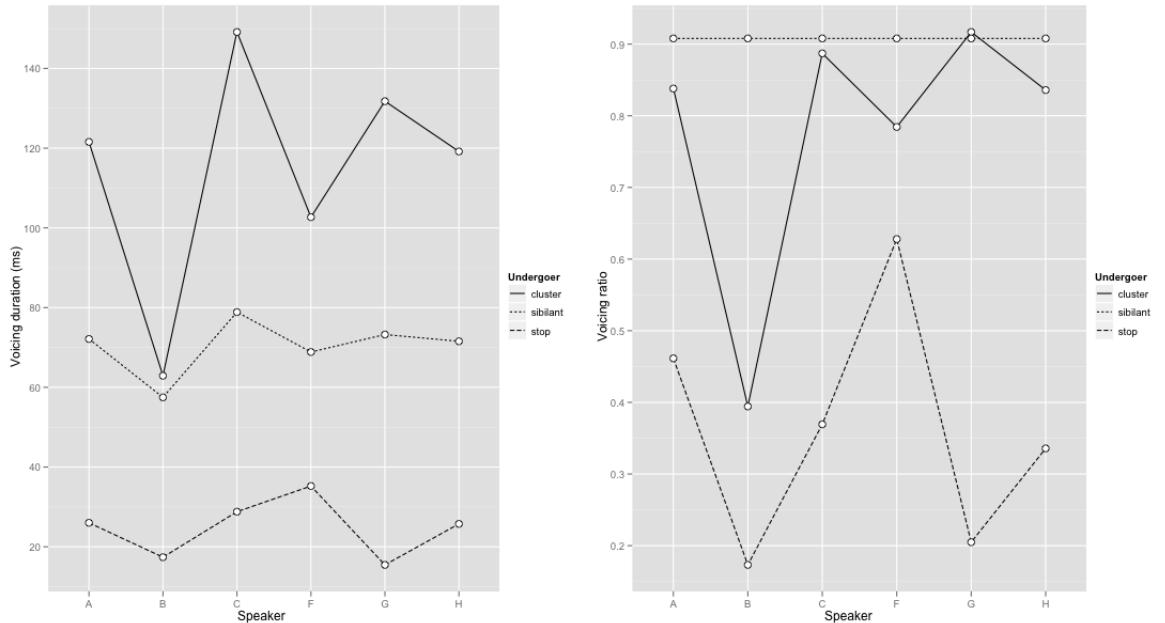


Figure 5.7: Random coefficients for undergoer's manner of articulation within speaker in a model of voicing duration in prevocalic obstruents (left), and voicing ratio in prevocalic obstruents (right).

The main effect of undergoer's manner of articulation on voicing duration shows increased

duration in clusters followed by a vowel compared to prevocalic sibilants, but the mean difference in the voicing ratio between these two conditions was not significant as indicated by the low t -value of -1.32. Stops showed significantly less voicing in terms of both voicing duration and voicing ratio. This general pattern was displayed by all the participants in the experiment with the exception of speaker B, who had a much lower voicing ratio in prevocalic clusters compared to prevocalic sibilants.

The models of voicing duration and ratio before a sonorant consonant also showed a significant improvement in fit when a random slope for undergoer's manner of articulation within speaker was added (voicing duration: $\beta=66.10$, $df=5$, $p < 0.001$; voicing ratio: $\beta=102.71$, $df=5$, $p < 0.001$). The fixed coefficients of the two models are in Table 5.8. Figure 5.8 represents the random coefficients for undergoer within speaker.

Table 5.8: Regression coefficients, with standard error and t values for models predicting voicing duration and voicing ratio in pre-sonorant obstruents. The intercept corresponds to a sibilant.

Voicing duration				
Term	Level	β	SE	t
(Intercept)		46.42	7.35	6.32
Undergoer	cluster	1.51	11.01	0.14
Undergoer	stop	-2.49	9.85	-0.25
Voicing ratio				
Term	Level	β	SE	t
(Intercept)		0.56	0.09	6.26
Undergoer	cluster	-0.23	0.09	-2.47
Undergoer	stop	0.04	0.14	0.25

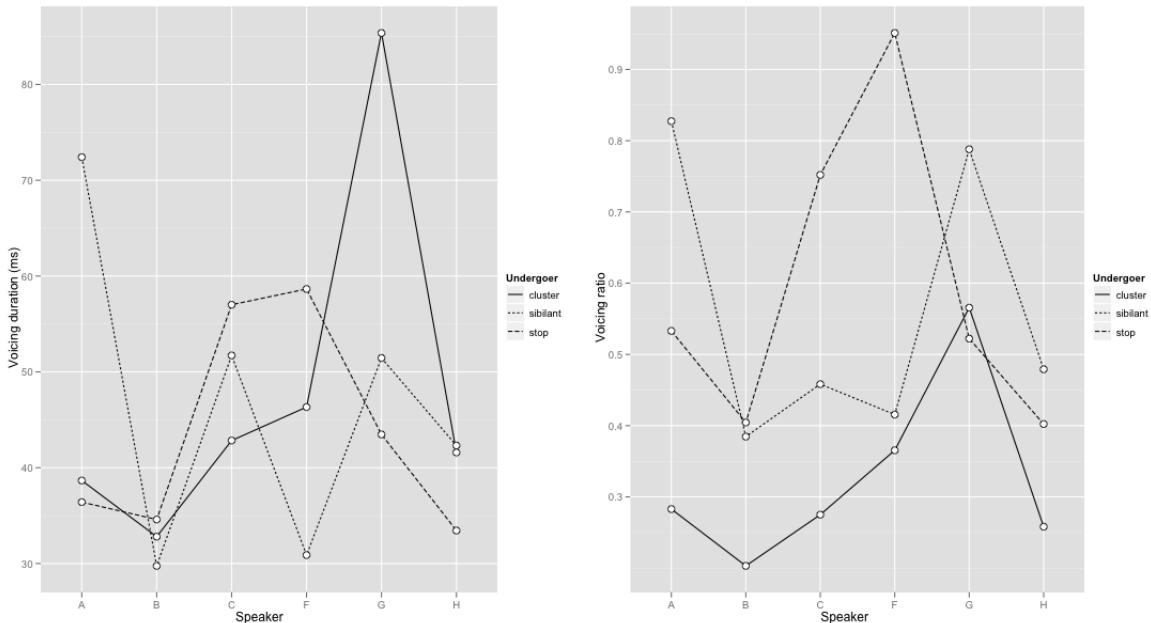


Figure 5.8: Random coefficients for undergoer's manner of articulation within speaker in a model of voicing duration in pre-sonorant obstruents (left), and voicing ratio in pre-sonorant obstruents (right).

The results of pre-sonorant modelling evade reliable generalisations per population. The only relatively stable effect involved lower voicing ratio in pre-sonorant clusters compared to pre-sonorant sibilants ($\beta=-0.23$, $SE=0.09$, $t=-2.47$), and compared to pre-sonorant stops ($\beta=-0.27$,

$SE=0.11$, $t=-2.35$). The difference between stops and sibilants was subject a large amount of inter-speaker variation (cf. Figure 5.8), and no stable generalisations for the population can be discerned, as confirmed by the large standard error and low t -values associated with the difference in means between voicing duration and ratio in pre-sonorant sibilants and stops.

5.4 Discussion

5.4.1 Undergoer asymmetry

The current findings confirm the existence of the effect of an undergoer asymmetry effect with respect to prevocalic voicing, which is problematic for just about any functional explanation of how voicing may come about. While competing hypotheses have been put forward to explain why fricative voicing may be preferred to stop voicing, none of them can straightforwardly account for the occurrence of stop+sibilant cluster voicing while prevocalic singleton stops remain voiceless. Instead, it appears that prevocalic cluster voicing is an arbitrary process that does not originate from functional influences on the language. At the same time, stop+sibilant cluster voicing is not unexpected in a system where voicing applies to sibilants before a vowel. Since the clusters do contain a word-final prevocalic sibilant, that sibilant might serve as an input to a prevocalic sibilant voicing rule, and that process coupled with voice assimilation in obstruents yields the voicing of the entire cluster. This scenario is straightforwardly predicted by any output-oriented formal analysis, such as Optimality Theory, and it is indeed easily modelled, but only if the sibilant voicing is analysed as prevocalic, and not intervocalic. Word-final stop+sibilant clusters do not constitute an input for an intervocalic sibilant voicing rule, and so intervocalic voicing is not expected to extend from singleton sibilants to clusters unless there is a case of rule extension, or analogical levelling. A prevocalic sibilant voicing rule, on the other hand, will produce voicing in word-final prevocalic sibilants and sibilant-final clusters alike.

A possible explanation for how the Catalan pattern might have originated is as follows. The pattern started off as intervocalic voicing that targeted delaryngealised sibilants. The connection between delaryngealisation and voicing is supported by a close match between the distribution of these two processes. While the coincidence between the environments of presonorant/prevocalic voicing and of delaryngealisation has been noted in other languages, Catalan provides a particularly striking example given the application of prevocalic sibilant voicing in prefix-final position (cf. (5)). The application of delaryngealisation and prevocalic/pre-sonorant voicing in the same prosodic positions could potentially be argued to follow directly from independent functional factors. For instance, one could hypothesise that there might be an aerodynamic explanation for why prevocalic voicing may apply word-finally, but not word-medially. One of the cues to voicelessness in fricatives is the presence of high-frequency friction noise, whose production requires high intraoral air pressure (cf. 3.4). Since air pressure is relatively lower at the end of the word, it might affect the production of high-frequency noise, which again impairs voiceless fricative perceptions. As for word-final delaryngealisation, it could be seen as a case of rule generalisation that develops from delaryngealisation in utterance-final position. This scenario is proposed by Blevins (2004, 2006), who also discusses the potential phonetic motivation behind utterance-final devoicing including the presence of laryngeal closing gestures, by the effect of final lengthening, and by poor perceptibility. Given the wide range of functional pressures on sound change modulated by prosody, the co-occurrence of delaryngealisation and prevocalic/pre-sonorant voicing in the same positions could be a mere coincidence. However, the case of Catalan goes against this hypothesis, since prosodically conditioned functional pressures do not apply directly to morphologically defined environments. Therefore, for both delaryngealisation and prevocalic sibilant voicing to apply in the same prefix-final domain only two diachronic scenarios are possible. Either the two processes

started off as prosodically conditioned to then become re-analysed as prefix-final independently of each other, or one of the processes developed from the other. The latter explanation appears more principled, as it accords well with the phonetic interpretation of voicing targets and their role in conditioning coarticulation (Jansen, 2004). Delaryngealised obstruents do not have their own voicing targets, and so they are predicted to be more susceptible to laryngeal coarticulation from the neighbouring sounds. Following the argument previously developed for West-Flemish in 3.4, passive voicing spilling over from the neighbouring sounds may be sustained for longer by delaryngealised obstruents, as no active devoicing gesture is executed to counteract the voicing. This extended portion of voicing is likely to be interpreted by listeners as intentional, and actively targeted in production once it is the listeners' turn to speak. This mechanism may yield a categorical voicing process whose distribution matches closely the environment for delaryngealisation. Passive voicing may also be less perceivable in stops than in sibilants, which could explain the observed synchronic asymmetry between prevocalic stop and sibilant voicing.

The next stage in the development of the Catalan undergoer asymmetries involves rule generalisation (Vennemann, 1972). The generalisation consists of a reinterpretation of intervocalic sibilant voicing as prevocalic. What is interesting about this process is that it involves a gradual loss of the functional motivation that instigated voicing in the first place. From the point of view of diachronic functionalism the presence of the vowel on the right-hand side is just one of the factors involved in triggering the change. The other necessary ingredients include the presence of a preceding vowel, as well as the undergoer being a singleton sibilant of limited duration. All of these play a part in the origin of aerodynamically conditioned passive voicing, as well as in the perception of passively voiced sibilants as voiced. In the synchronic grammar of Catalan, however, the role of the preceding vowel seems to have been reduced, or else it would block voicing in clusters. This development reflects a diachronic transition from a phonetically conditioned process in an environment where it is functionally motivated to a partially arbitrary, but nonetheless productive synchronic generalisation.

Loss of phonetic conditioning as part of rule generalisation has been observed in numerous instances of sound change, going back to Schuchardt (1885 [1972]). It is in fact relatively uncommon for phonological processes to develop in a way that exactly mirrors the phonetic environment where they were motivated in the first place. An example of this kind of generalisation comes from palatalisations. Based on experimental data, Cole & Iskarous (2001) argue that palatalisations are perceptually best motivated in VCV sequences where an obstruent is flanked by two high vowels. However, this conditioning is rarely reflected in the output of sound changes, which tend to only preserve the following high vowel as the palatalising context. Further examples of rule generalisation involve for instance the extension of /l/-darkening from a syllable-based process to a foot-based process (Bermúdez-Otero, 2012, and references therein).

The final stage in the development of the Catalan prevocalic voicing involves rule telescoping (Hyman, 1975), as the output of the voicing rule (voiced prevocalic sibilant) becomes an input to a voice assimilation rule that operates independently in language. In a discussion of rule telescoping, Hyman (1975) gives the example of labial stop spirantisation in Proto Bantu, as illustrated in (9).

- (9) $p \rightarrow s / _ i$

The rule in (9) is synchronically unnatural, but according to Hyman it may result from telescoping involving a series of natural sound changes, as shown in (10).

- (10) $p_i > p^{hi} > p^{si} > t^{si} > si$

The first step of the change involves an aspiration of the labial stop before a high vowel, followed by affrication with a [s] release. This change is followed by place assimilation of the closure to

the release of the affricate, and deaffrication. As the changes feed one another in a serial fashion, the phonetic motivation that led to the individual developments becomes obscured. In that way, rule telescoping may lead to unnatural phonological process that are removed from their original synchronic motivation.

Cases of rule generalisation and telescoping present an interesting problem in reflecting a mixture of phonetic and phonological influences. This is also the case with Catalan prevocalic voicing. On the one hand, the voicing pattern shows traces of functional motivation, as it descends from coarticulation. On the other hand, the synchronically unnatural aspects of prevocalic voicing (partial loss of the conditioning environment and cluster voicing in the absence of singleton stop voicing) indicate an influence of abstract phonological generalisation shaping the process. This brings about the question of where the locus of phonetic and phonological influences is. The diachronic explanation for the Catalan sound pattern proposed above is couched within an evolutionary approach which divorces diachronic and synchronic functionalism. A central idea is that functional factors influence language use, but they do not affect grammars directly. Instead, the primary task of synchronic grammar is seen as producing generalisations that closely match the perceived sound patterns in the ambient language. This task can be affected by internal pressures, such as the cognitive nature of category building and restructuring, as well as by extragrammatical factors that shape the content of the linguistic input, such as lexical frequency. As a result, any synchronic grammar may deviate from its diachronic predecessor on the level of abstract generalisation.

How phonological re-interpretation instantiated by numerous cases of rule generalisation happens exactly is a vexing question. A potential solution is proposed by Hayes (1999), who takes up the issue of indirect phonology-phonetics mappings in relation to natural class behaviour in sound patterns. Hayes considers the relative difficulty in sustaining stop voicing depending on the place of articulation and environment, and argues that the complexity of aerodynamic conditioning of voicing can be translated into an intricate difficulty landscape. Hayes further observes that a direct interpretation of phonetic difficulty thus conceived could give rise to a language with the following potential phonological constraints on voicing.

- (11) A hypothetical phonological constraint (Hayes, 1999)
 - *post-obstruent voiced stops,
 - *[d,g] in initial position,
 - *[g] after oral sonorants

Such restrictions, however, are virtually never found in natural language phonologies, which tend to employ considerably more coarse-grained constraints. To account for this fact, Hayes (1999) proposes that constraint inventories are restricted by functional factors on the one hand, and formal simplicity on the other. In consequence, synchronic grammars are neither entirely functional, nor formal. Instead, grammars reflect phonetic naturalness mediated by an abstract generalisation biased towards structural simplicity.

Interestingly, Hayes's proposal comes with a computational solution concerning the *inductive grounding* of markedness constraints, i.e. discovery of the relevant constraints on the part of the learner. This process is proposed to involve creating a phonetic difficulty map based on the learner's experience with language, followed by a selection of the most effective constraints to generalise over the observed language variation. This mechanism allows for creation of markedness constraints with partial functional grounding. For instance, inductive grounding could give rise to markedness relationship conditioning prevocalic sibilant voicing, similar to the constraints proposed by the Optimality Theoretic analyses of Catalan discussed in 5.1.1.

The proposal by Hayes (1999) that synchronic grammars may be removed from their functional

origins via the influence of formal generalisation is conceptually close to the evolutionary view on the development of grammars pursued in this work. Both approaches require an explanation for the mechanism responsible for the re-analysis of intervocalic voicing as prevocalic. I have suggested that the re-analysis may be partially driven by lexical and cognitive factors. The latter may possibly involve a simplicity bias, as proposed by Hayes (1999), given the frequency with which rule generalisation occurs in sound change. In addition to this, any account of Catalan prevocalic voicing seems to require a mechanism for inductive learning. Strictly functional synchronic solutions cannot account for the synchronic Catalan pattern by interpreting aerodynamic, perceptual and articulatory constraints as active influences in the synchronic grammar formalised as markedness constraints. The reason for this is that in Catalan no convincing case can be made about the actual motivation for the relevant constraints. The issue becomes especially visible when it comes to motivating prevocalic sibilant voicing. Just like the diachronic scenario presented above, a synchronic formal analysis of Catalan would have to account for a pattern of prevocalic, and not intervocalic, sibilant voicing in order to explain why the voicing also applies to clusters. However, it is intervocalic rather than prevocalic voicing that is functionally motivated.

In a diachronic analysis the partial loss of phonetic motivation for voicing can be explained via re-analysis. In a formal synchronic OT grammar capturing the Catalan pattern requires positing a markedness constraint that conditions prevocalic sibilant voicing. Such an analysis is formally possible, but it brings about the question of where markedness constraints come from. The lack of direct functional motivation for prevocalic sibilant voicing goes against the assumption that markedness constraints are grounded and innate. Instead, cases like Catalan seem to necessitate a mechanism for inductive grounding, where the relevant markedness constraints are discovered in the process of acquisition. However, once inductive grounding is allowed to play a role in shaping constraint inventories, the need arises to re-evaluate the role of markedness in phonological grammars. If markedness constraints are discovered in acquisition, then it is likely that the particular set of constraints that each speaker constructs will depend on her own experience. Thus, the membership of the constraint set is likely to vary from individual to individual, and may even be formally unbounded. Consequently, markedness cannot fulfil the role of limiting the number of possible phonological grammars based on constraint permutation (cf. 7.5 for further discussion on this point). This, in turn, forces the question of whether markedness genuinely drives the development of grammar. This assumption does not appear necessary from the point of view of inductive learning. For instance, Boersma (2008) shows that markedness effects may emerge in language from a learning algorithm oriented at closely recreating the faithfulness constraint ranking as present in the input to a child's grammar without stipulating any prior knowledge of markedness or phonetic difficulty. This result squares with the idea of Evolutionary Phonology that sound patterns are most accurately explained in the context of diachronic processes that led to their development coupled with an abstract mechanism used by learners to generalise over the perceived variation, where markedness effects may emerge from the system, but they are not a primary driving force.

5.4.2 Trigger asymmetry

Aside from the issue of manner asymmetries with respect to prevocalic voicing, Central Catalan also presents the problem of the difference between prevocalic and pre-sonorant voicing. This difference manifested itself in three ways: 1) Prevocalic voicing applied to sibilants, and (in all but one speaker) also to clusters, but failed to apply to stops. Pre-sonorant voicing showed no systematic differences depending on the undergoer's manner of articulation. 2) In the environments where it did apply prevocalic voicing was categorical, occurring in the vast majority of potential environments and typically resulted in fully voiced obstruents. Pre-sonorant voicing was more

variable, as indicated by low mean values for voicing duration and ratio compared to obstruents followed by a voiced obstruent. 3) There was relatively little inter-speaker variation with respect to prevocalic voicing, but pre-sonorant voicing varied from speaker to speaker. The variation included both the overall mean of voicing duration per speaker, as well as the within-speaker mean as a function of the undergoer's manner of articulation.

The differences between prevocalic and pre-sonorant voicing confirm the prediction that follows from the three formal analyses of Catalan presented in Section 5.1.1 (Bermúdez-Otero, 2001; Wheeler, 2005; Jiménez & Lloret, 2008) which treat voicing before a sonorant consonant and a vowel as conditioned by a different set of factors. From the point of view of diachronic functionalism however, the vowel-sonorant asymmetry is puzzling. As previously explained in Section 5.1.1, a perception driven re-analysis account of pre-sonorant/prevocalic voicing proposes that the sound change originates as the voicing from the preceding vowel is extended into a word-final obstruent due to the lack of the obstruent's voicing target. The right-hand environment has an effect on the duration of passive voicing, but the influence is mostly due to the absence or presence of a voicing target (e.g. the presence of a following voiceless obstruent will quench passive voicing). Under this view no difference is expected between sonorant consonants and vowels in the conditioning of word-final obstruent voicing. The presence of a robust difference in this respect in Catalan necessitates reconsidering the role that the right-hand environment might play in the initiation of voicing.

Considering the degree of inter- and intra-speaker variation observed in pre-sonorant voicing vis-à-vis the relatively established pattern of voicing before a vowel, it is likely that prevocalic voicing is the older pattern of the two. This allows us to reformulate the question concerning the role of trigger's manner of articulation as a question of why voicing might have applied only before a vowel at some stage, but not before a sonorant. An answer to this question would require an investigation into how the manner of articulation of the sound following a word-final obstruent affects the obstruent's duration and the air pressure. It would need to be conducted in a language that has no manner asymmetries of the kind found in Catalan, as otherwise the results could be confounded by voicing which affects both air pressure and segmental duration. Vowel articulations are relatively more open compared to sonorant consonants, and so their production may involve lower intraoral pressure. In consequence vowels might be potentially more conducive to passive voicing. However, the question of whether such differences indeed exist, and if so, whether they are big enough to trigger an asymmetry of the kind we see in Catalan yet awaits a systematic investigation.

Meanwhile, the acoustic data on pre-sonorant and prevocalic voicing obtained in the current study can still be used to address the one remaining piece of puzzle in Catalan voicing, namely the role of laryngeal specifications, and whether or not the trigger asymmetry could lend itself to a successful representational solution. This issue has been previously raised in the analysis of West-Flemish and Poznań Polish data, and it has to do with the question of whether pre-sonorant or prevocalic voicing involves analysing non-contrastive sonorant sounds as having a voiced target which corresponds to the presence of a [voice] feature in the formal representation. This solution has been proposed in all the consulted formal analyses of pre-sonorant voicing including Catalan (Bermúdez-Otero, 2001; Wheeler, 2005; Jiménez & Lloret, 2008). The basic idea of the proposal is that all phonetically voiced sounds in Catalan are laryngeally specified as [(+)voice], which enables them to spread voicing to the preceding obstruent, subject to a number of positional and manner restrictions.

The proposals concerning redundant [voice] specifications had been made under the assumption that sonorant consonants trigger voicing in a preceding obstruent to the same extent that voiced obstruents do. This assumption is not confirmed by the current data: voicing before voiced obstruents had the characteristics of a categorical process with relatively little variation.

Obstruents in the context of a following voiced obstruent surfaced systematically with an extended period of vocal fold vibration, and they displayed little variation conditioned by the undergoer's manner of articulation or by the speaker. Voicing before sonorant consonants, on the other hand, was on average more limited, and it was highly variable. The robust difference between voiced obstruents and sonorant consonants as triggers of voicing undermines the rationale behind postulating a common laryngeal specification between the two groups. It would perhaps be possible to reconcile a common specification with the observed difference, if the voicing rule made reference to manner of articulation. For instance, voicing before sonorants could be analysed as applying 50% of the time, whereas voicing before obstruents would exceed 90%. However, referencing trigger's manner of articulation within the rule formulation makes the step of assigning laryngeal features redundant. A rule that requires voicing of delaryngealised obstruents before a following sonorant will act identically regardless of whether or not the sonorant is specified for [(+)voice]. Thus, while there might be some theory-driven motivation to equip sonorants with a [voice] feature, such a specification is not required from the empirical point of view.

It is also worthwhile to consider the possibility that laryngeal specifications may play a role in conditioning the asymmetry observed between sonorant consonants and vowels as triggers of voicing. For instance, vowels could be analysed as [(+)voice], which would be consistent with the fact that for the most part they pattern phonetically with voiced obstruents in triggering voicing in preceding sibilants and clusters. Sonorants, on the other hand, would be analysed as laryngeally unspecified, or specified for a different feature, for instance a sonorant specific laryngeal feature [Sonorant Voice] (Avery, 1996; Currie-Hall, 2006; Blaho, 2008). Arguments against pursuing such a solution include phonetic arbitrariness and a learnability problem, as neither phonological contrast nor intrinsic phonetic characteristics motivate this feature system. Sonorant consonants and vowels in Catalan both share the properties of being laryngeally not distinctive, but phonetically voiced. Thus, there are no factors that could motivate a representational asymmetry between sonorants and vowels independently of the aim to capture the pattern formally.

5.5 Conclusion

Central Catalan is unique in the typology of pre-sonorant voicing languages studied in this work in displaying a range of interactions between the undergoer's and trigger's manner of articulation, some of which are subject to inter-speaker variation. These interactions have consequences for an analysis of how the Catalan pattern might have developed diachronically, as well as for how the pattern may be synchronically analysed by the learner. An extension of word-final sibilant voicing before a vowel to stop+sibilant clusters, but not to singleton stops, supports the hypothesis that the pattern developed through an interaction of phonetic and phonological pressures, specifically a phonetic bias for intervocalic sibilant voicing and a phonological reanalysis of this pattern as voicing before a vowel. I have also argued that the product of this sound change is arbitrary from the point of view of synchronic functionalism, and hence it does not lend itself to a non-arbitrary synchronic analysis. This in turn questions the rationale behind the proposals that functional influences or markedness constraints do indeed play a role in the synchronic generalisation of Catalan voicing. Another puzzling aspect of Central Catalan is the asymmetry between vowels and sonorant consonants with respect to how much voicing they trigger in preceding word-final obstruents. Some aspects of this asymmetry have been previously acknowledged in the literature, including the lack of prevocalic voicing in stops. The current data reveal even more complexities with respect to the difference between sonorants and vowels as triggers of voicing, as the two processes apply at different rates to sibilants and clusters. A successful representational solution to this problem is not readily available, at least not without posing further issues that involve

lexical and phonetic arbitrariness in the development of Catalan feature specifications. However, as no direct explanation follows from the proposed diachronic model in its current shape, the observed asymmetry necessitates more laboratory research into the functional factors which may condition different sound change pathways for prevocalic and pre-sonorant voicing environments.

CHAPTER 6

QUITO SPANISH

'How did you go bankrupt?' Bill asked.
'Two ways,' Mike said. 'Gradually and then suddenly.'

