

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI[®]
800-521-0600

Sandhi of /rs/ Sequences in North Norrland Swedish

by

Vern Myron Lindblad

**A dissertation submitted in partial fulfillment of the
requirements for the degree of**

Doctor of Philosophy

University of Washington

1999

Program Authorized to Offer Degree: Linguistics

UMI Number: 9952862

Copyright 1999 by
Lindblad, Vern Myron

All rights reserved.

UMI[®]

UMI Microform 9952862

Copyright 2000 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© Copyright 1999
Vern M. Lindblad

Doctoral Dissertation

In presenting this thesis in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to UMI Dissertation Services, 300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Vern M. Lundblad

Date December 7, 1999

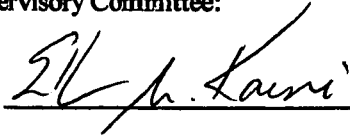
University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Vern Myron Lindblad

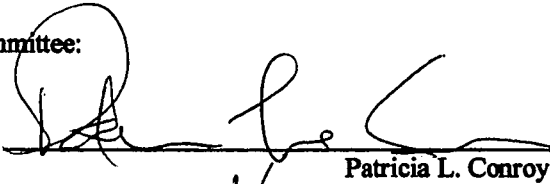
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

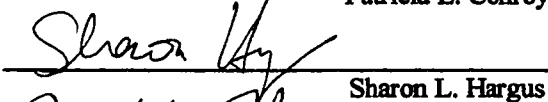


Ellen M. Kaisse

Reading Committee:



Patricia L. Conroy



Sharon L. Hargus



Joseph B. Voyles

Date:

December 7, 1999

University of Washington

Abstract

SANDHI OF /RS/ SEQUENCES IN NORTH NORRLAND
SWEDISH

by Vern Myron Lindblad

Chairperson of the Supervisory Committee: Professor Ellen M. Kaisse
Department of Linguistics

This dissertation investigates the phonology of /rs/-sandhi in North Norrland, the northernmost 40% of Sweden. North Norrland is of particular phonological interest because it is the locus of a neutralization between the phonetic product of /rs/-sandhi, [ʂ], and the phoneme of Swedish known as *sje* or /ʃ/, which in this region is most often also pronounced as [ʂ]. The phonology of /rs/-sandhi includes four overlapping areas of variability: the choice of variants of *sje* (Chapter 3), the occurrence of internal /rs/-sandhi (Chapter 4), the occurrence of external /rs/-sandhi (Chapter 5), and the production of the copula *är* 'is/are/am' with an unreduced coda-/r/ (Chapter 6).

I collected new data in North Norrland with two stories, and examined it using three population factors (the age, gender and dialect area of speakers) and several linguistic factors (including the lexical items, syntactic categories of words, position and types of adjacent stress, and c-command relationship between words). This dissertation focuses on describing the patterns of variation in this data, but also raises many issues of theory.

I found a wide range of variation between the various lexical items for each of the four areas of interest, and sought to explain them by other factors. The population factors of age, gender and dialect played a strong role in regard to the choice of variants of *sje*, a smaller role in regard to the unreduced copula *är*, and had little influence on internal or

external /rs/-sandhi. Both internal and external /rs/-sandhi appear to be influenced by adjacent segments, although data limitations reduce confidence levels. Closeness of the two elements favored sandhi. C-command by the word on the right favored sandhi. For each of the four sets of data, stress was a major factor. Primary stress on the following vowel had a range of influences; it facilitated the choice of dark-*sje* variants, favored internal /rs/-sandhi, disfavored external /rs/-sandhi, and disfavored the maintenance of coda-/r/ on the copula *är*. There was no evidence that frequency of occurrence of the words was a meaningful factor in any choice of variants.

TABLE OF CONTENTS

List of Figures	iv
List of Tables	vi
Glossary	viii
List of Abbreviations	ix
Preface	x
Introduction.....	1
Chapter 1: Background.....	2
1.0 The problems to be addressed	2
1.1 A sketch of Swedish segments & suprasegmental phonology	2
1.1.0 Vowels.....	3
1.1.1 Consonants.....	4
1.1.2 Stress	6
1.1.2.0 Stress and length.....	8
1.1.2.1 Stress: lexical and postlexical	9
1.1.2.2 Stress of function words	11
1.2 Previous studies	12
1.3 Rules.....	17
1.4 Summary of issues.....	21
1.5 Summary of contents	22
Chapter 2: Factors in the statistical analysis of the data.....	24
2.0 The factors.....	24
2.1 Demographic factors.....	24
2.1.0 The Gender factor	24
2.1.1 The Age factor	25
2.1.2 The Dialect factor	26

2.2 Linguistic factors	29
2.2.0 The lexical factor	29
2.2.1 The stress factor	29
2.2.2 The syntactic factor	30
2.3 The Contexts.....	30
2.3.0 Geographical context: North Norrland.....	30
2.3.1 Historical context: Mining.....	34
2.3.2 Historical context: Radio	35
2.3.3 Historical context: World War II	36
Chapter 3: Underlying <i>sje</i>	39
3.0 Some questions about <i>sje</i> in NN.....	39
3.1 Neutralization of <i>sje</i> in NN	40
3.2 Previous studies of <i>sje</i> in NN	42
3.3 The new data	45
3.4 The stress factor and <i>sje</i>	46
3.5 Hypotheses to be tested re <i>sje</i>	49
3.6 Tables and graphs of results	51
3.7 Summary of dark- <i>sje</i> usage	66
Chapter 4: Internal sandhi of /rs/ sequences	70
4.0 Overview of internal sandhi	70
4.1 Previous studies	71
4.2 Hypotheses to be tested re internal sandhi of /rs/	71
4.3 Factors in internal /rs/-sandhi	74
4.4 Tables and graphs of results	76
4.5 Summary and conclusions for internal sandhi.....	95
Chapter 5: External sandhi of /rs/ sequences.....	98
5.0 Overview of external sandhi.....	98
5.1 Previous studies	99
5.2 Hypotheses to be tested re external sandhi of /rs/	100

5.3 Factors in external /rs/-sandhi.....	105
5.4 The external /rs/-sandhi data	107
5.5 Tables and graphs of results	110
5.6 Summary and conclusions for external /rs/-sandhi.....	135
5.6.1 Summary of results for hypotheses.....	135
5.6.2 Summary of favorable and unfavorable factors.....	139
Chapter 6: Prosodic factors in coda retention by <i>är</i> ‘is/are/am’	141
6.0 Introduction: The problematic data.....	141
6.1 Eurhythmy, and a prosodic lapse analysis of <i>är</i> ‘is/are/am’	143
6.2 New data, and its relevance to the prosodic lapse analysis.....	146
6.3 Tables and graphs of results	147
6.4 Lapses with other function words.....	157
6.5 Summary and conclusions.....	158
Chapter 7: Summary and conclusions	160
Bibliography.....	164
Appendix A: Min familj och Norrland (numbered).....	170
Appendix B: Min familj och Norrland (interlinear translation).....	172
Appendix C: Min styvfars idol (numbered).....	184
Appendix D: Min styvfars idol (interlinear translation).....	188
Appendix E: Population data for MfoN and Msi.....	207
Appendix F: Lexical items for dark- <i>sje</i>	208
Appendix G: Lexical items for internal /rs/-sandhi.....	209
Appendix H: Lexical items for external /rs/-sandhi.....	210
Appendix I: Lexical items for <i>är</i> ‘is/are/am/’	214

LIST OF FIGURES

<i>Number</i>	<i>Page</i>
Figure 1. Standard Swedish vowel system	3
Figure 2. Consonant inventory of standard Swedish.....	4
Figure 3. Moraic representation of Swedish stressed syllables (after Riad 1992).....	9
Figure 4. Sandhi contexts (after Eliasson 1986: 278)	15
Figure 5. Assimilation, then Alveolarization (Hellberg 1974: 166)	18
Figure 6. Assimilation, then Reduction (Eliasson 1986: 286).....	19
Figure 7. One-step retroflexion (Hovdhaugen 1969: 91)	19
Figure 8. Retroflexion (Lindblad 1989)	20
Figure 9. Map of North Norrland dialect areas.....	28
Figure 10. Dark- <i>sje</i> pronunciation by dialect and gender	52
Figure 11. Dark- <i>sje</i> pronunciation by gender and age (in decades).....	54
Figure 12. Dark- <i>sje</i> by gender and age group (in half-decades).....	55
Figure 13. Dark- <i>sje</i> pronunciation by lexical item and gender.....	58
Figure 14. Dark- <i>sje</i> pronunciation by lexical item and dialect.....	63
Figure 15. Dark- <i>sje</i> pronunciation by lexical item and age group.....	65
Figure 16. Internal /rs/-sandhi, by gender and age in decades.....	76
Figure 17. Internal /rs/-sandhi, by gender and age in half-decades.....	78
Figure 18. Internal /rs/-sandhi, by gender and lexical item	79
Figure 19. Internal /rs/-sandhi, by lexical item and age group	83
Figure 20. Internal /rs/-sandhi, by gender and dialect area	85
Figure 21. Internal /rs/-sandhi, by dialect and stress on preceding vowel.....	87
Figure 22. Internal /rs/-sandhi, by dialect and stress on next vowel.....	89
Figure 23. Internal /rs/ sandhi, by dialect and type of internal boundary.....	93
Figure 24. External /rs/-sandhi, by gender and age in decades.....	111

Figure 25. External /rs/-sandhi, by gender and dialect area	112
Figure 26. External /rs/-sandhi, by first lexical item.....	114
Figure 27. External /rs/-sandhi, by preceding syntactic category	116
Figure 28. External /rs/-sandhi, by second lexical item	119
Figure 29. External /rs/-sandhi, by following syntactic category	120
Figure 30. External /rs/-sandhi, by direction of c-command across dialects.....	122
Figure 31. External /rs/-sandhi, by ‘closeness’ and dialect	124
Figure 32. /rs/-Sandhi, by dialect and boundary type (internal & external).....	125
Figure 33. External /rs/-sandhi, by gender and the vowel preceding /rs/	127
Figure 34. External /rs/-sandhi, by gender and the vowel following /rs/	129
Figure 35. External /rs/-sandhi, by gender and segment following /s-/	130
Figure 36. External /rs/-sandhi, by age and stress on the preceding vowel.....	131
Figure 37. External /rs/-sandhi, by age and stress on the following vowel.....	133
Figure 38. External /rs/-sandhi, by stress on preceding and following vowels	134
Figure 39. The Principle of Rhythmic Alternation (Selkirk 1984: 52)	144
Figure 40. Lapse resolution by beat addition.....	144
Figure 41. Lapse contexts requiring added stress (adapted from Selkirk 1984: 363)....	145
Figure 42. Unreduced <i>är</i> ‘is/are/am’ by lexical item.....	148
Figure 43. Unreduced <i>är</i> , by age and gender.....	149
Figure 44. Unreduced <i>är</i> , by dialect and gender.....	150
Figure 45. Unreduced <i>är</i> , by gender and following word class.....	151
Figure 46. Unreduced <i>är</i> , by age and following word class.....	153
Figure 47. Unreduced <i>är</i> , by age and following stress.....	154
Figure 48. Unreduced <i>är</i> , by dialect and following stress.....	156

LIST OF TABLES

<i>Number</i>	<i>Page</i>
Table 1. Stress patterns in Swedish: some variant notations	7
Table 2. Dark- <i>sje</i> pronunciation by dialect and gender.....	51
Table 3. Dark- <i>sje</i> pronunciation by gender and age (in decades).....	53
Table 4. Dark- <i>sje</i> pronunciation by gender and age group (in half-decades).....	55
Table 5. Dark- <i>sje</i> pronunciation by lexical item and gender.....	57
Table 6. Dark- <i>sje</i> pronunciation by lexical item and dialect.....	63
Table 7. Dark- <i>sje</i> pronunciation by lexical item and age group.....	65
Table 8. Internal /rs/-sandhi, by gender and age in decades.....	76
Table 9. Internal /rs/-sandhi, by gender and age in half-decades.....	77
Table 10. Internal /rs/-sandhi, by gender and lexical item.....	79
Table 11. Internal /rs/-sandhi, by lexical item and age group	82
Table 12. Internal /rs/-sandhi, by gender and dialect area.....	85
Table 13. Internal /rs/-sandhi, by dialect and stress on preceding vowel.....	87
Table 14. Internal /rs/-sandhi, by dialect and stress on next vowel	89
Table 15. Internal /rs/ sandhi, by dialect and type of internal boundary	92
Table 16. Eva Strangert's measurements of pause length in MfoN	108
Table 17. External /rs/-sandhi, by gender and age in decades.....	111
Table 18. External /rs/-sandhi, by gender and dialect area.....	112
Table 19. External /rs/-sandhi, by first lexical item and gender	113
Table 20. External /rs/-sandhi, by preceding syntactic category and gender	116
Table 21. External /rs/-sandhi, by gender and second lexical item.....	118
Table 22. External /rs/-sandhi, by following syntactic category and gender.....	120
Table 23. External /rs/-sandhi, by direction of c-command across dialects.....	121
Table 24. External /rs/-sandhi, by 'closeness' and dialect	123

Table 25. /rs/-Sandhi, by dialect and boundary type (internal & external combined) ...	125
Table 26. External /rs/-sandhi, by gender and the vowel preceding /rs/	127
Table 27. External /rs/-sandhi, by gender and the vowel following /rs/.....	128
Table 28. External /rs/-sandhi, by gender and segment following /s-/.....	130
Table 29. External /rs/-sandhi, by age and stress on the preceding vowel.....	131
Table 30. External /rs/-sandhi, by age and stress on the following vowel	133
Table 31. External /rs/-sandhi, by stress on preceding and following vowels.....	134
Table 32. Unreduced <i>är</i> , by age and gender	149
Table 33. Unreduced <i>är</i> , by dialect and gender	150
Table 34. Unreduced <i>är</i> , by gender and following word class	151
Table 35. Unreduced <i>är</i> , by age and following word class	152
Table 36. Unreduced <i>är</i> , by age and following stress	154
Table 37. Unreduced <i>är</i> , by dialect and following stress	156

GLOSSARY

External sandhi. Sandhi between two sounds across a word boundary

Internal sandhi. Sandhi between two sounds inside the same word

North Norrland. The northernmost 40% of the territory of Sweden

/rs/-Sandhi. Sandhi of an underlying /rs/ sequence, resulting in phonetic [ʂ]

Sandhi. A phonological blending of two sounds

Sje. /ʃ/, a fricative phoneme in Swedish with widely varying pronunciations

Tje. /ç/, a fricative phoneme in many dialects of Swedish, and an affricate in others

LIST OF ABBREVIATIONS

- DA.** Dialect Area (see Figure 9. Map of North Norrland dialect areas)
- MfoN.** Min familj och Norrland 'My family and Norrland' (see Appendixes A & B)
- Msi.** Min styvfars idol 'My stepfather's idol' (see Appendixes C & D)
- NN.** North Norrland (the northernmost 40% of Sweden's land area)
- SAOB.** *Svenska Akademiens Ordbok* (see Bibliography)
- U.** Uppland (a Dialect Area)
- WWII.** World War II (1939-1945)

PREFACE

The major part of this dissertation is organized by topics, with a large part of the input on each of those topics consisting of pronunciation data from two stories I wrote that were read by hundreds of people. To help the reader keep a perspective on the whole enterprise, let me first give a quick sequential run-through of the past few years as I experienced it, highlighting the questions that arose along the way. Then, in the main body of the text, we will deal with those questions in the appropriate sections, cutting across that timeline.

Many foreign students when beginning to learn Swedish are no doubt surprised, as I was, to hear the given name *Lars* pronounced in a way that an English speaker might want to spell as 'Losh.' That experience could be said to have planted the seed of this project in my mind, though it took several years for that seed to sprout.

The next step on the road to this dissertation was Linguistics 550, a class on the interface between Phonology and Syntax that I took from Ellen Kaisse. While researching my paper for that class, I became aware of some of the numerous unresolved questions regarding the various factors that govern the occurrence or nonoccurrence of sandhi (combinatorial interaction) between the phonemes /r/ and /s/ in Swedish. In the process, I visited Sweden and learned from Eva Ejerhed in Umeå about the Talspråksyntax (Syntax of spoken language) project that had been conducted at Lund University as a joint endeavor of the Linguistics and Sociology Departments. Jan Einarsson in Lund made available to me their tapes and transcripts from the interviews conducted in Borås. An important additional lesson I learned from working with the Borås materials was that potential /rs/-sandhi contexts could be found only a couple times per page of transcript, so that if I wanted to gather significant amounts of data in a reasonable amount of time, I would need to find a way to compress the process. This was the impetus that ultimately

led me to the idea of writing stories for people to read with a wealth of sandhi contexts, so that I could gather a lot of data in a relatively short time, and with the added advantage of having comparable data for all the people I interviewed, which would facilitate statistical analysis of the different population groups.

Soon after I finished that paper, I obtained Stig Eliasson's (1986) concise overview of sandhi in Norwegian and Swedish, which made a number of claims that agreed with my findings, but which also made other claims that I questioned, leading me to think that further exploration of the topic might be warranted. Many of his claims became hypotheses for my investigations, and when I was awarded a Fulbright grant to conduct research in Sweden during the academic year 1990-91, I was honored that he agreed to serve as my advisor there.

As I became interested in investigating /rs/-sandhi further and was trying to formulate a narrower topic, I focused on a map (Lindblad 1980: 141) showing an area in northern Sweden where the product of /rs/-sandhi is identical in pronunciation to an underlying fricative sound of Swedish known as /ʃ/ or *sje*, which is given a variety of pronunciations in other dialects; I decided to concentrate my research on that region. With my topic thus narrowed, another study by Per Lindblad (1982) on *sje* (plus another phoneme known as /ç/ or *tje*) in that part of northern Sweden also became highly germane. He described an intricate pattern of variation for the phonemes *sje* and *tje* in six towns in northern Sweden, and formulated a whole series of rather specific hypotheses about their patterning.

At the start of my 1990-91 Fulbright year in Sweden, I wrote a short story containing dozens of potential sandhi contexts that was designed to test a variety of hypotheses simultaneously and to allow for additional serendipitous discoveries about patterns of sandhi. In September/October 1990 I traveled through many parishes of Sweden's northernmost provinces, Västerbotten and Norrbotten, interviewing almost 200 people and recording their readings of my story in order to collect a body of data for my study.

While listening to those tapes and doing a preliminary analysis of the data later during the winter, I was struck by an unanticipated pattern of variation for the copula *är* ('are/is/am') that appeared to depend on prosodic context. I wrote a second story in 1991 for purposes of gathering additional data, including some that was designed to test hypotheses based on the patterns for the copula *är* that I thought I had detected in the data from the first story. I traveled around in the north of Sweden again in the summer of 1991 to interview over 150 people and collect an additional body of data by taping readings of this second story. I also traveled to northern Sweden on my honeymoon in 1996, and used the first story to collect a small body of supplemental data to broaden my base of dialects for that story. The largest part of this dissertation is based on the presentation and analysis of the data that I collected from people reading those two stories, viewed within a framework of hypotheses arising from previous studies.

ACKNOWLEDGMENTS

The author wishes to acknowledge that this dissertation was made possible in part by a Fulbright grant that allowed him to spend a year in Sweden, and by the granting of leaves of absence for that and other phases of the study by his employer, the United States Postal Service. It was also made possible by the help and support of hundreds of people, including some whose names are unknown to him. First and foremost, thanks are due to his two principal advisors, Ellen Kaisse at the University of Washington and Stig Eliasson at Uppsala University. Besides teaching the course on the syntax-phonology interface that provided the initial impetus toward the research that eventually resulted in this dissertation, Professor Kaisse has made a major contribution both with her editorial efforts and by discussing the issues involved. Others at the University of Washington who have also given of their time and wisdom include his committee members, Lee Osterhout, Pat Conroy, Sharon Hargus, and Joe Voyles, several members of the Scandinavian Department including especially Lars Warne and Ia Dubois, and many members of the Linguistics Department including Alicia Beckford Wassink, Siri Tuttle, Alice Taff, Maria Galvao, and Lorna Rozelle.

He is also indebted for the generous hospitality of the Departments of Linguistics at Uppsala University and Umeå University. At Uppsala University, besides Stig Eliasson who has been most helpful both during and since the 1990-91 Fulbright year there, mention must be made of help that was rendered by Anders Holmberg, Gunnel Björnhag, Alejandro Engelmann, Caroline Liberg, John Swedenmark, John-Sören Pettersson, Jan-Erik Widell and Ingrid Björk. At Umeå University his research benefited from assistance by Eva Ejerhed, Claes-Christian Elert, Eva Strangert, Ola Wennstedt and others. Elsewhere in Sweden, he received excellent counsel and other aid from Tomas Riad at Stockholm University and Per Lindblad at Gothenburg University.

He has also received valuable assistance from the able staffs of two archives of dialectal materials, Dialekt- och folkminnesarkivet in Uppsala and Dialekt- ortnamns- och folkminnesarkivet in Umeå. Finally, he is very specifically indebted to each of almost 400 people for reading one or the other of the two stories, most of them life-long residents of Sweden's two northernmost provinces, Västerbotten and Norrbotten, as well as a few residents of Uppland; without their efforts he would have had no data to work with and none of this would have been possible.

DEDICATION

The author wishes to dedicate this thesis to his wife Yumin Xie Lindblad.

INTRODUCTION

Phonological assimilation processes are widespread in the languages of the world. At the most detailed end of the spectrum, the realities of the articulatory act are such that it is virtually inevitable that in the transition between two adjacent sounds there will occur some degree of blending of their features, although it may take a very sharp ear or even an analysis employing sophisticated technical instruments to reveal these effects. At the other more overt end of the spectrum, a sound may be changed so radically by an adjacent sound that it can become difficult or impossible to recognize its original qualities. Phonological alternations of this more easily noticeable sort are commonly referred to by the Sanskrit grammarians' term, *sandhi*. When sandhi occurs within words it is known as *internal sandhi*, and when it occurs between words it is known as *external sandhi*.¹ Sandhi can usefully be further analyzed on a different dimension, namely whether it is a purely phonological rule dependent only on speech rate, syllabification, and the features of the focus and determinant (a "fast speech rule"), or whether it is sensitive to the syntactic or morphological environment in which its phonological terms appear (a "phonological rule of connected speech") (Kaisse 1985: 1).

¹ Most English speakers are familiar with external sandhi in the form of palatalization, which changes "t, d, s, and z into c, ĵ, š, and ž, respectively, when the next segment is y."(Kaisse 1985: 35) This is also reflected in common non-standard spellings for certain word-pairs such as 'did you' → 'didja', 'what you' → 'whatcha', 'bet you' → 'betcha', etc.

CHAPTER 1: BACKGROUND

1.0 THE PROBLEMS TO BE ADDRESSED

What happens when /r/ is followed by /s/ in Swedish? Under what conditions do these two sounds retain their separate identities, and when do they instead merge into a single sound that takes certain features from each? These questions about /rs/-sandhi are the central focus of this dissertation. In addition other closely related sounds will be touched on, for they are relevant to establishing the broader context in which the phonological blending of the /rs/ sequence occurs. The principal additional restriction imposed on the scope of this study is that most of the data that we will be looking at was collected from speakers who live in the northern 40% of the geographical area of Sweden. A special characteristic of the dialects in this region is that the sound that is the product of /rs/-sandhi, [ʂ], is also the most common pronunciation for another underlying sound (or phoneme) known as *sje* or /ʃ/, which arose historically from the palatalization of /s/. This raises additional questions about whether the identical pronunciation of these two different underlying representations for [ʂ] leads speakers into confusion about their identities and gives speakers any difficulties in keeping them separate, so that speakers can ever end up substituting another variant of /ʃ/ such as [f] for /rs/.