Ernest Hemingway
THE SUN ALSO RISES

This chapter presents experimental data on /s/-voicing in Quito Spanish, with a focus on categoricity and gradience. /s/-voicing has attracted considerable interest from theoretical phonologists concerning the overapplication of voicing in word final prevocalic /s/. This process cannot be straightforwardly accounted for as a correspondence relationship to a base form, and hence it provides a major challenge to output-output correspondence as a theory of phonological opacity (Bermúdez-Otero, 2011). However, the argument only holds insofar as word-final prevocalic /s/-voicing is considered a phonological process, as OO-correspondence can account for /s/-voicing under the assumption that it only applies in the phonetics (Colina, 2009). The diverging predictions concerning categoricity and gradience in Quito Spanish /s/-voicing are evaluated against acoustic data collected from seven native speakers in a production study involving speech rate manipulations. As some speakers show increased voicing duration at slower speech rates, while their voicing ratio remains constant, I argue that for these speakers /s/-voicing is a categorical process, and so it ought to be analysed as phonological. This result presents a challenge to the OO-correspondence approach, but can be accommodated within one assuming cyclicity. I then turn to discussing further alternatives to synchronic generative models in accounting for the distribution of /s/-voicing.

6.1 Quito /s/-voicing

The variety of Spanish spoken in Highland Ecuador, including Quito, has been reported by Robinson (1979) and Lipski (1989) to voice word-final /s/ when followed by a vowel in the next word. Another process affecting word-final sibilants, as well as other types of final codas, is resyllabification into an onset. In the word-medial position prevocalic sibilants surface as voiceless. The relevant examples are in (1).

- (1) Prevocalic /s/ in Quito Spanish¹:
- | | | |
|------------|-------------|---------------|
| /gas#akre/ | [ga.za.kre] | 'acrid gas' |
| /gasita/ | [ga.si.ta] | 'gauze, dim.' |

In addition, similarly to other varieties of Spanish, the Quito dialect has coda /s/ voicing before sonorants and before voiced obstruents as shown in (2).

¹All the data in this section come from Bermúdez-Otero (2011).

- (2) /s/ before sonorants and voiced obstruents:

/plasma/	[plaz.ma]	'plasma'
/gas#noble/	[gaz.no.ble]	'noble gas'
/rasgo/	[raz.uqo]	'feature'
/gas#blaNko/	[gaz.βlaŋ.ko]	'white gas'

In prepausal position and before a voiceless obstruent a coda /s/ surfaces as voiceless, as exemplified in (3).

- (3) /s/ word-finally and before voiceless obstruents:

/gas/	[gas]	'gas'
/gas#karo/	[gas.ka.ro]	'expensive gas'

6.1.1 Opacity

The unique and theoretically complex aspect of the Quito system is the positional restriction on pre-vocalic /s/-voicing; pre-vocalic voicing applies in word-final position, but it fails to apply to word-medial tokens. Another way of phrasing this generalisation is in terms of syllable structure, as voicing affects derived word-initial onsets, but not word-medial onsets. Even though word-final resyllabification is not obligatory (it might be blocked in slow speech, or where there is glottal insertion, cf. Lipski (1989)), it applies in most cases. This generalisation is supported by numerous descriptions (Harris 1969, 1983, *inter alia*), and by restrictions on some phonological processes. For example, emphatic trilling (Harris, 1983) is blocked from applying in word-final rhotics followed by a vowel, e.g. *mar Egeo* [ma.re.xe.o], but *[ma.re.xe.o] 'Aegean Sea', while it may optionally apply to rhotics in canonical coda positions, e.g. *mar* [mar]~[mar] 'sea'; *mar Negro* [mar.ne.uqro]~[mar.ne.uqro] 'Black Sea'.

Resyllabification of word-final voiced pre-vocalic sibilants into onsets creates a complex phonological problem, especially to monostratal models of phonology such as Classic Optimality Theory (Prince & Smolensky, 2004 [1993]). The crucial assumption of Classic OT is that phonological processes reflect interactions of constraints on phonological output. One consequence of this approach is that phonological generalisations must be surface-true. This is not the case with Quito /s/-voicing. The generalisation that coda /s/ undergoes voicing before a sonorant overapplies in the case of a final /s/ which surfaces in an onset². One possible way of modelling Quito /s/-voicing in Classic OT is by relegating it from the domain of phonology, as proposed by Colina (2009). Colina argues that word-final sibilants undergo delaryngealisation, to satisfy the restriction against laryngeal licensing in the coda position. As a result, a coda /s/ surfaces as a delaryngealised archiphoneme S in word-final position, as illustrated in the tableau in (4).

- (4) Delaryngealisation of coda /s/ in order to satisfy the coda licensing restrictions³

/los/	LICENSE[lar]	*S	*z	*s	IDENT(voice)
a. loz	*!		*		*
b. los	*!			*	
c. loS		*			*

The delaryngealised archiphoneme has no voicing target in the output of phonology, and thus its voicing is supplied by the phonetics, in accordance with Phonetic Underspecification Theory (Keating, 1988, 1990). The voicing targets are supplied depending on the neighbouring segments.

²For a detailed discussion of how a classic OT analysis fails to derive the Quito Spanish facts the reader is referred to Colina (2009).

³The constraints used in this and the subsequent tableau reflect Colina's (2009) analysis and are used for expository purposes.

In final position, the default realisation will be voiceless, whereas when a laryngeally unspecified sibilant is surrounded by phonetically voiced segments, the voicing might spill over, yielding a voiced sibilant realisation.

The analysis so far accounts for the voiceless realisation of prepausal sibilants and sibilants followed by voiceless stops, and for voiced realisation of coda /s/ followed by sonorants and voiced obstruents. What remains to be explained is why the voicing applies in word-final sibilants followed by a word-initial vowel. Delaryngealisation is superfluous in these cases, as the final sibilant surfaces in an onset, vacuously satisfying the restriction against laryngeal licensing in codas. Colina (2009), however, proposes that in these cases the sibilant undergoes delaryngealisation as well, in order to satisfy a high-ranked output-output constraint which requires identity to the base form (e.g. identity of *los* in *los otros* ‘the others’, to the base form *los*). The tableau in (5) provides an illustration.

- (5) Output-output conditioned delaryngealisation of a word-final /s/ followed by a vowel.

	IDENT-OO (voice)	LICENSE[lar]	*S	*z	*s	IDENT(voice)
a. lo.zo.troS	*!		*	*		
b. lo.so.troS	*!		*		*	*
c. lo lo.So.troS			**			*

A pre-vocalic delaryngealised S is expected to be realised with some voicing through interpolation from the neighbouring vowels, leading to voiced perceptions. Colina (2009) argues that the two mechanisms assumed in her analysis, phonetic underspecification and output-output correspondence, explain the voicing patterns found in Quito Spanish, and account for the ‘gradient and variable’ voicing as observed in the language. However, as pointed out by Bermúdez-Otero (2011), Colina’s (2009) analysis not so much explains the gradient and variable character of /s/-voicing, as crucially predicts it by virtue of assumed representations. The OO-correspondence analysis presented above relegates /s/-voicing from the phonological domain due to theoretical rather than empirical considerations, as phonologically opaque interactions cannot be easily handled within a strictly monostratal approach.

Bermúdez-Otero (2011) argues that whether Quito /s/-voicing is a phonetic process or not is down to empirical evidence. Should empirical evidence show that /s/-voicing is indeed categorical, an output-output relationship between phonetically underspecified members of a phonological paradigm would not suffice to fully account for the Quito Spanish facts. As a potential alternative which could handle categorical /s/-voicing, Bermúdez-Otero (2011) considers a derivational analysis based on lexical strata (Kiparsky, 1982, 1985). Under this approach, the rule of /s/-voicing applies postlexically. Bermúdez-Otero (2011) proposes that coda /s/ undergoes delaryngealisation at the Word Level, conditioned by the inability of codas to license voicing. This process turns the sibilant into a delaryngealised archiphoneme S. This output is further submitted to the Phrase Level of derivation. At this level the delaryngealised S is voiced before a vowel (or any other voiced sound), but interpreted as voiceless before a pause. Resyllabification also operates on the Phrase Level. Crucial to the analysis is the fact that a word-medial pre-vocalic /s/, as in *gasa*, is not a coda at any level of derivation. Consequently, it does not undergo coda delaryngealisation, and thus it does not meet the conditions for Phrase Level voicing, which only applies to the laryngeally unspecified S. The two types of derivation (for word-medial and word-final) /s/ are illustrated in (6).

- (6) A cyclic derivation of /s/ voicing in Quito Spanish (Bermúdez-Otero, 2011)

	[PL[WLGasa]]	[PL[WLGas][WLakre]]	[PL[WLGas]]
Word Level (coda delaryngealisation)	[ga.sa]	[gaS][a.kre]	[gaS]
Phrase Level (assimilation and default)	[ga.sa]	[ga.za.kre]	[gas]

The diverging assumptions made by output-output correspondence and cyclicity with respect to serial derivations lead to diverging empirical predictions in the case of Quito Spanish /s/-voicing. Colina's (2009) analysis crucially predicts /s/-voicing to be gradient, whereas the cyclical account put forward by (Bermúdez-Otero, 2011) can accommodate categorical /s/-voicing. These predictions from both analyses follow straightforwardly from the modular view of the phonetics-phonology interface (cf. Chapter 2), which both authors assume (if the analysis allows for an exclusion of /s/-voicing from the phonological domain, it follows that phonology is considered as distinct from phonetics).

6.1.2 Gradience and categoricity

The existing data on Quito /s/-voicing, being discrete, do not allow us to conclude whether the process is categorical or gradient. Robinson's (1979) data come from interviews he conducted during fieldwork in Ecuador. The participants read a word list containing tokens of pre-vocalic /s/ in different positions with respect to word, and morpheme boundary. The tokens were then auditorily classified by the investigator as either voiced or voiceless, and transcribed. Robinson's description was later supplemented by Lipski's fieldwork which was also based on auditory transcriptions. Prevocalic /s/-voicing in Quito Spanish has also been recently studied by Chappell (2011), who recorded 404 tokens of pre-vocalic /s/-voicing in different prosodic and morphological environments. Chappell's data come from Quito radio stations' archives, and feature local speakers discussing local issues and events. The data were analysed acoustically and labelled as voiced based on the presence of a strong voicing bar. If no such bar was present, the tokens were categorically labelled as voiceless.

Meanwhile, researchers working on Quito /s/-voicing have considered the process gradient. Lipski (1989) comments that the /s/-voicing is 'variable and gradient'. Chappell (2011) notes that /s/-voicing at word-boundaries in Quito Spanish is 'not as categorical as previously claimed'. However, although neither Lipski (1989) nor Chappell (2011) are explicit about this, they seem to use 'gradience' as a cover term for both quantitative and qualitative variation. This kind of ambiguous terminological approach to gradience is sometimes used in the literature, but the difference is crucial. Chappell (2011) presents firm evidence of variability of /s/-voicing, but not of phonetic gradience. In her study, the presence of a voicing bar was recorded for 98% of word-final sibilants followed by a vowel in the next word. When a sibilant was followed by a vowel in the same word, the voicing bar was typically absent; this was the case for 89% of /s/+vowel sequences word-medially, and in 94% of such sequences at the beginning of a word. Since voicing was variably realised in different environments, the numbers confirm that there was optionality. What they do not tell us is whether there were cases of partial voicing, where the voicing bar was present during just some part of the sibilant, which would be symptomatic of a phonetically gradient process. Due to these issues, I distinguish between gradience and optionality in this work, following Pierrehumbert (2006), who also points out that empirical methods used for identifying phonetic gradience critically rely on the use of continuous phonetic data.

6.1.3 The current study

This study sets out to settle the issue of the categorical or gradient nature of Quito Spanish /s/-voicing with two aims in mind. First, the empirical evidence of categoricity or gradience is used to discriminate between two theoretical approaches to phonological opacity, based on their diverging

empirical predictions. Second, in accordance with the goals of this thesis previously laid out, I consider the role of phonetic and phonological factors considered in the diachronic development and synchronic generalisation of /s/-voicing.

The key test employed for distinguishing gradient and categorical phenomena is based on speech rate manipulation, following Solé (1992, 1995). Solé argues that controlled and mechanical properties in speech can be distinguished by comparing the realisation of acoustic cues across different speech rates (cf. 2.6). In this study I use the speech rate test to compare the effect of speech rate manipulation on different phonetic exponents of voicing, including the duration of glottal fold vibration, as well as the effect of speech rate on the voicing ratio. If Quito /s/-voicing is a gradient phenomenon, and the observed voicing is due to gestural overlap with the neighbouring voiced sound, approximately the same amount of voicing is expected to be found across different speech rates. Additionally, there should be observable differences in voicing ratio depending on speech rate, as the ratio will increase with a decrease in segment duration brought about by faster speech. If /s/-voicing is categorical, however, the voicing ratio is expected to remain relatively stable across different speech rates, in which case the voicing duration varies proportionally to the variation in segment duration. I test these predictions on the Quito Spanish data, paying close attention to how the result is related to individual variation.

6.2 Materials and method

The test items included sequences of word final /s/ followed by a vowel in the next word, as well as of /s/ followed by a sonorant consonant within the same word and across a word boundary (nasals were used in all cases). I shall henceforth refer to these environments as the Quito /s/-voicing contexts. Even though it is only the word-final pre-vocalic /s/-voicing that distinguishes Quito Spanish from other Spanish dialects, the pre-vocalic and the pre-sonorant environment are similar in that they involve voicing before a segment which is not itself contrastively voiced. There were six tokens per condition.

In addition to the test items, the realisation of /s/ was studied in a range of baseline contexts. The Quito dialect, like most other varieties of Spanish, does not have a lexical voicing contrast in sibilants (Robinson, 1979). In order to establish a voicing baseline the realisation of /s/ in the test items was compared to the realisation of /s/ before a voiced and voiceless stop within the same word. The latter condition involved only 5 items (as opposed to 6) due to lexical limitations. In addition, 6 items of word-final /s/ followed by a voiced obstruent in the following word were included, as well as six items of a word-medial pre-vocalic /s/. Example of the test and baseline items are in Table 6.1.

Table 6.1: Example test and baseline items. The baseline items are indicated by shading.

	Word-internal		Word-final	
Prevocalic	1. <i>gasita</i>	‘gauze (dim)’	3. <i>gas acre</i>	‘acrid gas’
Pre-sonorant	2. <i>entusiasmo</i>	‘enthusiasm’	4. <i>gas noble</i>	‘noble gas’
Pre-voiced obstruent	5. <i>esbozo</i>	‘plan, sketch’	7. <i>marchas buenas</i>	‘good marches’
Pre-voiceless obstruent	6. <i>obispo</i>	‘bishop’	—	—

All the test and baseline items were embedded within a carrier phrase, as exemplified in (7).

- (7) The carrier phrase
Diga ‘gas acre’ otra vez.
‘Say ‘acrid gas’ one more time.’

Eight speakers participated in the experiment: four males aged 16-25, and four females aged 21-28. They were all born in Quito, where they also lived at the time the experiment was conducted. None of the speakers reported a history of speech or hearing deficiency. The data from one of the speakers (speaker 6) were discarded, as she produced multiple pauses and hesitations, and was judged to be markedly less fluent than others during the analysis of the recordings.

The recordings were made in a quiet room on a Marantz PMD 620 solid state recorder, with a Sennheiser PC 131 headset. The sampling rate was 48 kHz.

Four repetitions were recorded per speaker, and each repetition was followed by a short break. The stimuli were randomised for each repetition, and presented to the speaker one at a time on a computer screen. For the first two repetitions the speakers were instructed to read at a comfortable pace, and encouraged to correct themselves if they were unhappy with their pronunciation. For the second two repetitions the participants were asked to say the sentences fast, as if they were saying them to someone who is about to leave the room. Altogether 41 (test items)*4 (repetitions) * 7 (speakers) = 1148 tokens were collected and analysed acoustically. 45 tokens (3.9% of the data) were discarded due to disfluencies, reading errors, glottalisations, etc., leaving 1103 tokens for statistical analysis.

The data were analysed acoustically in Praat (Boersma & Weenink, 2010) on a 5ms Gaussian window. The sibilant tokens were segmented and annotated manually. The following measurements were taken based on the annotation.

1. Duration of vocal fold vibration during frication (voicing duration). An extended period of vocal fold vibration is a common voicing cue in fricatives. The degree of fricative voicing can be quantified in terms of the duration of vocal fold vibration. This approach has been previously used by Chappell (2011), who reports that varying degrees of voicing duration are found in the /s/-voicing context in Ecuadorian Spanish. The duration of voicing during frication was identified based on the presence of the voicing bar on the spectrogram. In ambiguous cases where the voicing was fading out gradually during frication, the offset of voicing was placed where the final pulse was detected by the Praat pitch tracker (using the standard settings).
2. Fricative duration. Voiced fricatives have been found to be shorter than voiceless ones in a number of languages, including Dutch (Slis & Cohen, 1969) and English (Crystal & House, 1988), *inter alia*. The potential effect of voicing on fricative duration in Ecuadorian Spanish was evaluated in the current data based on comparing fricative duration in various voicing contexts. Measurements of fricative duration were also used to calculate the ratio of voicing to frication, with a view to assessing the effect of speech rate on voicing duration as opposed to voicing ratio. Fricatives were identified based on the presence of high frequency noise.
3. Intensity. Gradoville (2011) reports that gradient differences in voicing in Argentinian Spanish are reflected in fricative intensity. I followed Gradoville's method in calculating the intensity difference which also serves as a normalisation strategy. Mean intensity was measured for the low frequency portion of the fricative (filtered from 0 to 900Hz) and for the unfiltered fricative. The total intensity value was then subtracted from the low-frequency intensity to calculate the intensity difference.
4. Duration of the carrier phrase. Duration of the latter portion of the carrier sentence (the duration of the phrase *otra vez*) was measured for each sibilant token in order to compare the speaking rate across different speakers, and to quantify the extent of experimental speech rate manipulation on the speakers' pronunciation.

All the duration measurements were taken in ms. Intensity was measured in dB. Unlike in other experiments on pre-sonorant voicing reported in this work, no measurements of preceding

vowel duration or formants were taken for the Quito Spanish data. The reason for this is that lexical restrictions limited the possibility to control for the preceding vowel, or even vowel height, which was likely to introduce confounds in vocalic measurements.

6.3 Results

6.3.1 Phonetic realisation

Initial exploration of the phonetic data reveals that there is a considerable degree of inter- and intra-speaker variation with respect to the duration of glottal fold vibration during frication in the /s/-voicing context. Individual pronunciations of /s/-voicing fricatives ranged from voicing present throughout the fricative to realisations with very limited voicing. The spectrograms in Figure 6.1 provide an example of fully voiced realisations with voicing present throughout frication.

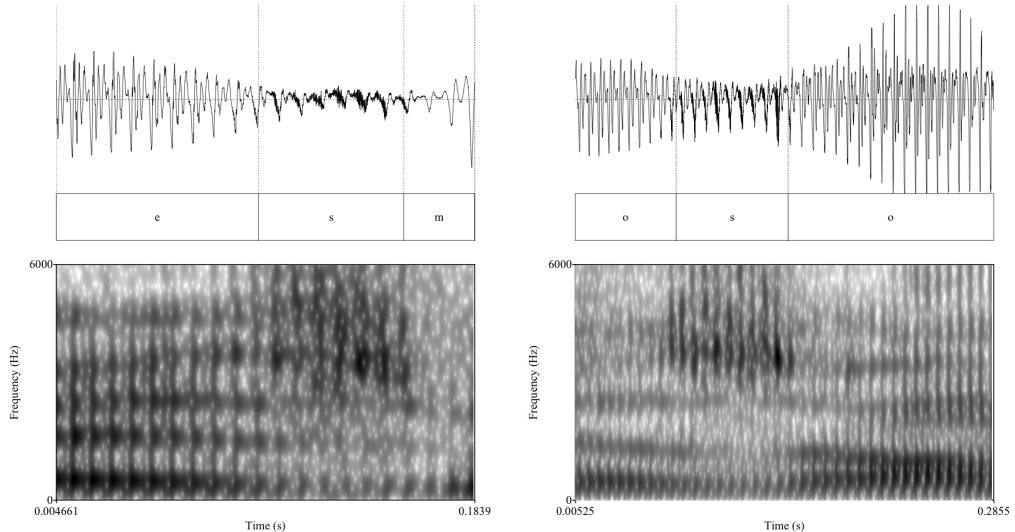


Figure 6.1: Fully voiced /s/ in *esmoquin* (left) and *muchos hombres* (right) pronounced by Speaker 1

However, instances of partial voicing were also found before a sonorant and word-finally before a vowel. In the partial voicing cases the sibilant was always realised with a short voicing tail continuing from the preceding vowel. The voicing would then stop and only resume with the onset of the following sonorant consonant or a vowel. Figure 6.2 illustrates two such realisations.

Realisations of coda /s/ followed by a voiced stop varied between partial and complete voicing, as exemplified in Figure 6.3. Before a voiceless stop, and in word-medial onsets there would be very little voicing, or no voicing during frication, as illustrated in Figure 6.4.

6.3.2 Statistical analysis

Quito /s/-voicing contexts vs. other environments

The first step in the statistical analysis consisted involved an exploration of how various acoustic correlates of voicing are realised depending on the prosodic and the segmental environment. Given the scarcity of experimental evidence on Quito Spanish /s/-voicing, the aim was to verify the empirical status of /s/-voicing before sonorants and prevocalically in the external sandhi context, compared to sibilant voicing elsewhere. Four acoustic correlates of voicing were analysed using conditional inference trees: voicing duration, voicing ratio, duration of the sibilant, and the

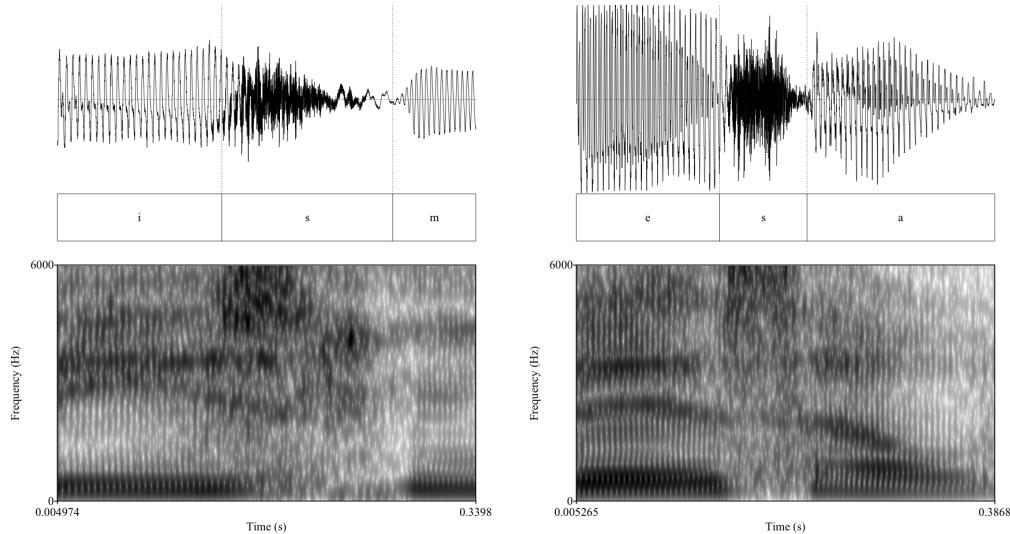


Figure 6.2: Partially voiced /s/ in *budismo* (left) and *tres autores* (right) pronounced by Speaker 5

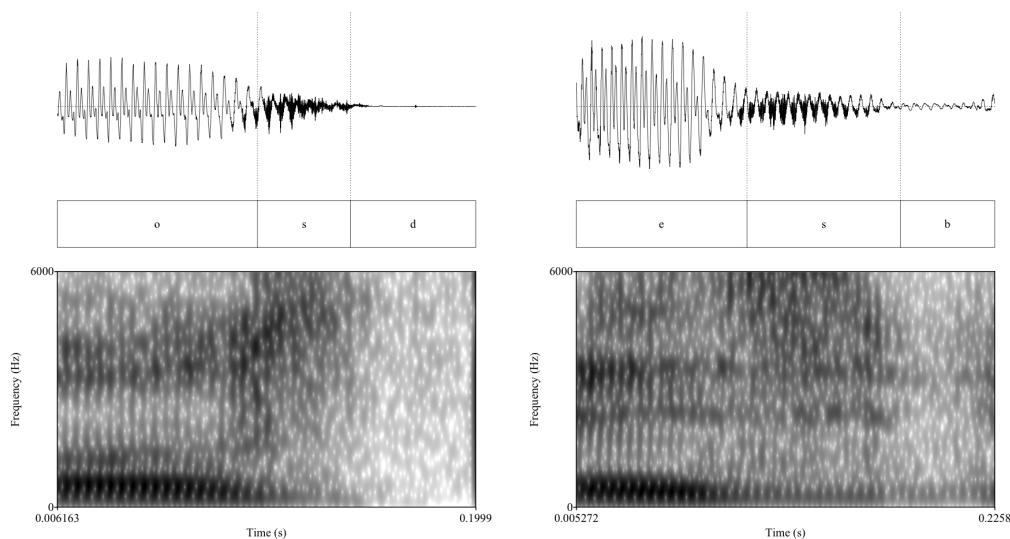


Figure 6.3: Partially voiced /s/ in *posdata* (left), and fully voiced *presbitero* (right) pronounced by Speaker 5

intensity difference. The analysis was performed in R (R Development Core Team, 2005), version 2.13.1, using the party package (Hothorn et al., 2006).

Conditional inference trees are a non-parametric recursive partitioning technique. The algorithm selects the predictor that provides the best binary split for the data, performs the split, and repeats the procedure until no further significant splits can be made. The output is a regression tree which provides a visualisation of how the value of the dependent variable can be predicted. The recursive aspect allows for insight into the hierarchical structure of predictors, whereas the binary splits allow one to identify clusters of data. This set of properties makes conditional inference trees a useful tool for discovering structure within the data.

Voicing duration Figure 6.5 presents a conditional inference tree for voicing duration measured in ms based on a model with two predictors: condition and rate. The first split separates Conditions

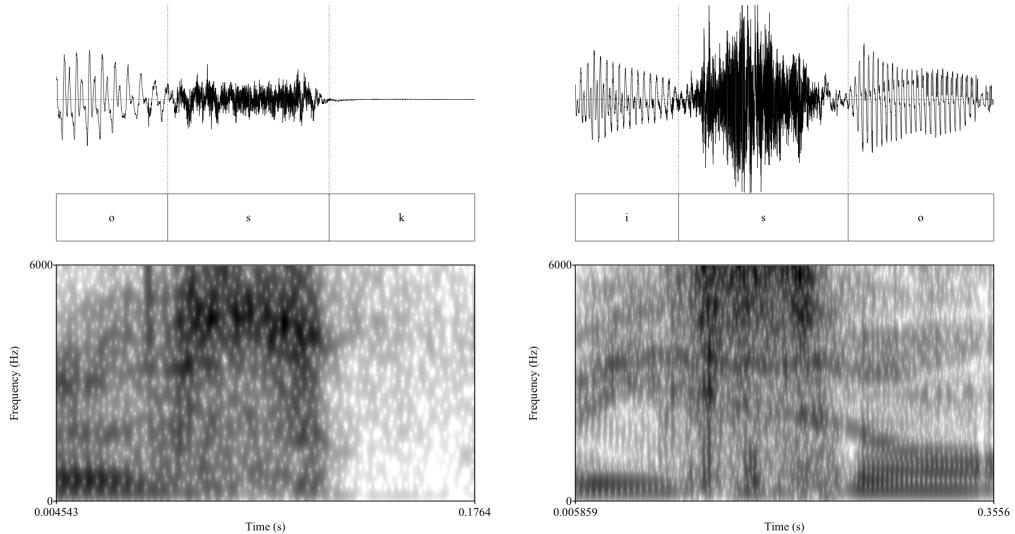


Figure 6.4: Limited voicing of /s/ in *microscopio* (left) and *aviso* (right) pronounced by Speaker 5

1 and 6 from the remaining contexts. Condition 1 involved word-medial onsets (*casita*), while Condition 6 involved sibilants followed by voiceless stops (*obispo*). Thus, the split confirms that these two conditions are characterised by significantly shorter vocal fold vibration compared to sibilants in the remaining contexts. The remaining contexts span the /s/-voicing cases, as well as sibilants followed by voiced stops. Within this group there is a significant split conditioned by speech rate, with relatively shorter voicing duration within the fast speech rate condition. Within the normal speech conditions there is a further split separating Conditions 5 and 7 from Conditions 2, 3, and 4. Conditions 5 and 7, where the voicing duration was relatively greater, had a voiced stop following a sibilant. Conditions 2, 3, and 4 are the /s/-voicing conditions, involving sibilants followed by sonorants (Condition 2: *entusiasmo*, Condition 4: *gas noble*) or a vowel in the following word (Condition 3: *gas acre*). The model confirms that the /s/-voicing conditions cluster together with respect to voicing duration, since no further significant splits were made separating Condition 2, 3, and 4.

Voicing ratio Figure 6.6 presents a conditional inference tree for the ratio of voicing duration to the duration of the sibilant. The results are very similar to those of the model of voicing duration as far as /s/-voicing is concerned. The first split separates the /s/-voicing contexts as well as sibilants followed by voiced stops from those sibilants where no voicing is expected, i.e. word-medially before a vowel (Condition 1: *casita*) and before a voiceless stop (Condition 6: *obispo*). The further splits in the left branch of the tree are identical as in the case of voicing duration. The fast speech rate was characterised by an overall higher ratio, and there were no further splits depending on condition within the speech rate. Within the normal rate, however, there was a split between sibilants followed by voiced stops (Condition 5: *esbozo*, Condition 7: *marchas buenas*) and the /s/-voicing environments (Condition 2, 3, and 4). The former were characterised by a relatively higher voicing ratio compared to the latter. The /s/-voicing conditions were again found to cluster together, with no further splits.

Fricative duration A tree for fricative duration (Figure 6.7) represents a different aspect of the data structure. The first split singles out Condition 1 (*casita*), which involved a word-medial onset, as having the relatively greatest duration. Within this condition there was a further split determined by speech rate, with increased fricative duration in normal rate as compared

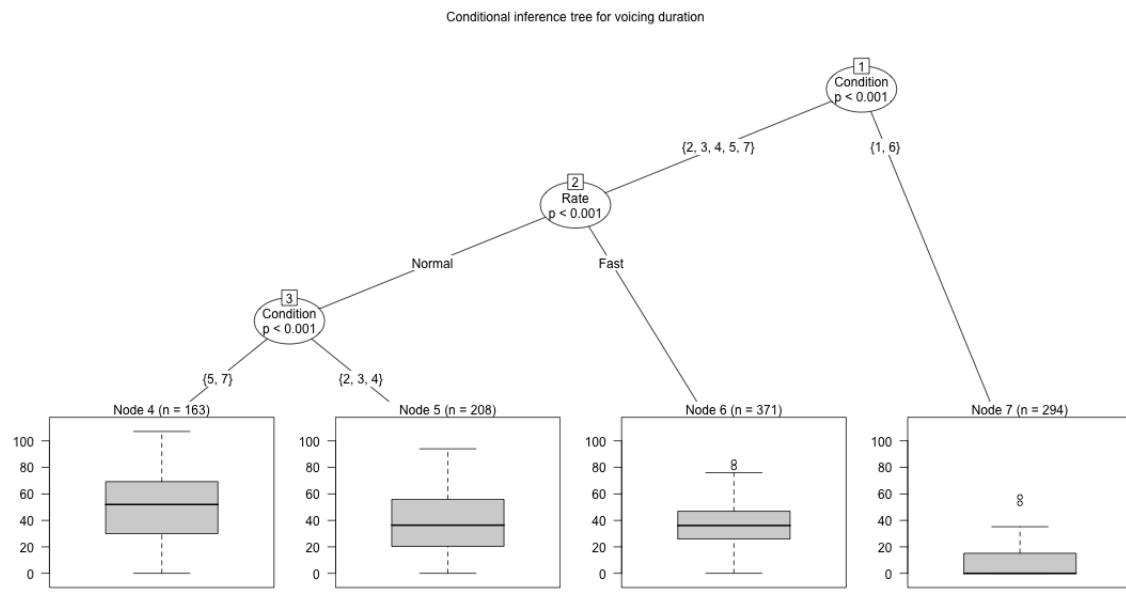


Figure 6.5: Conditional inference tree for voicing duration measured in ms based on a model with two predictors: condition and speech rate.

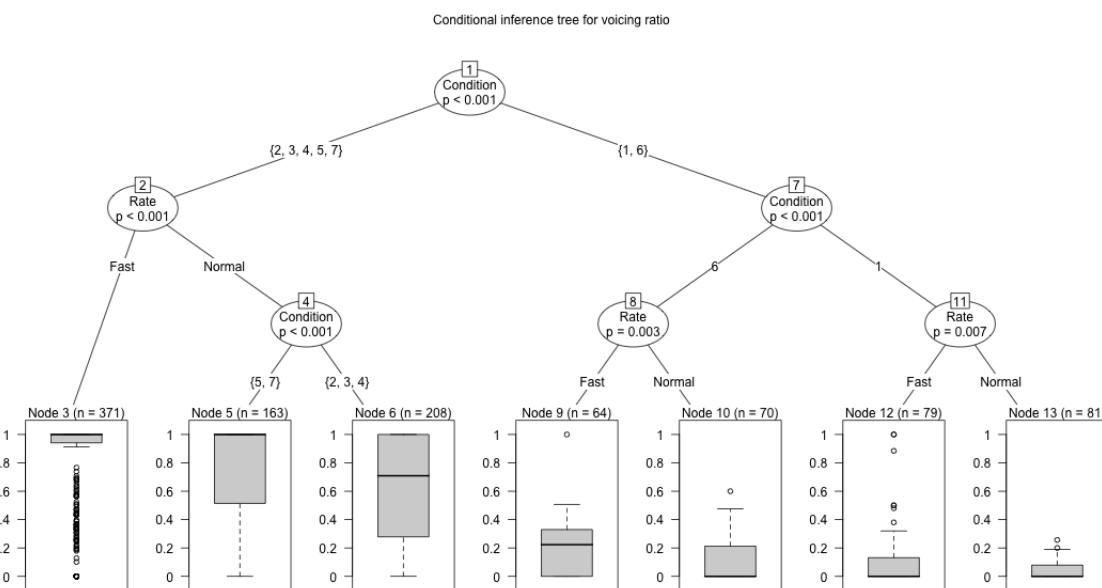


Figure 6.6: Conditional inference tree for voicing ratio based on a model with two predictors: condition and speech rate.

to fast. Similarly, speech rate was responsible for the second split within the remaining conditions. However, the algorithm did not converge at that point, producing further splits. Conditions 4 and 7 clustered together within fast and normal speech rate. Both of these conditions involved word-final sibilants followed by a consonant (sonorant in Condition 4, voiced stop in Condition 7). Word-final pre-consonantal sibilants were relatively shorter than word-medial sibilants as well as sibilants in derived onsets (Condition 3: *gas acre*). Furthermore, sibilants in Condition 6 (*obispo*) were relatively longer than sibilants in Condition 2 (*entusiasmo*), 3 (*gas acre*) and 5 (*esbozo*) within normal speech rate.

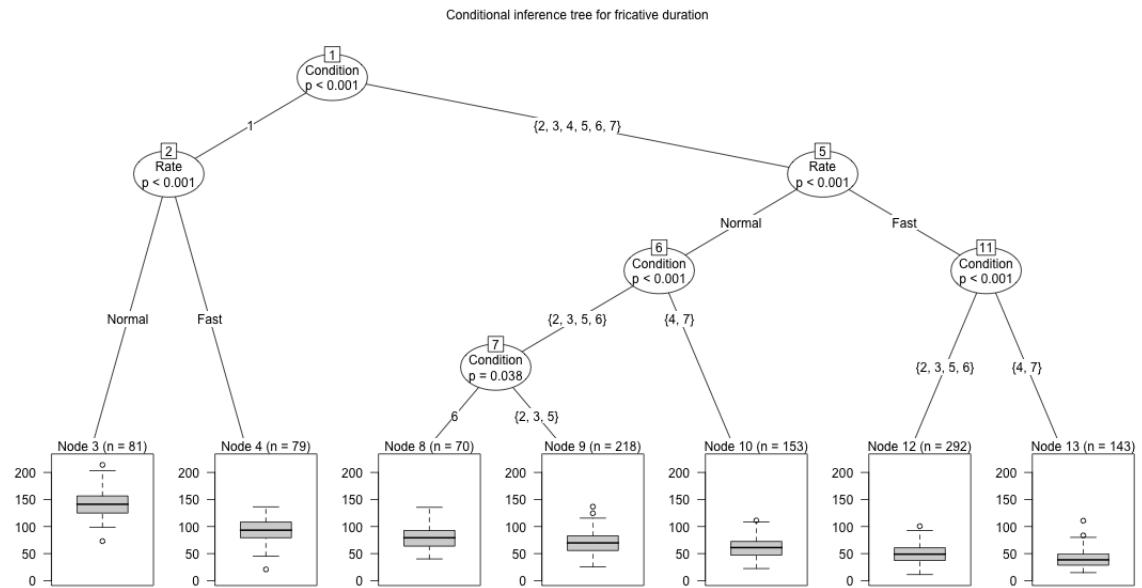


Figure 6.7: Conditional inference tree for fricative duration measured in ms based on a model with two predictors: condition and speech rate.

Intensity difference Figure 6.8 shows a tree for intensity difference measured in dB. The difference was calculated by subtracting mean intensity of the fricative from the mean intensity of its low frequency portion. Relatively smaller differences are expected in fricatives with a high intensity low-frequency portion, which is associated with voicing. This prediction is confirmed by the first split in Figure 6.8, which indicates that word medial prevocalic sibilants (Condition 1: *casita*) and sibilants followed by voiceless stops (Condition 6: *obispo*) had a relatively larger intensity difference compared to all other sibilants. Within the remaining conditions there was a further split which singled out word-medial presonorant sibilants (Condition 2: *entusiasmo*). Within this condition there was a difference between normal and fast speech condition, with smaller intensity difference in the former case. Within the remaining conditions there was no effect of speech rate.

Summary The partitioning analysis reveals a number of important generalisations concerning voicing in Quito Spanish sibilants. First, the partitioning confirms that the /s/-voicing environments involve increased voicing duration and voicing ratio when compared to the baseline involving word-medial prevocalic sibilants, or sibilants followed by voiceless stops. At the same time, however, the /s/-voicing contexts did not pattern consistently with sibilants followed by voiced stops. The two groups showed similar effects on voicing duration and voicing ratio, but only within fast speech rate. In the normal speech rate condition, sibilants followed by voiced stops showed higher voicing ratio and greater voicing duration than sibilants in the /s/-voicing contexts.

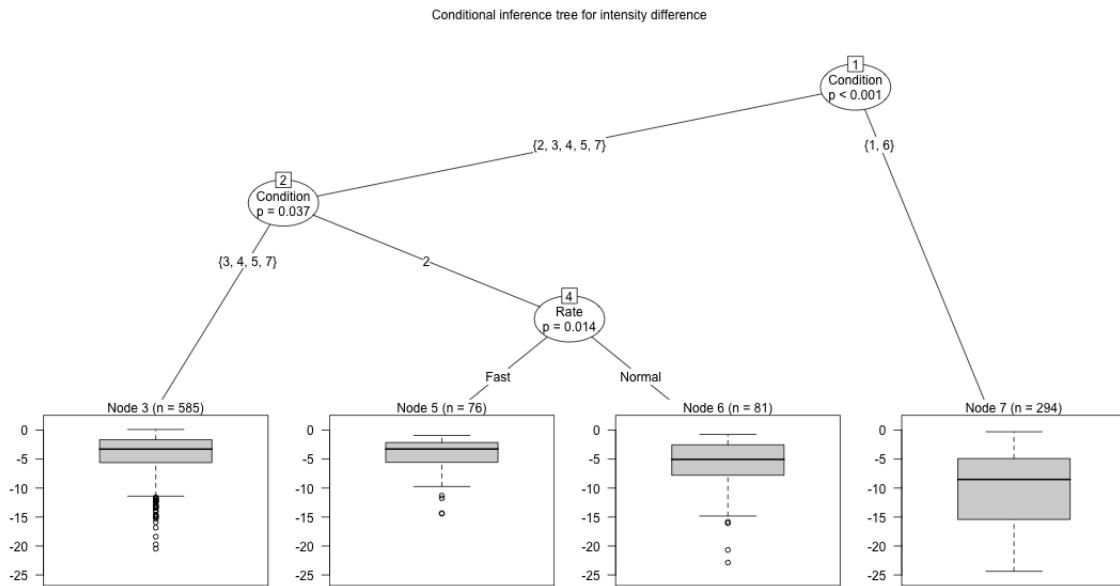


Figure 6.8: Conditional inference tree for intensity difference measured in dB based on a model with two predictors: condition and speech rate.

The effect of speech rate confirms that /s/-voicing is sensitive to speech rate manipulations. A comparison of nodes 3 and 6 in Figure 6.6 suggests that the voicing ratio increases for the /s/-voicing cases with an increase in speech rate. Comparing nodes 4 and 6 in Figure 6.5, on the other hand, it appears that the duration of voicing is relatively unaffected by speech rate manipulations.

Condition and speech rate patterned somewhat differently in the models of fricative duration and intensity difference. Fricative duration appears to be mostly conditioned by prosodic factors, as the successive splits separate word-medial onsets from the remaining cases, and further reveal the clustering of word-final coda sibilants. In addition, the model of fricative duration confirms that fricatives were relatively longer in normal speech rate, as compared to fast. The intensity measure confirmed a difference between sibilants where no voicing is expected (word-medial prevocalic and pre-voiceless stops) from the remaining cases, but it did not produce a cluster of the /s/-voicing cases, and it was largely insensitive to speech rate manipulations.

Speech rate effects and individual variation

The results considered thus far invite a gradient interpretation of /s/-voicing. The voicing duration and ratio are shorter in the case of /s/-voicing than when a voiced stop follows within normal speech rate. In addition, the /s/-voicing items appear to maintain a relatively stable voicing duration across different speech rates. This, coupled with fricative shortening conditioned by faster speech, leads to an increase in voicing ratio in the /s/-voicing contexts in the fast speech rate condition, as compared to normal speech rate. As a result, the /s/-voicing cases pattern with sibilants followed by voiced stops in fast speech. However, it remains to be seen whether this result obtained for the population of participants in the experiment also holds for the individual speakers separately. In order to assess the effect of inter-speaker differences in the current data, a mixed-effects regression analysis was performed on the /s/-voicing data.

The analysis of individual variation was run on a subset of the data involving the /s/-voicing environments ($n=428$) only. The partitioning analysis presented in Section 6.3.2 confirms that /s/-voicing shows unique speech rate effects compared to other contexts tested in the study, and so subsetting the data for analysis of speech rate was done to avoid confounds from non-/s/-voicing

environments. A series of models were fitted to the /s/-voicing data, using the lme4 package (Bates & Maechler, 2009). The dependent variables in the subsequent models were: voicing duration, voicing ratio, fricative duration and intensity difference. The fixed predictors included in the two models were the condition (word-medial pre-sonorant, word-final prevocalic, word-final pre-sonorant) and speech rate (fast vs. normal). Condition was not expected to interact with speech rate within the /s/-voicing cases, so no such interactions were left in the final models presented below. However, in all cases it was verified whether the inclusion of an interaction improved the fit of the model according to the log-likelihood test. All the models had random intercepts for speaker and item, allowing for speaker-specific and item-specific variation from the overall mean. In addition, the inclusion of a random slope for speech rate within speakers was considered. For every dependent variable a model with a random slope for rate within speaker was compared to a model with a random-intercept only model in a log-likelihood test. In cases where random slopes were found to improve the fit they were retained in the model. It was further considered whether the fit would improve further given an inclusion of a random effect of condition within speaker and random effect of rate within item.

Voicing duration For voicing duration a model with a random effect of rate within speaker was found to provide a significantly better fit than a model with a random intercept only, when compared in a log-likelihood test ($\chi^2=19.36$, $df=2$, $p < 0.001$). The intercept corresponded to fast speech rate. The improvement in fit with an inclusion of a random slope confirms that speakers varied not only with respect to voicing duration in the fast rate, but also with respect to the adjustment they made in voicing duration when switching between normal and fast speech. The effects for individual speakers are presented in Figure 6.9. Two opposing tendencies can be observed in the population. Speakers 1, 2, 7 and 8 produced more voicing at normal as opposed to fast speech rate. The effect was greatest for speaker 1, and smallest for speaker 7. For speakers 3, 4, and 5 the duration of voicing increased at fast speech rate. This effect was largest for speaker 5, and smallest for speaker 4.

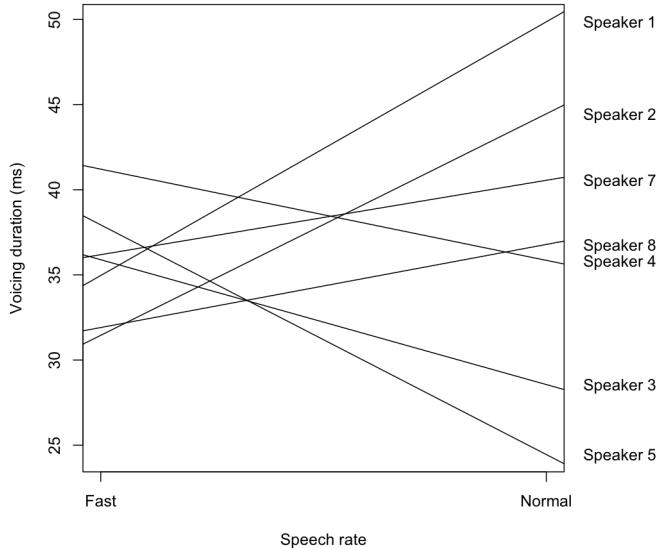


Figure 6.9: Coefficients of rate effects by speaker in a model with voicing duration (in ms) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

No further improvements in the model fit were obtained by adding random slopes for condition

Table 6.2: Regression coefficients, with standard error, and t values for a model predicting the duration of voicing in the /s/-voicing environments, measured in ms. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. *entusiasmo*) at a fast rate.

Term	Level	β	SE	t
Intercept		35.65	2.47	14.40
Rate	Normal	1.57	4.78	0.33
Condition	3 (<i>gas acre</i>)	3.31	2.36	1.41
Condition	4 (<i>gas noble</i>)	-5.04	2.28	-2.21

Table 6.3: Regression coefficients, with standard error, and t values for a model predicting the duration of voicing in the /s/-voicing environments, measured in ms. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. *entusiasmo*) at a fast rate.

Term	Level	β	SE	t
Intercept		0.81	0.06	13.22
Rate	Normal	-0.23	0.09	-2.41
Condition	3 (<i>gas acre</i>)	0.004	0.06	0.08
Condition	4 (<i>gas noble</i>)	0.05	0.05	0.95

within speaker and rate within word. Thus, only the random slope for rate within speaker was retained in the final model. Table 6.2 presents a summary of the fixed effects. Voicing duration was greater word-medially before a sonorant (*entusiasmo*) than in word-final pre-sonorant fricatives (*gas noble*). Word-finally before a vowel the voicing duration was greater still, although the relatively low t -value ($t=1.41$) suggests that the difference between that condition and word-medial pre-sonorant fricatives might not be significant. There was no significant effect of speech rate for the population, given that the value of the standard error exceeded the value of the β coefficient. This finding is consistent with the large amount of individual variation found in the model; once the individual opposing tendencies are taken into account, no significant generalisation can be made for the population.

Voicing ratio Similarly to the case of voicing duration, the modelling of voicing ratio showed a significant improvement to the fit once speech rate was considered as a random effect within speaker ($\chi^2=31.06$, $df=2$, $p < 0.001$). The two models compared had rate and condition as fixed effects, and random intercepts for speaker and item. Figure 6.10 presents the slopes and intercepts for individual speakers' voicing ratio. For speakers 3, 4 and 5 there was a large difference between normal and speech rate, with a higher voicing ratio in the latter case. Speakers 1, 7 and 8 also had a higher voicing ratio in fast speech rate, but for those speakers the adjustment between fast and normal speech rate was considerably lower, with a smaller difference in voicing ratio between fast and normal speech. Speaker 2 was atypical in producing a higher voicing ratio at normal speech rate as opposed to fast.

No further improvements in the model fit were obtained by adding random slopes for condition within speaker and rate within word. Table 6.3 presents the fixed effects structure of the model with fixed effects of rate and condition, and random effects of speaker, word and rate within speaker. There was a negative effect of rate, with higher voicing ratio at fast rate compared to normal. There was very little effect of condition on voicing ratio, as indicated by small β - and t -values.

Fricative duration Similarly to the case of voicing duration and ratio, individual effects were considered in mixed-effects modelling of fricative duration for the /s/-voicing data. Just like in

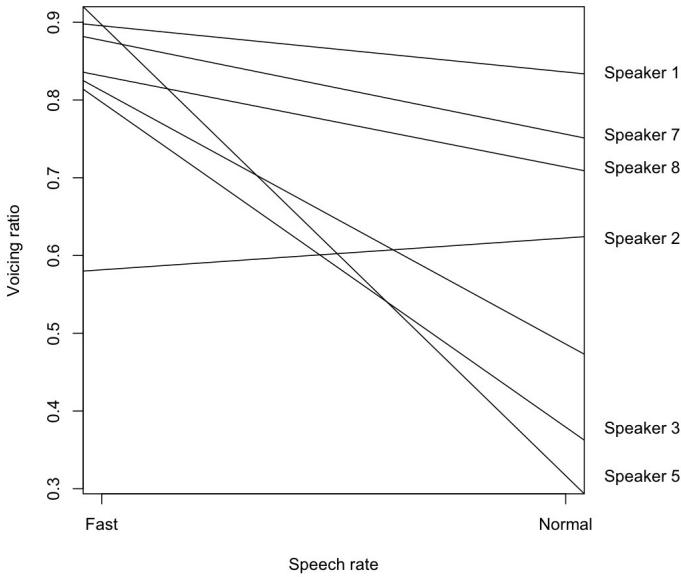


Figure 6.10: Coefficients of rate effects by speaker in a model with voicing ratio as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

Table 6.4: Regression coefficients, with standard error, and t values for a model predicting fricative duration in the /s/-voicing environments, measured in ms. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. *entusiasmo*) a fast rate.

Term	Level	β	SE	t
Intercept		48.31	4.69	10.31
Rate	Normal	22.86	3.82	5.98
Condition	3 (<i>gas acre</i>)	3.09	4.19	0.73
Condition	4 (<i>gas noble</i>)	-10.39	4.17	-2.49

previous cases, inclusion of a random slope for rate within speaker significantly improved the fit of the model compared to a model with random intercepts only ($\chi^2=30.48$, $df=2$, $p < 0.001$). However, no further significant improvements were found with the inclusion of random slopes for rate within item, or condition within speaker. The results of the model comparison show that there were significant individual differences with respect to how fricative duration was affected by speech rate manipulation. A plot of slopes and intercepts for individual speakers is in Figure 6.11. All the participants had longer fricative duration at normal speech rate as compared with fast, which is expected. However, speakers 3, 4, and 5 had both higher intercepts and higher slopes than speakers 1, 7, and 8. Speaker 2 had a relatively high intercept, but a shallow slope, again patterning differently from the remaining speakers.

Table 6.4 shows a summary of the fixed effects in the model. Fricative duration was greater in Condition 2 (*entusiasmo*), than in Condition 4 *gas noble*. There was, however, little difference between fricative duration in Condition 2 and in Condition 3 (*gas acre*, $t=0.73$). Fricative duration was also greater in Condition 3 than in Condition 4. There was a relatively large effect of speech rate, with an increase in fricative duration at normal speech rate ($\beta=22.86$, $SE=3.82$, $t=5.98$).

Intensity difference Individual differences with respect to speech rate effect were found in a model of intensity difference. A model with fixed effects of condition and rate, random intercepts

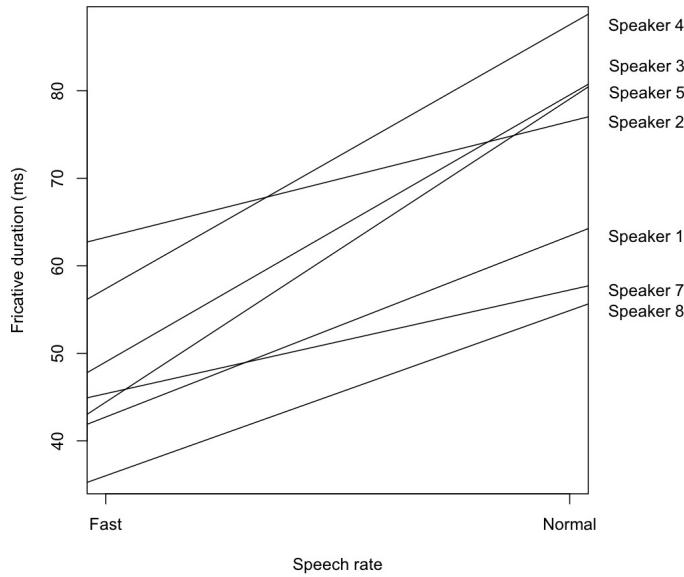


Figure 6.11: Coefficients of rate effects by speaker in a model with fricative duration (in ms) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

Table 6.5: Regression coefficients, with standard error, and t values for a model predicting intensity difference between mean low-frequency intensity and total mean intensity in the /s/-voicing environments, measured in dB. The intercept corresponds to a word-medial /s/ before a sonorant (e.g. *entusiasmo*) at a fast rate.