1.1 A SKETCH OF SWEDISH SEGMENTS & SUPRASEGMENTAL PHONOLOGY

An issue in the phonological analysis of Swedish which has not been completely resolved (and which is somewhat muddled by dialectal variations) is the inventory of phonemes, both vowels and consonants.² Neither the vowels nor the consonants of Swedish can be

² Resolving some of these issues is the principal concern of Eliasson's (1970) dissertation.

presented in a simple chart without omitting many details. However, not all the details are significant for the purposes of this dissertation, so I will simplify where possible in order to present a clearer picture.

1.1.0 VOWELS

The Swedish vowel system has nine vowels, each of which has the possibility of taking on length if it is stressed.³ In Figure 1 the vowels are represented with their normal orthographic representation, although that does not closely correspond to the International Phonetic Alphabet, because normal orthography will be used throughout this dissertation. The two high front rounded vowels are distinguished by very different lip shapes, among other articulatory details (see Wood 1982 for a detailed discussion of the acoustical and articulatory differences between /u, y/). These two vowels will be of special interest to us later, when we will see that they seem to behave as polar opposites, in terms of being conditioning factors that favor or disfavor /rs/-sandhi.

	front		back	
	unrounded	rounded	unrounded	rounded
high	i	y		o
		u		
mid	e	ö		å
low	ä		a	

Figure 1. Standard Swedish vowel system

³ The long and short variants of these nine vowels are often articulated at locations in the mouth that are far enough apart to be easily heard as different (cf. Elert 1964: 14, Elert 1970: 67, Eliasson 1970: 32, Hammarberg & Svensson 1975: 65, Garlén 1988: 61) This has led some people to analyze Swedish as having an 18 vowel system. Eliasson & La Pelle (1973) have convincingly refuted that theory.

1.1.1 CONSONANTS

It is generally agreed that Swedish has a dental series of consonantal phonemes: /t, d, n, l, s/. Paralleling this (but not in the south of Sweden where /r/ is back, nor in most Swedish spoken in Finland) is a series of consonants which normally result from the assimilation of /r/ to a following dental and which have variously been termed ‘retroflex,’ ‘supradental,’ ‘postalveolar,’ or ‘apico-alveolar’: [t̠ d̠ n̠ | s̠]. However, “[s]ome dialects have two series, an apico-alveolar (supradental) after r and an apico-palatal (‘cacuminal’) after l” (Haugen 1976: 275). Although they omit those cacuminals, Hammarberg & Svensson (1975: 71) present a chart that includes all the consonants commonly found in dialects of Swedish. Since this chart agrees generally with my analysis of the standard Swedish consonant system, I present my translation of it here as Figure 2, with only the slight additional modification of added square brackets [] to specify those segments that occur on the surface but are not underlying.

		labial	dental	retroflex	palatal	velar	laryngeal
stops	voiceless	p	t	[t̠]		k	
	voiced	b	d	[d̠]		g	
fricatives	voiceless	f	s	[s̠]	ç	ʃ	h
	voiced	v		[(ɹ)]	j	[(ʁ)]	
nasals		m	n	[n̠]		ŋ	
laterals			l	[l̠]			
r-sounds				r		[R]	
				[ɹ]		[ʁ]	

Figure 2. Consonant inventory of standard Swedish

It should be pointed out that /f, v/ are actually labiodental, not simply labial, and the back /r/-sounds vary from velar to uvular to pharyngeal (p.c., Per Lindblad, 1999). It is also important for us to note that the phoneme that we will be referring to in this dissertation as *sje* or /ʃ/ (and that will be a major topic in Chapter 3) is not always realized as the velar fricative [ʃ] (see Lindblad 1980 for a detailed discussion). Furthermore, most dialects of Swedish do not use all four of the phonetic realizations of the /r/-phoneme shown in this

chart. In southern Sweden, most speakers will have a uvular variant for the phoneme /r/, they will lack all the other consonants in the retroflex series, and there is a good chance that they will pronounce the phoneme /ʃ/ as the velar fricative [ɣ] or give it another articulation in the back of the mouth. On the other hand, in the northern dialect area that we will be primarily concerned with, all the consonants in the retroflex series occur, uvular /r/ is not the norm, and a velar realization of the phoneme /ʃ/ is not the standard, although it is sometimes found, as we will see in Chapter 3 below.

Introductory textbooks and student grammars of Swedish often list the sounds of the language in the first chapter, and they include the sound that will be our main focus, the retroflex voiceless fricative [ɕ], along with the other four retroflex consonants produced by sandhi [t̪ d̪ n̪ l̪] (and usually an example word for each), but they do not go into any details about their geographical, lexical or syntactic domains and other limiting factors.

The best concise treatment of these issues in such a work that I have encountered is Björkhagen (1962) who went to the unusual lengths of having three phoneticians (Daniel Jones, Lilius E. Armstrong, and Claes-Christian Elert) as consultants. He begins by stating, “When the dental sounds *t, d, n, l, s* are immediately preceded by *r* in the spelling, they become post-alveolar instead of dental, i. e. their point of articulation is moved further back, approximately to the *r*-position, further back than English *t, d, n, l, s*. The tip of the tongue is curled somewhat backwards to the back part of the teeth-ridge. The resulting sounds might be described as “retroflex” *t, d, n, l, s*. Phonetic symbols: [t̪, d̪, n̪, l̪, ɕ]. Orthographically these sounds are represented by the spellings *rt, rd, rn, rl, rs*. In the pronunciation the *r* is silent and the following consonant acquires the retroflex character. Practise the following word pairs: ...” (Björkhagen 1962: 6). After giving minimal pairs for each of the dentals with its corresponding retroflex, he goes on to make several other important points, interspersed with examples: “Retroflex [ɕ] resembles the English *sh*-sound in ‘she’, but it is formed further back. The tip of the tongue is curled

back behind the ridge of the upper teeth. The lips are considerably rounded. (...) In groups of words, spoken without a pause between them, initial dental sounds become amalgamated with a final *r* in a preceding word to form the retroflex variety. (...) Dental sounds following immediately after a retroflex sound become retroflex by attraction. (...) Retroflex [ʂ] is also used as the ordinary sh-sound. (...)."(Björkhagen 1962: 7) The complexities that have begun to become apparent here (although as they stand, each of these statements is overly simplified and potentially misleading) may be good motivation for sparing beginning language students the bewilderment that could be created by a detailed treatment of the subject. Those same complexities are a part of what makes /rs/-sandhi a subject that deserves a more detailed analysis than it has received until now.

1.1.2 STRESS

Stress in Swedish is characterized by several features. One is an increased degree of expiratory force, which is similar to the principal feature of stress in English (though on average possibly somewhat weaker). A second feature that is found on all syllables with primary stress and some syllables with secondary stress is length, which is realized either on the vowel or on the consonant immediately following the vowel. A third feature that is often associated with syllables with primary and secondary stress is an intonation curve such that a pitch peak lies on or near the stressed syllable, although the precise details of the shape and placement of that curve vary considerably between dialects.

Each word can have only one syllable with primary stress. Normally all major-category words (or open-class words, i.e. nouns, main verbs, adjectives, and adverbs derived from adjectives) will have a syllable with primary stress, but minor-category words (also known as closed-class, function or grammatical words, and including prepositions, conjunctions, pronouns and other adverbs) need not have any stressed syllables, depending on their context in the sentence. There is considerable variation in the treatment of secondary stress in the literature. Factors that may be taken into consideration in distinguishing among weaker degrees of stress include the presence or

absence of length and the special pitch contour associated in some dialects with Accent 2. Elert (1964: 17) has combined a variety of possible combinations of stress patterns in a table, given below as Table 1, which shows some of the main variants involved.

Table 1. Stress patterns in Swedish: some variant notations

Examples		A. Noreen's system	SAOB	I.P.A.	Syllabic structure
<i>boken</i>	'the book'	30	40	['bu:ken]	'SS
<i>stenar</i>	'stone'	31	32	[^x ste:naɾ]	^x SS
<i>bönbok</i>	'prayer book'	32	32	[^x bøn, bu:k]	^x S,S
<i>böckerna</i>	'the books'	300	400	['bøk:ɛŋa]	'SSS
<i>stenarna</i>	'the stones'	302	302	[^x ste:naŋa]	^x SSS
<i>filosofi</i>	'philosophy'	1003	1004	[filɔsɔ'fi:]	SSS'S

It is generally agreed that the primary stresses in Accent 1 and Accent 2 words are equivalent (as in Noreen's system), and that the special character of the intonation pattern on Accent 2 words in most central and northern Swedish dialects is connected with the secondary stress that falls on another syllable later in the word. When *SAOB* and some other dictionaries mark the primary stress in Accent 1 and Accent 2 words differently (e.g., *SAOB* marks primary stress in Accent 1 words as '4' and primary stress in Accent 2 words as '3'), that is to be understood as primarily a convenient way of distinguishing the two word accents, rather than as saying that their primary stresses are different. Some authors have noted that the distinctive pitch pattern of Accent 2 words tends to fade away if they are not stressed above the average level of the sentence; Gösta Bruce (1977) has interpreted this to mean that Accent 2 is a manifestation of sentence stress, and not a lexical phenomenon. However, the choice of the syllable on which the second stress is realized is clearly specific to the lexical item involved, so I believe that the lack of a distinct second peak on unstressed Accent 2 words is best treated as a matter of destressing.

The primary and secondary stresses that are associated with pitch peaks and/or syllabic lengthening are clearly lexical stresses, and vary in their location between different occurrences of the same word only for a few unusual lexical items. Other weaker degrees of stress also exist in Swedish. Many compounds contain three or more roots, and in such compounds even those roots that do not receive either primary or secondary stress usually retain some additional portion of stress and length beyond what is accorded unstressed syllables.

A further type of stress may be seen in cases where otherwise unstressed syllables occur in sequences of three or more, whether within a single word or across a series of words. This type of stress arises to fill the unacceptable gap in the stress pattern left by a prosodic lapse (see Bruce 1984, Nespor & Vogel 1989), and is postlexical in nature (see Kaisse & Shaw 1985) so it comes and goes depending on the presence or absence of adjacent lexical stress. Its effects can be seen in the numerous stressing alternatives given by *SAOB* for many polysyllabic words. It also accounts for some instances of alternation between full and reduced forms of function words (see discussion of [e] ~ [er] alternation in the pronunciation of the copula *är* 'is/are/am' in Chapter 6 below).

1.1.2.0 Stress and length

In Swedish the primary phonetic feature that correlates to phonological stress is a lengthening of the syllable's duration. Secondary features reflecting stress are the intonation pattern and intensity, though these vary considerably in their details between dialects, and will not concern us here. The standard phonological notation for syllabic length is the mora, and this notation with all the attendant principles that have been developed around it fits Swedish data nicely, as Riad (1988, 1992) has shown from a diachronic perspective. Riad (1992) presents a moraic formalism for several words essentially identical to their modern values, including *falla* /fal:a/, *tak* /ta:k/ and *hand* /han:d/ as seen in Figure 3.

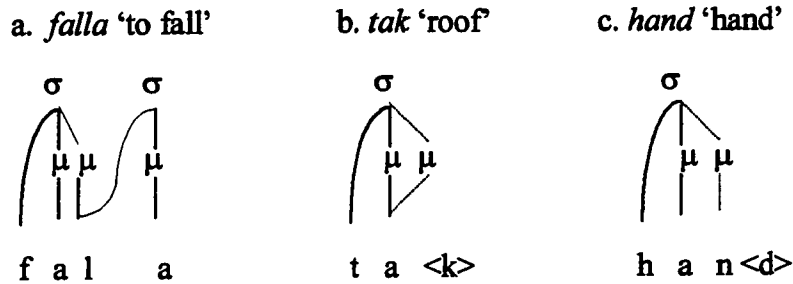


Figure 3. Moraic representation of Swedish stressed syllables (after Riad 1992)

Since all of Riad's examples are major-category words, they illustrate a feature common to all major-category words, as well as to the strong forms of function words in Swedish. That feature is a long syllable, or, as it may alternatively be expressed in terms of phonological weight, a heavy syllable or a bimoraic syllable. In a bimoraic syllable, the first mora regularly associates with the vowel, and the second mora associates either with that same vowel or with the following consonant, in a way that cannot be determined by rule, but that rather requires listing in the lexicon of the moraic status of consonants following directly after stressed vowels (Riad 1992). Unstressed syllables, on the other hand, are only one mora in length⁴, and that mora always associates with the vowel.

1.1.2.1 Stress: lexical and postlexical

One general framework assumed here is lexical phonology (Kaisse & Shaw 1985, Mohanan 1986). I assume that all major-category words, including monosyllables, receive stress in the lexicon by virtue of their major-category status before insertion into the syntax, and therefore can never have all of their syllables be totally unstressed in the same way that monosyllabic function words can. I also assume that all words of two or more syllables, including function words, must go through one or more stress cycles before insertion into the syntax, and thereby receive lexical stress on (at least) one of their syllables (cf. Selkirk 1984: 35 for an alternate account employing destressing). This significant role played by word-internal stress patterns is similar to the invariant patterns

⁴ Unstressed syllables on average probably also contain a smaller number of segments, though I am not aware of any studies of this topic.

of polysyllabic words in poetry discussed by Hayes (1988). He notes that monosyllables often are assigned a weak or strong stress in the metrical schema that does not correspond to a normal expectation for how they should be stressed; but if a bisyllabic word has primary stress on the first syllable in the lexicon, it will never be aligned with two spots in the verse that are labeled weak-strong, instead of strong-weak. Although major-category words have a greater degree of stress inherently than function words, this does not prevent function words from taking on extra stress in special circumstances; syllables (including monosyllabic function words) that have not received stress in the lexicon can still receive it postlexically, after they enter the prosodic component of the grammar.

A similar distinction was made by Sweet already in 1876, when he described the stress on major-category words as a “purely logical stress,” while the stress that sometimes falls on function words is “a purely metrical stress.” (Sweet 1913: 12) This resembles Hayes’ (1984: 33) proposal (cited by Nespor & Vogel 1989: 96) that rhythmic structure and linguistic stress are distinct phenomena and should be handled by metrical grids and metrical trees, respectively. Given that erasure of morphological boundaries and treating all syllables solely on the basis of their phonological components is characteristic for postlexical rules (Kaisse & Shaw 1985: 4-5), there seems to be the basis for a consensus that the locations of lexical stresses are fixed, but postlexical stresses can vary in placement depending on the prosodic environment.

Variability in the location of the secondary stress in some longer Swedish words with only one lexical stress has been discussed by Elert (1970: 38) and Bruce (1984), and can also be noted in many of the polysyllabic entries in *SAOB*. Although Elert does not make any claim about the conditioning factors, in my analysis it is the prosodic context that determines which of the syllables at or near the lexically unstressed end of such a word will receive postlexical stress. Similarly, whether or not a particular function word such as *är* will receive postlexical stress also depends on the prosodic context, and more specifically on whether or not it needs to be stressed to repair a lapse. An account along the lines of Nespor & Vogel (1989), using additional stresses to repair lapses, is able to

provide a single, unified explanation for these two similar phenomena of alternating stresses.⁵

1.1.2.2 Stress of function words

A primary distinction that must be made for lexical stress is between major-category words and minor-category words. Major-category words are generally agreed to include nouns, verbs, and adjectives, and sometimes adverbs are included as well. However, our principal interest now is in minor-category words. Sweet, in his 1876 article "Words, Logic, and Grammar" (included in Sweet 1913), refers to minor-category words as 'half-words,' but they are more commonly called function words. Function words include "auxiliary verbs, modals, prepositions, determiners, conjunctions, personal pronouns, etc." (Selkirk 1984: 335) Some scholars (cf. Hayes 1995: 377, f.n. 1; Halle & Vergnaud 1987: 264) mention function words, but describe them as unstressed and come close to implying that they are unstressable, an idea that I will argue is refuted by MfoN-39 and similar evidence. That English function words can take at least some degree of stress is shown by the evidence provided by Selkirk (1984), especially Chapter 7 "Function Words: Destressing and Cliticization" (cf. also her examples in (6) below).

Like English, Swedish has a number of monosyllabic function words that have both strong and weak forms. However, the phonologies of the two languages are quite different, so the phonetic realizations of reduction in the weak forms of function words are different in several ways. For one, several common English function words such as *are*, *is*, *will*, *would*, *had*, *has*, *have*, *in* and *and* are regularly reduced to a single consonant, and completely lose the vowel of the strong form. Forms of these and other English function words where the vowel reduces to schwa are also very common. In *Svenska ord* (1984) no Swedish word is shown as having a reduced form without a vowel; but most of the words that do have reduced forms are function words. What I

⁵ Selkirk's (1984, Chapter 7) destressing account of function words does not as easily generalize to polysyllabic words containing sequences of syllables without lexical stress.

would like to argue here is that these Swedish function words all have only one mora when they are in their weak form⁶, and in this respect they are like other unstressed syllables in Swedish.

1.2 PREVIOUS STUDIES

Dahlstedt & Ågren (1954) provide a survey of the phonetic traits associated with and distinguishing the various traditional rural dialects of Övre Norrland,⁷ several maps of isoglosses, and texts from a scattering of locales in each dialect area recorded from speakers, most of whom were born in the 19th century. Their transcription of those texts assumes that retroflexion always occurs whenever there is a sequence of /r/ + dental and that they have annotated all exceptions (Dahlstedt & Ågren 1954: 10), an assumption that suggests that we should find sandhi in most contexts both word-internally and externally. I have not attempted to make use of their dialect texts for this study. Their primary contribution to my concept of this study is their delineation of the dialect areas within NN. Our dialect areas 1-8 in Map 1 correspond to the traditional Swedish dialect areas on their overview map (Dahlstedt & Ågren 1954: 230), and our dialect areas 9 and 10 correspond to the peripheral areas that they designate as traditionally Finnish and Sami speaking, respectively.

Per Lindblad's (1980) dissertation provides a comprehensive overview of the phonetic variability of *sje* and *tje* in Swedish. Since he follows the standard practice in Sweden of referring to the Swedish phonemes that originated historically from palatalized /s, t/ as *sje* and *tje* respectively, and writing them in phonemic notation as /ʃ, tʃ/, those conventions

⁶ Furthermore, some of my data can be interpreted as suggesting that syllable geometry plays a role in determining what the results of that loss of the second mora will be (cf. the distinction between non-moraic and moraic coda consonants in 3b and 3c, respectively). However, I must leave that topic for further research.

⁷ Their Övre Norrland is essentially identical to the traditionally Swedish speaking parts of our North Norrland, and omits our dialect areas 9 and 10 where Finnish and Sami were spoken.

will be followed here as well. The inventory of underlying fricatives in the dialects of central Sweden (e.g. Stockholm) is standardly analyzed as consisting of /v, j, f, s, ç, ʃ, h/.⁸ In addition, the retroflex fricative [ʂ] occurs frequently in the central Swedish dialects, but for most speakers there it is best analyzed as being the product of (possibly optional) sandhi of the underlying pair /rs/. This picture is complicated in south and central Sweden by the extreme variability of the phonetic realizations of phonemic /ʃ/, which include [ʃ, ʂ, ʂ̥, f, x] (discussed in detail throughout Lindblad 1980). Lindblad (1980) argues for sibilants as a special class of sounds, distributed along its own scale that need not directly correspond to the distribution of points of articulation found for other classes of consonants such as stops or nasals. He argues that the Swedish fricative series includes three sibilants /s, ç, ʃ/ which are articulated with somewhat different shapes of the top of the tongue and constriction points, but whose true distinctive feature is probably the size of the cavity under the tip of the tongue: none for /s/, small for /ç/, large for /ʃ/ (Lindblad 1980: 76, illustrated in Figure 4.5 on p. 81, English summary p. 195). Lindblad (1980, 1982) was primarily interested in the phonetic variability of *sje*, but he also looked at the variability of *tje* because these two fricatives form an articulatory and acoustic continuum; however *tje* is outside the scope of this study.

Per Lindblad (1982) looked at two large sets of data collected during the 1970's in six towns in northern Sweden (Umeå, Lycksele, Skellefteå, Arvidsjaur, Luleå, Malmberget) as part of a larger study of northern Swedish in those towns (Elert & Fries 1982). Although [ʂ] is produced by /rs/-sandhi in those towns, as in most of the rest of central and northern Sweden, the main focus of his attention was on *sje*, and he found that the predominant realization of *sje* there clearly was [ʂ], but other variants were also found.

⁸ The Swedish phoneme /j/ is a voiced palatal fricative or approximant, sometimes referred to as a 'yod', and similar to the sound often written as /y/ by American phonologists. To avoid confusion with the Swedish phoneme /y/, a high front rounded vowel, I have decided to follow Swedish usage here.

He was able to make a number of generalizations about the distribution of [ʂ] and the other variants of *sje*, using sex, age group (children vs. parents), hometown, and the prosodic structure of words as factors. We will use many of his generalizations to form hypotheses that we can test with our new data. The main exception he found to the regular realization of *sje* as [ʂ] is in the twin mining towns Gällivare and Malmberget, where adult males had a ‘darker’⁹ norm for the pronunciation of *sje*, more akin to the norms in central and southern Sweden, from where many miners had immigrated during the early decades of this century. This darker norm for *sje* also influenced in lesser degrees the speech of younger boys and girls there, and, in a diluted form, the speech of males and children of the nearest city, Luleå. He also found that *sje* in a prosodic context following a stressed vowel was virtually guaranteed realization as [ʂ] in all six towns and for all speakers, a pattern that he described as unique to northern Sweden. We will formulate specific hypotheses based on these generalizations in Chapter 3, where we will investigate in more detail the patterns of variation of *sje* in our data.

Kaisse (1985) looked at various forms of phonology/syntax interaction, including a number of examples of sandhi in a variety of languages, though not at Swedish. Her general argument was that the syntactic relation between two words governed their interaction. However, in my very first bit of data collection on /rs/-sandhi, when a native speaker from central Sweden read a list of word pairs, he produced opposite results for the two items *har sett* (has seen) and *har sökt* (has sought). Since these presumably must have identical syntactic relationships, I have worked under the hypothesis that if syntactic relationship plays a role, it is probably a less dominant role than other factors play, though that does not mean it has no role, as is known for example from the role played by lesser ranked constraints in Optimality Theory. In any case, we will test hypotheses regarding syntactic factors when we look at external sandhi in Chapter 5.

⁹ ‘Dark’ fricatives are ones that give an auditory impression of being lower because the sharp cut-off of their strong high frequency energy occurs at a relatively lower frequency (Lindblad 1980: 193).

Eliasson (1986) provides the most comprehensive overview of sandhi in Norway and Sweden available. It is based primarily on previously published sources in the linguistic literature, but also draws on his own judgements and intuitions as a phonologist and native speaker from central Sweden. His main emphasis is on “how extremely regular and systematic the assimilation can be” (Eliasson 1986: 279). He presents the range of contexts in which sandhi occurs rather clearly by means of a chart, showing various types of internal sandhi as well as external sandhi (Eliasson 1986: 278), of which I will replicate here as Figure 4 the parts exemplifying /rs/-sandhi (and omit those dealing with sandhi with the other dentals).