Term	Level	β	SE	t
Intercept		-4.38	0.73	-6.05
Condition	3 (<i>gas acre</i>)	0.36	0.46	0.79
Condition	4 (<i>gas noble</i>)	0.61	0.45	1.36
Rate	Normal	-1.30	0.65	-2.00

for speaker and item and random slopes for rate within speaker was found to have a better fit than a corresponding model without random within-speaker effects ($\chi^2=17.66$, $df=2$, $p < 0.001$). No significant improvement in the fit of the model was found with an inclusion of random effects of rate within item, or condition within speakers. Figure 6.12 illustrates the effect of speech rate manipulation on intensity difference for individual speakers. Speakers 1, 2, 7, and 8 maintained a fairly steady intensity difference regardless of the speech rate manipulation. In addition to this, the overall values of intensity difference were closer to 0 for speakers 1, 7 and 8 than for the remaining speakers. Speakers 3, 4, and 5, on the other hand, had a negative slope, with larger difference between low-frequency intensity and overall mean intensity in normal speech rate, compared to fast.

Table 6.5 shows the coefficients for the fixed part of the model. The three /s/-voicing conditions did not differ significantly with respect to the intensity difference between low- and total frequency intensity. There was, however, a main effect of rate ($\beta=-1.30$, $SE=0.65$, $t=-2.00$), with a larger intensity difference at normal speech rate.

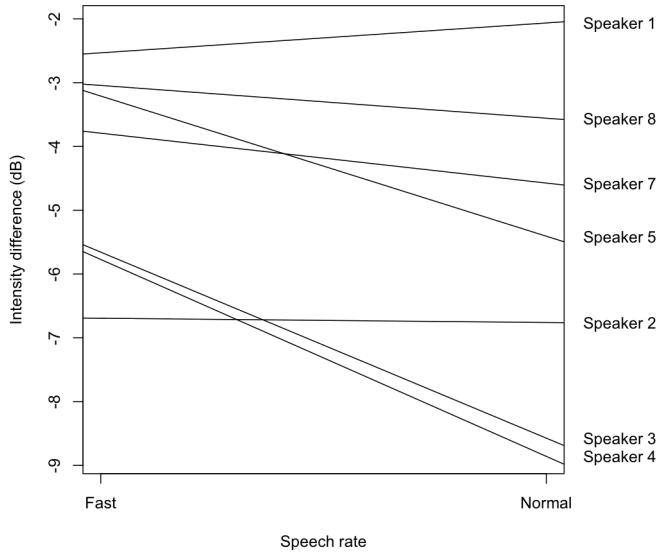


Figure 6.12: Coefficients of rate effects by speaker in a model with intensity difference (in dB) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

Analysis of speech rate

Since durational measurements, including voicing duration and fricative duration, are influenced by speech rate, an analysis of speech rate was performed to help contextualise previous findings concerning duration, with a view to determining which durational adjustments are likely to be an effect of individual speech rate, as opposed to individual voicing strategy. Speech rate was quantified as the duration of the phrase *otra vez* measured in ms. A linear mixed effects model was run on the data from all the environments ($n=1103$) with the duration of the phrase *otra vez* as a dependent variable. The effects of speaker and item were treated as random, and rate was included as a fixed effect, and as a random effect within speaker. This model was found in a log-likelihood test to provide a significantly better fit of the data in comparison with a model where the rate was included as a fixed effect only ($\chi^2=128.18$, $df=2$, $p < 0.001$). This shows that individual speakers differed not only in the intercept, but also in how much faster their speech was when they were trying to speak fast, as compared to the speech rate they found comfortable.

Figure 6.13 illustrates speech rate coefficients for individual speakers. At the intercept Speaker 2 was by far the slowest, using over 100 ms more than others to produce the phrase *otra vez*. This speaker also showed the smallest difference between the two speech rates. Speaker 5 showed the relatively greatest difference between fast and normal speech rate, being relatively slowest when speaking normally, but patterning with others in the fast rate condition. The remaining speakers showed some differences with respect to slope and intercept, but no distinctive individual strategies emerge from the data.

Summary

A partitioning analysis of voicing-related acoustic measurements confirms that there are differences between contexts reported as environments for Quito /s/-voicing, and baseline conditions. Effects of phonetic voicing were observed in sibilants followed by sonorants (within the same word as well as across a word-boundary), and in word-final sibilants followed by a vowel in the next word. Voicing consisted in increased duration of glottal fold vibration and voicing

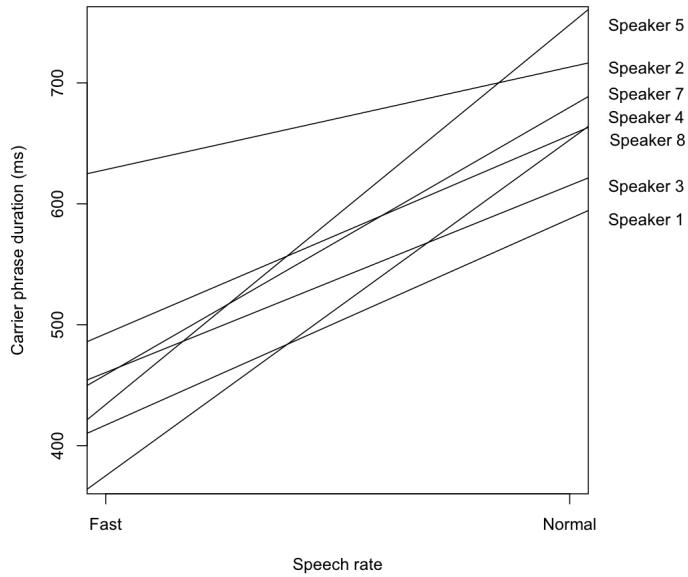


Figure 6.13: Coefficients of rate effects by speaker in a model with the duration of the phrase *otra vez* (in ms) as a response variable, rate and condition as fixed effects, random intercepts for speaker and item, and a random slope for rate within speaker.

ratio, compared with sibilants followed by a vowel or a voiceless stop within the same word. In addition to this, the degree of voicing found in the /s/-voicing contexts varied depending on the speech rate. In normal speech, voicing duration and voicing ratio were lower in the /s/-voicing environments than in sibilants followed by voiced stops. In fast speech, on the other hand, the two groups patterned together, typically with full voicing during frication.

The effect of speech rate on voicing duration and ratio within the /s/-voicing contexts uncovered by the partitioning analysis could be taken as an indication of a gradient status of Quito Spanish /s/-voicing. This generalisation is supported by the fact that voicing ratio was higher in fast speech than in normal speech in the /s/-voicing contexts, whereas there was a limited effect of rate on voicing duration. However, a closer analysis of the rate effects on Quito voicing reveals that this generalisation does not apply to all the participants in the study. While some speakers behaved consistently with the direction of the effect found by the partitioning analysis, others did the opposite.

Models of voicing duration, voicing ratio, fricative duration and intensity difference all showed a significantly improved fit once speech rate was considered as a random effect within speaker. This confirms that individual speakers responded differently to speech rate manipulations. Four out of seven participants in the experiment (1, 2, 7 and 8) showed increased voicing duration in normal as compared with fast speech rate within the /s/-voicing environments. The remaining three speakers (3, 4, and 5) showed the opposite tendency, producing more voicing in the fast speech rate, as compared with normal. Those speakers who produced more voicing in the fast rate condition also had a greater difference in the duration of fricatives produced at normal and fast rate. For speakers 1, 2, 7, and 8 the difference was relatively smaller, even though there was also an increase in the duration of a fricative from fast to normal speech condition. Finally, distinct individual strategies emerge from the data with respect to how rate affected voicing ratio and the difference between mean low- and total frequency intensity in fricatives. For speakers 3, 4, and 5 a large difference was observed between normal and fast rate in models of both voicing ratio and intensity difference. For these speakers an increase in speech rate brought about an increase in

voicing ratio, and the value of intensity difference was closer to 0. Speakers 1, 2, 7 and 8, on the other hand, maintained a similar voicing ratio and intensity difference across varying speech rates.

Finally, an analysis of speech rate was performed in order to verify whether any previously observed durational trends were also reflected in the duration of the phrase *otra vez*. The analysis showed that speaker 2 had a relatively small adjustment in the duration of the phrase between normal and fast speech rate. Speaker 5 showed a relatively large difference in the duration of the carrier phrase between normal and fast speech rates, consistent with the previously noted large adjustment in fricative duration between the two conditions. However, speakers 3, and 4, who had patterned with 5 with respect to large difference in fricative duration, did not also show the same effect with respect to overall speech rate, as they showed a relatively small difference in the duration of the phrase *otra vez* between normal and fast speech rate. Unlike in the case of voicing duration and fricative duration, no two distinct trends emerged from the measurements of speech rate.

6.4 Discussion

6.4.1 /s/-voicing – categorical or gradient?

The interpretation of the current data that I shall pursue is that Quito /s/-voicing is optional but categorical for some speakers and gradient for others. All speakers showed variation in their realisation on /s/-voicing in producing a varying duration of vocal fold vibration during the fricative. At first blush, the intra-speaker variation could be taken as supporting the hypothesis that /s/-voicing is gradient, i.e. it consists in varying degrees of laryngeal overlap between the fricative and the neighbouring segment. However, some aspects of the variation that we find in the data do not straightforwardly support the gradient interpretation.

An argument against the hypothesis that /s/-voicing in Quito Spanish is gradient for all speakers comes from speech rate effects. Drawing on previous work by Solé (1992, 1995) I hypothesised that a gradient process of /s/-voicing should involve either no variation in voicing duration across different speech rates, or potentially increased voicing at higher speech rates due to increased gestural overlap. The productions by speakers 3, 4, and 5 fit this description. Speaker 4 produced a similar duration of voicing across different speech rates, while for speakers 3 and 5 the duration of voicing increased somewhat in the fast speech condition. This effect, coupled with fricative duration shortening at fast speech rate, led to a significant increase in the voicing ratio from normal to fast speech rate for all three speakers. The increase in ratio was reflected in the increase in intensity difference between normal and fast speech rate. Partially voiced fricatives are expected to have relatively higher intensity difference than fully voiced ones, as the presence of voiceless frication will both reduce the low-frequency intensity and increase the mean total intensity. In that way, intensity difference can be said to be negatively correlated with voicing ratio. What is also interesting is that speakers 3, 4, and 5 showed a relatively large difference in the duration of fricative between normal and fast speech rate. This effect does not follow straightforwardly from individual speech rate, as the difference in duration of the phrase *otra vez* between normal and fast rate was not greater for speakers 3 and 4 than for other speakers. Therefore, the relatively large adjustment in fricative duration could potentially be secondary to increased voicing, which is consistent with the prediction by Jansen (2004) that fricative duration is mechanically linked to voicing, and so increased duration of voicing may shorten the fricative.

For categorical voicing processes no variation was predicted to occur in ratio across different speech rates, whereas voicing duration was expected to increase at normal speech rate. This was observed for four speakers: 1, 2, 7, and 8. Speaker 2 was somewhat exceptional in showing speech rate effects consistent with a categorical voicing hypothesis, while mostly producing partially voiced

tokens. However, evidence from speech rate effects has to be taken with caution in case of speaker 2, as the speech rate analysis shows an exceptionally small difference in the carrier phrase duration, which raises the question whether the speaker really completed the speech rate manipulation task. Speakers 1, 7, and 8 showed a clear preference for fully voiced fricatives regardless of speech rate, and they differed from speakers 3, 4, and 5 in producing the majority of fricatives as fully voiced at normal speech rate. The effects of a stable voicing ratio across speech rates was reflected for speakers 1, 7 and 8 in intensity difference, which was similar in fast and normal speech for these speakers.

There are no phonetic reasons for why gradient coarticulation should involve more voicing at slower speech rates. However, an increase in the voicing duration coupled with an increase in duration of the fricative makes sense as a strategy to realise a specific phonetic target. Therefore, the realisation of /s/-voicing by speakers 1, 7, and 8 is consistent with that of an optional phonological rule, which applies only to a subset of the /s/-voicing cases. When it does apply, however, the rule is categorical and reflected in full phonetic voicing.

6.4.2 Consequences for formal models

The categorical status of /s/-voicing, even if optional and subject to individual variation, presents a serious challenge to Colina's (2009) analysis, which, as previously discussed, crucially relies on /s/-voicing being gradient. Colina (2009) models /s/-voicing as a strictly phonetic effect which requires word-final prepausal and word-final pre-vocalic sibilants to have identical representations in terms of phonological features. However, if word-final pre-vocalic sibilants are optionally voiced at the phonological level, as follows from our data, their phonological form cannot be said to be identical to the sibilants in the word-final prepausal context where voicing does not apply.

Importantly, the significance of the Quito Spanish data goes beyond challenging a single analysis, as it shows the limits of modelling opaque patterns as surface correspondence effects. The optional but categorical voicing of /s/ in cases like [ga.za.kre] cannot be straightforwardly relegated from the domain of phonology. However, it cannot be analysed as a morphophonological correspondence effect either, as there is no principled means of selecting a base form that could serve as a source of output-output voicing correspondence. The word final sibilant in *gas* surfaces as voiced only in phrasal contexts, such as [gaz.no.ble], and there is no theoretical reason for why these forms should be considered the base.

In comparison, neither optionality nor gradience pose a problem to the cyclical analysis. Phrase-level /s/-voicing may be optional (in fact optionality of post-lexical rules is widely observed), and gradient effects follow from a modular view of the phonetics-phonology interface, where phonetics operates on the output of phonological computation. The fundamental insight of the cyclical analysis is that /s/-voicing applies to those sibilants that had been in a coda at some stage of the derivation. While the derivational character of this generalisation makes it incompatible with monostratal models, it can in fact emerge from a series of natural and well attested sound changes. Table 6.6 presents a succession of phonetic and phonological changes through which /s/-voicing can evolve, drawing on the insights and generalisations of the life cycle of a phonological process (Bermúdez-Otero, 2007; Bermúdez-Otero & Trousdale, 2012). At stage 1 the voicing contrast is lost in sibilants by way of devoicing. Early Spanish had a complex inventory of sibilant phonemes, consisting of several pairs of voiced and voiceless members, but by the sixteenth century various developments had reduced this inventory to a single sibilant: /s/. At stage 2 syllable codas are interpreted as laryngeally unspecified (cf. discussion of causes for coda delaryngealisation below). The pre-vocalic voicing at stage 3 in Table 6.6 is where the cyclic architecture becomes crucial, as this change is not phonetic, but phonological and it proceeds by input restructuring. In cyclic

terms, the process of coda delaryngealisation climbs up from the phrase level to the word level, as the learners reinterpret the output of the phrase level (in our case coda delaryngealisation), as being already present in the input, i.e. at the output of the word level. With delaryngealisation applying at word level, all word-final fricatives lose their voice specifications, becoming amenable to phrase level voicing before consonants and vowels alike. This type of voicing originates as gradient, but later it undergoes stabilisation, represented as stage 4 in Table 6.6.

Table 6.6: The life cycle of /s/-voicing

Stage	Change	Examples		
1	Initial stage	[gas]	[gas.no.βle]	[ga.sa.kre]
2	Coda delaryngealisation	[gaS]	[gaS.no.βle]	[ga.sa.kre]
3	Analogical change	[gaS]	[gaS.no.βle]	[ga.Sa.kre]
4	Categorical reinterpretation	[gaS]	[gaz.no.βle]	[ga.za.kre]

6.4.3 Further alternatives

The cyclical explanation for /s/-voicing is theoretically elegant. The basic architecture of the system accommodates both the synchronic and the diachronic aspect of /s/-voicing, and allows for a correct and succinct generalisation concerning the distribution of the pattern. However, in a pursuit of an adequate model of phonological competence, one needs to distinguish between generalisations that data lend themselves to, and generalisations that learners truly need to make in order to master the synchronic pattern. The two phonological models that have been discussed in this chapter strive to provide a generalisation for the opaque application of /s/-voicing. However, the opacity problem hinges on the assumption that all the environments for /s/-voicing must be captured by means of a single generalisation. This assumption is not inherently present in the model I have developed throughout this work which takes it that the synchronic grammar is not concerned with explaining sound patterns, but with providing abstract generalisations that match the data in the input. If the goal of synchronic grammar is redefined in these terms, there is nothing in the way of learners generalising a process over a list of environments where it applies, i.e. in word-internal codas and at the end of a word (cf. Nerset (2008, Chapter 3) for a discussion of how such lists may be cognitively represented in a form of phonological schemas).

A potential counterargument to forfeiting a common synchronic generalisation for all /s/-voicing environments comes from the fact that /s/-voicing appears fairly uniform across different environments. This generalisation is statistically supported by the clustering of the /s/-voicing environments in the partitioning models of the phonetic correlates of voicing (cf. 6.3). Therefore, positing a different status of lexical and postlexical /s/-voicing within phonological knowledge arguably involves a loss of generalisation. However, the uniform behaviour of /s/-voicing in different environments may also be a reflection of common diachronic origins. I have argued throughout this thesis that pre-sonorant and prevocalic voicing is conditioned by delaryngealisation, defined phonetically as absence of an active devoicing target. A similar assumption concerning delaryngealisation, albeit defined in representational terms, is also made by Colina (2009) and Bermúdez-Otero (2011). Delaryngealisation makes obstruents more susceptible to passive voicing from the preceding vowel, which may give rise to perception of a voicing target, and in turn also to voiced productions on the part of speakers (cf. 3.4). This scenario predicts that the voicing will develop in similar ways across all positions where the sibilant had previously been delaryngealised.

The account suggested above is potentially problematised by the fact that it assumes a diachronic development where delaryngealisation precedes resyllabification. This order of events, however, cannot be straightforwardly asserted for Quito Spanish. Bermúdez-Otero (2011) argues

that the diachronic development was in fact the opposite, i.e. resyllabification took place prior to delaryngealisation (cf. Table 6.6). A cyclic architecture can model his kind of diachronic development by analysing delaryngealisation as progressing from the Phrase Level to the Word Level, whereby delaryngealisation begins to affect also those word-final codas which undergo resyllabification. However, the very same process can also be modelled as analogical change that does not crucially require a cyclic interpretation. An example of such analogy is the spread of t-glottalisation from pre-consonantal to the pre-vocalic environment in Southern British English (Williams & Kerswill, 1999). Rácz (2011) argues that t-glottalisation is phonetically motivated before a consonant, but not before a vowel. Thus, pre-vocalic glottalisation emerges not due to phonetic conditioning, but as an analogical extension from the pre-consonantal environment. Assuming a rich memory model, in which phonetic detail is stored together with individual productions of lexical entries, and a categorisation model in which a learner classifies tokens as belonging to a specific category based on their similarity to another tokens, Rácz (2011) models a learner who begins glottalising before a vowel due to the influence of pre-consonantal t-glottalisation and 10% production noise, where pre-vocalic tokens are erroneously coded by the learner as pre-consonantal and *vice versa*. The same solution could potentially be applied to the extension of delaryngealisation from word-final sibilants followed by a sonorant in Quito Spanish to word-final prevocalic sibilants. The categorisation used by Rácz (2011) is not sensitive to subsegmental phonetics, only to token frequency, and the presence or absence of the conditioning environment, i.e. the following segment. Therefore, the conclusion that the pre-consonantal pattern may extend to the pre-vocalic environment is also plausible in the case of Quito /s/-voicing, assuming that tokens of word-final /s/ followed by a sonorant are more frequent than tokens of word final sibilants followed by a vowel. I do not attempt to provide an equivalent explicit model of analogical extension for Quito Spanish. I do, however, mention what such an account could look like to bring out the fact that there are alternatives to the two models I had originally set out to compare. Consequently, the finding that the cyclical model achieves descriptive adequacy unparalleled by the rival correspondence-based account, does not allow one to conclude that cyclicity is the only solution for the problem of Quito /s/-voicing.

6.5 Conclusion

The acoustic data collected in the current study show that the phonetic production of Quito /s/-voicing involves categorical behaviour on the part of some speakers, whereas other speakers show characteristics of phonetic gradience. I have argued for an interpretation of these facts consistent with a modular view of the phonetics-phonology interface, and I consider the categorical behaviour to signify an optional assignment of a phonological category. Rejecting the hypothesis that all /s/-voicing is necessarily gradient is of crucial consequence to the correspondence-based analyses of phonological opacity; although there are forms where the realisation of word-final /s/ as voiced is transparent, there is no principled way of selecting those forms as the base for the opaque word-final prevocalic sibilants. The case of synchronic phonological opacity we find in Quito Spanish appears to have originated from analogy between lexical forms which are not in a transderivational correspondence relationship. As a consequence of the analogical change, two steps of derivation seem necessary for a formal model to capture the resulting synchronic pattern, insofar as the goal of the model is to provide a common generalisation for all the environments where /s/-voicing applies. However, a multi-level derivation is not indispensable in a model where similar phonetic phenomena are attributed to common diachronic origins, as opposed to a common synchronic generalisation on the part of language users.

CHAPTER 7

EXPLAINING PRE-SONORANT VOICING

Having a personal philosophy is like having a pet marmoset, because it may be very attractive when you acquire it, but there may be situations when it will not come in handy at all

Lemony Snicket
THE GRIM GROTTO

This chapter brings together the findings on pre-sonorant voicing presented in this work. The recurrent empirical observations are identified in 7.1.5. In 7.2 I argue for the existence of a diachronic link between delaryngealisation and pre-sonorant voicing, where delaryngealisation is defined phonetically as absence of a voicing or devoicing target. I then discuss the role of the right-hand environment in conditioning the occurrence of pre-sonorant voicing, and the role of other modulating factors in 7.3. I propose that pre-sonorant voicing may emerge diachronically in positions of delaryngealisation, due to a combination of articulatory, aerodynamic and acoustic pressures without an intermediate step of assigning laryngeal specifications to sonorants. In 7.4 and 7.5 I consider the phonological pressures involved in the evolution and synchronic generalisation of pre-sonorant voicing. Finally I discuss the significance of the current findings to the debate on the distinction between phonetics and phonology in 7.6.

7.1 Empirical properties of pre-sonorant voicing

One of the aims of this thesis has been to systematically investigate the empirical properties of pre-sonorant voicing. Previous typological observations point toward pre-sonorant voicing being constrained by positional and segmental factors. In addition, existing descriptive and theoretical sources make predictions, explicit or otherwise, about the categorical or gradient status of pre-sonorant voicing. The emerging typological generalisations are crucial to understanding what factors are involved in the diachronic origin of pre-sonorant voicing, and its synchronic representation. The following summary presents the empirical findings of the current study. A summary of the main results on a language-by-language basis is presented in 7.1.1 – 7.1.4. A synopsis of the recurrent patterns is in 7.1.5.

7.1.1 West-Flemish

In West-Flemish word-final fricatives were typically found to undergo full voicing when a sonorant followed in the next word. However, there were some exceptions to the application of pre-sonorant voicing, evidenced by a bimodal distribution of voicing duration and voicing ratio in pre-sonorant fricatives. A similar voicing effect was not found in word-final fricatives in the same context produced by East-Flemish speakers. No increased phonetic voicing was found in West-Flemish pre-sonorant stops, compared to pre-sonorant stops in East-Flemish. In West-Flemish,

pre-sonorant stops patterned with stops followed by voiceless stops in the next word with respect to their voicing duration and ratio. Pre-sonorant fricatives patterned with fricatives followed by voiced stops in their mean values of voicing duration and ratio. This finding, coupled with the bimodal distribution of voicing duration and ratio, is consistent with the interpretation of the pre-sonorant fricative voicing as categorical, but optional. In addition, instances of gradient phonetic voicing were found in some of the pre-sonorant fricatives which did not undergo full voicing, as well as in fricatives followed by voiceless stops, and in stops followed by sonorants and voiced obstruents. The gradient voicing invariably involved a short voicing tail continuing from the preceding vowel. There were no systematic differences between sonorant consonants and vowels as triggers of voicing in the preceding fricatives in West-Flemish. There was no effect of underlying voicing on any voicing-related acoustic parameters in the word-final position.

7.1.2 Poznań Polish

The data on word-final pre-sonorant stops in Poznań Polish collected in two experiments show a varying degree of the effect of underlying voicing. In the first experiment, pre-sonorant voicing was strongly influenced by the stop's underlying voicing; underlyingly voiced stops frequently surfaced as voiced (with increased voicing duration and full voicing during closure), but hardly any cases of pre-sonorant voicing were found within underlyingly voiceless stops. Importantly, the effect of underlying voicing interacted with dialect and the right-hand environment in models of voicing duration and ratio; pre-sonorant voicing was limited to underlyingly voiced stops in the Poznań dialect. In another experiment, however, Poznań speakers showed pre-sonorant voicing also in underlyingly voiceless stops. In the same experiment, a significant three-way interaction was found between speaker, dialect and the right-hand environment in an analysis of voicing duration and voicing ratio in word-final fricatives. Voicing duration and ratio were relatively higher in the Poznań dialect than in the control group consisting of Central dialect speakers. Within the Poznań dialect, voicing duration and ratio were higher before a sonorant in underlyingly voiced fricatives, compared to underlyingly voiceless ones.

In addition to the robust effects of underlying voicing on voicing duration and ratio in pre-sonorant stops and fricatives in the Poznań dialect, an effect of underlying voicing on vowel duration was found in both dialects. Vowel duration was increased before underlyingly voiceless stops, and before underlyingly voiced fricatives. The size of the effect was small, not exceeding 7.28 ms.

Voicing duration and ratio followed a bimodal distribution in pre-sonorant stops and fricatives in the Poznań dialect, which is consistent with a categorical but optional interpretation of pre-sonorant voicing. There were also distributional differences conditioned by the underlying voicing. For instance, underlyingly voiced pre-sonorant fricatives were more likely to surface with full voicing than underlyingly voiceless fricatives in the Poznań dialect.

Voicing before sonorants was found to differ consistently from voicing before a voiced obstruent. Pre-sonorant stops and fricatives had a significantly lower voicing duration and voicing ratio than obstruent followed by a voiced obstruent. This effect was observed within both the Warsaw dialect and the Poznań dialect.

7.1.3 Central Catalan

The findings on word-final obstruents in Central Catalan confirm previously established generalisations concerning obstruent voicing before a vowel. Full voicing was found in most word-final prevocalic sibilants and stop+sibilant clusters, with the mean values of voicing duration and ratio similar to those of sibilants and clusters followed by a voiced stop. One of the speakers differed from the rest in voicing prevocalic sibilants, but not prevocalic stop+sibilant clusters. Very limited levels of voicing were found in word-final prevocalic stops, which patterned with stops followed by

voiceless fricatives in terms of voicing duration and ratio. Voicing before sonorant consonants affected all classes of word-final obstruents, but the degree to which it applied was subject to considerable inter-speaker variation. Pre-sonorant voicing was relatively limited, compared to voicing preceding a voiced obstruent. The phonetic correlates of voicing were not affected by the underlying voicing value of the obstruent.

7.1.4 Quito Spanish

In Quito Spanish the environments for /s/-voicing (word-medially and word-finally before a sonorant, word-finally before a vowel) consistently clustered together, with no differences depending on the position (word-medial vs. word-final), or the following segment's manner of articulation (sonorant vs. vowel). The /s/-voicing environments clustered with the environments which involved a following voiced stop, but only at a fast speech rate. At a normal speech rate, the /s/-voicing environments involved a relatively shorter voicing duration and a lower voicing ratio than the environments with a following voiced obstruent. The effect of speech rate within the /s/-voicing environments was subject to individual variation. Three speakers were found to produce slightly increased voicing duration in sibilants uttered at a fast rate compared to normal, which was accompanied by a relatively large increase in voicing ratio. This effect of speech rate is consistent with a gradient interpretation of /s/-voicing. Three other speakers, however, showed an increase in voicing duration at a normal speech rate when compared with a fast rate, whilst maintaining a steady voicing ratio across different speech rates. This behaviour is consistent with a categorical interpretation of /s/-voicing, although some variation within the categorical speakers points to optionality.

7.1.5 Emerging generalisations

Some of the typological generalisations previously made for pre-sonorant voicing are supported by the instrumental data collected in this project. For instance, the existing observation that fricatives are more likely to undergo pre-sonorant voicing than stops is confirmed by the data from West-Flemish and from Catalan prevocalic voicing. In both cases the asymmetry between fricatives and stops was reflected in a categorical opposition: while fricatives were most of the time found to undergo full voicing in the relevant position, stops showed no increase whatsoever in phonetic voicing.

The current data support the generalisation that pre-sonorant voicing targets obstruents in positions of delaryngealisation which is understood as the absence of a voicing or devoicing target. In the case of West-Flemish and Central Catalan, the observation concerning the role of delaryngealisation is confirmed by the lack of an effect of the underlying voicing value on the phonetic realisation of voicing in the word-final position. The lack of such an effect indicates that the underlying voicing contrast is lost on the surface, which is consistent with a delaryngealisation hypothesis. However, the role of delaryngealisation is complicated in the case of Poznań Polish due to two factors. First, the Poznań data showed a non-neutralisation effect, with a higher voicing ratio for underlying voiced stops in the pre-sonorant environment, as well as a higher voicing ratio and duration in pre-sonorant fricatives. Second, there was an effect of near-neutralisation reflected in a small-scale difference in vowel duration depending on the underlying voicing of the obstruent. The effect of underlying voicing could be taken as *prima facie* evidence against the hypothesis that delaryngealisation is a pre-requisite to the occurrence of pre-sonorant voicing. However, the significance of the non-neutralisation is less clear considering that the Poznań pattern can also follow from optionality. The discussion in Section 4.4.1 shows that a stochastic model where delaryngealisation feeds pre-sonorant voicing, and both processes apply optionally, can result in a variable application of pre-sonorant voicing combined with incomplete neutralisation, as observed

in Poznań Polish. Consequently, the Poznań data could be seen as consistent with the hypothesis that pre-sonorant voicing is conditioned by delaryngealisation, insofar as both delaryngealisation and voicing are allowed to apply variably.

When it comes to the role of the right-hand environment, the current data consistently reveal discrepancies between sonorants and voiced obstruents as triggers of voicing. Some of the discrepancies are seen in the phonological patterning. In West-Flemish, voiced obstruents were found to trigger voicing in stops and fricatives and alike, whereas pre-sonorant voicing was limited to fricatives. The same asymmetry was observed with respect to vowels as triggers of voicing in Catalan; categorical prevocalic voicing was found in sibilants, but it was absent from stops. In addition to these categorical asymmetries, sonorants and voiced obstruents were also differentiated by effects on the degree of phonetic voicing, or frequency of voicing. In Central Catalan and in Poznań Polish voiced obstruents and sonorants (including vowels in Polish, but limited to sonorant consonants in Catalan) did not pattern together. Voiced obstruents were associated with significantly higher mean values of voicing duration and voicing ratio. In Poznań Polish these mean differences corresponded to different levels of application, with full voicing being more frequent preceding voiced obstruents than preceding sonorants. A difference between voiced obstruents and sonorants as triggers of voicing was also found in the Quito Spanish data, as the two environments patterned differently depending on speech rate. At fast speech rates, the participants produced similar amounts of voicing in sibilants followed by sonorants and voiced obstruents, with no significant differences between the two environments. At a normal speech rate however, voiced obstruents triggered significantly higher voicing duration and voicing ratio values than sonorants.

Another feature of pre-sonorant voicing related to the right-hand environment concerns the voicing trigger's manner of articulation. Central Catalan provides an example of a robust difference between vowels and sonorant consonants as triggers of voicing. Prevocalic voicing was generally more advanced than pre-sonorant voicing, applying with a higher frequency and less inter-speaker variation. At the same time, however, vowels triggered less voicing than sonorants in word-final stops, where prevocalic voicing did not apply at all. No systematic differences between vowels and sonorant consonants as triggers of voicing were found in West-Flemish, Poznań Polish, or Quito Spanish.

As far as the question of categoricity vs. gradience is concerned, the current data confirm that pre-sonorant voicing need not necessarily be gradient. On the contrary, mostly categorical patterns were found in the studied cases. In pre-sonorant position in West-Flemish and Quito Spanish, as well as prevocally in Central Catalan, fricatives surfaced with very high mean values for voicing duration and voicing ratio, which was also reflected in the fricative's relative intensity. In West-Flemish and in Central Catalan, the values were high enough for the pre-sonorant/prevocalic fricatives to pattern statistically with fricatives followed by voiced obstruents. Further instances of categorical pre-sonorant voicing were evidenced by bimodality. In West-Flemish and Poznań Polish, obstruents in the pre-sonorant voicing contexts were realised variably with either full voicing or a very limited voicing tail continuing from the preceding vowel, forming a bimodal distribution in each case. This kind of distribution indicates categorical variation between two distinct options, rather than gradient variation conditioned by purely phonetic factors. Finally, evidence for categoricity is supplied by speech rate manipulations in the case of Quito Spanish. Some speakers were found to increase the duration of voicing in pre-sonorant sibilants when speaking at a normal rate, as opposed to fast, while maintaining a steady voicing ratio across speech rates. An increase in voicing duration at relatively slower speech rates does not follow directly from any constraints on coarticulation, and it is most straightforwardly explained as a phonetic strategy to maintain a consistent voicing target regardless of speech rate manipulation. Instances of gradient voicing were more limited in the current data, and are only supported for individual speakers of Quito Spanish, specifically the ones who showed a large increase in the voicing ratio of pre-sonorant sibilants at a

faster speech rate, while maintaining a constant voicing duration across different speech rates.

7.2 The role of delaryngealisation

The model developed throughout this work, and whose basic ideas were laid out in Chapter 3, constitutes an evolutionary approach, where the properties of pre-sonorant voicing are explained by the processes of change leading to it. Based on the empirical observations collected in this project and drawing on previous work on the phonetics and phonology of voicing I have been arguing that a diachronic link exists between delaryngealisation and pre-sonorant voicing. Central to this approach is an understanding of delaryngealisation as involving the loss of an active voicing or devoicing target, as proposed by Jansen (2004).

Jansen distinguishes between active and passive voicing and devoicing, which modulate the pressure drop across glottis during articulation. Sounds are said to be passively voiced if the supraglottal pressure involved in their articulation is low enough to generate vocal fold vibration without the involvement of additional voicing gestures. Sonorants are an example of passively voiced sounds, as their relatively open articulation is associated with a low intraoral pressure conducive to voicing. In comparison, obstruent articulation is relatively closed, with either closure or partial constriction in the vocal tract, which raises the pressure above the glottis. In this situation vocal folds are not expected to vibrate, unless an additional gesture is made to initiate or maintain voicing. The execution of such gestures on the part of the speaker constitutes a case of active voicing, and it is associated with the presence of a voicing target. In a neutral state of the glottis, the high intra-oral pressure associated with the production of a word-initial obstruent favours passive devoicing, but a speaker may actively voice an obstruent in this position, by means of voicing gestures, such as lowering the larynx, raising the soft palate and advancing the tongue root (Ladefoged, 1973; Stevens, 1998). Conversely, in some position obstruents may be subject to passive voicing. For instance, passive voicing may follow from transitions between adjacent vocalic and obstruent articulations. Once the vocal folds have been set into motion, they will continue to vibrate until the pressure rise above the glottis quenches voicing. Thus, a short period of voicing can be seen in postvocalic obstruents despite the inherent conflict between voicing and obstruency (Westbury & Keating, 1986). The portion of vocal fold vibration may also be facilitated by a truncation of glottal opening for the obstruent articulation conditioned by a reduced glottal opening gesture when a voiceless obstruent is flanked by voiced sounds (Beckman et al., 1992). This kind of passive voicing may be actively counteracted by devoicing gestures (e.g. raising the larynx, tensing the vocal tract walls, glottalisation, or glottal abduction) which are expected to occur in the case of actively devoiced sounds.

A distinction between active and passive mechanisms involved in voicing and devoicing provides an insight into the nature of the relationship between delaryngealisation and coarticulatory voicing. Word-final delaryngealisation, as seen for instance in Dutch, is viewed by Jansen (2004), as well as Ernestus (2000), as a case of phonetic underspecification, i.e. the absence of a voicing or a devoicing target. Support for such a specification comes from the phonetic properties of final segments, which are distinct in their realisation of voicing from both actively voiced and actively devoiced segments. The significance of underspecification for pre-sonorant voicing lies in the susceptibility to passive voicing. Underspecified sounds do not involve the production of an active devoicing gesture, and hence they are not expected to counteract passive voicing from neighbouring sounds. I have proposed that pre-sonorant voicing originates as passive voicing due to vocal fold vibration continuing from the preceding vowel. In the absence of an active devoicing target, this passive voicing is expected to extend for a relatively long time. This, in turn, may be perceived by listeners to involve an active voicing target, implemented in their own production, when it is their

turn to speak.

The involvement of devoicing gestures in counteracting passive voicing provides a phonetically motivated account of positional asymmetries seen in pre-sonorant voicing. The model predicts that pre-sonorant voicing applies only to obstruents without an active voicing target. Thus, asymmetries in pre-sonorant voicing may occur in obstruents within an identical segmental and syllabic environment, subject to whether or not delaryngealisation occurs in a given position. This case is illustrated by prevocalic sibilants in Central Catalan and Quito Spanish. In both languages, word-final prevocalic sibilants surface as voiced, but there is no prevocalic sibilant voicing in the word-medial position. Importantly, word-medial and word-final sibilants are not distinguished by their position in the syllable, as word-final consonants surface as onsets when a vowel follows in the next word (Harris, 1983; Wheeler, 2005). Thus, pre-sonorant voicing cannot be favoured in one case over another due to functional factors associated with the position in the syllable, or due to the influence of the following segment. However, as the word-medial and word-final environments are differentiated by whether or not they are positions of delaryngealisation, the observed distinction in pre-sonorant and prevocalic voicing follows straightforwardly from an account in which laryngeal underspecification is a prerequisite to voicing.

While the positional restrictions on prevocalic voicing cannot be seen as being directly motivated by functional pressures on voicing, they reflect the influence of various factors conditioning the occurrence of delaryngealisation. In a discussion of final devoicing, Blevins (2004, 2006) notes that it is common for devoicing to initially affect sounds in the phrase-final environment, as seen for instance in Gulf Arabic (Holes, 1990), Southern Luri (Anonby, 2003), and Nigerian Arabic (Owens, 1993). Devoicing in this position may follow from a number of functional pressures. For instance, phrase-final lengthening may trigger passive devoicing, due to an aerodynamic conflict between voicing and high intraoral pressure associated with a closure in the vocal tract (Blevins, 2004, pp. 104–105). Lengthening may also contribute to the voiceless perception of phrase-final sounds as increased duration is a common devoicing cue. These, along with other functional factors including laryngeal boundary marking and lack of audible release, may contribute to the initiation and phonologisation of phrase-final devoicing. In the course of phonologisation devoicing may be re-analysed by listeners as delaryngealisation. Having stabilised, the process is likely to undergo rule generalisation from the phrase-final to the word-final, and syllable-final position. Blevins (2006) proposes that overgeneralisation from the phrase-final to the word-final domain is likely to arise in language acquisition, driven by the prevalence of single-word utterances in child-directed speech. While the same explanation does not extend directly to the case of devoicing climbing down to the syllable-final domain, re-analysing final devoicing as syllable-final may be facilitated by phonetic differences between onset and coda obstruents. For instance, onset obstruents may be fortified due to syllable-initial strengthening (Keating et al., 2003) with increased duration and greater articulatory magnitude. One can hypothesise that such strengthening may make voicing and devoicing targets relatively more perceptible in onsets than in codas, increasing the chance that onset-coda asymmetries may arise with respect to the occurrence of voicing targets. Apart from rule generalisation to different prosodic domains, final devoicing may also undergo analogical extension to morphologically defined environments. An example of this is Catalan where delaryngealisation is seen in codas as well as prefix-finally (cf. 5.4.1).

Apart from motivating positional restrictions seen in pre-sonorant and prevocalic voicing, the phonetic underspecification proposal has important implications for the relationship between pre-sonorant voicing, delaryngealisation and lexical contrast. As previously mentioned in the discussion of the work by Ernestus (2000) and Jansen (2004), underspecified segments are defined in terms of surface phonological specifications setting out phonetic targets, rather than in terms of lexically contrastive specifications. Consequently, the number of voicing categories relevant in a language based on the phonetic diagnostics need not necessarily match the number of categories defined on

the basis of phonological contrasts. On the contrary, active and passive voicing and devoicing may in principle be seen in a language with a two-way phonological voicing contrast. This conception of categoryhood is supported by the existence of systems such as Quito Spanish, where pre-sonorant voicing occurs in the absence of a voicing contrast in sibilants. Pre-sonorant voicing in Quito Spanish affects obstruents in the word-final positions, as well as in word-medial codas. Both of these are common positions for delaryngealisation, as seen for instance in Catalan and Dutch. If delaryngealisation consists in the loss of voicing and devoicing gestures in a specific position, it can in principle co-occur with a devoicing process in other positions, which involves a re-interpretation of word-initial and word-medial onsets as actively devoiced. Actively devoiced and underspecified obstruents are defined phonetically based on the presence or absence of devoicing gestures, so the two categories can co-exist in a language, but only one of them is expected to feed pre-sonorant voicing. This in turn can give rise to the distribution of pre-sonorant voicing that we see in Quito Spanish, where pre-sonorant voicing is confined to cross-linguistically common positions of delaryngealisation, while word-medial and word-initial onsets remain invariably voiceless.

The Quito Spanish system poses a significant challenge to any feature systems without a ternary distinction in the laryngeal feature series Lombardi (1995); Iverson & Salmons (1995, 2006); Jessen & Ringen (2002); Honeybone (2005); Beckman et al. (2011), as well as to systems where laryngeal features are defined top-down based on the presence or absence of contrast (Avery et al., 2008; Currie-Hall, 2006; Dresher, 2009; Stevens et al., 1986; Stevens & Keyser, 1989; Keyser & Stevens, 2006; Stevens & Keyser, 2010). An account involving phonetic underspecification allows for a well-motivated explanation of the positional asymmetries seen in Quito Spanish /s/-voicing. If /s/-voicing is fed by delaryngealisation resulting in phonetic underspecification, and if underspecification is different from active devoicing, then word-final prevocalic sibilants are expected to undergo prevocalic voicing, while word-medial prevocalic sibilants are not, even though in both cases the sibilant is voiceless in some sense at the level of input. Formally, this can be straightforwardly captured by means of a distinction between [-voice] and [0voice], where word-final obstruents and word-medial codas are [0voice], whereas all non-derived onsets are specified as [-voice]. This kind of feature system is used by Wetzels (1997) to account for the distributional restrictions on laryngeal features in Bakairi roots. However, most phonological accounts argue against a distinction between [-voice] and [0voice] based on the fact that it is not motivated by the existence of a lexical contrast. This argument features in proposals concerning laryngeal realism (Iverson & Salmons, 1995, 2006; Jessen & Ringen, 2002; Honeybone, 2005; Beckman et al., 2011). Laryngeal realism holds that laryngeal contrasts rely on different features, depending on what the phonologically active member of the fortis-lenis pair is in a given language, and on how the voicing contrast is realised phonetically. As proposed by Iverson & Salmons (1995), central to the argument is a link between the two. For instance, in languages where the voicing contrast is phonetically expressed by means of an opposition between negative and short lag VOT ('true voice' languages), such as Dutch, the [voice] feature is phonologically active spreading in obstruent clusters. In aspirating languages on the other hand, where the laryngeal contrast is expressed by means of an opposition between short- and long-lag VOT, the active feature is [spread glottis]. Reflections of this are found for instance in English, German, Swedish, and Danish assimilatory devoicing. Phonetic underspecification forms an important part of the laryngeal realism approach, but the distinction involved is always binary; it is either between [spread glottis] and an unspecified sound, or between an underspecified sound and [voice]. Unfortunately, neither of these oppositions reflects the contrast I have argued to be active in Quito Spanish, at least not under the assumption that the content of the features is phonetically defined. The contrast in Quito Spanish appears to be the one of voicelessness vs. underspecification, which is more consistent with a [spread glottis] system over a [voice] system. However, Spanish is not an aspirating language, and it shows evidence of regressive [voice] assimilation from voiced stops. Thus, the Quito Spanish appears to be a case

where the evidence from phonological activity and from the phonetic realisation on voicing do not converge on either [spread glottis] or a [voice] specification, contrary to the predictions of the laryngeal realism approach.

The Spanish case is equally problematic to approaches which divorce feature specifications from their phonetic content entirely. This idea features in a number of formal phonology analyses, but is perhaps most explicitly articulated by Substance-Free Phonology (Hale & Reiss, 2008; Blaho, 2008), and the Toronto School (Avery et al., 2008; Currie-Hall, 2006; Dresher, 2009). Both programmes advocate a strictly symbolic approach to phonological representations, defined top-down on the basis of phonological contrast and phonological behaviour. A similar view of what constitutes the domain of phonology follows from the Quantal Theory (Stevens et al., 1986; Stevens & Keyser, 1989; Keyser & Stevens, 2006; Stevens & Keyser, 2010), as expressed in the following statement by Stevens & Keyser (2010, pp. 18).

“The model toward which we are moving, then, may be characterized as one in which:

- a. Underlying representations are entirely feature-based and contain only distinctive features.
- b. Differences between underlying and surface representations are mainly due to strategies of enhancement and overlap, which introduce, delete or extend gestures, but do not operate on features.”

The problem that Quito Spanish presents has to do with the opposition between word-final prevocalic and word-medial prevocalic onset sibilants, which is impossible to derive based on contrast alone due to the absence of any lexical voicing contrast in Spanish sibilants. A featural distinction could perhaps be posited based on the evidence of phonological behaviour. Currie-Hall (2006) gives an example of such an analysis based on the data from Czech and Slovak, where sonorants are equipped with different laryngeal features to account for some phonological patterns despite the absence of contrast. If Quito-Spanish /s/-voicing is analysed as conditioned by delaryngealisation, the evidence from where it applies could motivate a laryngeal distinction between word-initial/medial and word-/syllable-final sibilants in Spanish. However, it is less clear how the same distinction could be posited in a top-down approach for the same distribution of /s/-voicing when the process is gradient. This kind of system is a diachronic predecessor to the categorical variant of Quito Spanish /s/-voicing in the model developed in this thesis, and has in fact been identified for a number of Quito speakers based on how speech rate manipulations affected their voicing duration and ratio. A gradient voicing process does not constitute phonological behaviour in any model that distinguishes between phonetics and phonology, and such a distinction is central to both substance-free and contrast-driven approaches. Thus, gradient /s/-voicing does not fulfil the criteria for a factor driving feature specification. Even more problematic is the yet earlier stage, where there is only passive voicing, but where the distinction between actively devoiced and underspecified sibilants already exists. The phonetic characteristics associated with active devoicing and passive voicing are the only cues by which sibilants in different positions may be distinguished in this situation, but such characteristics entirely evade a top-down system.

The criticism above may not seem immediately relevant, considering that Quito Spanish /s/-voicing is not a phonological process insofar as phonology is defined by contrast, and so accounting for it does not at all fall within the scope of contrast-driven theories of phonology. The consequence of this, however, is that the theory entirely loses sight of the distinction between categorical and gradient voicing. Relegating Quito /s/-voicing from phonology altogether would entail that both categorical and gradient voicing should happen in the phonetics, which in turn necessitates allowing phonetics to be categorical. However, even if such concessions were to be made, we arrive at a model which can potentially accommodate the phonetic realisation and the synchronic distribution

of /s/-voicing, but does not explain the source of either. If categorical phonetic behaviour does not reflect the formation of an abstract category in the speakers' grammar, then the rise of categorical /s/-voicing is difficult to explain. If /s/-voicing positions are not phonologically distinct from non-/s/-voicing positions in their featural specifications, then the distribution of /s/-voicing is entirely arbitrary, both in the synchronic, *and* the diachronic perspective. In the absence of underspecification (which does not involve phonological contrast) it is not at all clear why the same voicing process would ever begin to affect one set of sibilants, but not another, especially considering that the synchronic naturalness of /s/-voicing is a difficult case to argue.

The underspecification proposal which I argued to be central to understanding pre-sonorant voicing does lend itself to an analysis in terms of abstract phonological features, but only insofar as the content of these features is informed by the phonetics. What this entails formally is essentially a ternary distinction, as explicitly argued for Bakairi by Wetzels (1997), and as embraced by a number of generative analyses of pre-sonorant voicing, including Gussmann (1992), and Rubach (1996, 2008) for Polish, Bermúdez-Otero (2001), Wheeler (2005), Jiménez & Lloret (2008) for Catalan, as well as Colina (2009) and Bermúdez-Otero (2011) for Quito Spanish. Theories which reject either ternarity, or phonetically-defined features, on the other hand, struggle to account for the entirety of the empirical facts found in pre-sonorant voicing, which is especially acute in the case of Quito Spanish where top-down and bottom-up diagnostics for features do not converge.

7.3 Modulating factors

In the current proposal phonetic underspecification is central to the initiation of passive voicing. However, whether or not passive voicing may be reinterpreted as involving a voicing target, thus giving rise to pre-sonorant voicing, depends on a number of additional factors. The factors that may play a role in the reinterpretation include:

1. The duration of passive voicing. Due to the Aerodynamic Voicing Constraint (Ohala, 1983), passive voicing will be counteracted in obstruents by the rise of intraoral pressure, and it will decline over time. However, the rate at which this happens may be constrained by a number of factors. For instance a high degree of turbulence will counteract passive voicing due to high intraoral pressure associated with the production of a narrow constriction. The degree of passive voicing may also be affected by place of articulation, due to the effect of place of constriction on the length of the vocal tract, which in turn affects air pressure. For instance, passive voicing in stops is expected to be greater in labials, than in coronals and velars (Keating, 1984).
2. Obstruent duration. Obstruent duration might influence reinterpretation of passive voicing in two ways. First, obstruent duration affects the voicing ratio generated by the presence of passive voicing; the longer the obstruent, the lower the ratio. Therefore, obstruents of relatively greater duration are less likely to be perceived as voiced by listeners. This is reinforced by the fact that increased obstruent duration is a potential cue for voicelessness in itself, for instance increased duration of frication is found in Dutch (Slis & Cohen, 1969) and English (Crystal & House, 1988) voiceless fricatives. Similarly, increased closure duration has been found in voiceless stops (Chen, 1970; Kluender et al., 1988).
3. The salience of glottal fold vibration as a voicing cue. Perceptual studies have shown that presence of vocal fold vibration during frication increases the chance of voiced percepts in listeners (Forrez, 1966; Stevens et al., 1992). It is conceivable that the importance of glottal fold vibration as a voicing cue varies depending on manner of articulation. Voicing in fricatives affects the intensity of high-frequency frication (Ladefoged & Maddieson, 1996),

and it is thus potentially salient. Passive voicing in the closure phase of stops, on the other hand, does not have a similar effect, due to overall low intensity of occlusion, which might make it less perceivable. In addition, initial voicing in stops does not affect the VOT, which is recognised as one of the main voicing cues in stops (Lisker & Abramson, 1964). The different ways in which vocal fold vibration interacts with other voicing cues in stops and fricatives may affect the degree to which passive voicing is perceived by listeners, making fricatives more likely to undergo perceptual reanalysis as voiced due to the presence of passive voicing.

The implications of the articulatory and perceptual factors discussed above for the initiation, maintenance and reanalysis of passive voicing grow out of the existing experimental findings, and are consistent with the empirical tendencies found in pre-sonorant voicing. For instance, the increased susceptibility of fricatives to pre-sonorant voicing, compared to stops, can be explained by asymmetries in perceptual cues to voicing in fricatives and stops.

Similarly, obstruent duration is an important predictor for the positions in which passive voicing is likely to undergo phonologisation. This last issue has to do with the role of the right-hand environment. Passive voicing seen in word-final obstruents followed by sonorants is crucially predicted to be sourced by the preceding, rather than the following vowel/sonorant sound. This prediction follows from the passive nature of sonorant voicing. Since the assumption is that passive voicing does not involve a voicing gesture, it is not expected to trigger any anticipatory laryngeal coarticulation. It may, however, be the source of perseverative coarticulation that extends to the following sound. This prediction is borne out by all the cases of gradient intersonorant voicing seen in the current data, which invariably takes the form of a short voicing tail continuing from the preceding vowel and extending through the initial part of the obstruent. At the same time, however, the typology of pre-sonorant voicing leaves no doubt that the right-hand context matters for the occurrence of voicing. For instance, passive postvocalic obstruent voicing does not seem to phonologise before a pause. One way to explain this tendency in an evolutionary perspective relies on the likelihood of the phonologisation of passive voicing in a prepausal position. Phrase-final environments are commonly associated with lengthening. Utterance-final lengthening in fricatives has been found by Cooper & Danly (1981) for American English fricatives, and for Dutch by Hofhuis et al. (1995). Lengthening is likely to counteract voiced perceptions of passive voicing, due to a lower voicing ratio, which is likely to block passive voicing from phonologising in the utterance-final position. Another factor potentially involved in counteracting utterance-final pre-sonorant voicing is the presence of boundary marking laryngeal spreading and closing gestures commonly found at the end of a phrase (Blevins, 2006). Both of these are devoicing gestures, which directly conflict with passive voicing, making pre-pausal voicing less likely to occur.

A more complex aspect of the right-hand environment's role is related to the effect of manner of articulation, as seen for instance in Catalan, where prevocalic voicing is more advanced in fricatives and clusters than when a sonorant consonant follows. This effect is not predicted by any of the modulating factors discussed thus far, and its full understanding requires further research. A potential direction for research in this respect involves the articulatory and aerodynamic characteristics of obstruent+sonorant and obstruent+vowel sequences. The aerodynamic properties of the vocal tract depend on its shape, which in turn is affected by the supralaryngeal gestures. Therefore, passive voicing may be affected by the degree of supralaryngeal coarticulation. For instance, obstruent+sonorant sequences may display a high degree of gestural overlap associated with a slower drop in air pressure relative to the onset of the consonantal gesture, which creates conditions for passive voicing. Gestural blending is also linked to the issue of duration discussed above, as increase in overlap has a limiting effect on duration. If any systematic articulatory and/or aerodynamic differences were found between obstruent+sonorant and obstruent+vowel sequences, this could indicate the presence of yet another type of functional

pressure of the evolution of pre-sonorant voicing.