A. Morpheme-internally				
<i>fors</i>	/fors/	[fɔ̯ʂ:]		‘rapids’
B. Across inflectional boundaries				
<i>förs</i>	/fɔ̯r+s/	[fɔ̯ʂ:] (pass.)		‘is brought’
C. Across derivational boundaries				
<i>försorg</i>	/fɔ̯r+sɔ̯rj/	[fɔ̯ʂ'ɔ̯r:j]		‘taking care’
D. In formal compounding ¹⁰				
<i>försal</i>	/fɔ̯r#sal/	[^ʔ fɔ̯ʂ:;sɔ̯:l]		‘entrance hall’
E. Across regular compound boundary				
<i>färsax</i>	/fɔ̯r#saks/	[^ʔ fɔ̯ʂ:;sɔ̯k:s]		‘sheep shears’
F. Across word boundary				
<i>för sen</i>	/fɔ̯r##sen/	[fɔ̯ʂ'se:n]		‘too late’

Figure 4. Sandhi contexts (after Eliasson 1986: 278)

However, Eliasson also lays out a number of exceptions to the regular pattern of assimilation, some described by phonological characteristics and others limited to certain

¹⁰ Formal compounds are like regular compounds in that they take Accent 2 and also have long segments in both halves of the compound; they are different because one half is not a major-category word so it would be expected to be unstressed.

specific lexical items. I incorporated many of these morphological patterns and specific words into my stories, and so tested my readers in northern Sweden for them. We will formulate specific hypotheses based on these generalizations in Chapters 4 and 5, where we will investigate in more detail the patterns of variation in our data for internal and external /rs/-sandhi, respectively.

In Lindblad (1987a), I did a small study of /rs/-sandhi in the central Swedish town of Borås, based on a set of tapes and their transcriptions, which were data collected by the Departments of Linguistics and Sociology at Lund University as part of a study of the syntax of spoken Swedish. These interviews had been conducted with thirty-two people, eight men and eight women from each of two social classes, one of them professionals and the other factory workers. One clear result was that 100% of the speakers had sandhi across the word boundary in the two fixed phrases *år sen* ('years ago') and *(i) stort sett* ('on the whole'), while some other pairs of words had varying lesser degrees of sandhi usage. It also was clear that different morphological boundaries within words affected sandhi production to varying extents, so we will hypothesize that the same should apply to our new data from the far north. A striking and unexpected discovery from the Borås material was that male factory workers had much less sandhi in morphologically complex words such as those where the first syllable was the prefix *för-* than did women of both social classes or professional men. Some of these findings resulted in hypotheses that we will test on my new data. However, I am unable to determine the social class membership of the people who read my stories, so I cannot test for comparable sociological patterns in my new data.

In Lindblad (1987b) I worked with an informant from the southern part of Norrland in central Sweden, using word lists that mostly consisted of auxiliary verbs followed by main verbs. Although this methodology left the results open to question, it appeared that the frequency of words bore no relationship to their probability of participating in sandhi, and it also seemed that words beginning with the sequence <sm-> were somewhat more prone to sandhi.

In a paper presented at Nordic Prosody VII in 1996 (Lindblad 1998), using data from my field work in northern Sweden, I described patterns of alternation between reduced and unreduced forms of certain minor category words. These include the copula *är* ('are/is/am'), which is most commonly pronounced as [ɛ] without an audible final /-r/ in the north, as in more southerly parts of Sweden. A surprising result that I reported there was that when the prosodic context did not allow reduction because failure to stress *är* would create a 'lapse' (a sequence of three or more unstressed syllables), the final /-r/ not only remained audible but also frequently participated in sandhi with a following /s/-initial word. Since I was under the impression that this pattern of failure to reduce was not known in central Sweden, I hypothesized that it was a feature specific to the dialects of the north. However, as we will see in Chapter 5, additional data that I collected in central Sweden (Uppland) in 1996 following Nordic Prosody VII suggest that this pattern may be more widespread than previously reported, and deserves further research.

1.3 RULES

When /r/ and the dentals /t, d, n, s, l/ co-occur in sequence, they may interact to produce new sounds which partake partially of the qualities of each, but which are not identical to either one. One way of describing them is as retroflexed counterparts of the dentals from which they are derived, with the phoneme /r/ being deleted. Another approach would be to say that the new sound which is produced takes its place features from the /r/ and its other features from the following dental consonant.

Although the supradental consonants have a sort of semi-phonemic status, it seems reasonable in most dialects to derive them by rule for a variety of reasons, including the fact that they can also arise in novel contexts such as when one word ends in /-r/ and the following word begins with a dental, that it can spread from a single /r/ through an uninterrupted series of following dentals, etc. Proposals for accomplishing this derivation by rule have in essence taken one of two approaches, either first assimilating

the manner features of the /r/ to the following dental and then alveolarizing the dental (this approach should result in two timing-tier slots), or instead retroflexing the dentals in the environment following an /r/ and then deleting the /r/ (this approach could be expected to result in a single timing-tier slot).¹¹

The former approach was taken in his doctoral dissertation by Hellberg (1974: 166), whose rules (10) (/r/ Assimilation) and (11) (Alveolarization) are repeated here as Figure 5 a. and b (he also assumes that a rule of degemination applies following these rules).

(a) /r/ Assimilation

$$\begin{array}{c} \left[\begin{array}{l} + \text{ voc} \\ + \text{ cons} \\ - \text{ ant} \end{array} \right] \rightarrow \left[\begin{array}{l} \alpha \text{ voc} \\ \beta \text{ son} \\ \gamma \text{ lat} \\ \delta \text{ nas} \\ \epsilon \text{ cont} \\ \zeta \text{ voice} \end{array} \right] / \text{ ______ } (\#(\#)) \left[\begin{array}{l} + \text{ cor} \\ \alpha \text{ voc} \\ \beta \text{ son} \\ \gamma \text{ lat} \\ \delta \text{ nas} \\ \epsilon \text{ cont} \\ \zeta \text{ voice} \end{array} \right]$$

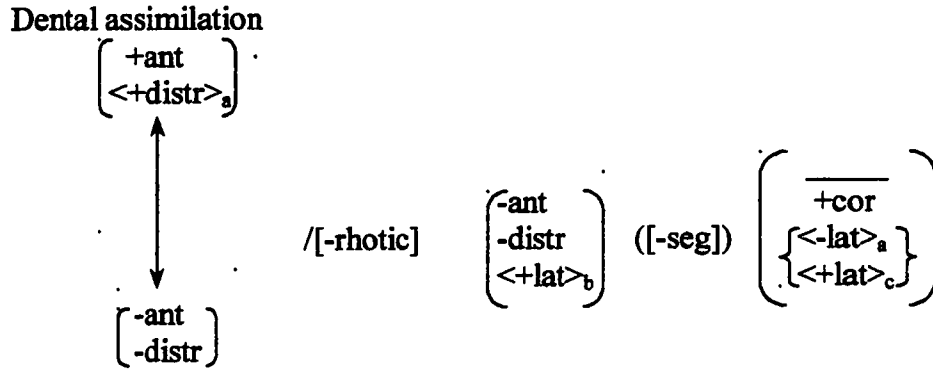
(b) Alveolarization

$$[+ \text{ cor}] \rightarrow [- \text{ ant}] / \left[\begin{array}{l} + \text{ cor} \\ - \text{ ant} \end{array} \right] (\#(\#)) \text{ ______ }$$

Figure 5. Assimilation, then Alveolarization (Hellberg 1974: 166)

Eliasson (1986: 286) has taken the latter approach, and formulated rules of Dental assimilation and r-Reduction, as in Figure 6.

¹¹ Although there are certainly ways in which the processes involved in retroflexion are rule-governed, and therefore amenable to formulation in ways such as those described, I do not thereby exclude the description of the same process in terms of constraints, as for example in the manner familiar from Optimality Theory.



- Conditions:
1. If *b*, the phonological input segment must contain *c*.
 2. Optional in the case of /r/.

b) r-Reduction

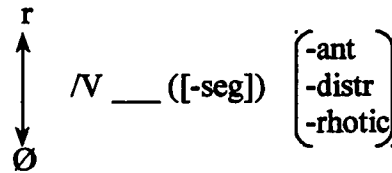


Figure 6. Assimilation, then Reduction (Eliasson 1986: 286)

Although Swedish linguists have traditionally expressed retroflexion as a two-step process, there are other possibilities. Adapting the formalism that Chomsky & Halle (1968: 360) employed for vowel contraction, Hovdhaugen (1969: 91) formulated the one-step rule given in for generating the retroflex consonants in Norwegian.

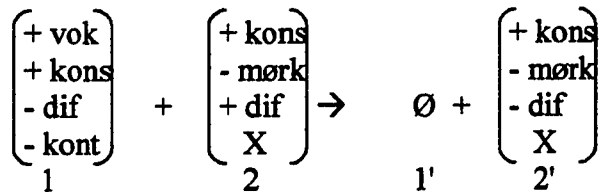


Figure 7. One-step retroflexion (Hovdhaugen 1969: 91)

This approach has the advantage of showing the process as a single step, which accords with the normal situation in which no trace of /r/ is audible as a separate sound. Although his formalism now appears outdated, it can be updated somewhat by employing the

insights and tier formalism of Autosegmental Phonology (Clements 1985, etc.). If we use this newer formalism together with the bundles of features employed by Hellberg (1974) to state the rules in Figure 5, we can express sandhi as a one-step process in a way that better fits with my intuitions of what occurs. Leaving aside the laryngeal feature [voiced] for the moment and dealing only with the manner and place features, we may depict the two rules in Figure 5 a. and b. as the single rule shown in Figure 8, which involves only the supralaryngeal tier (and its subsidiaries), as I have proposed elsewhere (Lindblad 1989).

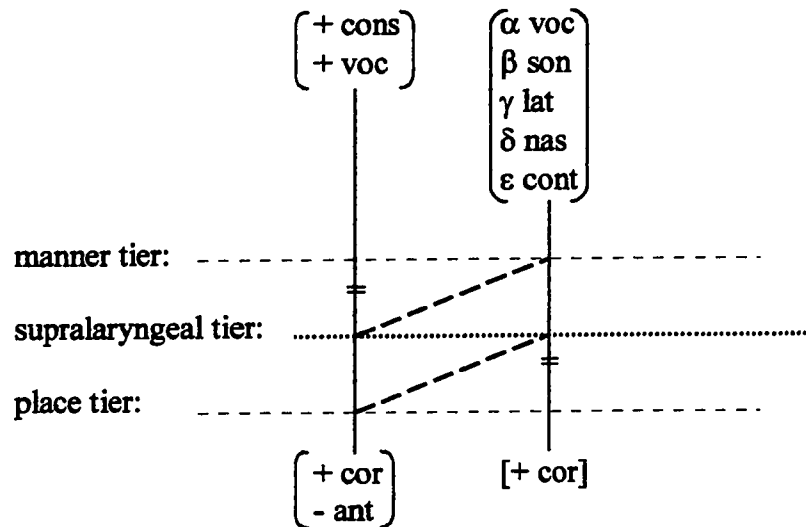


Figure 8. Retroflexion (Lindblad 1989)

Since we are primarily interested in /rs/-sandhi, but are also looking at retroflexion of the other dentals, we should mention one way in which the voiceless supradentals [ʃ, ʈ] differ systematically from the voiced supradentals [dʒ, ɳ, ʈ], namely their relationship to the length of preceding vowels. Without getting into all the exceptions, the general pattern in Swedish is for vowels followed by two or more consonants to be short, while stressed vowels followed by one short consonant (or none) are long. Should we expect the supradentals to behave in this regard as single consonants, or do they reflect their origin as /r/ + dental pairs? Curiously, they do not all behave the same. The voiced

supradentals [d, n, ɲ] regularly follow long vowels, while the voiceless supradentals [s, t] tend to follow short vowels, although not necessarily so if there is a morphological boundary between the /r/ and the /s/ or /t/. Eliasson & La Pelle (1973: 145) handle this situation by positing a rule lengthening vowels before /rd, rn, rl), while Eliasson (1986: 287-8) handles the complexities involving the crucial influence of morpheme boundaries in /rs/ and /rt/ sequences by rule ordering. We will not be concerned with the formulation of such rules here. What is of most importance for us is to recognize the generalization that /rs/ has a complex nature; it often is pronounced as a single phonetic segment [ʃ], but it usually continues to behave phonologically as though it were two consonants because it causes the preceding vowel to be short.

Rules for generating /ʃ/ from various underlying sequences have been proposed as well, but do not have the same solid support in synchronic variability that those for the supradentals have. It is clear that *ʃe* originated historically from palatalization of /s/ that was triggered by the fronting influence exerted by any of various following segments, especially front vowels and /j/, but this is now largely opaque synchronically. Eliasson (1970: 104-127) reviewed the literature on /ʃ/ as a complex segment, examined its nature from a wide variety of angles, and concluded that it is best treated as a single segment. Nonetheless it is true that *ʃe* rarely occurs in consonant clusters (never in native words), it is primarily found in syllable onsets, and when *ʃe* does occur in syllable codas it behaves as though it by itself were a consonant cluster and because it requires the preceding vowel to be short.

1.4 SUMMARY OF ISSUES

Without going into all the specific assertions, proposals, and hypotheses that can be found in those previous studies, let us now briefly review the types of phonological issues that arise and will be looked at in this dissertation.

1. Is there any consistent difference maintained in northern Sweden between the two types of sounds that can be/are pronounced as [ʂ], (1) the underlying sound (phoneme) /ʃ/ (*sje*) and (2) the product of /rs/-sandhi?
2. Is there evidence in that region for the reversal of the neutralization as [ʂ] of the merger between /ʃ/ and the product of /rs/-sandhi?
3. Which lexical and prosodic contexts favor the realization of /ʃ/ as [ʂ]?
4. Which contexts within words favor /rs/-sandhi (their realization as [ʂ])?
5. Which contexts between words favor /rs/-sandhi (their realization as [ʂ])?
6. Which demographic contexts (e.g., age, gender, dialect areas) favor the realization of /ʃ/ and /rs/-sandhi as [ʂ]?
7. Which contexts impede the dropping of the coda consonant /r/ of the copula *är*?

Given the total lack of previous studies on this topic that empirically test specific hypotheses, the major thrust of this study will be toward a description of the state of affairs in the current moment, as well as its development through apparent time. In addition, notice will also be taken of theoretical issues raised, unresolved questions, points deserving further research, and possibilities for the analysis of my data within other current theoretical frameworks such as Optimality Theory.

1.5 SUMMARY OF CONTENTS

In Chapter 2 we will look at the factors that we will use to analyze our data for patterns of variation of all sorts. In Chapter 3 we will look specifically at patterns of variation of *sje* in NN between the normal [ʂ] and the less common variants. Then in Chapter 4 we will

look at the patterns seen for /rs/-sandhi within words, and make a comparison of them with those seen for *sje*, which will test the extent to which the reported neutralization of the two is a reality, and whether there is evidence that the situation for them is changing, as well as investigate the significance of word-internal morphological boundaries on internal sandhi. In Chapter 5 we will look at our last set of data, which consists of contexts that go across word boundaries, to investigate what affects the possibility that external sandhi can apply to /rs/ sequences. In Chapter 6 we will look at the copula *är* 'is/are/am' and the contexts where it drops its coda segment /r/. Finally, in Chapter 7 we will summarize our results and conclusions; the clearest result is that stress is the strongest factor in all the areas we looked at, although it interacts differently with sandhi in different contexts. The first four appendixes contain two copies of each of the two stories that people read for me in northern Sweden to generate my basic data, one copy of running text with the different contexts I looked at underlined and numbered, and the other copy annotated with parts of speech and interlinear translation. The final four appendixes contain summaries of the four sets of data that were used as the basis for Chapters 3-6, namely those for *sje*, for internal sandhi, for external sandhi, and for the copula *är* 'is/are/am'.

CHAPTER 2: FACTORS IN THE STATISTICAL ANALYSIS OF THE DATA

In order to investigate the questions that have been raised above, I will use descriptive statistics to elucidate the distribution of the data. This chapter will seek to explicate the factors that will be used in that descriptive statistical analysis, and why they might be expected to be of interest.

2.0 THE FACTORS

This data has been partitioned according to three demographic factors or speaker variables, and three linguistic factors that describe the items read:

- a. Gender of the speaker (male or female),
- b. Age expressed in decades (e.g., any age from 30-39 is labeled as '3'),
- c. Dialect area (numbered 1-10, plus Uppland),
- d. Identity of the lexical item itself,
- e. Positional relationship between /rs/ or /ʃ/ and main stress,
- f. Syntactic relationship between words containing -/r/ and /s/- .

2.1 DEMOGRAPHIC FACTORS

2.1.0 THE GENDER FACTOR

Lindblad (1982) reported significant differences in the behavior of the sexes in terms of both of the fricatives he studied, sje and tje, especially in Malmberget and Luleå, so the

gender factor needs to be examined. It is also general knowledge in Sweden that the [ʂ] pronunciation of /ʂ/ in positions other than after a stressed vowel is referred to as “*fruntimmers sje*,” a somewhat denigrating way of saying that it has long been perceived in central and southern Sweden as characteristic of feminine speech (Elert 1981: 120). This could perhaps be stretched to suggest that men from NN spoke effeminately in regard to this sound (though they have a very virile reputation in regard to outdoor activities such as moose hunting, for example). Although such attitudes in central Sweden might be conveyed by the media, they are unlikely to have any effect of the pronunciation of /ʂ/ by men in NN unless those men have spent considerable time in southern or central Sweden and been subjected to direct social pressure from teasing or the like.

2.1.1 THE AGE FACTOR

Lindblad (1982) refers to age groups more than once. With regard to dark-*sje* in Malmberget and Luleå, he notes that men use it frequently, women almost never, and boys and girls tend in the same directions as their elders without showing as clear a differentiation (Lindblad 1982: 109-112). With regard to *tje*, the other fricative that he studied, Lindblad (1982: 115) reports that he listened to tapes of people born before 1900, and that those who lived in rural areas at that time used uniformly affricate pronunciations of /tʂ/. As far as the living population sampled for his study was concerned, the affricate pronunciation seemed to be progressively fading away, being replaced almost totally with a pure fricative among the youngest group. Although we will not be looking at *tje* here, this does give us additional reason to anticipate a possibility of changes between different age groups.

2.1.2 THE DIALECT FACTOR

The question of how to divide up NN into dialect areas is complex, and has been solved here in a way that follows Dahlstedt & Ågren (1954) quite closely. Their work shows that the old rural dialects of NN varied in certain ways that might affect /rs/-sandhi, so we should investigate the possibility. It is generally known that dropping of weak final segments or syllables is frequent in NN. Dahlstedt & Ågren describe the dropping of final /-r/, both when it is the present tense ending and when it is the plural marker on nouns, in a complicated pattern that breaks up both of those categories into subcategories, with different conditions applying in different geographical locales.

The arrangement in the list below goes generally from south to north to the extent possible, so that higher numbers are further north, as can be seen on the map in Figure 1 further below (n.b.: Uppland is not in NN, nor is it on the map of NN in Figure 9, and its northern boundary is about 180 miles (300 km) south of the southernmost of the other dialect areas under consideration here).

Dialects:

- U. Uppland (this region is far outside NN, and is at the southern extreme of the very broadest definition of Norrland; data was collected only with MfoN, and is used only in comparisons where each dialect is a separate item, so this data is not included in two-way analyses of other categories, e.g. gender vs. age groups, etc.)
- 1. Ångermanländska mål ([approx. northern half of] Ångermanland dialect)
- 2. Övergångsmålen mellan Ångermanland och Västerbotten (Transitional dialects between Ångermanland and Västerbotten)
- 3. Sydvästerbottniska mål (Southern Västerbotten dialect)
- 4. Nordvästerbottniska mål (Northern Västerbotten dialect)
- 5. Pitemål (Piteå dialect)
- 6. Lulemål (Luleå dialect)
- 7. Kalixmål (Kalix dialect)

8. Nybyggarmål (New settlers' dialect)¹²
9. Torneå valley dialect
10. Mining district dialect

Some towns represented in our data from these dialect areas:

- U. Uppsala, Rimbo, Storvreta, Björklinge, Älvkarleby, Skutskär, Sätra, Tierp
1. Örnsköldsvik, Husum, Gideå, Björna, Själevad, Domsjö, Kroksta, Bredbyn, Högländ, Dorotea, Åsele, Vilhelmina
2. Nordmaling, Hörnefors, Rundvik, Bjurholm, Fredrika
3. Umeå, Bygdeå, Lycksele, Stensele, Sorsele, Storuman, Gargnäs, Vindeln, Vännäs, Sörmjöle
4. Skellefteå, Löfvånger, Burträsk, Byske, Norsjö, Jörn, Malå, Arvidsjaur, Boliden
5. Piteå, Norrfjärden, Älvsbyn, Arnemark
6. Luleå, Råneå, Harads, Boden
7. Kalix, Töre, Överkalix, Morjärv
8. Arjeplog, Porjus, Jokkmokk¹³
9. Haparanda, Övertorneå, Pajala
10. Malmberget, Gällivare, Kiruna

A map depicting the boundaries of these dialect areas may be seen in Figure 9 below, along with some of the towns represented by speakers that recorded my stories. (The Arctic Circle crosses just south of Jokkmokk, and near the southern boundary of DA-10.)

¹² Here 'new' means after about 1600 AD. Jokkmokk parish was established in 1607 and Arjeplog parish in 1641 (Dahlstedt & Ågren 1954: 287).

¹³ According to the map in Dahlstedt & Ågren (1954: 230), Jokkmokk belongs to dialect area #8, but they grouped the sample text from Jokkmokk together with #6 Luleå for reasons that are unclear to me. In this dissertation all data from Jokkmokk have been assigned to dialect area #8.