7.4 Constraints on pre-sonorant voicing in a diachronic and a synchronic perspective

An idea central to the discussion of factors modulating pre-sonorant voicing within the model pursued in this work is that functional pressures on the development of pre-sonorant voicing are extragrammatical, and they are most accurately interpreted as factors affecting language use. This differs from the approach adopted by a number of formal synchronic models presented in this work, which propose that the cognitive synchronic generalisation involved in pre-sonorant voicing includes not only the distribution, but also some of its conditioning. The perceptual reanalysis scenario which I put forward is inherently a multi-step process, where initiation of voicing, its perception on the listeners' part, re-analysis and stabilisation apply in a sequence. As such, the perceptual account is not compatible with a full explanation from synchronic grammar alone. The existing synchronic grammatical accounts of pre-sonorant voicing rely on an altogether different concept that the conditioning of voicing can be accounted for based on functional pressures on production, and on the analysis of sonorants as actively voiced. Both of these, however, involve a number of empirical predictions that are not borne out by the data.

Let us first consider synchronic functionalist models, which propose that grammars reflect an interaction of direct phonetic pressures on language production and perception (Flemming, 2001; Hayes & Steriade, 2004). This notion is of little help in explaining the undergoer manner asymmetries seen in pre-sonorant voicing, as they appear synchronically unnatural. One of the asymmetries concerns the greater susceptibility of fricatives to undergo pre-sonorant voicing, compared to stops. Production related proposals to account for this asymmetry have suggested a potential aerodynamic influence. For instance, Recasens (2002) notes that since stop articulations involve a complete closure in the vocal tract, the intraoral pressure is expected to drop faster than in the case of fricatives where there is only partial obstruction, and therefore stop voicing may be dispreferred compared to fricative voicing. However, this aerodynamic implication is contested by Ohala (1981), as it is not supported by typological evidence that voiced fricatives are relatively less frequent than voiced stops (Ladefoged & Maddieson, 1996). Ohala (1981) observes that turbulence, characteristic of fricative production, requires high intra-oral pressure, which conflicts with voicing. Ohala & Solé (2010) support this prediction with a typological overview of cases where voicelessness favours obstruency, including spirantisation of devoiced fricatives, emergence of 'buccal' fricatives and voiceless stop assimilation in palatalisation contexts. In the light of these arguments it becomes difficult to argue the case for the relative ease of fricative voicing, which undermines potential effort-based accounts of undergoer manner asymmetries in pre-sonorant voicing.

Even more problematic for a synchronic functionalist account is the case of prevocalic voicing in Catalan, which affects sibilants and stop+sibilant clusters, but not singleton stops. Cluster articulation is expected to combine the aerodynamic properties of stops and sibilants that conflict with voicing. In addition, clusters are longer than singleton consonants, which is yet another factor counteracting voicing. I proposed in Chapter 5 that Catalan cluster voicing is a synchronically unnatural product of rule telescoping, where intervocalic fricative voicing originates from a re-analysis of the rule's environment as prevocalic, whereupon the prevocalic voicing rule extends over all preceding obstruents due to the role of regressive voice assimilation in obstruents clusters, which operates independently in the language. Rule telescoping can be formally incorporated, for instance in an OT analysis, as an interaction of constraints necessitating prevocalic fricative voicing, and AGREE-like constraints to account for the spread of voicing in clusters. However, since a functional motivation for the former is dubious, the analysis could not claim to reflect the

activity of independently motivated pressures on language.

The existence of unnatural patterns in language has of course long been observed (Kiparsky, 1972; Anderson, 1981), and it is not in itself an argument against synchronic functionalism. Hayes (1999) takes up this issue using the example of /s/-voicing in Northern Italian which applies intervocally but not post-nasally. This kind of voicing distribution is not predicted by the relative contextual difficulty of voicing quantified by Hayes based on simulations of aerodynamic conditions in the vocal tract, and it is indeed cross-linguistically rare. The Northern Italian case is considered to have arisen historically through a loss of pre-sibilant nasals, followed by intervocalic /s/ voicing, and a reintroduction of nasal+sibilant sequences through borrowing from Latin (Maiden, 1995). Synchronously, however, the pattern is unnatural, and thus does not lend itself to an analysis based on functional ('grounded' in Hayes' terms) constraints. According to Hayes, cases like the Italian /s/-voicing indicate that ungrounded constraints also have a role in synchronic grammars. Such constraints may be posited by learners based on exposure to a given sound pattern. Insofar as functional analyses admit the possibility of arbitrary pattern acquisition based on induction, pre-sonorant voicing does not appear to challenge the basic principles involved in synchronic functionalism, even though it may undermine specific proposals, such as the idea that manner asymmetries in pre-sonorant voicing follow from constraints on aerodynamic difficulty. What sets pre-sonorant voicing, and the Italian /s/-voicing case apart, however, is that pre-sonorant voicing is not an isolated historical accident. On the contrary, cases of pre-sonorant voicing are typologically well attested. What is more, pre-sonorant voicing shows some recurrent characteristics, such as targeting the positions of neutralisation, or affecting specific manners of articulation, and there is little reason to believe that any of those descend from language contact. This is problematic to the hypothesis that recurrent sound changes result from learners' knowledge of phonetically grounded constraints, as no such relevant constraints are easily identifiable for conditioning pre-sonorant voicing in a synchronic context, yet the process displays similar characteristics in a number of languages. In contrast, as most of the restrictions on pre-sonorant voicing make sense from a more universal functional point of view in a diachronic perspective, the cross-linguistic similarity of pre-sonorant voicing is consistent with an evolutionary view of phonology (Blevins, 2004, 2006), where recurrent sound changes are understood as resulting from the activity of common influences on language use, including phonetic factors.

While synchronic functional approaches may struggle with providing a well grounded explanation of the typological tendencies observed in pre-sonorant voicing, strictly formal analyses do not necessarily fare better. As far as constraint-based approaches are concerned, the issues associated with defining synchronically grounded constraints that condition manner asymmetries in pre-sonorant voicing also extend to markedness-based approaches. Coming back to the Catalan example, a pattern where voicing affects sibilants and top+sibilant clusters before a vowel to the exclusion of singleton stops could fall out from a constraint conditioning prevocalic fricative voicing interacting with a constraint on [voice]-agreement. Once again though, the asymmetry between fricatives and stops as pre-sonorant voicing undergoers is questionable. Since the asymmetry is not synchronically grounded, it does not warrant positing a universal markedness hierarchy that favours fricative voicing to stop voicing. On the other hand, if constraints enforcing stop and fricative voicing were in free ranking, we would expect to find a language that is a mirror image of Catalan, where prevocalic voicing affects stops to the exclusion of fricatives. The existence of such a system would falsify my own proposal, while giving some support to formal accounts. However, based on the typological data on pre-sonorant voicing obtained so far, formal accounts leave the asymmetry bias unexplained.

Another aspect of pre-sonorant voicing that is potentially problematic to formal accounts is the source of voicing in pre-sonorant obstruents. Voicing involves the insertion of a marked feature

that is not underlyingly present, thus incurring a markedness¹ as well as a faithfulness violation simultaneously (in Optimality Theoretic terms). However, as non-contrastive voicing tends to be observed when there is another voiced sound in the vicinity, it has come to be understood as feature spreading between neighbouring segments, rather than feature insertion. Feature spreading has formed a central part of the existing accounts of pre-sonorant voicing, including including Bethin (1984), Gussmann (1992) and Rubach (1996, 2008) for Polish, Bermúdez-Otero (2001), Wheeler (2005), Jiménez & Lloret (2008) for Catalan, as well as Currie-Hall (2006) and Blaho (2008) for Slovak. All these approaches rely on positing laryngeal features for sonorants (including vowels), so that sonorants become a source of laryngeal spreading. Within these proposals there are two types of justification for assuming such representations in the absence of laryngeal contrast in sonorants. Bethin (1984), Gussmann (1992) and Rubach (1996) talk about ‘default filling’, an idea that non-contrastive but predictable features may be inserted in the course of phonological derivation by means of a redundancy rule, which is not specific to pre-sonorant voicing languages. Pre-sonorant voicing is then implemented by means of a specific rule of [voice] spreading from sonorants to preceding delaryngealised sounds. The proposals by Bermúdez-Otero (2001), Wheeler (2005), Jiménez & Lloret (2008), Currie-Hall (2006) and Blaho (2008) are constraint- rather than rule-based, and thus do not involve default filling, which is inherently derivational. Instead, the authors propose that sonorants may be laryngeally specified in some languages, and that pre-sonorant voicing constitutes evidence for such specifications. In addition to this, Currie-Hall (2006) and Blaho (2008) argue for sonorant-specific laryngeal feature [Sonorant voice], which is different from [voice] in obstruents.

The proposal that sonorants are laryngeally specified for [voice] predicts that there should not be any differences between [voice] spreading from obstruents and sonorants, as the output of the process is in both cases phonologically identical, i.e. an obstruent with a [voice] specification. This is falsified by the current data, particularly in the case of languages where pre-sonorant and pre-obstruent voicing apply with a different frequency, including Central Catalan, Poznań Polish, and Quito Spanish. In all those languages pre-sonorant voicing was found to be relatively less frequent than voicing before obstruents (cf. Section 7.1.5). A descriptively adequate grammar should be able to accommodate the differences in rate of application of categorical processes. In order to achieve that, a model would need to distinguish between pre-sonorant and pre-obstruent voicing, effectively referencing the manner of articulation of the trigger in the formalisation of pre-sonorant voicing. The consequences of this are as follows. First, the step of positing laryngeal sonorant specifications becomes redundant; if the grammar must place explicit constraints on voicing before sonorants and obstruents, it will work much to the same effect whether or not the voicing is formalised as spreading, or as feature insertion. Apart from that, the asymmetry between pre-sonorant and pre-obstruent voicing undermines the argument for assuming laryngeal specifications for sonorants based on similar phonological behaviour (i.e. voicing before sonorants and obstruents), as the similarity is only apparent. Finally, assuming laryngeal specifications for sonorants adds to the degree of arbitrariness in the synchronic grammar, as synchronic factors do not immediately explain why there should be any difference in the level of application of pre-sonorant and pre-obstruent voicing given identical specifications. In comparison, different rates of application between pre-sonorant and pre-obstruent voicing are readily explicable from a diachronic perspective, as the two processes develop largely independently of each other, being conditioned by different factors, and therefore they may be at different stages of stabilisation at any point in time. Interestingly, a discrepancy between pre-sonorant and pre-obstruent voicing is predicted by the proposal that sonorants and voiced obstruents are specified for different laryngeal features

¹The markedness violation follows the assumption that [\pm voice] is indeed the relevant feature at work in Catalan, which is supported by the phonetics of Catalan voicing, by evidence of regressive [voice] assimilation, and from the asymmetry between voiceless and laryngeally underspecified obstruents in their susceptibility to prevocalic voicing (cf. Section 7.2).

(Currie-Hall, 2006; Blaho, 2008). Trigger asymmetries, however, are not limited to an opposition between contrastively voiced obstruents and non-contrastively voiced sonorants. In Catalan there is a further split, where prevocalic voicing is more frequent than voicing before a sonorant consonant in sibilants and stop+sibilant clusters, whereas in stops, only pre-sonorant voicing is observed. If this discrepancy was also to be reflected in separate laryngeal features for sonorants and vowels, the distinction would have to be posited solely based on exposure to the relevant phonological pattern, as sonorant consonants and vowels are both phonetically voiced, but phonologically not contrastive. This is formally possible, but somewhat circular: feature systems are derived based on the existence of phonological patterns, and then proposed to account for the very same patterns that motivated the features in the first place.

7.5 Arbitrariness and the role of phonology

What emerges from the considerations above is that pre-sonorant voicing is not an impossible process to formalise in synchronic terms, but any such formalisations are necessarily arbitrary. This need not necessarily be a problem, although arbitrariness has frequently been evoked as criticism of phonological analyses and entire frameworks have been postulated in partial motivation of improving on this issue. Re-write rules descending from Chomsky & Halle (1968) have been particularly singled out for their arbitrariness. Since re-write rule formalism is virtually unrestricted, it does not provide any inherent links between processes and the environments in which they occur. This makes the formalism vulnerable to overgeneration, because both frequent and rare, and even entirely non-existent sound patterns are formally equally likely. Various proposals have been put forward to make grammars more restrictive. One approach is to limit the scope of phonological operations by introducing formal restrictions on the number of units, structural properties and formal relationships computed by the phonological components, as advocated by the proponents of Government Phonology (e.g. Kaye (1989), Harris (1990)). Another approach, associated with the framework of Optimality Theory (Prince & Smolensky, 2004 [1993]) is to incorporate the notion of phonological markedness into the grammar, so that frequent sound patterns follow from restrictions against marked structures.

Attempts to restrict phonological grammars in the ways discussed above have faced at least two kinds of problems. One is that of undergeneration, as there are numerous examples of unnatural synchronic processes that defy a narrow set of markedness restrictions (Vaux, 2008). Another problem concerns defining the content of markedness constraints. A number of relevant defining criteria for markedness have been proposed in the literature, including typological frequency, order of acquisition and substitution patterns in speech disorders (see Rice (2007) for an overview). However, a frequent practice of introducing new markedness constraints is to use the existence of a particular sound pattern as evidence, where the necessary markedness restrictions are evoked as required to generate the observed generalisation. Although such practice may be justifiable as a tentative exploration into how a particular analysis may work, it is symptomatic of a larger issue associated with using markedness to account for linguistic typology: if constraints cannot be motivated independently of the existence of a particular pattern, they fail to achieve one of the goals driving the development of the very concept of markedness, i.e. reducing arbitrariness in grammar. Data from pre-sonorant voicing are relevant to this debate in providing an example of a typologically quite regular pattern that is formally expressible in terms of markedness influences, but there is no evidence for the relevant constraints outside of the typology. This is instantiated by the case of manner effects in pre-sonorant voicing, which I previously discussed in the context of synchronic functionalism. Formally, it is possible to posit a markedness constraint which favours pre-sonorant voicing in fricatives to a similar process in stops. However, since it is difficult to

provide an independent motivation for such a constraint (cf. 7.4), there is no genuine explanatory value to such a formalisation.

Another response to the issue of arbitrariness in the phonological literature has been to accept arbitrariness as an aspect of formal grammar. It is largely undisputed that arbitrary processes do form a part of phonological grammars, and hence they must be learnable. The question that arises at this point, however, is whether unnatural sound patterns exist at the periphery of phonological grammars subject to a special learning process (Hayes, 1999), or whether arbitrariness is in fact a core property of abstract phonology². The latter stance has been taken by a number of phonologists representing a wide range of approaches, including Evolutionary Phonology (Blevins, 2004, 2006), Exemplar Theory (Silverman, 2006b), Substance-Free Phonology (Hale & Reiss, 2008; Blaho, 2008), and Rule-Based Phonology (Vaux, 2008). What these proposals largely share is the idea that synchronic phonology does not directly encode the factors responsible for conditioning sound patterns. Instead, the explanation for typological frequencies is sought in extra-grammatical factors. These are typically understood in the context of influences on sound change, the main idea being that while synchronic grammars may produce arbitrary patterns, the development of such patterns will be limited by external pressures, including phonetic factors, speaker-hearer interaction in a speech community, and language acquisition.

An important corollary of forfeiting typology as a phonological explanandum is the need to re-evaluate the role of abstract phonology in our understanding of sound patterns. Associated with this is the issue of how to constrain formal grammars. Typological predictions have a long standing status of an important criterion for discriminating between phonological analyses, certainly within Optimality Theory, where constraint permutations are used to derive factorial typologies. A close mapping between the generative power of the formal tools employed and the observed language variation has been considered a mark of descriptive adequacy of grammars. Giving up on the idea of explaining sound patterns from within the grammars considerably expands generative power, thus also limiting the role of descriptive adequacy in evaluating grammars, insofar as descriptive adequacy is directly linked to typology. This outcome may not be altogether undesirable. The effectiveness of typologies in evaluating grammatical power is seriously limited by the availability and quality of data, as well as accidental gaps (Hale & Reiss, 2000, 2008; Reiss, 2008; Buckley, 2009). However, there is another way of looking at descriptive adequacy, which does not involve typology. A descriptively adequate grammar could be seen as one where the proposed abstract representations map onto the observed variation in speech, and where the assumed categories make predictions about the potential directions for sound change. The role of phonological grammar thus defined may seem relatively modest compared to the goal of generating linguistic typologies. Nevertheless, a careful consideration of evidence reveals that phonological generalisations still have a crucial role to play in our understanding of sound patterns and sound systems.

In 2.1 I had considered whether we need to posit a level of phonological knowledge at all, given that language users show highly detailed and sophisticated subphonemic behaviour which does not seem to be mediated by abstract categories. Some of the rationale for considering phonological abstraction as a part of linguistic competence comes from sound change, as sound patterns frequently stabilise in a way which suggests a re-analysis of the observed variation in terms of abstract units on the part of the learners. Further support for abstract phonology is provided by phonology-morphology interactions and salience of phonological categories. All of these arguments are further supported by data presented in this thesis, as phonological influences are evident throughout the conditioning and development of pre-sonorant voicing. One instance of this is delaryngealisation. The data strongly support the case of a conditioning relationship between delaryngealisation and pre-sonorant voicing. Although delaryngealisation as such may be expressed

²Cf. 5.4.1 for more discussion on this point.

phonetically (as absence of a voicing or devoicing gesture), the environment in which it applies is defined in abstract prosodic or morphosyntactic terms. This is evidenced by the analogical extension of delaryngealisation from the domain where it is directly phonetically motivated (phrase-final position) to narrower domains where some of the direct phonetic conditioning is lost (word-final and syllable-final environment). What is more, delaryngealisation shows reflections of phonology interacting with morphology, as shown by the opaque application of word-final delaryngealisation in the prevocalic environment in Catalan and Spanish. Another example of phonological behaviour in pre-sonorant voicing is stabilisation of voicing itself, i.e. the stage when the process moves beyond passive voicing and starts exhibiting categorical characteristics. This, again, is associated with a loss of some original phonetic motivation that drives the origins of the process. For instance, upon stabilisation pre-sonorant voicing becomes less sensitive to obstruent duration. This is illustrated by the speech rate manipulations seen in Quito Spanish /s/-voicing. The speakers for whom /s/-voicing was gradient showed increased voicing ratio in relatively shorter obstruents, in accordance with the hypothesis that faster speech involves relatively more laryngeal overlap. Those speakers for whom /s/-voicing had stabilised, on the other hand, showed a steady voicing ratio in the /s/-voicing context regardless of the differences in obstruent duration induced by speech rate manipulations. This kind of behaviour is consistent with an analysis where voicing is interpreted by the speaker as involving the presence of an abstract category which becomes orthogonal to the presence of functional pressures involved in the origin of the process. Finally, the most striking example of phonological generalisation influencing sound change comes from the extension of intervocalic sibilant voicing in Catalan to prevocalic clusters. The direction of change shows a widening of the conditioning environment from a narrow domain where the process is phonetically motivated (intervocalic sibilants) to a broader domain where some of the phonetic conditioning is lost (sibilants before a vowel), coupled with the effect of rule telescoping whereby the output of a new sound pattern begins to serve as an input to other process operating independently within the language (voicing in obstruent clusters). The resulting pattern is synchronically unnatural, as voicing is found in prevocalic stop+sibilant clusters, but not in prevocalic singleton stops, which suggests the influence of abstract generalisations formed by language users as driving the direction of change beyond its initial stages.

7.6 The relationship between phonology and phonetics

In my analysis of pre-sonorant voicing I have assumed a separation between phonology and phonetics, operationalised as a distinction between empirically defined categoricity and gradience. Such a distinction is supported by the current data on a descriptive level, as instances of both gradient and categorical patterns are found in the typology of pre-sonorant voicing. However, as strict separation between phonology and phonetics has been contested in the literature, it is important to re-assess this assumption in the light of its validity and its usefulness in gaining an insight into sound patterns and their cognitive representation in the mind of language users.

One of the arguments against viewing phonology and phonetics as separate modules of grammar is based on the existence of considerable duplication between the two components, as phonetic and phonological patterns show many similarities. I have discussed in 2.3.1 why duplication between phonology and phonetics may not be a problem, but in fact a intrinsic property of grammars shaped by phonological stabilisation of phonetic patterns. This position is further supported by pre-sonorant voicing data, especially by the systematic differences observed between gradient and categorical patterns, with the former showing a tendency for phonetically natural behaviour, whereas the latter may develop phonetically idiosyncratic properties (cf. the discussion on the role of phonetic duration in gradient and categorical /s/-voicing in Quito Spanish, and rule telescoping

in Catalan in 7.5). In addition, pre-sonorant voicing data problematise the proposal put forward as an alternative to a modular distinction, namely the idea that phonetic substance forms part of the speakers' cognitive knowledge, and that this knowledge is then directly reflected in recurrent sound changes. Contrary to this proposal, pre-sonorant voicing shows the characteristics of a recurrent sound change (being attested in a number of languages), but whose exact phonetic motivation is only transparent in a diachronic perspective, and so it does not translate directly into synchronic constraints.

A more serious challenge to the phonetics-phonology distinction assumed in the current approach is the issue of simplification, which is associated with the existence of 'grey areas'. Pierrehumbert et al. (2000) and Scobbie (2005) point out that while phonetics and phonology may not be the same thing, it is difficult to draw a line between the two. This problem is also present in the current data, which show cases where categorical and gradient patterns may coexist within a single system, or even within a single speaker, as well as cases where patterns cannot be straightforwardly classified as either gradient or categorical. The latter problem surfaces with respect to obstruent voicing before sonorant consonants in Catalan, where the high degree of inter- and intra-speaker variation defy a straightforward classification of the process. Indeed, the point that not every process can always be labelled as categorical or gradient can hardly be contested. However, while a strict categoricity-gradience separation may not always be viable, attempting it still appears informative enough to be considered a useful enterprise. From the point of view of phonologisation and stabilisation assumed in this work, processes undergoing stabilisation may simultaneously display properties of categoricity and gradience. This may manifest itself for instance as inter-speaker variation between categorical and gradient realisations (e.g. West-Flemish), as intra-speaker variation (Quito Spanish), or other possible kinds of variation, such as lexical diffusion. Importantly, although none of these cases are canonically categorical or gradient, they still lend themselves to a relatively informative description in terms of where they fall in the continuum. This may be further illuminated with the use of multiple empirical diagnostics. Speech rate manipulations, as seen in the case of the Quito Spanish data, are particularly promising in this respect, giving us a better insight into whether the speaker's behaviour is consistent with the effect of being modulated by the presence of an abstract category.

As a final point, while a distinction between categorical phonology and gradient phonetics may not translate into two clearly separate empirical categories, it feeds into a conceptually clear distinction. Individual patterns may simultaneously show categorical and gradient characteristics, but the characteristics themselves are distinguishable as belonging to either phonology or phonetics. Similarly, phonetic and phonological influences on sound change can be separated into two distinct categories, as the former represent articulatory, acoustic and aerodynamic pressures on language, whereas the latter reflect abstract generalisations that speakers make about sound patterns. As sound patterns continue to change due to the influences exerted by both phonetic and phonological factors, it is not surprising that phonetics-phonology distinctions may at times appear blurred from an empirical point of view. This, however, does not automatically render it incoherent to understand phonology and phonetics as distinct levels of representation with the possibility of individual processes to move gradually from one level to another.

CHAPTER 8

CONCLUSIONS, MAIN FINDINGS, DIRECTIONS FOR RESEARCH

‘Well, in my book, Jeeves, I’ll simply pose these Questions. I won’t come up with any Answers. But that’s all right because you don’t have to be conclusive in novels about the human condition.’

*Joanathan Aimes
WAKE UP, SIR*

From the empirical point of view the data presented in this thesis deliver a number of generalisations relevant to an understanding of pre-sonorant voicing, and voicing processes in general. One of the chief empirical findings that emerge from the data concerns positional co-occurrence between laryngeal neutralisation and pre-sonorant voicing. Another important observation has to do with manner effects that modulate the application of pre-sonorant voicing: fricatives are more likely to undergo the process than stops, and fricative voicing shows a tendency to be more likely before a vowel than before a sonorant consonant. In contrast, the current data do not offer support to another generalisation on pre-sonorant voicing previously proposed in the literature, namely that pre-sonorant voicing is a case of [voice]-assimilation which happens to involve sonorants as triggers due to redundant laryngeal specifications. In all languages examined in this work pre-sonorant voicing and voicing before obstruents show systematic differences with respect to environments where they apply, and with respect to rates of application. This finding motivates an analysis of pre-sonorant voicing as a distinct process from voice assimilation to obstruents. Finally, instances of both categorical and gradient pre-sonorant voicing were found in the data, with categorical cases being considerably more common.

Findings concerning categoricity and gradience highlight the need for future analyses of external sandhi cases to distinguish between different types of variation in speech patterns, including the opposition between gradient and categorical but optional processes. This distinction deserves attention if only for the sake of descriptive adequacy, and it requires employing a range of relevant empirical diagnostics. In this study I have focused on two specific criteria in this respect: the bimodality criterion and duration effects conditioned by speech rate manipulations. Although not unproblematic or universally applicable, both methods offer a closer insight into the type of variation characterising sound patterns. Another methodological finding emerging from this work concerns the need to consider individual variation in quantifying empirical results from phonetic research. This is especially true of external sandhi cases, as they are inherently variable, but the type of variation may differ from speaker to speaker.

When it comes to explaining the typological tendencies found in pre-sonorant voicing, I have argued that they are most accurately understood as synchronic reflections of a range of phonetic and phonological factors influencing the initiation and development of pre-sonorant voicing at different diachronic stages. I have questioned the hypothesis that pre-sonorant voicing involves anticipatory

laryngeal coarticulation between an obstruent and the following laryngeally specified sonorant. Instead, I have argued for a scenario where pre-sonorant voicing emerges from a series of sound changes, including delaryngealisation, passive voicing, and re-interpretation of the passive voicing as involving a laryngeal gesture on the part of the speaker. Under this view, pre-sonorant voicing is in essence a perception-driven change. This might explain some of the positional asymmetries and manner asymmetries found in pre-sonorant voicing, as re-interpretation of passive voicing is hypothesised to be more likely in positions of delaryngealisation due to the extended duration of passive voicing. At the same time, some positions, such as phrase-final environments, are expected to resist re-interpretation due to lengthening effects and laryngeal boundary marking. Manner asymmetries may also emerge from acoustic influences on perception on voicing, as the acoustic cues to voicing in fricatives may be more affected by passive voicing than cues to stop voicing.