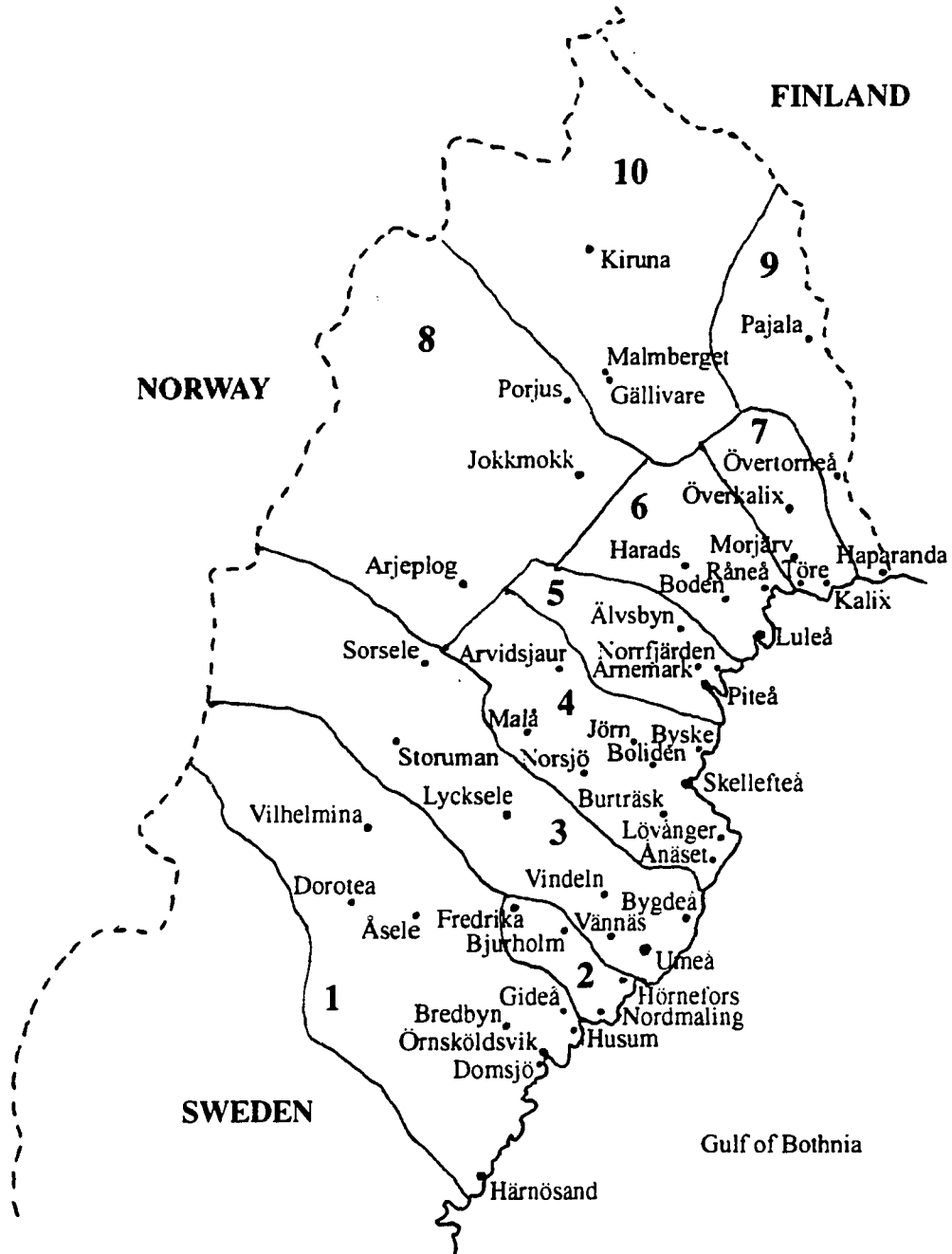


Figure 9. Map of North Norrland dialect areas

2.2 LINGUISTIC FACTORS

2.2.0 THE LEXICAL FACTOR

Since some sound changes have been reported to proceed at different speeds for different lexical items, we need to examine individual words to see if there is evidence for such a pattern, or whether these sound changes are operating smoothly and equally across the lexicon in a Neogrammarian fashion (cf. Labov 1994, Part D).

2.2.1 THE STRESS FACTOR

Standard analyses, as in *SAOB*, distinguish four degrees of stress in Swedish. However, as discussed in the introduction, we will collapse the distinctions between the two tone accents, so that primary stress with either accent will be considered primary, and secondary stress with either accent will be considered secondary. This still leaves us with three possibilities for the syllable before an /rs/-sandhi context (primary / secondary / unstressed) and the same three possibilities for the syllable after the sandhi context. The permutations that arise from the possible relationships between these stresses on the two adjacent syllables create nine possible stress contexts for /rs/-sandhi.

Lindblad (1982: 103), citing two earlier studies of the patterns of pronunciation of *sje* in Swedish (Sellberg et al. 1975 and Hammermo 1976) and his own preliminary study, states that it suffices to distinguish two positions, namely (my translation, with underlining of *sje* and stress marks on vowels added):

- “1. Position alone or in a consonant group immediately following a stressed vowel
(dúsch, lúnch, hássj^á, kánsk^è)
2. Other positions, with the main categories
 - a. Position before a stressed vowel (sj^ó, st^{at}ión)

- b. Position between two unstressed vowels and initially before an unstressed vowel (traditionéll, regissórr, sjalétt)”

We will use this schema as a basis for formulating hypotheses, but when we examine the data we will find reasons to refine these definitions of the two positions.

2.2.2 *THE SYNTACTIC FACTOR*

Kaisse 1985 found evidence for syntactic relationships as a crucial factor in determining the applicability of sandhi in some languages, and we will examine the possibilities for this being the case for external /rs/-sandhi in Swedish as well.

2.3 THE CONTEXTS

In order to understand the background behind some of these factors, it may be helpful to go into some detail about the general contexts of place and time in which the persons have lived who have participated in this research by reading my two stories, and a little about how the dialects in NN developed historically.

2.3.0 *GEOGRAPHICAL CONTEXT: NORTH NORRLAND*

The region of northern Sweden under consideration here, roughly speaking, is the part of Sweden from Örnsköldsvik north, and is slightly larger at its southernmost extent than the diocese of Luleå (founded 1904). It is essentially coterminous with Sweden's two northernmost counties, Norrbottens län and Västerbottens län, plus Örnsköldsvik kommun, the northernmost municipality in Västernorrlands län. The total population of this region on January 1, 1990, was some 572,000 people (6.7% of Sweden's total population of 8,526,000) (Öberg & Springfieldt 1991: 15), in an area covering about 160,730 square kilometers, or 39% of the area of Sweden and extending from approximately 63° N up to 69° N latitude (a little north of the Arctic Circle) (Castensson 1992: 112-126).

Until the diocese of Härnösand was created in 1772, the diocese of the Bishop of Uppsala was responsible for all of Sweden from Uppland north (i.e. most of Sweden north of a line running northwest from Stockholm)(Castensson 1992: 145). Thus there is some basis for suggesting that in the popular mind Norrland still starts at the northern outskirts of Stockholm. By this old definition, over 60% of the area of Sweden would be included in Norrland (though only about 17% of Sweden's population). However the area that concerns us is smaller, being approximately the northernmost 40% of Sweden. Dahlstedt & Ågren (1954) termed this region of interest to us Övre Norrland, and that is also the term used by Elert & Fries (1982), but I have instead chosen to refer to it as North Norrland (NN).

For me the distinguishing characteristic of NN is that it is the area where the standard realization for /ʃ/ is [ʂ], regardless of its phonotactic position within the syllable. I have defined the bounds of this region that is the focus of my investigations based on my interpretation of a map of the variants of /ʃ/ (Lindblad 1980: 141), and intend for it to include the entire area where the normal realization of /ʃ/ is [ʂ]. I am not aware of any careful survey of the southern boundary of NN, as defined here. Indeed, since /ʃ/ is also frequently realized as [ʂ] south of this 'boundary' (especially in the coda of syllables), it may be impossible to define its location precisely, in a manner similar to that in which Sjöstedt (1936) and Elert (1981) were able to establish rather precisely where the northern boundary for back /ɪ/ runs across southern Sweden. I have collected data primarily (except for a small sample in Uppland) within the bounds of NN as defined by a small-scale map (Lindblad 1980:141). That definition of NN very closely approximates the definition of *Övre Norrland* as used by Dahlstedt & Ågren (1954), so our NN is essentially identical to the region that they were describing, except that here we have also included the areas further north in Sweden where Finnish and Sami formerly were the dominant languages. Further south in southern Norrland and central Sweden [ʂ] also emerges as the realization of /ʃ/, generally following stressed vowels, but in those areas it

is only one of many competing variants with widely disparate articulations (though somewhat more similar acoustic features), and [ʂ] has a different phonological status there, being found in many people's speech only as the product of rs-sandhi, or if it does occur as an allophone of /ʃ/ it is commonly limited to codas.

Another characteristic of NN is a stubborn local pride of place and dialect, which finds expression, for example, in a phrase that forms part of a chapter title in Elert & Fries (1982), "Vi flytt int" ('We won't leave'). This differs from standard Swedish "Vi flyttar inte" through the dropping of weak final syllables where the preceding syllable is long, widely recognized as being characteristic of many of the traditional dialects of NN (cf. Dahlstedt & Ågren 1954: 237-50, Pamp 1978: 30). A sociolinguistic study of the strong emotional ties to local dialect in a nearby part of northern Norway was conducted by Blom & Gumperz (1972).

The remoteness and isolation of this region before the development of the modern communications network contributed to the evolution of highly divergent dialects. As an example we may take the Kalix valley, which was the furthest to the north and east of the areas to speak Swedish in the pre-modern era (since to its east the adjacent Torne valley, through the center of which the current Swedish-Finnish border runs, was traditionally Finnish-speaking). Not only is the upriver Överkalix dialect considered one of the three most highly divergent dialects of Swedish¹⁴ (Dahlstedt & Ågren 1954: 218), but also it is reported (Pamp 1978: 142) that dialect speakers from Kalix cannot even readily understand speakers of the upriver Överkalix dialect, even though their two main population centers and parish churches are only about 50 miles (80 km) apart.¹⁵ Despite

¹⁴ The others are Älvdals dialect in Dalarna and certain dialects on the island of Gotland.

¹⁵ Although the Kalix dialect is regarded as closer to standard Swedish, in my personal experience, the only person in NN that I attempted (with only limited success, even with help from his family) to interview and have read my story, but whom I essentially could not understand and who was not able to switch dialects to a variant of standard Swedish, was an old farmer in his 70's who lived out in the country a few miles from Kalix.

the known differences between Kalix and Överkalix dialects, Dahlstedt & Ågren (1954) treat the two together as one dialect area, and we will follow their lead (it is our dialect area 7). Although perhaps the effects further south were somewhat less dramatic, still geographical isolation (and the harsh winter climate) of NN has played a role in the historical development of rather distinct dialects throughout the area. Whether all these dialects in NN are distinct in regards to the various phonological factors that we are interested in is something that we must examine.

Of course severe winter storms force people to seek shelter to survive, but even in the Arctic, normal winter weather need not be a bar to travel. Since prehistoric times, many people (especially men) in NN and nearby parts of present-day Norway, Finland and Russia have done a certain amount of travel as a regular part of their way of life. Winter could put a hard surface of ice on rivers and marshes, and thereby facilitate some kinds of travel, trade, hunting and trapping. The traditional life of many Sami included seasonal migrations with their reindeer up to the mountains in summer, then back down to the shelter of lowland forests in the winter. In the second half of the 19th century (after steam-driven saws came onto the scene beginning in 1848) lumber mills were built at the mouths of all the major rivers of NN (Scott 1977: 447), and it became common for men to travel inland in the winter to cut down trees and drag them to the edge of rivers or even onto the ice of rivers, then float the logs downstream in the spring. This logging and floating activity probably had a reinforcing effect on the tendency of early modern dialects to be coterminous with river valleys, a tendency whose roots may plausibly be traced back both to the difficulties of travel between valleys and to the church's earlier assignment of parish boundaries with geographical limits that frequently corresponded to valleys. Nonetheless, even if some people may have traveled hundreds of miles annually in pursuit of these various activities, it still occurred within the bounds of northern Scandinavia, so that exposure to other dialects was largely limited to neighboring ones.

2.3.1 HISTORICAL CONTEXT: MINING

Toward the end of the 19th century, major iron ore deposits at Malmberget and Kiruna finally began to be worked seriously. These deposits had been discovered much earlier in the 18th century, but could not be economically exploited when the only way to transport the ore out from the Arctic interior to the coast was by reindeer sleighs. This transportation problem was partially solved when the railroad from Luleå to Gällivare was completed in 1888, and extended to Kiruna a few years later. However, because the northern Gulf of Bothnia and the port at Luleå are often frozen over for several months in the winter, thereby blocking water-borne shipping, in 1903 a railroad link was completed across the mountains to the Norwegian port of Narvik, which is kept open all winter by the relatively warm water of the north Atlantic, so that year-round shipments became possible (Lundholm et al. 1996: 78). All of these activities led to the opening up of the area to outside influences.

What effects might this history be expected to have on the dialects and languages of NN? The opening of the iron ore mines in Kiruna and Malmberget, and the attendant inflow of many miners from areas far away in south and central Sweden, could potentially introduce a significant element of dialect features from outside NN into this area above the Arctic Circle where Sami had previously been the main language, and Swedish had never established itself. Lindblad (1982: 113) argues that in essence this is what occurred in the mining localities (which will be our dialect area #10) in regard to the darker variants of *sje*, though its influence is seen primarily in the speech of males, because immigrants from outside NN were almost exclusively male, and because the women who settled in these towns came predominantly from within NN. Once the large-scale immigration of males from outside NN ceased, this feature of male speech could be expected to gradually fade away in the direction of the norm of the surrounding dialects. Since there is a reported merger in NN of phonemic /s/ with [ʂ], the phonological product of rs-sandhi, it would be unexpected and noteworthy to find that dark-*sje* pronunciations

are spreading, even in the context of the immigration of a considerable number of men from other dialect areas, because this could imply the reversal of a reported merger.

2.3.2 HISTORICAL CONTEXT: RADIO

The first radio broadcast in NN came from Boden in July 1921 in connection with the city of Luleå's 300th anniversary.¹⁶ Interest in radios was widespread in NN in the mid-1920's, and ownership of the expensive new apparatus conferred considerable prestige. There are reports of solemn gatherings of dozens of people in private homes in small towns to listen to broadcasts of church services in that era, which is understandable in view of the great difficulties then of traveling long distances to the parish church. The national telegraph management's proposals for siting of broadcast stations were debated in the national press in 1924, and eventually they were persuaded to add one station in NN, in Boden. Private radio clubs formed in many towns in NN, and the clubs in Umeå, Kiruna, Malmberget, and Örnsköldsvik built stations in 1925-27, with all these transmitters except Kiruna's being taken over by the state in 1935. Since most of the programs came from the central, national level, it became unusual to hear voices speaking local dialects on the radio. This norm created such strong expectations that some letters to the editor from that era express the shame some people felt at hearing their own dialect spoken on the radio. As a final blow to the use of local dialects on the radio, in 1940 security concerns led to the elimination of all local programs for the duration of World War II, except for farming programs and certain important messages.

This history suggests a clear opportunity for standard Swedish to make a breakthrough in the direction of replacing local dialects in NN via the modern medium of radio, beginning on a small scale in the 1920's and reaching greatest intensity during World War II. Other complementary developments during the same period may also have contributed to this same trend of changes in language norms to greater or lesser degrees, including the

¹⁶ Information for this paragraph is based on Nordberg 1995, which draws on a dissertation being written at Umeå University about radio and adult education .

practice of punishing children for speaking their dialect in school, a rise in the standard of general education, and the advent of “talkies”, i.e. movies with soundtracks (p.c., Ola Wennstedt 1998).

When Inga-Britt Lindblad (1985) compared news broadcasts of Radio Västerbotten with those of two other local radio stations in central Sweden, Radio Gotland and Radio Östergötland, she found that Radio Västerbotten had less regional accent in speech, and that its news language closely resembles national broadcasting language. Thus it would seem that the tendency toward avoiding local dialects on the radio in NN that was already established by the time of World War II has been continued up to the present. It is clear that the phonological systems of speakers in NN continue to be under considerable pressure from standard Swedish, which many of them now hear every day on the broadcast media. Nevertheless, Labov (1994) and others have shown that phonological developments in this era of modern broadcast media often are in a direction away from the national norm, as has been seen for example with the various changes in vowel quality in parts of the northeastern USA known collectively as the Northern Cities Shift. Thus there is no necessity for the merger in NN of the phoneme *sje* with [š] (the output of *rs-sandhi*) to be reversed just because that would bring the realization of the phoneme /š/ in the dialects of NN closer to its (widely varying) realization in the standard Swedish heard in the media (and in a plurality of central Swedish dialects) in this one feature. Similarly, the effect of the orthography, which is based on a conservative norm, is not enough to account for the reversal of the merger, although it certainly could help facilitate it in a society that supports a high standard of universal education. The reversal of mergers in dialects is never predicted or expected, regardless of national norms.

2.3.3 HISTORICAL CONTEXT: WORLD WAR II

It could reasonably be argued that all of these factors coalesced to bring an unusual degree of outside pressure to bear on the linguistic environment in NN during World War II. Although Sweden was officially neutral, major actions in the war took place nearby to

NN. For instance, the USSR's army first entered Finland at Petsamo in the Arctic and began the Winter War on November 29, 1939 (Wuorinen 1965: 349). During the war, some 70,000 Finnish children were sent to Sweden for their safety (Scott 1977: 507), and older residents of NN have spontaneously mentioned those Finnish children to me more than once, suggesting that it must have been a monumental disruption of everyday life at the time it occurred if it was still such a noteworthy event in their lives more than 50 years later. Because of its key place in the export of the iron ore from Kiruna and Malmberget, Narvik became the site of a major battle between the Allied forces and the Germans in late May 1940 (Larsen 1948: 544). Sweden allowed certain German supply and troop trains to cross northern Sweden during the war, including the highly controversial movement of the entire German Engelbrecht Division from Norway to Finland in June 1941 (Scott 1977: 506). In 1944 when the German troops retreated northward through northern Finland, their ravaging tactics caused 55,000 Finns to flee across the border into NN, bringing their cattle and belongings with them (Andersson 1970: 453). After heavy fighting, the last German troops were finally ousted from northern Finland on April 27, 1945 (Wuorinen 1965: 350). No doubt all this was sensational international news with an unmistakable local angle for most people in NN during WWII. Interest must have been high, with many people regularly listening attentively to authoritative voices on the radio speaking in outside dialects about these momentous events. It is hard to prove exactly how much this affected the dialect features of young people throughout NN at the time, but we will see suggestive evidence below that something significant occurred, and what we know about openness of children to outside influences during the various developmental stages of language acquisition will suggest that it was at its maximum during WWII.

In this chapter we have looked at the factors that can be separated out for a statistical analysis of our data. In Chapter 3 we will take one segment of our data, the words containing the phoneme *sje*, and look at them from various angles. Then in Chapter 4 we will look in a similar way at words containing the underlying sequence */rs/*. By

comparing the two, we will address the question of the extent to which the underlying distinction between the two phonological sources of phonetic [ʂ] is neutralized in the dialects of NN, and whether that neutralization is being reversed.

CHAPTER 3: UNDERLYING *SJE*3.0 SOME QUESTIONS ABOUT *SJE* IN NN

In this chapter we will focus on the varieties of *sje* in the geographical context of North Norrland. Two major contributions to describing *sje*, a historical and phonological discussion by Stig Eliasson (1970: 104-127) and an extensive analysis of its phonetic variants by Per Lindblad (1980), deal primarily with central standard Swedish and Swedish as a whole, and have little to say specifically about the dialects of NN. The only study to look at *sje* in a geographical area close to our definition of NN is by Per Lindblad (1982), who described the variation of the two Swedish phonemes *sje* and *tje* that he found in two sets of data collected in six towns in northern Sweden from 1973-76. However, as a matter of scientific methodology, he did not formulate explicit hypotheses prior to examining the data, so the generalizations that he extracts from the data that he examined cannot be considered as proof of any hypothesis, but only as providing guideposts for us to keep in mind when we attempt to formulate plausible hypotheses on this topic.

Here we will formulate a series of such hypotheses based on the results in Lindblad (1982), and then test those hypotheses to the extent possible by using two new sets of data that I collected for other purposes in 1990-91 and 1996 in all of those six towns, plus many more towns between and around them in a slightly expanded geographical area. Our more inclusive geographical and age-group coverage allows us a better overview of the whole regional situation than his data from children and their parents in six towns was able to achieve, although our new data does not include nearly as many speakers from those two generations in those six towns. We are able to further expand on his

description of the geographical distribution of dark-*sje*,¹⁷ thereby casting some doubt on his account of its origin. Also, this new data on the distribution of dark-*sje* raises questions about the standard assumption that phonological mergers are not reversible (cf. Labov 1994: 311). In addition, our more inclusive age group coverage has allowed us to discover an interesting anomaly in the trends of usage of *sje* across some of the age groups of older speakers that was not previously hypothesized or reported, and that may become untestable if not further investigated soon. Furthermore, an examination of individual lexical items in light of their stress patterns allows us to refine the stress-based generalizations of Lindblad (1982), with implications for the analysis of lexical stress in Swedish.

3.1 NEUTRALIZATION OF *SJE* IN NN

Our concern here is with *sje* in the north of Sweden, where the situation is different from that in central Sweden in several respects. First of all, this geographical region appears to be the locus of a neutralization between /*j*/ and [ʂ]; i.e. it is the area where the normal realization of underlying *sje* is [ʂ], even though here as in central Sweden [ʂ] may also be the product of /*rs*/-sandhi (Lindblad 1980: 141). Some of the other darker variants of *sje*, primarily [ʌ, ʃ], have also been reported by Lindblad (1982) as occurring in this northern region as less frequent variants of /*j*/, and we will look at the contexts in which they occur in our data and attempt to discern the factors that govern their provenance.

From the point of view of a speaker of central standard Swedish, there is a neutralization in NN of a distinction that is made by many other Swedes. Historically, however, the pronunciation of *sje* as [ʂ] may be the earlier one, so that central and southern Swedish dialects have innovated in the creation of their various darker realizations for *sje*. Thus,

¹⁷Recall that 'dark' fricatives are ones that give an auditory impression of being lower because the sharp cut-off of their strong high frequency energy occurs at a relatively lower frequency (Lindblad 1980: 193).

when we describe the pronunciation of *sjē* as [ʂ] instead of [ʌ, fj] as being a merger with the pronunciation of the product of /rs/-sandhi as [ʂ], we must keep in mind the possible historical sequences. In this scenario, before the development of /rs/-sandhi in the phonology of Swedish, the /rs/ phoneme sequence was distinct from the /ʃ/ phoneme because they were pronounced [rs] and [ʂ] respectively, and the two had a tendency to phonetic merger only after the sandhi process was introduced. In all dialects of Swedish with apical /r/, the introduction of /rs/-sandhi into the phonology produced the possibility of the confusion of /rs/ with /ʃ/ and their phonetic merger (or near-merger?). In central Sweden when the introduction of /rs/-sandhi produced the possibility of confusion with /ʃ/, this may have been the impetus for the drift toward the current situation where the phoneme /ʃ/ is pronounced in the many different and widely divergent ways that Lindblad (1980) describes. On the other hand, in that scenario in NN the introduction of /rs/-sandhi did not lead to any change in the phonetic realization of /ʃ/. For those who would argue that languages avoid phonological mergers to prevent confusion between homophones, a question that remains unanswered is why central Swedish resisted the phonetic merger of /ʃ/ and /rs/ by finding another place of articulation for /ʃ/, while the dialects of NN did not? Also if avoiding merger was the motivation, why did southern Swedish change its place of articulation for /ʃ/ even though those dialects with their uvular pronunciation of /r/ do not produce [ʂ] from /rs/-sandhi? Since there are so few possibilities for confusion between homophones due to their relatively low frequency of occurrence and the different phonotactic distribution of /ʃ/ and /rs/, a more plausible answer to these questions might draw a connection between the historical tendency of /ʃ/ to drift toward the back of the mouth in southern Swedish and the known historical introduction of uvular articulations of /r/ from the continent into southern Sweden, which could be argued to have resulted in a generally greater tendency to pronounce sounds in the back of the mouth. Another plausible reason for the backward drift of /ʃ/ could be

pressure exerted by the change of *tje* from the affricate [tʃ] to a fricative [ç] that is articulated between /s/ and /ʃ/, a change that began in southern Sweden earlier than further north (p.c. Per Lindblad, 1999).

In any case, for those speakers in NN for whom the pronunciation of /ʃ/ has gone to some other darker pronunciation such as [fj], this is the reversal of a merger. However, this is a reversal of a progressive kind, with /rs/ having taken over the pronunciation [ʃ] as its own, and /ʃ/ going elsewhere, rather than being a case where the two returned to their earlier values. If this can indeed be properly described as a case of a merger being reversed, it is of interest for the study of this rarely found and much disputed linguistic phenomenon (cf. Labov 1994, Part C). Looking at this sound change in progress in NN may also contribute to our understanding of how changes in phonemes occur, whether in a uniform Neogrammarian fashion, or one at a time as individual changes in the underlying representation of each lexical item; the variation that I found between words with similar morphological and stress patterns suggests the possibility that individual lexical items may have unique histories.

3.2 PREVIOUS STUDIES OF *SJE* IN NN

Throughout his dissertation, Per Lindblad discusses in great detail the remarkable allophony of the /ʃ/ phoneme in Swedish (summarized by Lindblad 1980: 11; 1982: 100). In his map of its regional variation in Sweden (Lindblad 1980: 141), the northern two-fifths of the country (our NN) is shown as an area where this underlying sound is realized with the same retroflex 's' sound [ʃ] that is found in most central and northern Swedish dialects as the sandhi product of /rs/ combinations. However, in his contribution to *Nordsvenska* (Elert & Fries 1982), a collection of studies that deal in particular with speakers from six towns in this region (Umeå, Lycksele, Skellefteå, Arvidsjaur, Luleå, Malmberget/Gällivare), Lindblad (1982) observed that to a certain extent the normal

northern variant [ʂ] is replaced by other darker variants of /ʂ/, such as are more commonly found in southern and central Sweden. He found this to occur principally in Malmberget and Luleå, and to a lesser extent in Umeå, and principally among the males in these towns. Lindblad (1982: 113) suggested that this is due to the influence of the dialects of workers who immigrated into these areas from southern Sweden, especially of men to work in the mines at Malmberget and Kiruna. Although this influx of miners began about a century ago, he suggested that the use of darker variants of /ʂ/ was still perceived as a male speech trait in the language of the present generation of native-born Norrlanders.

The corpus of dialect material that forms the basis for the majority of the studies in *Nordsvenska* is based on dialect samples from 72 primary informants. These 72 people were evenly apportioned over the six towns, with 3 men, 3 women, 3 boys and 3 girls from each. The majority of the adults were born in the 1930's, with the oldest born in 1923 and the youngest born in 1944. The children were all born between 1954 and 1962 (Lindblad 1982: 117-8). Most of those people had lived in one (the same one) of those six towns all their lives, and many of their parents also lived their entire lives in that same town.¹⁸

In addition to this main corpus of dialect material collected for *Nordsvenska* in 1973-1976 by others and utilized by many of the contributors to that volume for their studies of various topics, Lindblad (1982) also made use of special additional data that he himself had collected in Autumn 1976, using as primary informants 88 seventh grade children, with between 10 and 20 of them being from each of those same six towns. All of these children had lived in their hometown since before the age of three, a somewhat looser

¹⁸ This is a higher standard of dialectal purity than I was able to apply when selecting informants to interview, since I had a very limited time to find them, and since these days so many people move around to go to school, to serve in the military, to find jobs, etc., so that such sedentary people probably no longer are typical in NN.

residency criterion (and therefore perhaps more comparable to criteria for the people who contributed to my database).

The main corpus collected data of five types from each of the 72 people (Lindblad 1982: 101)(my translation):

- “A. Conversation between Åke Jonsson and (normally) two informants lasting ca. 30 minutes.
- B. Conversation between two informants lasting ca. 10 minutes.
- C. Reading of two short texts.
- D. Reading of a list of ca. 90 words and ca. 35 phrases.¹⁹
- E. Responses to a series of questions.”

In his recordings of the 88 seventh grade children from those six towns, Lindblad (1982: 102) collected the following five types of data (my translation):

- “A. Answers to several questions, where words with *sje*-sounds are certain to be a part of the answer.
- B. Reading of a short text with relatively many *sje*-sounds, of which several in a less prominent position than in the rest of the material.
- C. Naming of ca. 10 pictures of objects whose names include *sje*-sounds.
- D. Reading of a list of words with some 80 *sje*-sound words, 10 *tje*-sound words and a few words with initial /s/.²⁰
- E. Isolated, drawn-out pronunciations of /s/, *sje* and *tje* (simultaneously with which the majority of the students from Umeå, Lycksele, Skellefteå and Luleå were photographed from the front.)”

¹⁹ None of the specific *sje* words included in C. and D. of the main corpus occur in my data.

²⁰ The words *mánniskà* (person) and *kánskè* (maybe) occur both in this list and in my story Msi.

Lindblad (1982: 111-113) examined in detail the stylistic differences in *sje*-usage between conversation and text-reading in both the main and his special materials, with a further distinction made in his special materials between text- and list-reading. He judged that the two materials were not totally comparable because the circumstances of the recordings of his special materials were less casual in several ways and had therefore called forth a greater tendency towards the [ʂ] pronunciation of *sje* (in Position 2, as defined below). However, within each material there was a clear tendency for more dark-*sje* usage to occur in conversation than in text-reading, and in his special material the lowest percentages of dark-*sje* usage were found in list-reading. Thus he found a clear stylistic formality hierarchy: list-reading > text-reading > conversation > emotional language. This accords with findings made elsewhere regarding stylistic formality hierarchies (cf. Labov 1972: 79-109; Trudgill 1974: 110-115). Lindblad (1982: 111) attributes this greater avoidance of dark-*sje* in more formal contexts to a perception of [ʂ] as having “a status of ‘finer’ or ‘more correct’ pronunciation than [ɫ, ʃ] in these dialects, except for males in Malmberget.” Since the main body of new data introduced in this study is derived solely from readings of texts, we can expect the frequency of dark-*sje* usage to be lower than what would occur in casual conversation, but higher than in list-reading.

3.3 THE NEW DATA

To test the hypotheses that Lindblad (1982) put forward to account for his data, I will analyze a new corpus of data that is based on two stories that I wrote, and then asked several hundred people in North Norrland to read during 1990-91 and 1996. These readings were recorded using a Marantz Model PMD 201 Portable Cassette Recorder and an Electro-Voice Model RE50 Omnidirectional Dynamic Microphone. Because my stories were not constructed with the specific intention of collecting data on *sje*, only a small number of examples of words with this sound occurred in the two texts. From one

of the stories, 'Min familj och Norrland' (My family and Norrland)(MfoN), I have extracted data from 232 recordings for two *sje*-words, *själv* (self) and *flygsj^uk* (airsick).²¹ From the other story, 'Min styvfars idol' (My stepfather's idol)(Msi), I have extracted data from 154 recordings for nine *sje*-words: *skölsk^öterska* (school nurse), *människ^ör* (people), *känsk^e* (maybe), *sköldig* (liable), *människ^a* (person), *själv^a* (self), *skicklig* (skillful), *skärpsk^yttarna* (sharpshooters-the), *vins^{ch}* (winch).

I have entered my data on the pronunciations of these words onto Microsoft Excel spreadsheets, and then analyzed this data using the Varbrul program for statistical analysis of linguistic data (originally developed by David Sankoff and others). Some of that output from Varbrul has in turn been entered onto Excel spreadsheets, and then transformed into graphs for easier visualization of the resulting patterns, as may be seen further below.

In order to analyze data with Varbrul, the data must be annotated and assigned a value for each of the factors to be used in the analysis. Each of the factors used here for the analysis of *sje* has already been discussed in separate sections above in Chapter 2, specifically: §2.1.0 Gender, §2.1.1 Age, §2.1.2 Dialect, §2.2.0 Lexical Items, and §2.2.1 Stress. However, the stress factor as it applies specifically to *sje* is discussed in some detail by Lindblad (1982), and requires further elucidation here.

3.4 THE STRESS FACTOR AND *SJE*

Lindblad (1982: 103), citing two earlier studies of the patterns of pronunciation of *sje* in Swedish (Sellberg et al. 1975, Hammermo 1976) and his own preliminary study, states that it suffices to distinguish two positions, namely (my translation, with underlining of *sje* and stress marks on vowels added):

²¹ Letters pronounced as *sje* are underlined.

“1. Position alone or in a consonant group immediately following a stressed vowel (dúsch, lúnch, hássjà, kánskè)

2. Other positions, with the main categories

a. Position before a stressed vowel (sjö, statión)

b. Position between two unstressed vowels and initially before an unstressed vowel (traditionéll, regissór, sjalétt)”

Lindblad (1982: 104) reports that in his data, in Position 1 speakers from NN almost uniformly use [ʃ] to realize /ʃ/, while in Position 2 there is a lot of variation, and the darker variants of /ʃ/ alternate with [ʂ]. It is not clear how his description of the two contrasting positions of *sje* can best be expressed in standard phonological terms. What the ‘other positions’ subsumed under Position 2 seem to have in common is that *sje* appears in the onset of a syllable, whether that syllable is stressed or not. However it is not so obvious how to analyze his examples of Position 1, since the first two examples have *sje* unambiguously in the syllable coda, but the last two have *sje* in between two vowels and the syllabification is not transparent.²² Note here that these are the two possibilities for the syllabification of the consonant sequence /rs/ as well: it may either be a coda or it may be split between two syllables, but it can never be an onset. This suggests the possibility of an analysis where [ʂ] is taken to be /rs/ when phonotactically plausible, and /ʃ/ otherwise. This closely resembles descriptions of the so-called Göta-rule for dialects of Swedish that have both apical and uvular /r/, whereby uvular /r/ is found only in places where /r/ cannot participate in sandhi in any dialect (cf. Elert 1981: 20).

While the primary stress falls on the first vowel in both *hássjà* and *kánskè*, both of these are Accent 2 words, which implies (at least some degree of) secondary stress on the second vowel. Some analyses would assign intervocalic consonants preferentially to

²² See Sigurd (1965: 166-174) for a review of the literature and discussion of Swedish syllabification.

onsets rather than codas on the basis of the cross-linguistic primacy of CV syllables (found in all languages) over CVC and VC syllables (not allowed in some languages), and the (probably related) fact that some languages require all syllables to have onsets (cf. Clements & Keyser 1983: 29), or for various other reasons. However, another possibility is for syllabification to follow the strategy of maximizing the number of consonants assigned to the syllable with primary stress (cf. Sigurd 1965: 166), which would bring the *sje* of both of these words into the coda of the first syllable (or at least make it ambisyllabic). In that case, Position 1 = coda and Position 2 = onset.

However, this dichotomous analysis does not cover all the cases in Lindblad's data, because he reported the exceptional appearance of dark-*sje* in the words *särskilt* ('special') (occurred one time) and *úsch* ('ugh') (5 times) (Lindblad 1982: 104). Although it would be possible to squeeze the *sje* of *särskilt* into either syllable depending on your analysis of the interacting effects of syllable structure and stress in Swedish, there is no ambiguity about *úsch*, where the *sje* can only be in the syllable coda. Lindblad (1982: 114) accounts for *úsch* by assigning it to a special category of emotionally laden words, together with *ásch* ('ugh!'), *sjúttön* ('17')²³, and *skít* ('shit'), which were often pronounced with dark-*sje* by speakers who otherwise had [s̥]. This claim seems plausible, since Labov (1994: 537) cites Jespersen's (1949: 308-12) example of "affective words (ah!, haha, mama)" retaining back vowels where most other words do not, which appears to be a comparable instance of emotions having an influence on the choice of phonetic features.

²³ It should be possible to test whether there is a difference in the variants of *sje* that are used for the word *sjutton* when it is used as purely the number '17' versus when it is used as a swearword. This could be an interesting way of producing a minimal pair where the only variable is pure affect, with no other phonetic differences between the words. Incidentally, I once saw a Volvo in Seattle with the license plate "XVII", which self-evidently had gotten past the censors at the Department of Motor Vehicles.

Even though the relatively high frequency of dark-*sje* found in Malmberget and Luleå might seem to make them similar to central Swedish dialects, the absence of dark-*sje* in Position 1 is distinctively different. This pattern was relatively infrequent in Hammermo's (1976) data from Eskilstuna, and apparently marks even the men of Malmberget, who are otherwise anomalously (within the context of NN) prone to using dark-*sje*, as belonging to the NN dialect area (Lindblad 1982: 104).

3.5 HYPOTHESES TO BE TESTED RE *SJE*

Based on the observations and generalizations in Lindblad (1982), it is possible to formulate a series of hypotheses regarding the distribution of the variant forms of *sje* (/ʃ/) in North Norrland. We can then go on to test these hypotheses using data that I collected for other purposes in the same region in 1990-91 and 1996.

Darker variants of /ʃ/ have replaced the former norm in NN, the retroflex [ʂ], only in certain cases where it has been imported from southern/central Sweden, especially by men in the mining industry in Malmberget and Kiruna (DA-10), and to a slightly lower degree downriver in their port city Luleå (DA-6); he also found it to a lesser extent in the university city of Umeå (DA-3). This can be reduced to two frequency hierarchies:

Hypothesis A

Men > women.

Hypothesis B

Malmberget/Kiruna > Luleå > Umeå > elsewhere.

In Malmberget (DA-10) and Luleå (DA-6), Lindblad (1982) was able to define a third, more detailed hierarchy, which was darkest at the adult male end in Malmberget and

lightest at the adult female end in Luleå. He interpreted this pattern as indicating a progressive breakdown in the younger generation of an earlier clear gender-based pattern:

Hypothesis C

Men > boys > girls > women (in Malmberget and Luleå only).

In all six towns Lindblad (1982) found a nearly exceptionless pattern, whereby the possibility of realizing *sje* with a darker variant was correlated to its positional relationship to the word stress:

Hypothesis D

/ʃ/ is always realized as retroflex [ʂ] immediately following the stressed vowel, whether alone or in a consonant cluster [with minor exceptions for two affective words that do not occur in our data].

Hypothesis E

Retroflex [ʂ] can alternate with darker variants of /ʃ/ elsewhere.

Lindblad (1982) also found evidence that more casual speech situations encourage greater usage of dark-*sje* in those phonological environments where it is free to occur, and that more formal situations inhibit it, which leads to a further hypothesized hierarchy:

Hypothesis F

Emotional language > conversation > text-reading > list-reading.

Other possibilities for describing the patterns of change must also be kept in mind, such as whether it is possible to describe either of them as Neogrammarian spread or as spread via lexical diffusion. Since Lindblad (1982) suggests no particular lexical items as behaving idiosyncratically (except for one occurrence of dark-*sje* in *särskilt*, which is

susceptible to prosodic reanalysis), it would essentially be an argument *ex nihilo* to interpret his study as supporting a claim either for or against Neogrammarian spread of dark-*sje*. I can see no basis in his research on these particular problems in NN for formulating a hypothesis on the typology of spreading prior to analyzing our data. Therefore we will use our data on this aspect of the situation only to form hypotheses, and future research will then be required in order to substantiate or disprove those hypotheses.

3.6 TABLES AND GRAPHS OF RESULTS

The data displayed in the following tables and their derivative graphic representations confirm many of these hypotheses, but show the need for refinement of others. The first three sets of comparisons in this chapter will involve the pairing of the gender of the speakers with other categories. Hypothesis A (Males > Females) and Hypothesis B (Malmberget (10) > Luleå (6) > Umeå (3) > elsewhere) can be considered in light of Table 2 and Figure 10, where we examine the gender of the speakers as it relates to the dialect areas.

Table 2. Dark-*sje* pronunciation by dialect and gender

Dialect Area	Male		Female	
U	14	100%	10	83%
1	26	14%	12	11%
2	12	18%	4	17%
3	24	16%	1	1%
4	13	7%	3	2%
5	15	12%	1	1%
6	32	55%	0	0%
7	29	35%	4	6%
8	22	34%	2	10%
9	2	6%	6	9%
10	28	39%	1	2%
Overall Totals	217	21%	44	6%

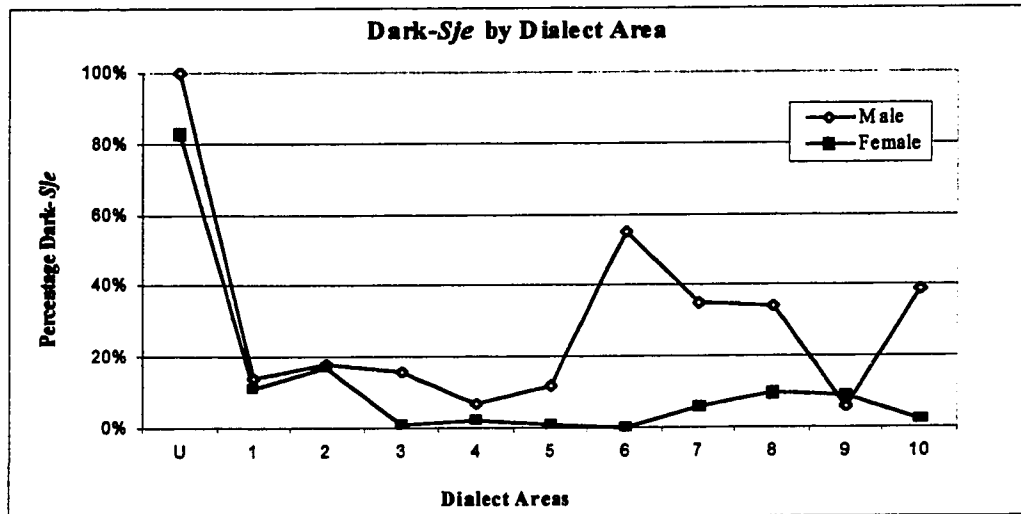


Figure 10. Dark-*sje* pronunciation by dialect and gender

It is clear that the pattern in Uppland (outside our core area) is for nearly pure dark-*sje* usage, while all NN areas have predominantly [ʂ].²⁴ It also is clear that males show a greater frequency of dark-*sje* pronunciations than females do overall in NN (20% vs. 4% in NN), although there are two dialect areas (2, 9) where females used slightly more dark-*sje* than males did, and it appears that there is very little difference between the sexes in all areas from Pitemål (DA-5) south. Gender differentiation is especially clear in the two principal dialect areas in NN where Lindblad (1982) found them, Luleå (DA-6) and the mining district (DA-10). Furthermore, we also find the complete absence of dark-*sje* among the women of Luleå that he reported. It appears that two other northern dialects that he did not investigate, Kalix (DA-7) and New Settlers' (DA-8), also have similarly high usage of dark-*sje* by males, although the Finnish-influenced dialects along the Finnish border in the Torneå valley (DA-9) do not. There may also be a slight tendency toward more dark-*sje* pronunciations in dialects toward the southern edge of NN (DA-1, DA-2). Our data show no noticeably greater usage in Umeå, which lies in our area (DA-

²⁴ Actually, the situation in Uppland may be even more categorical than these data suggest, since the one female speaker who used [ʂ] instead of a darker variant for /s/ specifically told me that her speech might not be typical for Uppland because her mother was from Skellefteå, in DA 4.

3), where it is lumped together with various other towns, including Lycksele where he found no such tendency, so that any effect might be diluted and obscured. In any case this was a relatively small factor for Lindblad (1982). We must remember that all of these numbers should be expected to be somewhat lower than would normally be found in casual speech, in light of his observation that there was less tendency to use *dark-sje* when reading (as in our data) than when conversing casually (cf. Hypothesis F). In any case, there appears to be sufficient basis in our data to propose the reformulation of Hypothesis B in order to take into account the wider perspective that our geographically more diverse database allows us, in comparison to his data from only six towns. The generalization seems to be that males in the dialect areas from Luleå (DA-6) north (except for DA-9 along the border with Finland) use *dark-sje* more than those in more southerly parts of NN. This can be formalized as a revised Hypothesis B'.

Hypothesis B'

North (DA 6-10 except 9) > South (DA 1-5) (male dialects).

Hypothesis A (Males > Females) and Hypothesis C (men > boys > girls > women) (in Malmberget and Luleå only) can be considered in light of Table 3 and Figure 11, where we examine the gender of the speakers as it relates to the age groups.

Table 3. *Dark-sje* pronunciation by gender and age (in decades)

Age	Male		Female		Both	
1	20	18%	2	2%	22	9%
2	37	12%	9	3%	46	8%
3	27	19%	1	1%	28	12%
4	32	20%	9	7%	41	14%
5	51	32%	1	1%	52	22%
6	18	17%	5	9%	23	14%
7	18	42%	7	33%	25	39%
Totals	203	20%	34	4%	237	13%

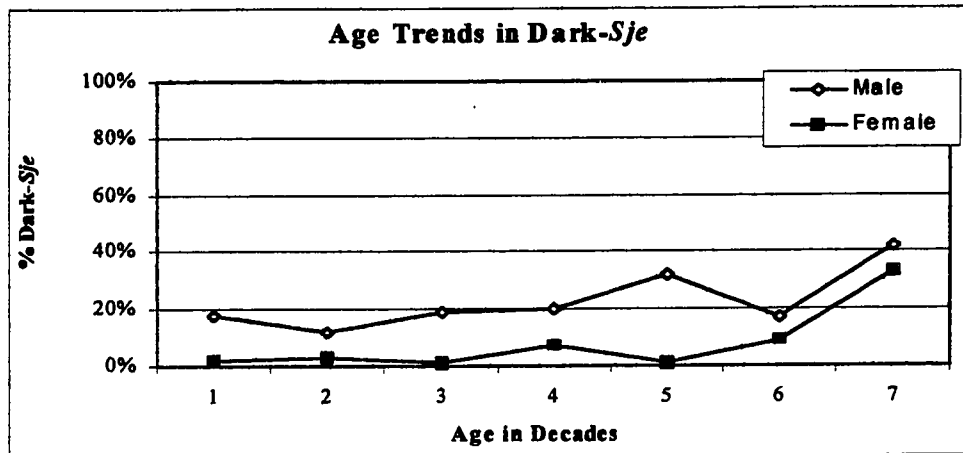


Figure 11. Dark-*sje* pronunciation by gender and age (in decades)

Since when all NN dialects are taken together, both men and women tend to use dark-*sje* less at younger ages, we have evidence for NN as a whole that Hypothesis C (men > boys > girls > women) is not confirmed. However, we cannot use this overall pattern to draw conclusions about the particular patterns found in Malmerget and Luleå. Still, it does appear that both gender and age are significant factors in the patterns of dark-*sje* usage when all dialects are lumped together. Men use dark-*sje* more at all ages, confirming again Hypothesis A (men > women). There is a clear gradual trend toward less usage of dark-*sje* by younger speakers, though with a very interesting deviation by men in their 60's who dropped off so quickly from their elders' level that men in their 50's appear as an unexpected peak in the downward trend, and even the youngest male speakers barely get below the level of men aged 60-69. Women aged 50-59 show a similar dip, but the overall rate for women is much lower, so this dip is a less pronounced deviation from the general trend, when compared to the trend line for other generations of women. (We will have occasion to discuss these patterns again later.)

Since this pattern for the various age groups seen in Figure 11 is so unexpected, we can look at it from another perspective by grouping the data in a different way to see if it is an obvious statistical aberration. Therefore we will break the decade groups in half, to the extent that our number of speakers will allow. Therefore Hypothesis A (Males >

Females) and Hypothesis C (men > boys > girls > women) can also be considered in light of Table 4 and Figure 12, where we divide each age group further by half.