The diachronic re-analysis scenario proposed provides explicit hypotheses which can hopefully be evaluated by further research into perception of voicing and reinterpretation. One research question that emerges in this respect concerns the conditions for passive voicing to be re-interpreted as intended by listeners. This includes establishing what threshold of voicing duration may facilitate categorically voiced percepts in listeners, and how this is modulated by obstruent duration and manner. Further avenues for research relevant to the current findings on pre-sonorant voicing involve an exploration of trigger effects observed in pre-sonorant voicing. One particularly interesting case is that of Catalan, where fricative voicing is more advanced and more frequent before vowels than before sonorant consonants, but stop voicing is only found before sonorant consonants, and not before vowels. A pattern as complex as this is likely to have originated from an interplay of effects, which may include articulatory and aerodynamic pressures. Our understanding of what factors are at work could benefit from an in-depth study of articulatory and aerodynamic characteristics of obstruent+sonorant and obstruent+vowel sequences, since the two types of sequences may reveal some systematic differences with a potential phonetic bias for voicing in one environment over the other. Another approach to exploring the manner asymmetries seen in Catalan could be driven by lexical effects, such as token frequency.

Further experimental evidence concerning the role of perceptual factors in re-interpretation of voicing would yield support to the vision of phonology pursued in this work, where sound patterns are primarily shaped through pressures external to grammar to then be perceived, encoded and reproduced by language users and language learners. I have considered this process to comprise two aspects: encoding in terms of abstract categories which form a part of the phonological knowledge, and learning transparently defined phonology-phonetic mappings which form a part of the phonetic knowledge. This notion brings forth the final and most complex research question that follows from the analysis pursued in this work, and which has to do with phonological category formation and category learning. Pre-sonorant voicing delivers evidence for categorical phonetic behaviour which is not in a one-to-one mapping with top-down diagnostics for phonological categories such as contrast. This suggests that phonological categories should be discoverable by learners based on bottom-up methods, for instance involving clustering of tokens in a continuous phonetic space. This kind of category formation has been modelled using artificial neural networks (Guenther & Gjaja, 1996; Boersma, 2012b), but the question still remains of how the category peaks that emerge from a phonetic space become linked with abstract phonological category labels (Boersma, 2012a). While this problem yet awaits a successful computational solution, the existence of mismatches between phonological categories as defined on top-down and bottom-up basis provides empirical motivation for the pursuit of this issue, and for further experimental explorations into the nature of phonological categories.

BIBLIOGRAPHY

- Anderson, S. R. (1981). Why phonology isn't "natural". *Linguistic Inquiry* 12, pp. 493–539.
- Anonby, E. J. (2003). *A phonology of Southern Luri*. Munich: Lincom.
- Avery, P., B. E. Dresher & K. D. Rice (eds.) (2008). *Contrast in Phonology: Theory, Perception, Acquisition*. Berlin: Mouton de Gruyter.
- Avery, P. J. (1996). *The representation of voicing contrasts*. Ph.D. thesis, University of Toronto.
- Baayen, R. H. (2011). *languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics"*. URL <http://CRAN.R-project.org/package=languageR>. R package version 1.2.
- Bárkányi, Z. & Z. G. Kiss (2012). On the border of phonetics and phonology: Sonorant voicing in Hungarian and Slovak. Paper presented at the Twentieth Manchester Phonology Meeting.
- Barry, M. C. (1992). Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication* 11, pp. 393–400.
- Barry, S. (1988). Temporal aspects of the devoicing of word-final obstruents in Russian. Holmes, J. N. & W. A. Ainsworth (eds.), *Speech'88. (Proceedings of the Federation of Acoustical Societies of Europe, August 1988)*, Edinburgh: Institute of Acoustics, pp. 81–88.
- Bates, D. & M. Maechler (2009). *lme4: Linear mixed-effects models using S4 classes*. URL <http://CRAN.R-project.org/package=lme4>. R package version 0.999375-32.
- Baumann, M. (1995). *The Production of syllables in connected speech*. Ph.D. thesis, University of Nijmegen.
- Beckman, J., P. Helgason, B. McMurray & C. Ringen (2011). Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics* 39, pp. 39–49.
- Beckman, M., K. De Jong, S. Jun & S. Lee (1992). The interaction of coarticulation and prosody in sound change. *Language and Speech* 35, pp. 45–58.
- Benni, T. (1959). *Fonetika opisowa języka polskiego*. Wrocław: Zakład Narodowy im. Ossolińskich.
- Bermúdez-Otero, R. (2001). Voicing and continuancy in Catalan: A nonvacuous Duke-of-York gambit and a Richness-of-the-Base paradox, URL <http://www.bermudez-otero.com/Catalan.pdf>. Manuscript, University of Manchester.
- Bermúdez-Otero, R. (2007). Diachronic phonology. de Lacy, P. (ed.), *The Cambridge handbook of phonology*, Cambridge: Cambridge University Press, pp. 497–517.
- Bermúdez-Otero, R. (2010). Morphologically conditioned phonetics? Not proven. Talk presented at 'On Linguistic Interfaces II', Belfast.

- Bermúdez-Otero, R. (2011). Cyclicity. van Oostendorp, M., C. J. Ewen, E. Hume & K. Rice (eds.), *Blackwell companion to phonology*, Chichester: Wiley-Blackwell, vol. IV, pp. 2019–2048.
- Bermúdez-Otero, R. (2012). Traces of change in synchronic phonology: English syllabification and the life cycle of lenition, URL <http://www.bermudez-otero.com/PortoAlegre.pdf>. Paper presented at the IV Seminário Internacional de Fonologia, Porto Alegre, 26 April 2012.
- Bermúdez-Otero, R. & G. Trousdale (2012). Cycles and continua: on unidirectionality and gradualness in language change. Nevalainen, T. & E. C. Traugott (eds.), *Handbook on the history of English: rethinking and extending approaches and methods*, New York: Oxford University Press, pp. 691–720.
- Bethin, C. Y. (1984). Voicing assimilation in Polish. *International Journal of Slavic Linguistics and Poetics* 29, pp. 17–32.
- Blaho, S. (2008). *The syntax of phonology: A radically substance-free approach*. Ph.D. thesis, Universitetet i Tromsø.
- Blevins, J. (2004). *Evolutionary Phonology. The emergence of sounds patterns*. Cambridge: Cambridge University Press.
- Blevins, J. (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32, pp. 117–166.
- Boersma, P. (2008). Emergent ranking of faithfulness explains markedness and licensing by cue, URL <http://roa.rutgers.edu/>. Available as ROA 954.
- Boersma, P. (2012a). Modelling phonological category learning. Cohn, A. C., C. Fougeron & M. K. Huffman (eds.), *Handbook of Laboratory Phonology*, Oxford: Oxford University Press, pp. 207–218.
- Boersma, P. (2012b). Phonological categories as attractors of behaviour. Paper presented at the Twentieth Manchester Phonology Meeting.
- Boersma, P. & D. Weenink (2010). Praat: doing phonetics by computer [Computer programme]. Version 5.1.12, retrieved 15 October 2009 from <http://www.praat.org/>.
- Bonet, E. & M.-R. Lloret (1998). *Fonologia catalana*. Barcelona: Ariel.
- Booij, G. (1995). *The Phonology of Dutch*. Oxford: Clarendon.
- Braver, A. (2011a). Incomplete neutralization in American English flapping: A production study. *Proceedings of the 34th Annual Penn Linguistics Colloquium*, Penn Linguistics Club, vol. 7 of *University of Pennsylvania Working Papers in Linguistics*.
- Braver, A. (2011b). Perception of incompletely neutralized /d/ and /t/ flaps in American English. Poster presented at NELS 42.
- Brownman, C. P. (1995). Assimilation as gestural overlap: comments on Holst and Nolan. *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence* pp. 334–342.
- Brownman, C. P. & L. Goldstein (1986). Towards an articulatory phonology. *Phonology yearbook* 3, pp. 219–252.
- Brownman, C. P. & L. Goldstein (1992). Articulatory phonology: An overview. *Phonetica* 49, pp. 155–180.

- Brownman, C. P. & L. M. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* 6, pp. 201–251.
- Buckley, E. (2009). Locality in metrical typology. *Phonology* 26, pp. 389–435.
- Byrd, D. (1994). *Articulatory timing in English consonant sequences*. Ph.D. thesis, UCLA.
- Chappell, W. (2011). The intervocalic voicing of /s/ in Ecuadorian Spanish. *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, Somerville, MA: Cascadilla Proceedings Project, pp. 57–64.
- Charles-Luce, J. (1985). Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics* 13, pp. 309–324.
- Charles-Luce, J. (1993). The effects of semantic context on voicing neutralization. *Phonetica* 50, pp. 28–43.
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, pp. 129–159.
- Cho, T. & P. Ladefoged (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27, pp. 207–229.
- Cho, Y.-m. Y. (1990). A typology of voicing assimilation. *Proceedings of the ninth West Coast Conference on Formal Linguistics*, Stanford: CSLI.
- Chomsky, N. & M. Halle (1968). *The Sound Pattern of English*. Harper & Row. New York.
- Coetzee, A. & J. Pater (2011). The place of variation in phonological theory. Goldsmith et al. (2011), pp. 401–434.
- Cohn, A. C. (1990). *Phonetic and phonological rules of nasalization*. Ph.D. thesis, UCLA.
- Cohn, A. C. (1993). Nasalization in English: phonology or phonetics. *Phonology* 10, pp. 43–82.
- Cohn, A. C. (2006). Is there gradient phonology? Fanselow, G., C. Féry, M. Schlesewsky & R. Vogel (eds.), *Gradience in Grammar: Generative Perspective*, Oxford: OUP, pp. 25–44.
- Cohn, A. C. (2007). Phonetics in phonology and phonology in phonetics. *Working Papers of the Cornell Phonetics Laboratory* 16, pp. 1–31.
- Cole, J. & K. Iskarous (2001). Effects of vowel context on consonant place identification: Implications for a theory of phonologization. Hume, E. & K. Johnson (eds.), *The role of speech perception in phonology*, San Diego: Academic Press, pp. 103–122.
- Colina, S. (2009). Sibilant voicing in Ecuadorian Spanish. *Studies in Hispanic and Lusophone Linguistics* 2, pp. 3–29.
- Collins, B. & I. M. Mees (1999). *The Phonetics of English and Dutch*. Leiden: Brill, 4th edn. First published in 1981.
- Cooper, W. E. & M. Danly (1981). Segmental and temporal aspects of utterance-final lengthening. *Phonetica* 38, pp. 106–115.
- Crystal, T. H. & A. S. House (1988). A note on the durations of fricatives in American English. *Journal of the Acoustical Society of America* 84, pp. 1932–1935.
- Cuartero Torres, N. (2001). *Voice assimilation in Catalan and English*. Ph.D. thesis, Universitat Autònoma de Barcelona.

- Currie-Hall, D. (2006). *The Role and Representation of Contrast in Phonological Theory*. Ph.D. thesis, University of Toronto.
- Cyran, E. (2012). Cracow sandhi voicing is neither phonological nor phonetic. It is both phonological and phonetic. Cyran, E., H. Kardela & B. Szymanek (eds.), *Sound, Structure and Sense. Studies in Memory of Edmund Gussmann*, Lublin: Wydawnictwo KUL, pp. 153–183.
- Daniloff, R. G. & R. E. Hammarberg (1973). On defining coarticulation. *Journal of Phonetics* 1, pp. 239–248.
- De Schutter, G. & J. Taeldeman (1986). Assimilatie van Stem in de Zuidelijke Nederlandse Dialekten. Devos, M. & J. Taeldeman (eds.), *Vruchten van z'n akker: opstellen van (oud-) medewerkers en oud-studenten voor Prof. V.F. Vanacker*, Ghent: Seminarie voor Nederlandse Taalkunde, pp. 91–133.
- Dejna, K. (1973). *Dialekty polskie*. Wrocław: Zakład Narodowy im. Ossolińskich.
- Dinnsen, D. A. & J. Charles-Luce (1984). Phonological neutralization, phonetic implementation, and individual differences. *Journal of Phonetics* 12, pp. 49–60.
- Dresher, B. E. (2009). *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.
- Ellis, L. & W. J. Hardcastle (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. *Journal of Phonetics* 30, pp. 373–396.
- Ernestus, M. (2000). *Voice Assimilation and Segment Reduction in Casual Dutch*. Ph.D. thesis, Free University of Amsterdam/Holland Institute of Generative Linguistics.
- Ernestus, M. & H. Baayen (2006). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. Goldstein, L. M., D. Whalen & C. T. Best (eds.), *Laboratory phonology 8*, Berlin & New York: Mouton De Gruyter, pp. 27–49.
- Fischer-Jørgensen, E. (1954). Acoustic analysis of stop consonants. *Miscellanea Phonetica* 2, pp. 42–59.
- Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18, pp. 7–44.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT press.
- Forrez, G. (1966). Relevante parameters van de stemhebbende fricatief /z/. *Inst. Perceptie Onderzoek Verslag (Eindhoven)*.
- Foulkes, P., J. M. Scobbie & D. Watt (2010). Sociophonetics. Hardcastle, W., J. Laver & F. E. Gibbon (eds.), *The Handbook of Phonetic Sciences*, Wiley Blackwell, pp. 703–764, second edn.
- Fourakis, M. & G. Iverson (1984). On the ‘incomplete neutralization’ of German final obstruents. *Phonetica* 41, pp. 140–149.
- Francard, M. & Y. Morin (1986). Sandhi in Walloon. Andersen, H. (ed.), *Sandhi phenomena in the languages of Europe*, Berlin: Mouton de Gruyter, pp. 453–74.
- Goldinger, S. D. (1997). Words and voices—perception and production in an episodic lexicon. Johnson, K. & J. W. Mullennix (eds.), *Talker Variability in Speech Processing*, San Diego: Academic Press, pp. 33–66.

- Goldinger, S. D. (2007). A complementary-systems approach to abstract and episodic speech perception. Trouvain, J. & W. J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken: Saarland University, pp. 49–54.
- Goldrick, M. & S. E. Blumstein (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes* 21, pp. 649–683.
- Goldrick, M., H. Ross Baker, A. Murphy & M. Baese-Berk (2011). Interaction and representational integration: Evidence from speech errors. *Cognition* 121, pp. 58–72.
- Goldsmith, J., J. Riggle & A. Yu (eds.) (2011). *The Handbook of Phonological Theory, 2nd Edition*. Oxford: Blackwell.
- Gradoville, M. (2010). A gradient approach to fricative voicing comparing acoustic measures to perception. Conference Handout. Laboratory Approaches to Romance Phonology.
- Gradoville, M. S. (2011). Validity in measurements of fricative voicing: Evidence from Argentine Spanish. Alvord, S. M. (ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, Somerville, MA: Cascadilla Proceedings Project, pp. 59–74.
- Guenther, F. & M. Gjaja (1996). The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America* 100, pp. 1111–1121.
- Gussmann, E. (1992). Resyllabification and delinking: the case of Polish voicing. *Linguistic Inquiry* 23, pp. 29–56.
- Gussmann, E. (2007). *The phonology of Polish*. Oxford: Oxford University Press.
- Hale, M. & C. Reiss (2000). Phonology as cognition. *Phonological knowledge: Conceptual and empirical issues* pp. 161–184.
- Hale, M. & C. Reiss (2008). *The phonological enterprise*. Oxford University Press, USA.
- Harris, J. (1990). Segmental complexity and phonological government. *Phonology* 7, pp. 255–300.
- Harris, J. & G. Lindsey (1993). There is no level of phonetic representation. *UCL Working Papers in Linguistics* 355–373, pp. 355–373.
- Harris, J. W. (1969). *Spanish phonology*. Cambridge, MA: MIT Press.
- Harris, J. W. (1983). *Syllable structure and stress in Spanish: a nonlinear analysis*. Cambridge, MA: MIT Press.
- Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31, pp. 373–405.
- Hayes, B. (1999). Phonetically driven phonology: The role of Optimality Theory and inductive grounding. Darnell, M., E. Moravscik, M. Noonan, F. Newmeyer & K. Wheatly (eds.), *Functionalism and Formalism in Linguistics*, Amsterdamy: John Benjamins, vol. 1, pp. 243–285.
- Hayes, B. & D. Steriade (2004). Introduction. the phonetic bases of phonological markedness. Hayes, B., R. M. Kirchner & D. Steriade (eds.), *Phonetically Based Phonology*, Cambridge: Cambridge University Press, pp. 1–33.
- Hayes, B., R. M. Kirchner & D. Steriade (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.

- Herd, W., A. Jongman & J. Sereno (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics* 38, pp. 504–516.
- Herrera Z., E. (2000). Amuzgo and Zapotec: two more cases of laryngeally complex languages. *Anthropological Linguistics* 42, pp. 545–563.
- Hofhuis, E., C. Gussenoven & T. Rietveld (1995). Final lengthening at prosodic boundaries in Dutch. Elenius, K. & P. Branderup (eds.), *Proceedings of 13th International Congress of Phonetic Sciences*, Stockholm, vol. 1, pp. 154–157.
- Holes, C. (1990). *Gulf Arabic*. London: Croom Helm.
- Holst, T. & F. Nolan (1995). The influence of syntactic structure on [s] to [ʃ] assimilation. Connell, B. & A. Arvaniti (eds.), *Phonology and phonetic evidence: papers in Laboratory Phonology IV*, Cambridge: Cambridge University Press, pp. 315–333.
- Honeybone, P. (2005). Diachronic evidence in segmental phonology: the case of obstruent laryngeal specifications. van Oostendorp, M. & J. van de Weijer (eds.), *The internal organization of phonological segments*, Berlin: Mouton De Gruyter, pp. 319–354.
- Hothorn, T., K. Hornik & A. Zeileis (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15:3, pp. 651–674.
- House, A. S. & G. Fairbanks (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25, pp. 105–113.
- Hyman, L. M. (1975). *Phonology: theory and analysis*. New York: Holt, Rinehart and Winston.
- Hyman, L. M. (1976). Phonologization. Julland, A. (ed.), *Linguistic studies presented to Joseph H. Greenberg*, Saratoga, pp. 407–418.
- Hyman, L. M. (2001). The limits of phonetic determinism in phonology: *NC revisited. Hume, E. & K. Johnson (eds.), *The role of speech perception in phonology*, San Diego: Academic Press, pp. 141–185.
- Hyman, L. M. (2008). Enlarging the scope of phonologization. *UC Berkeley Phonology Lab Annual Report* pp. 382–409.
- Iverson, G. K. & J. C. Salmons (1995). Aspiration and laryngeal representation in Germanic. *Phonology* 12, pp. 369–396.
- Iverson, G. K. & J. C. Salmons (2006). On the typology of final laryngeal neutralization: Evolutionary Phonology and laryngeal realism. *Theoretical Linguistics* 32, pp. 205–216.
- Jakobson, R. (1929). Remarques sur l'évolution phonologique du russe comparée à celle des autres langues slaves, travaux du cercle linguistique de prague 2. *Selected writings, vol. I: Phonological studies*, The Hague: Mouton & Co., pp. 7–116. Reprinted in 1962.
- Janda, R. D. (2003). “Phonologization” as the start of dephoneticization – or, on sound change and its aftermath: Of extension, generalization, lexicalization, and morphologization. Joseph, B. D. & R. D. Janda (eds.), *The handbook of historical linguistics*, Oxford: Blackwell, pp. 401–422.
- Jansen, W. (2004). *Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English, Hungarian, and Dutch*. Ph.D. thesis, University of Groningen.
- Jassem, W. & L. Richter (1989). Neutralisation of voicing in Polish obstruents. *Journal of Phonetics* 17, pp. 317–325.

- Jessen, M. & C. Ringen (2002). Laryngeal features in German. *Phonology* 19, pp. 189–218.
- Jiménez, J. (1999). *L'estructura sil·làbica del català*. València, Barcelona: IIFV, Publicacions de l'Abadia de Montserrat.
- Jiménez, J. & M.-R. Lloret (2008). Asimetrías perceptivas y similitud articulatoria en la asimilación de sonoridad del catalán. *Cuadernos de Lingüística del I.U.I. Ortega y Gasset* 15, pp. 71–90.
- Johnson, K. (1997). The auditory/perceptual basis for speech segmentation. *Ohio State University Working Papers in Linguistics* pp. 101–113.
- Johnson, K. (2007). Decisions and mechanisms in exemplar-based phonology. Solé, M.-J., P. S. Beddor & M. Ohala (eds.), *Experimental approaches to phonology*, New York: Oxford University Press, pp. 25–40.
- Jongman, A., J. Sereno, M. Raaijmakers & A. Lahiri (1992). The phonological representation of [voice] in speech perception. *Language and speech* 35, pp. 137–152.
- Kaye, J. (1989). *Phonology: A cognitive view*. Hillsdale, NJ: Lawrence Erlbaum.
- Keating, P., T. Cho, C. Fougeron, C. Hsu et al. (2003). Domain-initial articulatory strengthening in four languages. *Papers in laboratory phonology VI* pp. 145–163.
- Keating, P. A. (1980). *A Phonetic Study of a Voicing Contrast in Polish*. Ph.D. thesis, Brown University.
- Keating, P. A. (1984). Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics* 59, pp. 29–34.
- Keating, P. A. (1988). Underspecification in phonetics. *Phonology* 5, pp. 275–292.
- Keating, P. A. (1990). Phonetic representations in a generative grammar. *Journal of Phonetics* 18, pp. 321–334.
- Keating, P. A. (1996). The phonology-phonetics interface. Kleinhenz, U. (ed.), *Interfaces in Phonology*, studia grammatica 41, Berlin: Akademie Verlag, pp. 262–278.
- Keyser, S. J. & K. N. Stevens (2006). Enhancement and overlap in the speech chain. *Language* 82, pp. 33–63.
- Kidda, M. E. (1993). *Tangale phonology a descriptive analysis*. Berlin: Reimer.
- Kingston, J. & R. L. Diehl (1994). Phonetic knowledge. *Language* 70, pp. 419–454.
- Kiparsky, P. (1972). Explanation in phonology. *Goals of linguistic theory*, Englewood NJ: Prentice-Hall, pp. 189–227.
- Kiparsky, P. (1982). From Cyclic Phonology to Lexical Phonology. van der Hulst, H. & N. Smith (eds.), *The structure of phonological representations*, Dordrecht: Foris, vol. 1, pp. 131–75.
- Kiparsky, P. (1985). Some consequences of Lexical Phonology. *Phonology Yearbook* 2, pp. 83–138.
- Kiparsky, P. (1995). The phonological basis of sound change. A., G. J. (ed.), *The handbook of phonological theory*, Cambridge MA: Blackwell, pp. 640–670.
- Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3, pp. 129–140.

- Kluender, K. R., R. L. Diehl & B. A. Wright (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics* 16, pp. 153–169.
- Labov, W. (1981). Resolving the Neogrammarian controversy. *Language* pp. 267–308.
- Labov, W. (1994). *Principles of linguistic change: Internal factors*, vol. 1. Oxford: Blackwell.
- Ladd, D. R. (2011). Phonetics in phonology. Goldsmith et al. (2011), pp. 348–373.
- Ladd, D. R. & J. M. Scobbie (2003). External sandhi as gestural overlap? Counter-evidence from Sardinian. Local, J., R. Ogden & R. Temple (eds.), *Phonetic interpretation: Papers in Laboratory Phonology VI*, Cambridge: Cambridge University Press, pp. 164–182.
- Ladefoged, P. (1973). Features of the larynx. *Journal of Phonetics* 1, pp. 73–83.
- Ladefoged, P. & I. Maddieson (1996). *The Sounds of the World's Languages*. Cambridge, MA: Blackwell.
- Lipski, J. M. (1989). /s/-voicing in Ecuadoran Spanish: patterns and principles of consonantal modification. *Lingua* 79, pp. 49–71.
- Lisker, L. & A. S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20, pp. 384–422.
- Lombardi, L. (1995). Laryngeal features and privativity. *The Linguistic Review* 12, pp. 35–59.
- Maiden, M. (1995). *A linguistic history of Italian*. London: Longman.
- Maye, J., J. F. Werker & L. Gerken (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, pp. B101–B111.
- Maye, J., D. J. Weiss & R. N. Aslin (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science* 11, pp. 122–134.
- McLennan, C. T., P. A. Luce & J. Charles-Luce (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29, pp. 539–553.
- Mielke, J. (2008). *The emergence of distinctive features*. New York: Oxford University Press.
- Nessett, T. (2008). *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface*, vol. 40 of *Cognitive Linguistics Research*. Mouton de Gruyter.
- Nolan, F., T. Holst & B. Kühnert (1996). Modelling [s] to [ʃ] accommodation in English. *Journal of Phonetics* 24, pp. 113–137.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 115.
- Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14.
- Ohala, J. J. (1981). The listener as a source of sound change. Masek, C., R. A. Hendrick & M. M. F (eds.), *Papers from the Parasession on Language and Behavior*, Chicago: Chicago Ling. Soc., pp. 178–203.
- Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. MacNeilage, P. (ed.), *The Production of Speech*, New York: Springer.

- Ohala, J. J. (1990a). The phonetics and phonology of aspects of assimilation. Kingston, J. & M. E. Beckman (eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech*, Cambridge: Cambridge University Press, pp. 258–275.
- Ohala, J. J. (1990b). There is no interface between phonology and phonetics: a personal view. *Journal of Phonetics* 18, pp. 153–171.
- Ohala, J. J. & M.-J. Solé (2010). Turbulence and phonology. Fuchs, S., M. Toda & M. Źygis (eds.), *Turbulent sounds. An interdisciplinary guide*, Berlin: Mouton deGruyter, pp. 37–97.
- Ostaszewska, D. & J. Tambor (2000). *Fonetika i fonologia współczesnego języka polskiego*. Warszawa: PWN.
- Owens, J. (1993). *A grammar of Nigerian Arabic*, vol. 10. Wiesbaden: Otto Harrassowitz.
- Perkell, J., S. Boyce & K. Stevens (1979). Articulatory and acoustic correlates of the [s-sh] distinction. Wolf, J. & D. Klatt (eds.), *Speech Communication Papers, 97th Meeting of the Acoustical Society of America*, pp. 109–113.
- Peterson, G. E. & I. Lehiste (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32, pp. 693–703.
- Pierrehumbert, J. (2001). Exemplar dynamics: word frequency, lenition, and contrast. Bybee, J. & P. Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, Amsterdam: John Benjamins, pp. 137–157.
- Pierrehumbert, J., M. Beckman & D. Ladd (2000). Conceptual foundations of phonology as a laboratory science. Burton-Roberts, P. Carr & G. Docherty (eds.), *Phonological Knowledge*, Oxford: Oxford University Press, pp. 273–303.
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Ph.D. thesis, MIT.
- Pierrehumbert, J. B. (1994). Knowledge of variation. Beals, K., J. Denton, R. Knippen, L. Mielnar, H. Suzuki & E. Zeinfeld (eds.), *Papers from the 30th Meeting of the Chicago Linguistics Society, ii. Papers from the Parasession on Variation*, Chicago: Chicago Linguistic Society, pp. 232–256.
- Pierrehumbert, J. B. (2002). Word-specific phonetics. Gussenhoven, C. & N. Warner (eds.), *Laboratory phonology 7*, New York: Mouton de Gruyter, pp. 101–139.
- Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics* 34, pp. 516–530.
- Port, R. & P. Crawford (1989). Pragmatic effects on neutralization rules. *Journal of Phonetics* 16, pp. 257–282.
- Port, R. & M. O'Dell (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics* pp. 455–471.
- Port, R., F. Mitleb & M. O'Dell (1981). Neutralization of obstruent voicing in German is incomplete. *The Journal of the Acoustical Society of America* 70, p. Supplement 1.
- Prince, A. & P. Smolensky (2004 [1993]). *Constraint interaction in generative grammar*. Malden, MA, and Oxford, UK: Blackwell.
- R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL <http://www.R-project.org>.