Table 4. Dark-*sje* pronunciation by gender and age group (in half-decades)

Age	Male		Female		Both	
<20	20	18%	2	2%	22	9%
20-24	22	15%	6	3%	28	8%
25-29	15	9%	3	3%	18	7%
30-34	14	20%	1	2%	15	12%
35-39	13	18%	0	0%	13	12%
40-44	13	13%	3	5%	16	10%
45-49	19	32%	6	10%	25	20%
50-54	20	22%	0	0%	20	15%
55-59	31	46%	1	3%	32	31%
60-64	8	12%	3	6%	11	10%
65-69	10	24%	2	29%	12	25%
70-74	6	67%	5	26%	11	39%
75-84	12	35%	2	100%	14	39%
Totals	203	20%	34	4%	237	13%

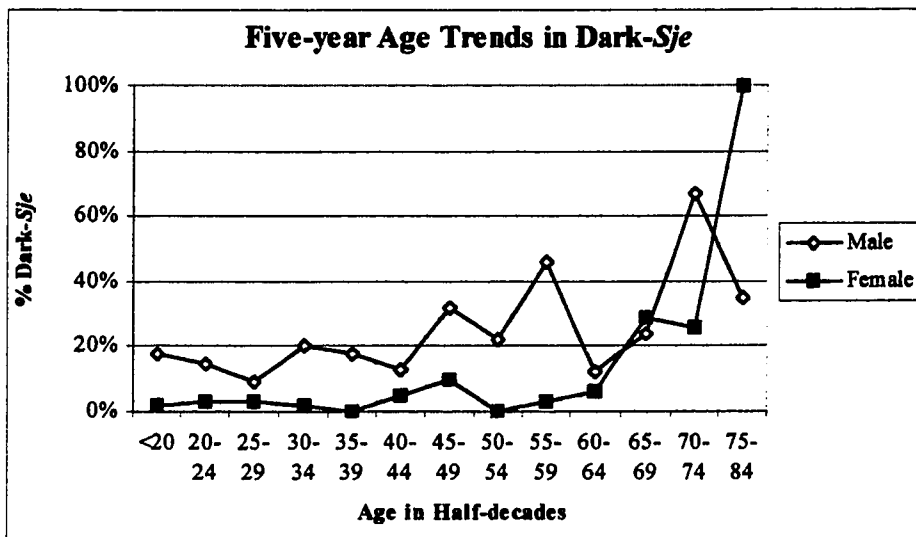


Figure 12. Dark-*sje* by gender and age group (in half-decades)

By breaking down the age groups in sets of five years instead of decades, we can localize the deepest trough of men's drop-off in dark-*sje* usage to those aged 60-64 (12%), and see that only those aged 25-29 (9%) went lower, although men aged 65-69 (24%) also

already used *dark-sje* at only about half the rate of their juniors aged 55-59 (46%). The curve for women across the relevant age groups becomes somewhat smoother than in Figure 2, but also shows a clear drop in usage of *dark-sje* already for the age group 60-64, bringing their pattern more in line with that of the males, for whom we noted a trough for ages 60-64. These groups with anomalously low usage of *dark-sje* correlate rather closely to those who were teenagers during World War II. I am inclined to think that that is not a coincidence, though it may be impossible to prove a causal relationship to radio listening or any other single factor.

Here we must also point out that when we look at the older age groups, we find two groups where women have a higher percentage of usage of *dark-sje* than do the men, contrary to most of our earlier findings. However when we look more carefully at the details for the group of women aged 75-84, we find that the two instances of *dark-sje* by the sole female speaker in that group produce a result of 100%, which is no doubt a statistically meaningless blip on the graph, and which should alert us to be cautious when looking at our data. Our results may be interesting, but in the technical sense they are only frequency tables (or “descriptive” statistics) and not the sort of “inferential” statistics that would allow us to reach any secure conclusions about the population that we are looking at.

Looking at the overall picture in Figure 11 and Figure 12, we can see a pattern in the results, regardless of whether we break down the data by five or ten year age cohorts, whereby the trend is clearly toward lower usage of *dark-sje* by younger speakers, except that those who were teenagers during World War II were far ahead of the trend. This can be formulated as a revised Hypothesis C' which relates only to age (with a special place for the WWII generation).

Hypothesis C'

Older > younger (but WWII generation far ahead of their time)

Examination of the dialects of the speakers as they relate to age groups revealed no new insights. Unfortunately not all age groups are represented in our data for Uppland (DA-U) and the Mining district (DA-10), and speakers over the age of 70 are not represented in all of the remaining dialect areas. Still, the pattern described above, of generally greater usage for the various dialect areas in the northern part of NN and a slight increase again for those at the southern edge of NN, did apply across a variety of age groups.

The remainder of the sets of comparisons in this chapter will involve the pairing of lexical items with other categories. In Table 5 and Figure 13 we will examine lexical items matched with the gender of the speakers.

Table 5. Dark-*sje* pronunciation by lexical item and gender

Lexical Item:	Male		Female		Both	
<u>kanske</u>	0	0%	1	1%	1	1%
<u>vinsch</u>	1	1%	0	0%	1	1%
<u>flygsjuk</u>	55	44%	12	14%	67	32%
<u>skolskötterska</u>	18	21%	1	1%	19	12%
<u>skarpskyttarna</u>	28	32%	4	6%	32	21%
<u>människor</u>	3	3%	0	0%	3	2%
<u>människa</u>	3	3%	0	0%	3	2%
<u>själv</u>	26	21%	6	7%	32	15%
<u>själva</u>	32	37%	6	9%	38	25%
<u>skyldig</u>	20	23%	1	1%	21	14%
<u>skicklig</u>	17	20%	3	4%	20	13%
Overall Totals	203	20%	34	4%	237	13%

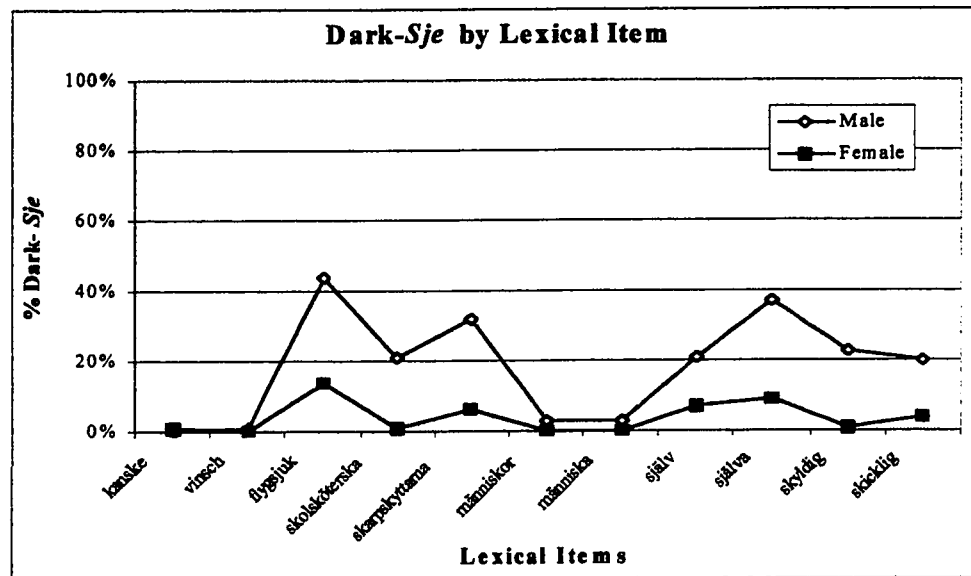


Figure 13. Dark-*sje* pronunciation by lexical item and gender

Here too we see that males consistently use dark-*sje* more than do females (Hypothesis A). We also see that there is a distinct variability related to the lexical items, which here have been specially arranged to reflect their stress patterns. Recall the discussion in §3.4, and Hypothesis D which says that /ʃ/ should always be realized as retroflex [ʂ] immediately following the stressed vowel, whether alone or in a cluster. Taken literally, this would apply to the first five lexical items, since they have primary stress on the first syllable and *sje* occurs in the following consonant cluster. This chart shows clearly that the data for items 1 and 2 fit the hypothesis nicely, while 3-5 do not. According to a common view, all of these words (3-5) accord with a literal interpretation of Lindblad's (1982:103) description of Position 1 (where NN speakers uniformly have [ʂ] for /ʃ/). However, 3-5 are transparently compounds, with /ʃ/ properly belonging to the start of the second element rather than the end of the first, so that there is a boundary (probably a word boundary and not just a morpheme boundary) between the stressed vowel and *sje*. A check of the specific lexical items used by Lindblad (1982: 120-121) reveals that some of them were compounds that correspond rather closely to these three compounds in terms of stress pattern and morphological composition, including *ljússkèn*, *hélskinnade*,

sjúkskötarska, *svårskötta* in the text, and *påssjuka*, *nåttskift*, *börtskämd*, *knäskydd*, *nórrskèn*, *kálvskinn*, *pípskäg*, *störtskön*, *júlskinka*, *bórdsskiva*, *úppsjö*, *vérsschema* in the word list. All of these are compounds that have a (word) boundary after a first syllable that takes primary word stress, and immediately following that boundary comes a syllable that both begins with a *sje* and takes secondary stress.

Since my results seem rather robust in showing darker variants of /s/ for the three lexical items 3-5, I suspected that Lindblad 1982 also got darker /s/ variants for the compounds above, and that he classified these words as having their /s/ sound before the (secondary) stressed vowel, presumably factoring in their internal boundaries and thereby disregarding the common view that the only lexical stress falls on the first morpheme of such words. Since Lindblad (1982) had included four compounds of exactly this pattern in the text and twelve more in the word list that the students read, I decided to contact him to verify his analysis of them. Lindblad (p.c. 1998) would assign them to his category 2a, which agrees with my results (an appreciable number of dark-*sje* pronunciations). This implies at the very least that the second stress of Accent II words is phonologically significant, like the first stress (although some analyses of the Swedish stress system have denied this). A possible additional implication is that the morpheme boundary is of crucial significance, but this cannot be definitively shown from the set of examples available in his or my data, since all of our data follows the normal pattern where the *sje* comes after the boundary between the two parts of the compound, and is in the onset of the following syllable. What could provide the needed test of the significance of the compound boundary is a form in which *sje* is the second member of the coda of a monosyllabic first part in a compound, followed by a second part of the compound that is both vowel-initial and initially-stressed. The only words with this unusual structure that I find in a search of Hedelin et al. (1989) are *luncheko* ('noon news broadcast') and *branschindex* ('branch index'). Neither of these words found their way into either of our databases, but the radio program *Luncheko* is probably heard by

millions of Swedes every day, and its name could easily be used in a new collection of data. Although words of this phonological structure have not been tested, I hypothesize that his description of Position 1 (following Sellberg et al. (1975) and Hammermo (1976)) which formed the basis for Hypothesis D needs to be restated in a way that recognizes the crucial significance of morphological boundaries for many phonological patterns in Swedish, including this one. On the assumption that both a word boundary before the *sje* and a secondary stress following it are required to produce the results in lexical items 3-5, I suggest that Hypothesis D could be tentatively revised as follows, pending further research to confirm the significance of the word boundary.

Hypothesis D' (first tentative revision)

/ʃ/ is always realized as retroflex [ʂ] immediately following the vowel with primary stress, whether alone or in a consonant cluster, unless that stressed vowel and the /ʃ/ transparently belong to different words that are combined to form an unambiguous compound.

The remaining lexical items 6-11 all seem to fall under Hypothesis E, so that dark-*sje* should alternate with [ʂ] for all of them. However, it seems clear that items 6 and 7 (essentially a single lexical item, being the plural and singular forms of the same noun) elicited virtually no dark-*sje* pronunciations, and therefore instead must pattern with the group described by Hypothesis D. Assuming that in rough outline the formulation of this hypothesis is correct, so that the disfavored position for dark-*sje* is one that follows a stressed vowel with no intervening transparent morphemic boundary, it would seem that one way to achieve this result would be by syncope of the /i/ vowel in the second syllable of *människà*. *SAOB*, Volume 17 (1945) lists an alternate pronunciation /män³ʃa²/, so a pronunciation with syncope of the /i/ in *människà* must have been around for considerably more than 50 years. However, after listening again to a sampling of twelve readings of this story on my tapes to test this idea, I find no evidence for syncope of the

/i/ of the second syllable in my database, and therefore believe that this analysis cannot account for the consistent pattern revealed in my data, and that some other explanation must be sought.

Lindblad (1982: 120) lists *människà* ‘person’ (=Item 7) as one of the words in the list that the 7th grade students read, and it clearly fits his definition of Position 2. Conceivably this discrepancy in our results could also be attributable to the difference in style between reading a text and reading individual words in a list, but this seems unlikely. Notice that since *människà* is Accent 2, with primary stress on the first syllable and secondary (or tertiary) stress on the final syllable, either possible analysis of the second stress (i.e., that it is significant or not) still puts the *sje* in this word in Position 2 (whether 2a or 2b), because no available analysis puts any stress on the vowel /i/ of the second syllable.²⁵ Lindblad (p.c. 1998) has confirmed that he regards the /s/ of *människà* as being in Position 2b. Since he expected variability in Position 2, it perhaps is not surprising that the extremely low frequency of dark-*sje* for this word in his data would not attract his attention, since he might attribute it to simply being one end of the normal range of variation. At this point it remains unclear whether this unexpected result should be attributed to this lexical item or to this stress pattern (primary – unstressed – secondary), which does not occur elsewhere in our corpus.

Although the secondary stress on the final syllable of *människà* / *människòr* is considered to be weaker than that on the second part of a compound because it is not accompanied by lengthening of the vowel, these cases could be accounted for by an alternative revision of Hypothesis D, such that the crucial factor is seen as being the secondary stress but word boundaries must also still be taken into account. One possible way to do this is given here in Hypothesis D".

²⁵ *SAOB* gives an earlier pronunciation (the citation is from the year 1716) that placed the stress on the middle vowel /i/, but as far as I can tell this is an obsolete pronunciation, unobserved in current usage.

Hypothesis D " (second tentative revision)

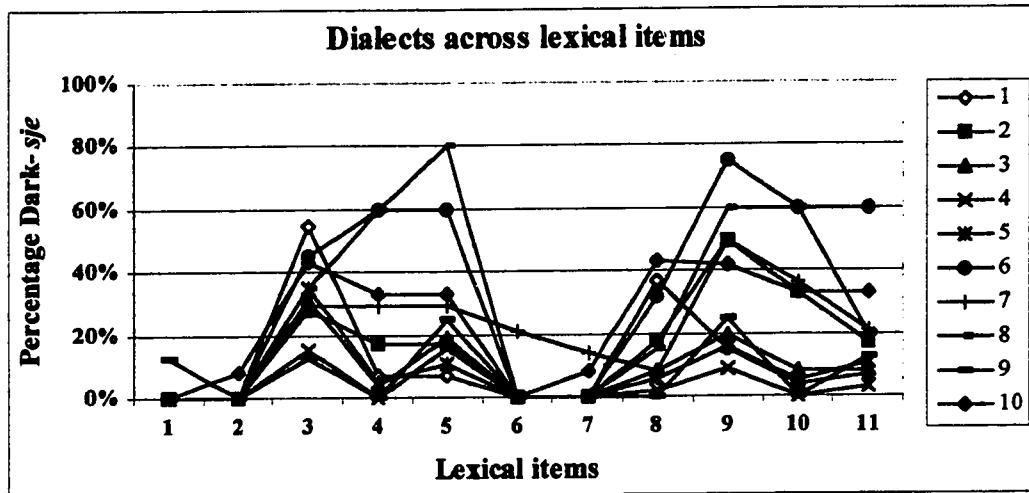
/ʃ/ is always realized as retroflex [ʂ] immediately following the vowel with primary stress, whether alone or in a consonant cluster, and in the onset of a syllable with secondary stress that follows an unstressed syllable, so long as /ʃ/ is not immediately preceded by a word boundary.

Although Hypothesis D" is preferable to Hypothesis D' because it can handle both the compounds handled by Hypothesis D' and the word(s) *människà* / *människòr*, our data include so few different patterns of word stress and morphological composition interacting with the location of the /ʃ/ phoneme that it is impossible for me to have a lot of confidence in this formulation of the hypothesis. One puzzling aspect of Hypothesis D" is that it lumps together some coda positions with some onsets. From our limited set of data, it would be possible to instead formulate Hypothesis E as being the onset of a syllable following a word boundary, leaving Hypothesis D as the elsewhere case. Unfortunately this is not in accord with the analyses presented by previous researchers on the various dialects of Swedish, so a final formulation of Hypothesis D for NN dialects in a way that accounts for all of our lexical items remains problematic, and the resolution of this conundrum must await further research.

In Table 6 and Figure 14 we will examine lexical items matched with the dialect of the speakers.

Table 6. Dark-*sje* pronunciation by lexical item and dialect

Lexical Item:	Dialect Areas										Comb.	
	Upp.	1	2	3	4	5	6	7	8	9		10
1. <u>kanske</u>		0%	0%	0%	0%	0%	0%	0%	0%	13%	0%	1%
2. <u>vinsch</u>		0%	0%	0%	0%	0%	0%	0%	0%	0%	8%	1%
3. <u>flygsjuk</u>	92%	55%	28%	31%	15%	35%	45%	29%	35%	13%	43%	35%
4. <u>skolsköterska</u>		7%	17%	4%	0%	5%	60%	29%	60%	0%	33%	12%
5. <u>skarpsskyttarna</u>		7%	17%	19%	16%	11%	60%	29%	80%	25%	33%	21%
6. <u>människor</u>		0%	0%	0%	0%	0%	0%	21%	0%	0%	0%	2%
7. <u>människa</u>		0%	0%	0%	0%	0%	0%	14%	0%	0%	8%	2%
8. <u>själv</u>	92%	37%	18%	8%	2%	6%	32%	8%	15%	0%	43%	20%
9. <u>själva</u>		15%	50%	20%	9%	16%	75%	50%	60%	25%	42%	25%
10. <u>skyldig</u>		4%	33%	8%	0%	5%	60%	36%	60%	0%	33%	14%
11. <u>skicklig</u>		7%	17%	8%	3%	11%	60%	21%	20%	13%	33%	13%
Overall Totals	92%	13%	18%	9%	4%	8%	36%	22%	28%	8%	24%	14%

Figure 14. Dark-*sje* pronunciation by lexical item and dialect

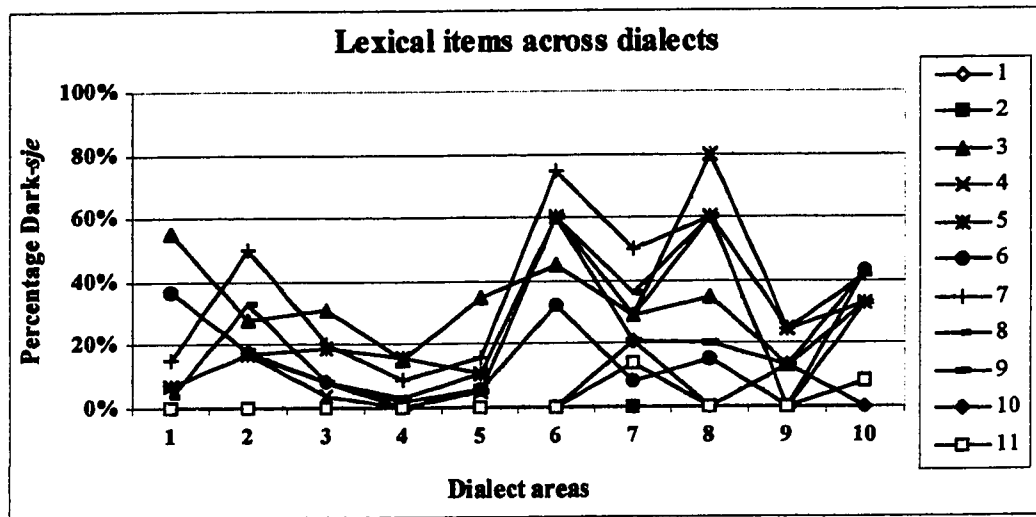


Figure 14. Dark-*sje* pronunciation by lexical item and dialect (continued)

Tabulation and graphing of lexical items and dialect areas against each other confirms the findings that we have made earlier. Lexical items 1-2 and 6-7 produce very few instances of dark-*sje*, which can be taken as support for the need to revise Hypothesis D. Dark-*sje* is most frequent in the northerly dialects except along the Finnish border (DA 6-10, except 9), supporting a modified Hypothesis B (north [except DA-9] > south). Examination of the pattern of individual lexical items across dialects shows some degree of continuity, and might suggest partial support for a theory of lexical diffusion. However, as Labov (1994: 483) has observed, “[m]ost of the studies that argue for lexical diffusion show a strong frequency effect, with the more frequent words favored in the change.” Consulting two standard frequency dictionaries of Swedish (Hassler-Göransson 1966, Allén 1970) reveals no discernable correlation between word-frequency and the frequency with which dark-*sje* occurs. Since the change also does not appear to be spreading in the smooth manner anticipated by the theory Neogrammarian change, it does not appear that any clarification for its quirkiness is to be found from these theories.

In Table 7 and Figure 15 we will examine lexical items matched with the age group of the speakers.

Table 7. Dark-*sje* pronunciation by lexical item and age group

Lexical Item:	Age in Decades							Combined
	1	2	3	4	5	6	7	
1. <u>kanske</u>	0%	0%	0%	4%	0%	0%	0%	1%
2. <u>vinsch</u>	0%	0%	0%	0%	5%	0%	0%	1%
3. <u>flygsjuk</u>	29%	24%	11%	44%	21%	53%	75%	32%
4. <u>skolskötterska</u>	5%	6%	17%	4%	33%	21%	25%	12%
5. <u>skarpskyttarna</u>	10%	7%	28%	26%	48%	21%	50%	21%
6. <u>människor</u>	5%	2%	0%	0%	5%	0%	0%	2%
7. <u>människa</u>	0%	2%	0%	0%	10%	0%	0%	2%
8. <u>själv</u>	12%	17%	8%	16%	9%	11%	54%	15%
9. <u>själva</u>	10%	19%	33%	22%	50%	36%	0%	25%
10. <u>skyldig</u>	5%	7%	22%	9%	38%	7%	25%	14%
11. <u>skicklig</u>	11%	7%	17%	13%	29%	0%	50%	13%
Overall Totals	9%	8%	12%	14%	22%	14%	39%	13%

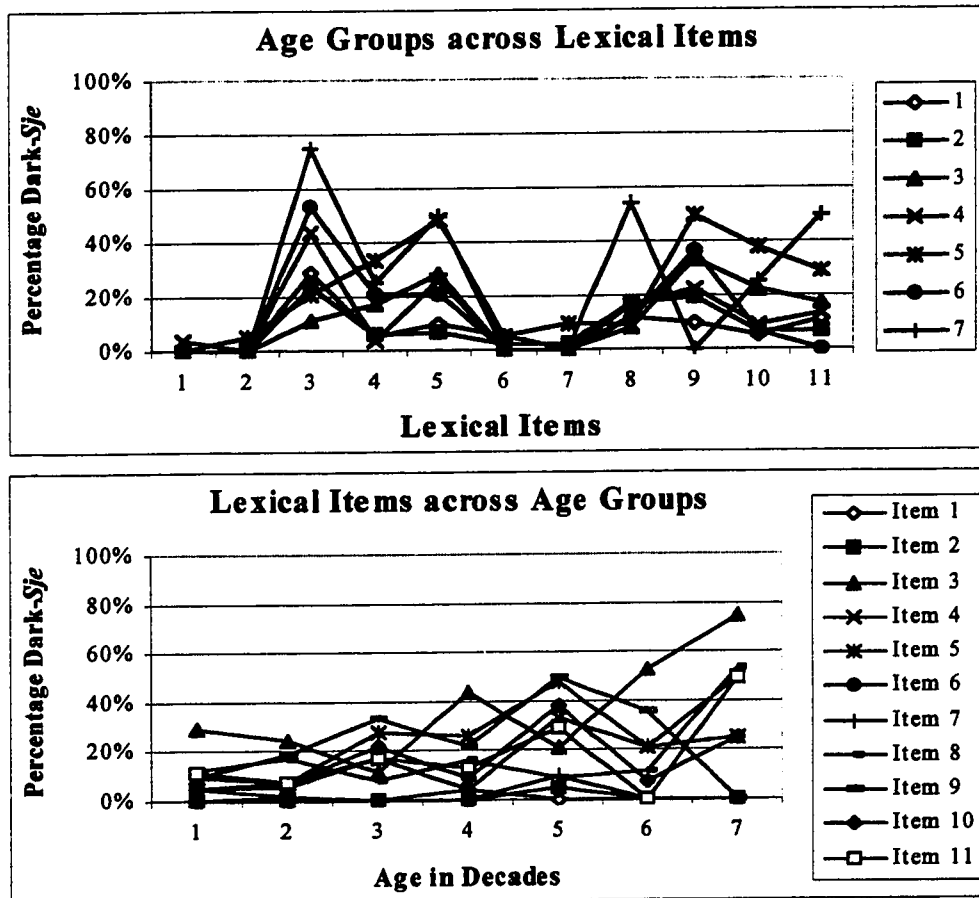


Figure 15. Dark-*sje* pronunciation by lexical item and age group

When we look at the graph of the various age groups across the lexical items, we can see that all age groups have low rates of *dark-sje* for lexical items 1-2 and 6-7. The graph of the various lexical items across the age groups fits the general pattern that we have noted of greater usage among older decades than younger ones, and suggests that each word may have its own trajectory as it passes through (apparent) time.

3.7 SUMMARY OF DARK-SJE USAGE

In this chapter we have used our new data to test a series of hypotheses based on the findings of Lindblad (1982), and where the new data showed that the original hypothesis was not adequate, we have formulated alternatives.

Hypothesis A

Men > women.

Hypothesis A is generally confirmed for all cross-groupings, including dialect area (Table 2), age in decades (Table 3), age in half-decades (with two minor deviations) (Table 4), and lexical item (with one minor deviation) (Table 5).

Hypothesis B

Malmberget/Kiruna > Luleå > Umeå > elsewhere.

Our data do not support Hypothesis B for Malmberget and Luleå as formulated, as can be seen most easily in Figure 10. However, there is clearly a tendency for males in the north more generally to use *dark-sje* more than those in the south. Therefore we formulated a revised Hypothesis B'.

Hypothesis B'

North (DA 6-10 except 9) > South (DA 1-5) (male dialects).

Hypothesis B' is generally confirmed by the cross-groupings with age in decades (not shown) and lexical item (Figure 14), although the inclusion of the data for both males and females together in those figures somewhat obscures the effect.

We also formulated another hypothesis that singled out Malmberget and Luleå, where Lindblad (1982) believed he saw evidence of a progressive breakdown in the younger generation of an earlier clear gender-based pattern:

Hypothesis C

Men > boys > girls > women (in Malmberget and Luleå only).

Although we did not attempt to test Hypothesis C for Malmberget and Luleå, due to our more limited set of data from those two towns, we found no evidence overall that girls had more dark-*sje* usage than adult women. Therefore we formulated a revised Hypothesis C', which relates only to age (with a special place for the WWII generation), but which applies to all of NN.

Hypothesis C'

Older > younger (but WWII generation far ahead of their time)

Hypothesis C' is generally confirmed for cross-groupings with gender (Figure 11 and Figure 12), and lexical item (Figure 15).

Based on earlier studies, Lindblad (1982) distinguished two positional relationships between *sje* and stress, which he claimed accounted for much of the distribution of dark-*sje*, and which we restated as Hypothesis D and Hypothesis E.

Hypothesis D

/ʃ/ is always realized as retroflex [ʂ] immediately following the stressed vowel, whether alone or in a consonant cluster [with minor exceptions for two affective words that do not occur in our data].

Hypothesis E

Retroflex [ʂ] can alternate with darker variants of /ʃ/ elsewhere.

Our data showed that both of these are imperfect statements of the actual distribution of dark-*sje*. Hypothesis D fails to mention the possibility of boundaries within compounds, which appear to be a crucial factor. Dark-*sje* can appear in a consonant cluster following a stressed vowel if there is a morpheme (word?) boundary between them. Hypothesis E is too broad because it includes the word *människà* / *människòr*, although this word does not share in dark-*sje* pronunciations, so it should somehow be accommodated under Hypothesis D. We suggested some possible alternatives to Hypothesis D and Hypothesis E, but none were wholly satisfactory because our list of lexical items is too small to allow us to see clearly what the prosodic conditioning factors may be. Nonetheless, there can be no doubt that stress placement in relation to *sje* is a very strong conditioning factor.

Lindblad (1982) also found a clear effect from the formality of the speech situation, forming a hierarchy that we represented as Hypothesis F.

Hypothesis F

Emotional language > conversation > text-reading > list-reading.

Since all of our data were from text-reading, we were unable to test this hypothesis, and we used it only to gain perspective when examining discrepancies between our results and his.

In the next chapter, we will look at /rs/ sequences within words, and then compare them to the results we have obtained in this chapter for underlying /ʃ/ to see if there is evidence for or against neutralization of the two, either now or in the past.

CHAPTER 4: INTERNAL SANDHI OF /RS/ SEQUENCES

Since no-one has studied /rs/ sequences in North Norrland, there are many unanswered questions about their patterning that are of interest for dialectologists and phonological theoreticians, among others. What governs the occurrence of sandhi in word-internal /rs/ sequences? Are there any similarities between /rs/ sequences and the underlying *sje* phoneme that we looked at in Chapter 3, so that we can say that there is a degree of neutralization between these two different phonological entities that both are normally realized as [ʂ] in NN?

In this chapter, after some initial remarks, we will present the hypotheses that we have extracted from previous studies and the factors that we have available to test them. Next we will present and discuss certain compilations of that data by Varbrul that are instructive for our purposes. Finally, we will summarize our results, the conclusions for our hypotheses that we can draw from our compiled data.

4.0 OVERVIEW OF INTERNAL SANDHI

In this chapter we will look at examples of the underlying sequence /rs/ that occur within words, i.e. in what are known as internal sandhi contexts. These word-internal contexts can be further subdivided according to whether the /rs/ sequence occurs within a morpheme or there is a morpheme boundary between the /r/ and /s/. In the case where there is such a boundary, we can distinguish various types of boundary, including those before inflectional and derivational suffixes and those between the parts of a compound. We can also distinguish them on the basis of the parts of speech involved, as well as the various possible combinations of stress on the preceding and following syllables in the prosodic environment.

Besides comparing each of these different types of internal sandhi to each other, there are also other comparisons to make. The first one, which we will be able to make already in this chapter, is the comparison between monomorphemic /rs/ sequences and the phoneme *sje* that we looked at in Chapter 3. The other /rs/ sequences, those that contain a morpheme boundary, encompass a range of variation in the degree of distance between their morphemes, and after we have looked at the similar range of variations in external sandhi contexts in Chapter 5, we will compare internal and external sandhi contexts in an attempt to blend the two together into a single continuum.

4.1 PREVIOUS STUDIES

Although no-one has looked at internal sandhi of /rs/ sequences in NN, Eliasson (1986) made a number of observations about it in Central Standard Swedish that I used as hypotheses to test when writing my stories. I have also used my results in Lindblad (1987a) to form one additional hypothesis (Hypothesis L).

4.2 HYPOTHESES TO BE TESTED RE INTERNAL SANDHI OF /RS/

Eliasson (1986) makes several claims regarding internal sandhi sequences of /r/ and a following (series of) dental(s). In formulating my hypotheses I have focused primarily on those claims that involve /s/ as the target segment, regardless of other intervening dentals.

Hypothesis G

Sandhi occurs regularly through sequences of dental consonants both within and between morphemes

This first and broadest of the claims states that sandhi regularly occurs and recurs both morpheme-internally and across all sorts of morphological and word boundaries. “[P]ostalveolarization is a recursive or iterative rule, i.e., it can in most cases apply and reapply as long as there occur dentals in a direct succession after an /r/ ... (a) morpheme-

internally, (b) in internal and (c) in external sandhi.” (Eliasson 1986: 279) Phonological adjacency is a normal requirement for sandhi processes such as this one, and it is to be expected that the inclusion of a non-dental consonant, a vowel or a pause between the trigger segment /r/ and the dental target segment will block sandhi. The only apparent claim to the contrary that I have found occurs in Sjögren (1978: 142), who illustrates his definition of the terms ‘*sandhi*’ and ‘*supradentalisering*’ ‘supradentalization’ with the phonetic transcription of a possible pronunciation of the name ‘Ernst Snellman’ as follows: [æ:ŋst̪ s̪nɛ[-man]. This seems to suggest that sandhi in Swedish can progress across vowels. However, Peter Sjögren (p.c. 1997) informed me that this was a misprint (that nobody had pointed out to him previously), and that the <ɪ> in ‘Snellman’ should be not a supradental but an ordinary <ɪ>.

Hypothesis H

Sandhi is favored: morpheme-internally > inflectional boundaries > derivational boundaries > formal compounds > regular compounds

Although he does not state it precisely this way, (Eliasson 1986: 282) suggests this more specific hypothesis about the hierarchy of probability of occurrence of sandhi both by the arrangement of his Table 3 in this sequence (Eliasson 1986: 278, relevant parts of which were presented above as Figure 4 in Chapter 1) and by his statement singling out compounds for a tendency to non-occurrence, “[t]one and stress patterns may [...] favor a separation, especially in true and formal compounds.”

Hypothesis I

Sandhi does not occur following geminate /rr/

Eliasson also lists three types of exceptions to the generalizations represented by Hypothesis G and Hypothesis H. The first involves geminate /rr/, which “generally does not cause postalveolarization of the following dentals.” (Eliasson 1986: 280) That /rr/

will not lose features to merge with a following dental is not surprising in view of the numerous reports of geminate inalterability that may be found in the linguistic literature. This prohibition on multiple linkings of phonological segments may be expressed using the formalism of Autosegmental Phonology.



(Cf. also the Linking Convention.)

Hypothesis J

Sandhi with non-laterals does not occur following /r/

A further observation is that the normally free recursion of sandhi through a string of dentals “does not ... extend to non-laterals that follow a lateral,” (Eliasson 1986: 280), for which he gives *pärsocker* ‘crushed loaf sugar’ as one example. This observation may possibly be related to the one embodied in Hypothesis I if the sandhi product of the /r/ sequence results in an entity that partakes of partial geminate inalterability while those of other /r/ plus dental sequences do not, and it suggests that the class of liquids consisting of /r/ and /l/ may form a significant entity in the phonology of Swedish. (It should also be noted here that this is not relevant to the word *värld* ‘world’, since its orthographic <l> is not pronounced and therefore plays no role in its phonological relations.)

Hypothesis K

Sandhi tends to not occur in the words *absurd, urdu, nokturn, Saturnus, Ursula*

Eliasson (1986: 281) notes a few words, which he observes to be mostly of foreign origin and to mostly contain /u/ immediately preceding the sandhi site, that “tend not to undergo the process.” Of those he lists, the only one containing the sequence /rs/ is the woman’s name *Ursula*.

Hypothesis L

/rs/-Sandhi tends to occur less frequently in Dialect Area 9 and among males, especially following the prefix *för*-

In Lindblad (1987), working class men in the central Swedish town of Borås were found to have less sandhi than other men and all women, most notably in words that begin with the prefix *för*-. Drawing on those results, I hypothesize that such words might also be treated differently in NN by men and/or in DA-9 along the Finnish border.

In the next section we will discuss briefly the factors that will be used to analyze the data and test these hypotheses.

4.3 FACTORS IN INTERNAL /RS/-SANDHI

The demographic factors in this chapter are the same as those described in Chapter 2, namely the gender (§2.1.0), age (§2.1.1), and dialect area (§2.1.2) of the speakers. The linguistic factors to be considered here are the morphemes, stress, and lexical items.

By preceding and following morphemes, we mean the grammatical class category of the morphemes containing respectively the /r/ and /s/ that potentially participate in /rs/-sandhi, and that immediately precede and/or follow its site.

By stress we mean the degree of stress on the syllabic nuclei immediately preceding and following the sandhi site. Because I have tested for the influence on the occurrence of sandhi of the stress levels of the preceding and following syllables, the question arises of what the stress level of the following syllable is in cases where the /rs/ sequence is word-final. Based on my experience with *är* (Lindblad 1998), I have decided that the following word may be a part of the prosodic environment, and therefore in such cases I have both tested them as being word-final and tested them as being under the influence of the first syllable of the following word. In the specific case of the word *hårs* 'hither', which is followed by the conjunction *och* 'and', I have assumed that this following minor

category word will consistently be destressed in this prosodic environment (between two stressed syllables), and therefore have evaluated its stress level as unstressed.

By lexical items we mean the words that contain the internal sandhi site, regardless of their morphological make-up. These lexical items are detailed below in Appendix G, including the total number of responses of each type that each received, even those that were not used for Varbrul calculations. In addition, it should be said that the requirement of the Varbrul program that a factor group have no more than 30 members played a constraining role in deciding how to code these items. The two occurrences of the words *surströmming* ‘fermented Baltic herring’, *världsmästare* ‘world champion’, *eftersom* ‘because’, *sorts* ‘of a sort’, *största* ‘largest’, and *Sigvarlds* ‘Sigvard’s’ had to be listed as one lexical item, even though in certain cases (*surströmming*, *eftersom*, *sorts*) there was an apparent difference between the two in their frequency of participation in sandhi. It may be uncontroversial to treat the word *världsspråk* ‘global language’ as a single lexical item since it only occurred once, but it should be pointed out that I tested separately for sandhi from its /r/ through /d/ (<l> is not pronounced) to the first /s/ (=MfoN 23) and for sandhi from that /s/ to the following /s/ (=MfoN 24). The words *skolsköterska* ‘school nurse’ and *bagerska* ‘female baker’ both contain the female agentive suffix *-erska*; I assume that we are justified in considering them to be the same lexical item because we are looking at two occurrences of the same morpheme, even if they are embodied in two different words. Finally, we have treated two different conjugated forms built on the verb root *försök-* ‘attempt’ as identical, which is justifiable because the sandhi context is entirely contained in the root itself, even though *för-* is clearly a derivational prefix, and as such was found to contribute significant morphology in Lindblad (1987a).

A few comments may help in the interpretation of Appendix G. You may use the Item # in the third column to see the context in which each lexical item occurred, by using it as a reference to its place in the numbered versions of the stories in Appendixes A (MfoN) and C (Msi). In my statistical analysis in Varbrul, I used only the data in the [s] (sandhi

occurred) and [rs] (sandhi did not occur) columns, and all other results have been discarded. It is not clear to me why there were so many cases in which I could not decide whether or not sandhi had occurred on the a. and b. lexical items, so that they ended up with an unusually large number of 'Maybe' results. However, it is clear why there are so many 'Omitted' results for lexical item *t. styvfars*; this word was part of the title of *Msi*, and those 27 people did not read the title before they started reading the story.

4.4 TABLES AND GRAPHS OF RESULTS

Let us now begin to examine our data for internal /rs/-sandhi in the light of these hypotheses, starting with the two genders observed across ten-year age groups in Table 8 and Figure 16.

Table 8. Internal /rs/-sandhi, by gender and age in decades

Age	Male		Female		Both	
<20	413	83%	417	80%	830	82%
20-29	888	83%	737	85%	1625	84%
30-39	477	78%	371	79%	848	78%
40-49	552	78%	399	80%	951	79%
50-59	369	75%	276	75%	645	75%
60-69	322	70%	142	74%	464	71%
70-84	171	68%	92	70%	263	69%
Totals	3192	78%	2434	80%	5626	79%

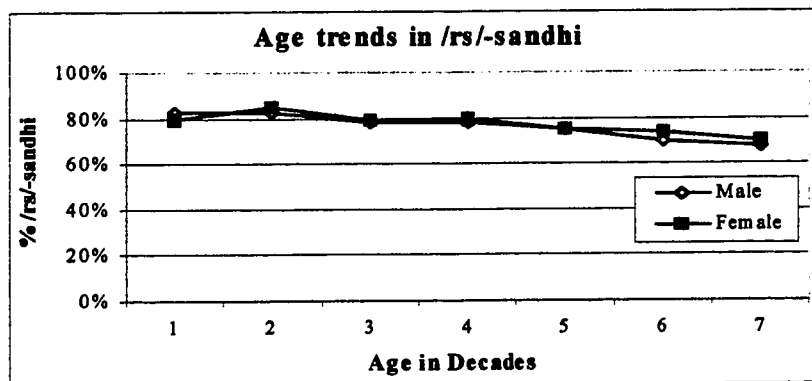


Figure 16. Internal /rs/-sandhi, by gender and age in decades

It appears from Figure 16 that there is very little difference between males and females at any age. Thus this pairing of factors does not support Hypothesis L, which foresaw a possibility of less usage by males. On the other hand, Figure 16 shows a clear trend of gradually increasing usage of sandhi by younger speakers, which none of our hypotheses predicted. This is interesting in light of the implication by Dahlstedt & Ågren (1954: 9-10) that the normal case was for sandhi to occur among the speakers that they were looking at, most of whom were of an earlier generation and many of whom were born in the nineteenth century. Although even our oldest speakers use sandhi more than two-thirds of the time, the trend line suggests that /rs/-sandhi production was not necessarily the predominant pattern earlier in the nineteenth century.

Although we found little difference between males and females by decadal age group, in the previous chapter we found it worthwhile to further subdivide the speakers into five-year cohorts, so let us look at that breakout here as well in Table 9 and Figure 17.

Table 9. Internal /rs/-sandhi, by gender and age in half-decades

Age	Male		Female		Both	
<15	45	76%	90	76%	135	76%
15-19	368	84%	327	81%	695	83%
20-24	394	83%	491	85%	885	84%
25-29	494	84%	246	86%	740	84%
30-34	249	75%	205	81%	454	78%
35-39	228	80%	166	77%	394	79%
40-44	309	77%	232	81%	541	78%
45-49	243	80%	167	80%	410	80%
50-54	219	82%	169	79%	388	81%
55-59	150	66%	107	71%	257	68%
60-64	186	70%	94	82%	280	74%
65-69	136	70%	48	62%	184	68%
70-74	73	72%	80	71%	153	71%
75-84	98	65%	12	63%	110	65%
Totals	3192	78%	2434	80%	5626	79%

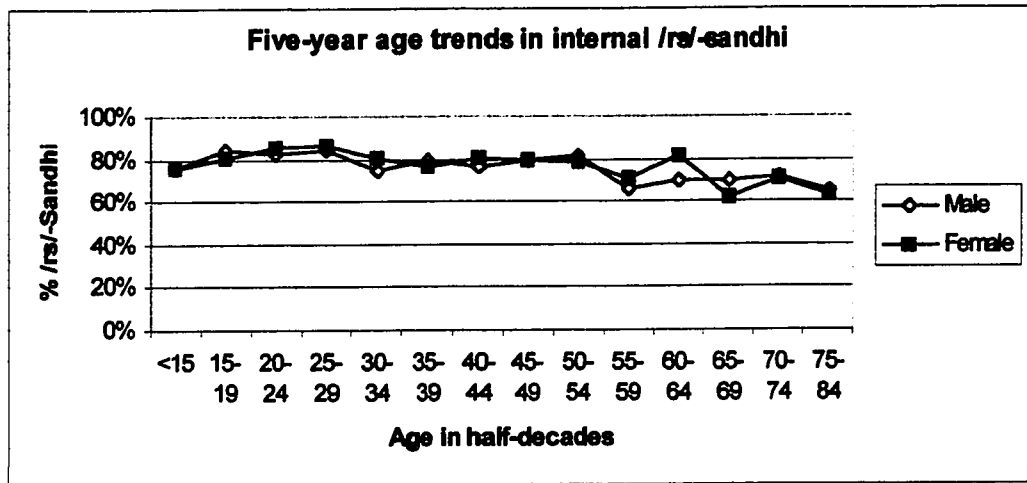


Figure 17. Internal /rʌ/-sandhi, by gender and age in half-decades

Although Figure 17 in most respects agrees with the contours of Figure 16, and shows little difference between Male and Female speakers at most ages, we do find here an interesting deviation in the World War II generation that did not show up in Figure 16, where the Females aged 60-64 and 65-69 were averaged together. This reinforces the impression we got earlier from looking at the variation of *sje* in Chapter 3, that something happened to disturb the continuity of the speech community in NN around the time of WW II, although in that case it was the Males who changed most rapidly, and here it is the Females.

Let us now look at the pronunciations by the two genders of individual lexical items in Table 10 and Figure 18, where the lexical items have been arranged in order of morphological closeness, ranging from monomorphemic to regular compounds.

Table 10. Internal /rs/-sandhi, by gender and lexical item

Lexical Item:	Male		Female		Both	
	Count	%	Count	%	Count	%
i. bagerska / skolsköterska	203	95%	151	96%	354	95%
m. Ursula	47	37%	33	37%	80	37%
p. Helsingfors	124	99%	91	100%	215	100%
a. hemortskänsla	101	90%	74	89%	175	90%
c. världsspråk	210	84%	170	94%	380	89%
d. världsmästare	167	96%	129	97%	296	96%
f. morfars	119	94%	87	97%	206	95%
j. systers	104	86%	83	91%	187	88%
n. sorts	196	95%	140	92%	336	94%
o. största	213	100%	157	99%	370	99%
t. styvfars	65	96%	56	95%	121	95%
u. sorters	44	52%	42	63%	86	57%
y. Sigvards	155	90%	130	98%	285	93%
l. sparsamt	103	82%	68	75%	171	79%
r. försöker / försöka	207	99%	154	98%	361	99%
w. härs	84	97%	67	100%	151	98%
x. tvärs	84	97%	66	99%	150	97%
z. försiktig	80	93%	64	96%	144	94%
3. förstod	84	97%	67	100%	151	98%
e. eftersom	153	74%	130	83%	283	78%
h. Persson	122	98%	90	99%	212	99%
4. urskog	76	90%	62	94%	138	92%
b. surströmming	81	34%	67	41%	148	37%
g. storstadsbo	122	98%	87	97%	209	97%
k. pärlsocker	3	2%	2	2%	5	2%
q. vintersemester	66	54%	38	42%	104	49%
s. barrskogen	2	2%	2	2%	4	2%
v. tvåmetersstaket	80	93%	58	87%	138	90%
1. sommarställe	37	43%	20	30%	57	37%
2. sommarstuga	60	69%	49	74%	109	71%
Overall Totals	3192	78%	2434	80%	5626	79%

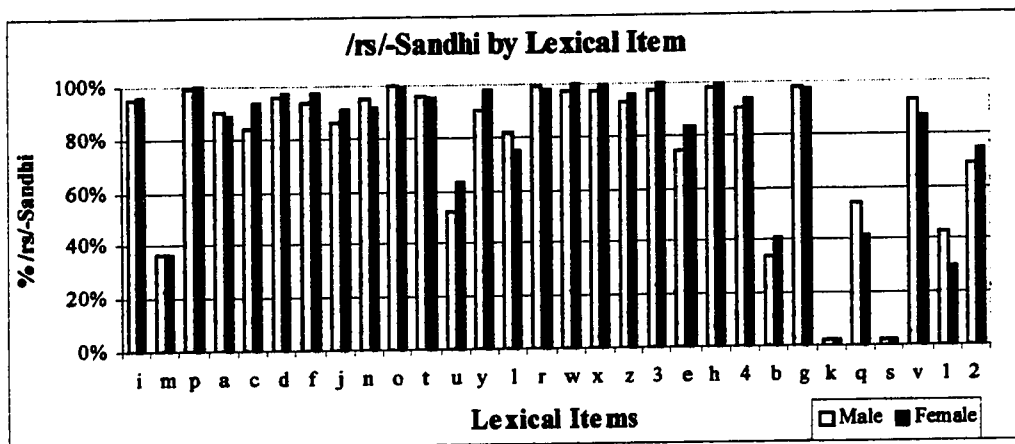


Figure 18. Internal /rs/-sandhi, by gender and lexical item

When we look at these results in Table 10 and Figure 18 in detail, we can observe several points. As with Figure 16 and Figure 17, we see here that there is very little difference between male and female speech in regard to internal /rs/-sandhi. Although more than half of the Lexical Items show /rs/-sandhi at a rate above 90%, it is clear that it does not apply regularly to all words. When we look for cases where another dental consonant is interposed between /r/ and /s/, we see that Lexical Items a., c., d., k., n. and y. fit that description, and all except Item k. *pär|socker* ‘crushed loaf sugar’ (which is an exceptional case covered by Hypothesis J) take sandhi at a rate between 89% and 96%. Thus the regularity of application postulated by Hypothesis G (Sandhi occurs regularly through sequences of dental consonants both within and between morphemes) is supported to the extent that intervening consonants in general seem to play no role in any irregularity, but sandhi applies in a quite high percentage of cases only for a small plurality of our Lexical Items.