- Rácz, P. (2011). To err is human – rich memory language models and t-glottalisation in English. Talk presented at OCP8.
- Recasens, D. (1993). *Fonètica i fonologia*. Barcelona : Encyclopèdia Catalana.
- Recasens, D. (2002). Weakening and strengthening in Romance revisited. *Rivista di Linguistica* 14, pp. 327–374.
- Reiss, C. (2008). Constraining the learning path without constraints, or the OCP and NOBANANA. Vaux, B. & A. Nevins (eds.), *Rules, Constraints and Phonological Phenomena*, Oxford: Oxford University Press, pp. 252–301.
- Rice, K. (2007). Markedness in phonology. de Lacy, P. (ed.), *The Cambridge handbook of phonology*, Cambridge: Cambridge University Press, pp. 79–97.
- Rice, K. D. (1993). A reexamination of the feature [Sonorant]: the status of ‘sonorant obstruents’. *Language* 69, pp. 308–344.
- Robinson, K. L. (1979). On the voicing of intervocalic *s* in the Ecuadorian highlands. *Romance Philology* 33, pp. 132–143.
- Romero, J. (1999). The effect of voicing assimilation on gestural coordination. ICPHS99.
- Rubach, J. (1996). Nonsyllabic analysis of voice assimilation in Polish. *Linguistic Inquiry* 27, pp. 69–110.
- Rubach, J. (2008). Prevocalic faithfulness. *Phonology* 25, pp. 433–468.
- Schilling, M. F., A. E. Watkins & W. Watkins (2002). Is human height bimodal? *The American Statistician* 56, pp. 223–229.
- Schuchardt, H. (1885 [1972]). On sound laws: against the neogrammarians. Vennemann, T. & T. H. Wilbur (eds.), *Schuchardt, The neogrammarians and the Transformational Theory of Phonological Change: four essays*, Frankfurt am Main: Athenäum Verlag, pp. 41–72. First published (1885), *Über die Lautgesetze: gegen die Junggrammatiker*, Berlin: Oppenheim.
- Scobbie, J. M. (2005). The phonetics-phonology overlap. *QMUC Speech Science Research Centre Working Papers* 1.
- Scobbie, J. M. & K. Sebregts (2010). Acoustic, articulatory, and phonological perspectives on allophonic variation of /r/ in Dutch. Folli, R. & C. Ulbrich (eds.), *Interfaces in Linguistics. New Research Perspectives*, Oxford: Oxford University Press, pp. 257–277.
- Scobbie, J. M., J. Stuart-Smith & E. Lawson (2008). Final report for “Looking variation and change in the mouth: developing the sociolinguistic potential of Ultrasound Tongue Imaging”. ESRC Research Grant RES-000-22-2032.
- Sebregts, K. & J. M. Scobbie (2011). Functional, non-functional or multi-functional? Of covert gestures and free variation. Talk presented at OCP8, Marrakesh.
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, Lawrence Erlbaum Hillsdale, NJ, pp. 295–342.
- Silverman, D. (2006a). *A critical introduction to phonology: of sound, mind, and body*. London: Continuum.

- Silverman, D. (2006b). The diachrony of labiality in Trique, and the functional relevance of gradience and variation. Goldstein, L. M., D. H. Whalen & C. T. Best (eds.), *Papers in Laboratory Phonology VIII*, Berlin: Mouton de Gruyter, pp. 133–154.
- Simon, E. (2010). *Voicing in Contrast. Acquiring a second language laryngeal system*. Ghent: Academia Press (Ginkgo series).
- Skousen, R. (1989). *Analogical modeling of language*. Springer.
- Slis, I. H. (1985). *The voiced-voiceless distinction and voice assimilation in Dutch*. Ph.D. thesis, University of Nijmegen.
- Slis, I. H. & A. Cohen (1969). On the complex regulating the voiced-voiceless distinction I. *Language and Speech* 12, pp. 80–102.
- Slowiaczek, L. & D. A. Dinnsen (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13, pp. 325–341.
- Slowiaczek, L. M. & H. J. Szymanska (1989). Perception of word-final devoicing in Polish. *Journal of Phonetics* 17, pp. 205–212.
- Solé, M.-J. (1992). Phonetic and phonological processes: the case of nasalization. *Language and Speech* 35, pp. 29–43.
- Solé, M.-J. (1995). New ways of analyzing sound change. speech rate effects. *Belgian Journal of Linguistics* 9, pp. 21–44.
- Solé, M.-J. (1995). Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech* 38, pp. 1–23.
- Solé, M.-J. (2007). Controlled and mechanical properties in speech: a review of the literature. Solé, M.-J., P. S. Beddor & M. Ohala (eds.), *Experimental approaches to phonology*, New York: Oxford University Press, pp. 302–321.
- Sproat, R. & O. Fujimura (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, pp. 291–311.
- Stevens, K. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- Stevens, K., S. Keyser & H. Kawasaki (1986). Toward a phonetic and phonological theory of redundant features. Perkell, J. S. & D. H. Klatt (eds.), *Invariance and variability in speech processes*, Hillsdale: Lawrence Erlbaum, pp. 426–449.
- Stevens, K. N. & S. J. Keyser (1989). Primary features and their enhancement in consonants. *Language* 65, pp. 81–106.
- Stevens, K. N. & S. J. Keyser (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics* 38, pp. 10–19.
- Stevens, K. N., S. E. Blumstein, L. Glicksman, M. Burton & K. Kurowski (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91, pp. 2979–3000.
- Strycharczuk, P. (2012). Sonorant transparency and the complexity of voicing in Polish. *Journal of Phonetics* 40, pp. 655–671.

- Stuart-Smith, J. (2007). Empirical evidence for gendered speech production: /s/in Glaswegian. Cole, J. & J. I. Hualde (eds.), *Laboratory Phonology 9: Phonology and phonetics*, New York: Mouton de Gruyter, pp. 65–86.
- Ternes, E. (1970). *Grammaire structurale du Breton de L'Île de Groix (dialecte occidental)*. Heidelberg: Carl Winter Universitätsverlag.
- Traill, A. (1990). Depression without depressors. *South African Journal of African Languages* 10, pp. 166–172.
- Tucker, B. V. & N. Warner (2010). What it means to be phonetic or phonological: the case of Romanian devoiced nasals. *Phonology* 27, pp. 289–324.
- Twaddell, W. (1938). A note on Old High German umlaut. *Monatshefte für deutschen Unterricht* 30, pp. 177–181.
- Vaux, B. (2008). Why the phonological component must be serial and rule-based. Vaux, B. & A. Nevins (eds.), *Rules, Constraints and Phonological Phenomena*, Oxford: Oxford University Press, pp. 20–60.
- Vennemann, T. (1972). Phonetic analogy and conceptual analogy. Vennemann, T. & T. H. Wilbur (eds.), *Schuchardt, the neogrammarians, and the transformational theory of phonological change*, Frankfurt am Mein: Athenäum Verlag, pp. 181–204.
- Warner, N., A. Jongman, J. Sereno & R. Kemps (2004). Incomplete neutralization and other subphonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32:2.
- Wedel, A. B. (2004). *Self-organization and categorical behavior in phonology*. Ph.D. thesis, University of California, Santa Cruz.
- Wedel, A. B. (2006). Exemplar models, evolution and language change. *The Linguistic Review* 23, pp. 247–274.
- Wedel, A. B. (2007). Feedback and regularity in the lexicon. *Phonology* 24, pp. 147–185.
- Weijnen, A. (1991). *Vergelijkende Klankleer van de Nederlandse Dialecten*. The Hague: SDU.
- Westbury, J. R. & P. A. Keating (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics* 22, pp. 145–166.
- Wetzel, W. L. (1997). Bakairi and the feature ‘voice’. *Revista da Associação Brasileira de Lingüística. Homenagem a Aryon Dall'Igna Rodrigues* pp. 21–34.
- Wheeler, M. (1979). *Phonology of Catalan*. Publications of the Philological Society XXVIII, Oxford: Basil Blackwell.
- Wheeler, M. W. (2005). *The Phonology of Catalan*. New York: Oxford University Press.
- Wierzchowska, B. (1980). *Fonetika i fonologia języka polskiego*. Wrocław: Zakład Narodowy im. Ossolińskich.
- Williams, A. & P. Kerswill (1999). Dialect levelling: change and continuity in Milton Keynes, Reading and Hull. Foulkes, P. & G. Docherty (eds.), *Urban voices. Accent studies in the British Isles*, London: Arnold, pp. 141–162.
- Wright, S. & P. Kerswill (1989). Electropalatography in the study of connected speech processes. *Clinical Linguistics and Phonetics* 3, pp. 49–57.

- Yu, A. (2004). Explaining final obstruent voicing in Lezgian: phonetics and history. *Language* 80, pp. 73–97.
- Yuan, J. & M. Liberman (2009). Investigating /l/ variation in English through forced alignment. *Tenth Annual Conference of the International Speech Communication Association*, pp. 2215–2218.
- Zilyns'kyj, I. (1979). *A phonetic description of the Ukrainian language*. Harvard University Press.
- Zonneveld, W. (1983). Lexical and phonological properties of Dutch voicing assimilation. van den Broecke, M., V. van Heuven & W. Zonneveld (eds.), *Studies for Antonie Cohen. Sound Structures*, Dordrecht: Foris Publications, pp. 297–231.
- Zsiga, E. (1995). An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. Connell, B. & A. Arvaniti (eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV*, Cambridge: Cambridge University Press, pp. 282–302.

APPENDIX A

TEST ITEMS USED IN THE EXPERIMENT ON BELGIAN DUTCH

A.1 Test items

In cases where there is only one test item per environment, the item was used twice.

Environment	Item	Gloss
/t#n/	<i>zout nootje</i>	‘salty nut, DIM.’
/t#m/	<i>zout muntje</i>	‘salty mint, DIM.t’
/t#r/	<i>fout riempje</i>	‘wrong strap, DIM.’
/t#l/	<i>fout lintje</i>	‘wrong ribbon, DIM.’
/t#w/	<i>zout worstje</i>	‘salty sausage, DIM.’
/t#j/	<i>fout jasje</i>	‘wrong coat, DIM.’
/t#ei/	<i>zout eitje</i>	‘salty egg, DIM.’
/t#ɛ/	<i>zout erwte</i>	‘salty pea, DIM.’
/d#n/	<i>koud nootje</i>	‘cold nut, DIM.’
/d#m/	<i>koud muntje</i>	‘cold mint, DIM.’
/d#r/	<i>koud riempje</i>	‘cold strap, DIM.’
/d#l/	<i>koud lintje</i>	‘cold ribbon, DIM.’
/d#w/	<i>koud worstje</i>	‘cold sausage, DIM.’
/d#j/	<i>koud jasje</i>	‘cold coat’, DIM.
/d#ei/	<i>koud eitje</i>	‘cold egg, DIM.’
/t#ɛ/	<i>koud erwte</i>	‘cold pea, DIM.’
/s#n/	<i>Fries nichtje</i>	‘Frisian niece, DIM.’
	<i>Fries neefje</i>	‘Frisian nephew, DIM.’
/s#m/	<i>Fries Mieke</i>	‘Frisian Mieke’
/s#r/	<i>Fries ratje</i>	‘Frisian rat, DIM.’
/s#l/	<i>Fries liefje</i>	‘Frisian boyfriend/girlfriend, DIM.’
/s#w/	<i>Fries wolfje</i>	‘Frisian wolf, DIM.’
/s#j/	<i>Fries Jantje</i>	‘Frisian Jantje’
/s#a/	<i>Fries aapje</i>	‘Frisian monkey, DIM.’
/z#n/	<i>vies nichtje</i>	‘dirty niece, DIM.’
	<i>vies neefje</i>	‘dirty nephew, DIM.’
/z#m/	<i>vies Mieke</i>	‘dirty Mieke’
/z#r/	<i>vies ratje</i>	‘dirty rat, DIM.’
/z#l/	<i>vies liefje</i>	‘dirty boyfriend/girlfriend, DIM.’
/z#w/	<i>vies wolfje</i>	‘dirty wolf’, DIM.
/z#j/	<i>vies Jantje</i>	‘dirty Jantje’
/z#a/	<i>vies aapje</i>	‘dirty monkey, DIM.’

A.2 Control items

Environment	Item	Gloss
/t#p/	<i>zout peertje</i>	'salty pear, DIM.'
/t#b/	<i>zout boontje</i>	'salty bean, DIM.'
/t#f/	<i>zout frietje</i>	'salty chip, DIM.'
/t#v/	<i>zout visje</i>	'salty fish, DIM.'
/d#p/	<i>koud peertje</i>	'cold pear, DIM.'
/d#b/	<i>koud boontje</i>	'cold bean, DIM.'
/d#f/	<i>koud frietje</i>	'cold chip, DIM.'
/d#v/	<i>koud visje</i>	'cold fish, DIM.'
/s#p/	<i>Fries poesje</i>	'Frisian cat, DIM.'
/s#b/	<i>Fries beertje</i>	'Frisian bear, DIM.'
/s#f/	<i>Fries fietsje</i>	'Frisian bike, DIM.'
/s#v/	<i>Fries vatje</i>	'Frisian barrel, DIM.'
/z#p/	<i>vies poesje</i>	'dirty cat, DIM.'
/z#b/	<i>vies beertje</i>	'dirty bear, DIM.'
/z#f/	<i>vies fietsje</i>	'dirty bike, DIM.'
/z#v/	<i>vies vatje</i>	'dirty barrel, DIM.'

APPENDIX B

TEST ITEMS USED IN THE EXPERIMENT ON POLISH

B.1 Experiment 1

B.1.1 Test items

Environment	Item	Gloss
/d#m/	<i>pośród miasta</i>	'amidst the city'
/d#n/	<i>poród naturalny</i>	'natural delivery'
/d#r/	<i>rozwód rodziców</i>	'parents' divorce'
/d#l/	<i>powód lęków</i>	'reason for anxieties'
/d#j/	<i>przeszkód jeźzieckich</i>	'show jumping obstacles, GEN.'
/d#w/	<i>zawód łowcy</i>	'hunter's profession'
/d#ɔ/	<i>swobód obywatelskich</i>	'civic rights GEN.'
/t#m/	<i>przewrót majowy</i>	'The May Coup d'État'
/t#n/	<i>przerzut narkotyków</i>	'illegal drug transfer'
/t#r/	<i>walut Rosji</i>	'currencies of Russia GEN. PL.'
/t#l/	<i>nawrót lęku</i>	'return of anxiety, GEN.'
/t#j/	<i>zarzut jest</i>	'an objection is'
	<i>debiut łódzkiego</i>	'debut of a Łódź-based
/t#w/	<i>(dokumentalisty)</i>	documentary-maker'
/t#ɔ/	<i>statut okresła</i>	'charter determines'

B.1.2 Control items

Environment	Item	Gloss
/t#p/	<i>minut po</i>	'minutes after, GEN.'
/t#b/	<i>dysput badaczy</i>	'scholarly disputes GEN.'
/t#s/	<i>wyrzut sumienia</i>	'remorse'
/t#z/	<i>pokut za</i>	'penance for, GEN. PL.'
/t#f/	<i>dysput filozoficznych</i> <i>minut filmu</i>	'scholarly debates, GEN.' 'minutes of the film, GEN.'
	<i>zarzut fałszerstwa</i>	'accusation of forgery'
/t#v/	<i>atut Watszawy</i>	'Warsaw's asset'
	<i>mamut włochaty</i>	'woolly mammoth'
	<i>obrót wydarzeń</i>	'turn of events'
/d#p/	<i>zachód Polski</i>	'Western Poland'
/d#b/	<i>dochód brutto</i>	'gross income'
/d#s/	<i>zachód słońca</i>	'sunset'
/d#z/	<i>ogród zimowy</i>	'winter garden'
/d#f/	<i>zachód Filipin</i>	'Western Philippines'
	<i>spośród fałszerzy</i>	'amongst the forgers'
/d#v/	<i>zagród wiejskich</i>	'rural pens GEN.'
	<i>dochód Wielkiej Brytanii</i>	'Great Britain's income'
	<i>powód wojny</i>	'reason for war'
	<i>wywód wodza</i>	'chief's address'

B.2 Experiment 2

B.2.1 Test items

Environment	Item	Gloss
/t#m/	<i>zarzut Marka</i>	‘Marek’s accusation’
/t#n/	<i>zarzut Neli</i>	‘Nela’s accusation’
/t#r/	<i>powrót Radka</i>	‘Radek’s return’
/t#l/	<i>powrót Lucka</i>	‘Lucek’s return’
/t#w/	<i>zarzut Łucji</i>	‘Łucja’s accusation’
/t#j/	<i>zarzut Janka</i>	‘Janek’s accusation’
/t#a/	<i>powrót Ani</i>	‘Ania’s return’
	<i>powrót Arka</i>	‘Arek’s return’
/d#m/	<i>dowód Majki</i>	‘Majka’s ID’
/d#n/	<i>dowód Nadii</i>	‘Nadia’s ID’
/d#r/	<i>zawód Rocha</i>	‘Roch’s profession’
/d#l/	<i>zawód Leny</i>	‘Lena’s profession’
/d#w/	<i>dowód Łucji</i>	‘Łucja’s ID’
/d#j/	<i>dowód Jagi</i>	Jaga’s ID’
/d#a/	<i>zawód Ani</i>	‘Ania’s profession’
	<i>zawód Arka</i>	‘Arek’s profession’
/s#m/	<i>status Marka</i>	‘Marek’s status’
/s#n/	<i>status Neli</i>	‘Nela’s status’
/s#r/	<i>status Radka</i>	‘Radek’s status’
/s#l/	<i>status Lucka</i>	‘Lucek’s status’
/s#w/	<i>status Łucji</i>	‘Łucja’s status’
/s#j/	<i>status Janka</i>	‘Janek’s status’
/s#a/	<i>globus Ani</i>	‘Ania’s globe’
	<i>globus Arka</i>	‘Arek’s globe’
/z#m/	<i>obóz Majki</i>	‘Majka’s camp’
/z#n/	<i>obóz Nadii</i>	‘Nadia’s camp’
/z#r/	<i>powóz Rocha</i>	‘Roch’s carriage’
/z#l/	<i>powóz Leny</i>	‘Lena’s carriage’
/z#w/	<i>obóz Łucji</i>	‘Łucja’s camp’
/z#j/	<i>obóz Jagi</i>	Jaga’s camp’
/z#a/	<i>powóz Ani</i>	‘Ania’s carriage’
/z#a/	<i>powóz Arka</i>	‘Arek’s carriage’

B.2.2 Control items

Environment	Item	Gloss
/t#f/	<i>powrót Fiony</i>	'Fiona's return'
	<i>powrót Franka</i>	'Franek's return'
/t#v/	<i>powrót Wojtka</i>	'Wojtek's return'
	<i>powrót Wandy</i>	'Wanda's return'
/d#f/	<i>zawód Fiony</i>	'Fiona's profession'
	<i>zawód Franka</i>	'Franek's profession'
/t#p/	<i>powrót Pawła</i>	'Paweł's profession'
	<i>powrót Piotra</i>	'Piotr's profession'
/t#b/	<i>powrót Basi</i>	'Basia's profession'
	<i>powrót Bartka</i>	'Bartek's profession'
/s#f/	<i>globus Fiony</i>	'Fiona's globe'
	<i>globus Franka</i>	'Franek's globe'
/z#v/	<i>powóz Wojtka</i>	'Wojtek's camp'
	<i>powóz Wandy</i>	'Wanda's camp'
/s#p/	<i>globus Piotra</i>	'Piotr's globe'
	<i>globus Pawła</i>	'Paweł's globe'
/s#b/	<i>globus Bartka</i>	'Bartek's globe'
	<i>globus Basi</i>	'Basia's globe'
/z#p/	<i>obóz Pawła</i>	'Paweł's camp'
	<i>obóz Piotra</i>	'Piotr's camp'
/z#b/	<i>obóz Basi</i>	'Basia's camp'
	<i>obóz Bartka</i>	'Bartek's camp'

APPENDIX C

TEST ITEMS USED IN THE EXPERIMENT ON CATALAN

C.1 Test items

Environment	Item	Gloss
/p#ə/	<i>drap anglès</i>	‘English cloth’
/p#i/	<i>drap històric</i>	‘historical cloth’
/p#u/	<i>drap opac</i>	‘opaque cloth’
/p#m/	<i>drap millor</i>	‘best cloth’
/p#n/	<i>drap noruec</i>	‘Norwegian cloth’
/p#r/	<i>drap romà</i>	‘Roman cloth’
/p#l/	<i>drap letó</i>	‘Latvian cloth’
/p#ʎ/	<i>drap llampant</i>	‘bright cloth’
/b#ə/	<i>sap anglès</i>	‘(s)he knows English’
/b#i/	<i>sap història</i>	‘(s)he knows history’
/b#ɔ/	<i>sap obrir</i>	‘(s)he knows how to open’
/b#m/	<i>sap mongol</i>	‘(s)he knows Mongolian’
/b#n/	<i>sap noruec</i>	‘(s)he knows Nowegian’
/b#r/	<i>sap remar</i>	‘(s)he knows how to row’
/b#l/	<i>sap letó</i>	‘(s)he knows Latvian’
/b#ʎ/	<i>sap llegir</i>	‘(s)he knows how to read’
/s#ə/	<i>pas audaç</i>	‘bold step’
/s#i/	<i>pas immens</i>	‘huge step’
/s#ɔ/	<i>pas obert</i>	‘open step’
/s#m/	<i>pas mandró</i>	‘tired step’
/s#n/	<i>pas nerviós</i>	‘nervous step’
/s#r/	<i>pas robust</i>	‘robust step’
/s#l/	<i>pas legítim</i>	‘legitimate step’
/s#ʎ/	<i>pas lleuger</i>	‘light step’
/z#ə/	<i>vas antic</i>	‘old glass’
/z#i/	<i>vas immens</i>	‘huge glass’
/z#u/	<i>vas oficial</i>	‘official glass’
/z#m/	<i>vas mullat</i>	‘wet glass’
/z#n/	<i>vas normal</i>	‘normal glass’
/z#r/	<i>vas rodó</i>	‘round glass’
/z#l/	<i>vas letó</i>	‘Latvian glass’
/z#ʎ/	<i>vas lleuger</i>	‘light glass’

/bz#ə/	<i>saps anglès</i>	'you know English'
/bz#i/	<i>saps història</i>	'you know history'
/bz#ɔ/	<i>saps obrir</i>	'you know how to open'
/bz#m/	<i>saps mongol</i>	'you know Mongolian'
/bz#n/	<i>saps noruec</i>	'you know Norwegian'
/bz#r/	<i>saps remar</i>	'you know how to row'
/bz#l/	<i>saps letó</i>	'you know Latvian'
/bz#f/	<i>saps llegir</i>	'you know how to read'

Forms with an underlying /ɔ/ in unstressed syllables, as in *obert*, or *obrir*, surface with a high vowel [u] in Catalan.

C.2 Control items

Environment	Item	Gloss
/p#s/	<i>drap suec</i>	'Swedish cloth'
/p#z/	<i>drap zambià</i>	'Zambian cloth'
/b#s/	<i>sap suec</i>	'(s)he knows Swedish'
/b#z/	<i>sap zoologia</i>	'(s)he knows zoology'
/s#t/	<i>pas tranquil</i>	'calm step'
/s#d/	<i>pas dinàmic</i>	'dynamic step'
/z#t/	<i>vas trencat</i>	'broken glass'
/z#d/	<i>vas daurat</i>	'gold glass'
/bz#t/	<i>saps tocar</i>	'you know how to touch/play (an instrument)'
/bz#d/	<i>saps daurar</i>	'you know how to gild'

APPENDIX D

TEST ITEMS USED IN THE EXPERIMENT ON QUITO SPANISH

Condition	Item	Gloss
1. /sV/	<i>gasita</i>	‘gauze (dim)’
	<i>casita</i>	‘home (dim)’
	<i>aviso</i>	‘notice’
	<i>gusano</i>	‘worm’
	<i>mayonesa</i>	‘mayonaise’
	<i>mosaico</i>	‘mosaic’
2. /sN/	<i>entusiasmo</i>	‘enthusiasm’
	<i>espasmo</i>	‘spasm’
	<i>budismo</i>	‘buddhism’
	<i>bautismo</i>	‘baptism’
	<i>esmoquín</i>	‘the tuxedo’
	<i>cosmólogo</i>	‘cosmologist’
3. /s#V/	<i>gas acre</i>	‘acrid gas’
	<i>palmas altas</i>	‘tall palm trees’
	<i>tesis obvia</i>	‘obvious thesis’
	<i>virus asnal</i>	‘brutal virus’
	<i>tres autores</i>	‘three authors’
	<i>muchos hombres</i>	‘many men’
4. /s#N/	<i>gas noble</i>	‘noble gas’
	<i>ropas negras</i>	‘black clothes’
	<i>croquis nuevo</i>	‘new foundation’
	<i>crisis mundial</i>	‘world crisis’
	<i>bases nuevas</i>	‘naval bases’
	<i>muchos monjes</i>	‘many monks’
5. /s+voiced obstruent/	<i>esbozo</i>	‘plan, sketch’
	<i>Lisboa</i>	‘Lisbon’
	<i>jurisdicción</i>	‘jurisdiction’
	<i>presbítero</i>	‘priest’
	<i>posdata</i>	‘postscript’
6. /s+voiceless obstruent/	<i>subasta</i>	‘auction’
	<i>carraspera</i>	‘transport’
	<i>obispo</i>	‘bishop’
	<i>crepúsculo</i>	‘dawn’
	<i>asbesto</i>	‘asbestos’
	<i>microscopio</i>	‘microscope’

7. /s#voiced obstruent/	<i>marchas buenas</i>	'good marches'
	<i>velas blancas</i>	'white sails'
	<i>brindis digno</i>	'dignified toast (fig.)'
	<i>cactus grande</i>	'big cactus'
	<i>grandes barcos</i>	'big ships'
	<i>gatos bellos</i>	'beautiful cats'