In Table 10 and Figure 18 we do find strong confirmation for Hypothesis I (Sandhi does not occur following geminate /rr/) in our results for Item s. *barr|skogen* ‘pine forest’, with only 2% sandhi overall. We also find strong confirmation for Hypothesis J (Sandhi with non-laterals does not occur following /rl/) in our results for Item k. *pär|socker* ‘crushed loaf sugar’, also with only 2% sandhi overall. We find some confirmation, though to a weaker degree, for Hypothesis K (Sandhi does not occur in the words [...] *Ursula*) in our results for Item m. *Ursula* (personal name), with 37% overall, although I am unwilling to put too much credence in that precise number because many readers seemed to stumble over the name when they read it, suggesting that they were unfamiliar with it and unsure of its pronunciation. *Ursula* is probably an uncommon name in NN, since on January 1, 1973, there were only 1385 women in all of Sweden who bore the name, and most of them lived in the far south of the country (Allén & Wåhlin 1979: 218). Although Item s. *barr|skogen* ‘pine forest’ and Item k. *pär|socker* ‘crushed loaf sugar’ clearly showed the least frequent usage of sandhi, there were two others as low as Item m. *Ursula*, namely Item b. *sur|strömming* ‘sour herring’ (37%) and Item l. *sommar|ställe* ‘summer place’

(37%). Others were nearly as low, including Item q. *vintersemester* ‘winter vacation’ (49%) and Item u. *sorters* ‘of sorts’ (57%). Although *Ursula* may be a rare name in NN, the same certainly cannot be said about *surströmming* ‘sour herring’, which might be called the regional specialty food of NN, with *Surströmming Premiär* ‘Sour herring première’ on the third Thursday of August being a major regional holiday. Thus pragmatics, in terms of frequency or familiarity effects, seems unable to explain the low rate of usage of /rs/-sandhi in certain cases, and we have no available hypothesis that accounts for it. Since *surströmming* ‘sour herring’ contains the vowel /u/ immediately preceding the /rs/ sequence, one might be tempted to treat it along with the words given by Eliasson (1986: 281) and cited in Hypothesis K, even though all of those differ from *surströmming* in being monomorphemic and of foreign origin. However, if one were to do that, then one would need to provide an explanation for why *urskog* ‘old growth forest’ has sandhi at the much higher and more typical rate of 92%, even though it too contains the sequence /urs/; clearly the credit cannot simply belong to *skog* ‘forest’, which had only 2% sandhi when it appeared in the word *barrskogen* ‘the pine forest’. Perhaps further research will find an explanation for this result.

Now let us examine our lexical items across the different age groups in Table 11, looking to see if we find the same trend toward greater usage in each of them that we noticed for the population overall in Figure 16, or whether each word is going its own way.

Table 11. Internal /rs/-sandhi, by lexical item and age group

Lexical Item:	Age in Decades							Combined
	1	2	3	4	5	6	7	
i. bagerska / skolsköterska	94%	96%	96%	97%	95%	94%	90%	95%
m. Ursula	28%	38%	42%	44%	21%	40%	44%	37%
p. Helsingfors	100%	100%	100%	100%	100%	95%	100%	100%
a. hemortskänsla	97%	98%	92%	92%	96%	65%	58%	90%
c. världsspråk	100%	99%	96%	83%	77%	67%	74%	89%
d. världsmästare	100%	99%	94%	100%	90%	89%	87%	96%
f. morfars	100%	100%	100%	93%	91%	89%	81%	95%
j. systers	97%	96%	89%	93%	88%	65%	63%	88%
n. sorts	92%	99%	93%	91%	98%	82%	89%	94%
o. största	100%	99%	100%	100%	100%	97%	100%	99%
t. styvfars	95%	100%	100%	89%	94%	86%	100%	95%
u. sorters	65%	65%	59%	41%	57%	36%	50%	57%
y. Sigvards	98%	100%	94%	91%	82%	79%	100%	93%
l. sparsamt	82%	85%	71%	80%	88%	63%	75%	79%
r. försöker / försöka	100%	100%	100%	100%	91%	97%	100%	99%
w. härs	95%	100%	94%	100%	95%	100%	100%	98%
x. tvärs	90%	100%	94%	100%	95%	100%	100%	97%
z. försiktig	95%	98%	100%	96%	81%	93%	67%	94%
3. förstod	100%	100%	100%	100%	90%	93%	100%	98%
e. eftersom	82%	76%	85%	85%	69%	63%	70%	78%
h. Persson	100%	100%	100%	97%	96%	100%	94%	99%
4. urskog	95%	100%	89%	86%	85%	86%	75%	92%
b. surströmming	50%	41%	26%	38%	30%	32%	39%	37%
g. storstadsbo	100%	98%	95%	100%	100%	90%	94%	97%
k. pärlsocker	6%	0%	8%	0%	0%	0%	0%	2%
q. vintersemester	56%	53%	49%	64%	35%	45%	13%	49%
s. barrskogen	6%	2%	0%	3%	0%	0%	0%	2%
v. tvåmetersstaket	90%	93%	89%	87%	81%	100%	100%	90%
1. sommarställe	30%	44%	50%	30%	29%	21%	50%	37%
2. sommarstuga	80%	76%	65%	74%	62%	71%	25%	71%
Overall Totals	82%	84%	78%	79%	75%	71%	69%	79%

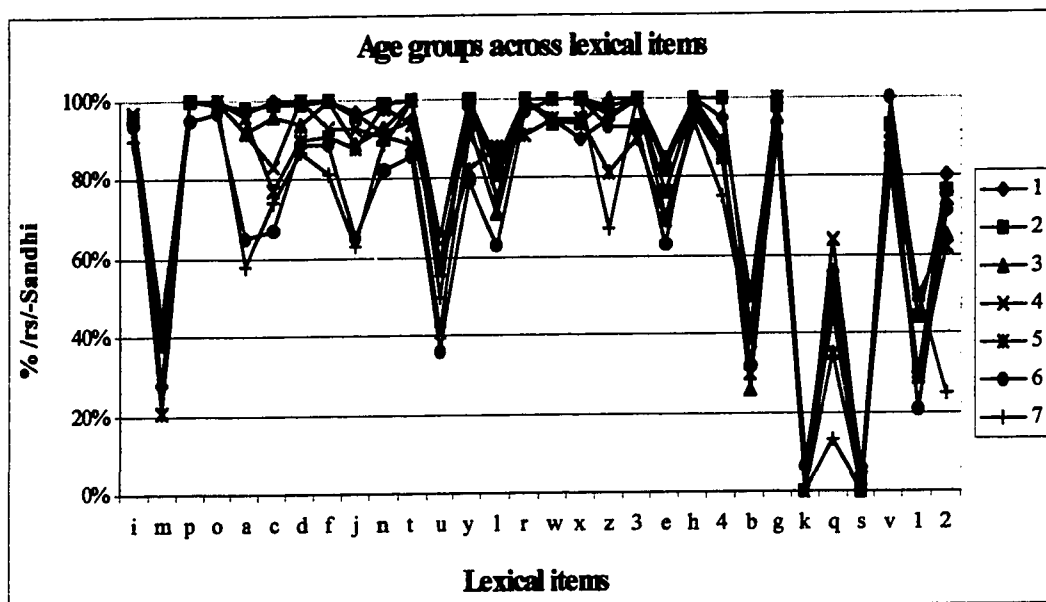


Figure 19. Internal /rs/-sandhi, by lexical item and age group

In Table 11 and Figure 19 we can see that there is a general tendency for all age groups to treat the various lexical items in the same way as other age groups do, although a few items behave somewhat more erratically. In Figure 19 the pattern for words with the genitive suffix <s> (a, c, d, f, j, n, t, u, y) looks rather more random than most, but Table 11 shows that most of the lower frequency results are due to older age groups. It appears that there may have been an earlier pattern where sandhi did not occur as regularly on this suffix, but now the norm seems to be firming up behind its use. The one word in this group that stands apart is *u. sorters* 'of sorts', perhaps attributable to the plural suffix *-er* behaving differently from noun stems, which constitute all the other examples.

We can see a general tendency for younger speakers to use sandhi more, with twenty of the thirty items converging above the 90% level, and most of the others trending upward, even if still at lower levels of usage. Comparison of the results for lexical items and age groups here with those for lexical items and gender above shows a tendency for the same words that have more sandhi for women to also be trending upward in usage for younger speakers, which is in line with a tendency that sociolinguists have often noted for women

to be in the vanguard of change toward more prestigious forms. The two words that seem to clearly be trending downward, at least within the group of people below the age of 40, are Item m. *Ursula* (personal name) and Item 1. *sommarställe* 'summer place'. As discussed above, it seems unwise to put much weight on the results for *Ursula*, since many readers seemed unfamiliar with it and stumbled over its pronunciation. As for *sommarställe* as contrasted with Item 2. *sommarstuga* 'summer cabin', although I find them listed as synonyms, side by side in Prisma's *Modern svensk engelsk ordbok*, and they are treated identically in *Svenska ord* where the pronunciation guide shows both of them with sandhi, still Hassler-Göransson (1966) has neither word in her main list of 400,000 words, collected in the 1920's and 1930's, and her list of 100,000 words collected in the 1950's has *sommarstuga* only twice and *sommarställe* only once. Allén (1970, vol. 3) in his much larger (and newer) corpus found *sommarstuga* eighteen times but *sommarställe* only six times. I believe that *sommarstuga* is the more common usage, but frequency has not appeared in any of our results to be a determining factor. As was the case with *Ursula*, a few readers here stumbled over *sommarställe*, casting some small additional doubt on its results. Although women and young people show a trend toward greater use of sandhi on many lexical items, there is no clear prognosis for the future of the others.

The remainder of the tables and figures that we will look at in this chapter will concern the variability of our several factors across the different dialects. First let us look at the patterning of gender across the dialect areas in Table 12 and Figure 20.

Table 12. Internal /rs/-sandhi, by gender and dialect area

Dialect Area	Male		Female		Combined	
	Count	%	Count	%	Count	%
U	95	68%	92	77%	187	72%
1	504	83%	382	78%	886	81%
2	141	67%	204	75%	345	72%
3	492	79%	306	85%	798	81%
4	634	81%	550	82%	1184	82%
5	388	82%	197	85%	585	83%
6	284	77%	142	80%	426	78%
7	263	79%	181	85%	444	81%
8	223	78%	137	69%	360	74%
9	113	57%	209	77%	322	69%
10	150	71%	126	81%	276	75%
Totals	3287	78%	2526	80%	5813	79%

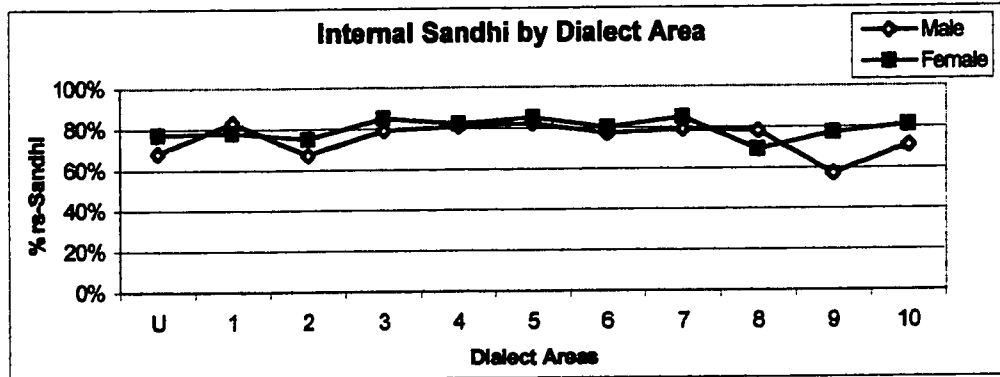


Figure 20. Internal /rs/-sandhi, by gender and dialect area

Here for the first time we find some small degree of confirmation for Hypothesis L, that /rs/-sandhi may occur less frequently in DA-9 and among males. Although these results do not specifically address the prefix *för-*, it is at least suggestive that the lowest levels of usage (and by a full 10%) occur among the males of DA-9. We can also observe that Dialect Areas 8, 9 and 10, all of which are beyond the traditional Swedish-speaking region and which could be argued to have a Sami or Finnish substratum, do not fit as smoothly into the regional pattern as do the areas further south.

I have looked at the patterning of ten-year age cohorts across the dialect areas, but the results are not clear-cut or dramatic enough to warrant presentation here. In brief, they show that the tendency for older speakers to use less sandhi is strongest in Dialect Areas 2, 3, and 9. Also, DA-2 has less usage of sandhi for all of the older age groups, suggesting that the somewhat deviant patterning of this dialect area that we saw for *sje* in Chapter 3 is long-standing, and that Dahlstedt & Ågren (1954) were justified in describing it as a separate dialect area despite its small size.

I have also looked at the patterning of lexical items across the dialect areas, but again the results are not clear-cut or dramatic enough to warrant presentation here. Dialect Areas 2 and 9 seem to break quite sharply downward from the patterns of their adjacent areas for many words. The Finnish substratum can be appealed to in the case of DA-9, but DA-2's patterning is an unexplained anomaly once again. In light of Hypothesis L, it is interesting to see that Lexical Items *r.*, *z.*, and *3.* do not behave identically in DA-9, even though they all begin with the sequence *för-*. Both Lexical Item *r.* *försöker / försöka* 'attempts / to attempt' and Lexical Item *3.* *förstod* 'understood' take sandhi at a 100% rate in DA-9, while Lexical Item *z.* *försiktig* 'cautious' takes sandhi at only a 75% rate. One possibility is that they analyze *försiktig* 'cautious' as a parallel form to English 'foresighted' and as containing a weakened variant of the morpheme *förr* 'before' instead of the common prefix *för-*, so that this lower usage could be attributed to a weakened variant of geminate inalterability, similar to Lexical Item *s.* *barrskogen* 'pine forest' but less categorical.

Next let us look at the influence of stress, beginning with the patterning of stress on the preceding vowel across the dialect areas in Table 13 and Figure 21.

Table 13. Internal /rs/-sandhi, by dialect and stress on preceding vowel

Dialect Area	Type of stress on vowel preceding -rs-						Total	
	Primary		Secondary		Unstressed			
U	98	63%	35	90%	54	83%	187	72%
1	415	77%	159	95%	312	80%	886	81%
2	176	67%	62	84%	107	74%	345	72%
3	376	78%	135	94%	287	81%	798	81%
4	556	76%	210	95%	418	84%	1184	82%
5	260	76%	103	98%	222	86%	585	83%
6	212	69%	79	96%	135	87%	426	78%
7	205	76%	82	99%	157	81%	444	81%
8	172	64%	72	99%	116	81%	360	74%
9	154	62%	61	85%	107	71%	322	69%
10	124	74%	51	91%	101	71%	276	75%
Totals	2748	73%	1049	94%	2016	81%	5813	79%

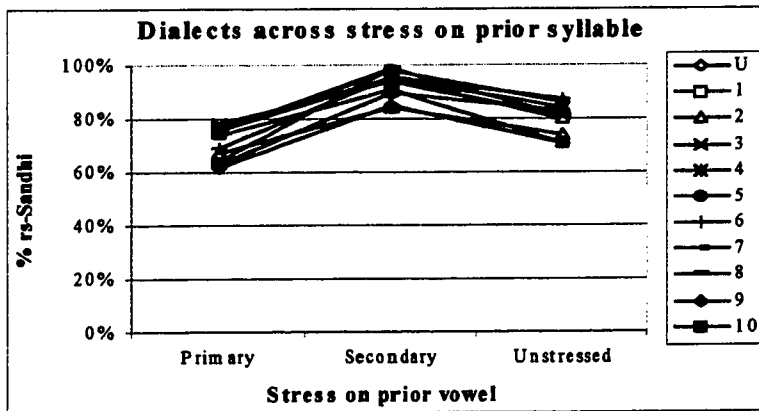
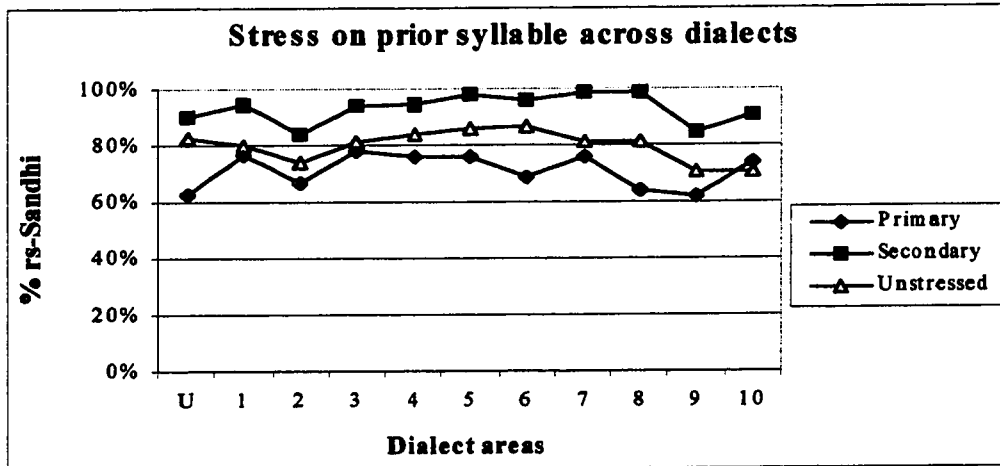


Figure 21. Internal /rs/-sandhi, by dialect and stress on preceding vowel

Secondary stress on the preceding vowel is the most favorable degree of stress in all dialects, with Primary stress the least favorable and with Unstressed falling in between in most cases. None of my specific hypotheses predicted this result, and our only previous hint about stress being a factor was the vague suggestion by Eliasson (1986: 282) that “stress patterns may ... favor a separation.” It seems counterintuitive that the middle grade of stress should rise above the higher and lower grades. Examination of the data reveals that all six contexts with secondary stress preceding the sandhi site are word-final. Five of the six involve nouns followed by the genitive suffix *-/s/*, and the other is the monomorphemic *Helsingfors* ‘Helsinki’. Since these are morphological contexts that arguably favor sandhi, it is unclear how much weight to put on the prosodic factor here.

Next let us look at the patterning of stress on the following vowel across the dialect areas. Drawing on Lindblad (1998) where prosodic environment was found to contribute to the presence or absence of the coda */r/* on the copula *är* ‘is/are/am’, we will look at stress within a prosodic context instead of merely in the isolated word. The patterning of stress on the following vowel within its prosodic context across the dialect areas is presented in Table 14 and Figure 22.

Table 14. Internal /rs/-sandhi, by dialect and stress on next vowel

Dialect	Type of stress on the vowel following /rs/						Total	
	Primary		Secondary		Unstressed			
U	50	96%	104	67%	33	63%	187	72%
1	315	94%	394	73%	177	79%	886	81%
2	108	88%	169	65%	68	71%	345	72%
3	283	93%	362	75%	153	79%	798	81%
4	402	95%	541	74%	241	81%	1184	82%
5	211	96%	246	72%	128	88%	585	83%
6	126	96%	215	70%	85	79%	426	78%
7	156	94%	196	73%	92	82%	444	81%
8	115	97%	182	67%	63	66%	360	74%
9	109	86%	148	60%	65	68%	322	69%
10	109	88%	110	65%	57	76%	276	75%
Totals	1984	93%	2667	71%	1162	78%	5813	79%

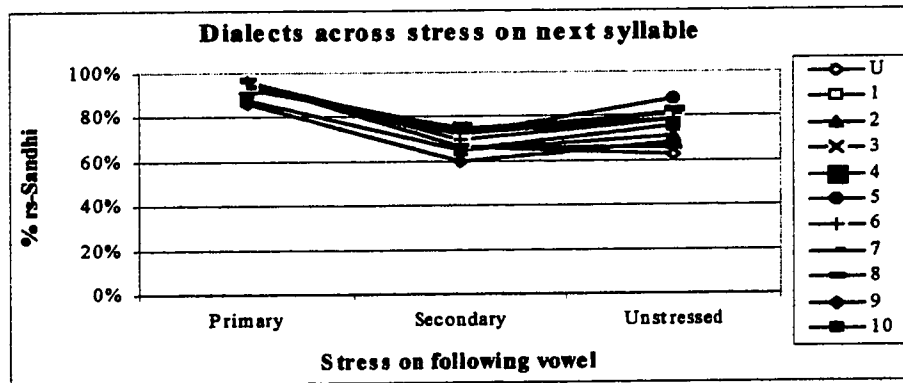
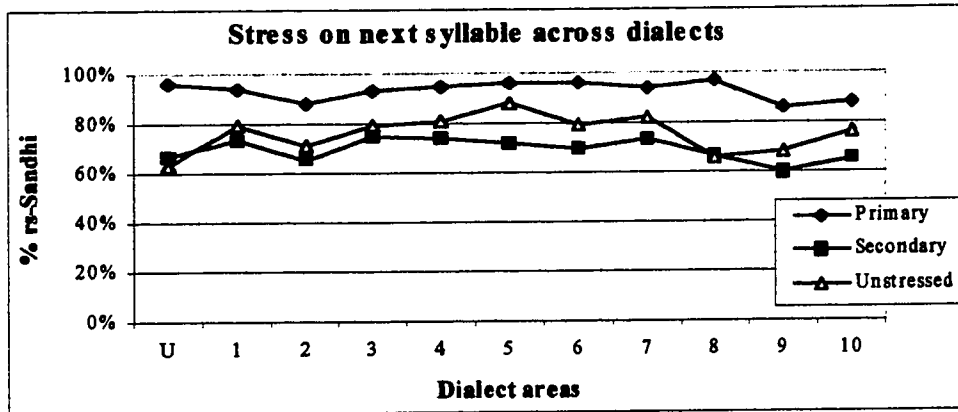


Figure 22. Internal /rs/-sandhi, by dialect and stress on next vowel

Here we see that in general secondary stress and unstressed vowels pattern together, which seems unexceptional. Primary stress on the following syllable appears to be most favorable to sandhi. However, inspection of the data reveals that the items with primary stress on the following syllable are primarily of two types, words that begin with the prefix *fōr-* and nouns that end with the genitive suffix /s/ (and incidentally are followed by a word with a stressed first syllable). Since both of these are morphological contexts where sandhi may be particularly likely, it is unclear how much credit should be given to the prosodic context.

I have looked at the grammatical category of the preceding and following morphemes across dialects as possible factors for sandhi, but also here the results are not clear. The set of items in each of the various categories is disparate, and does not seem representative. As one example, there is only one item where the preceding morpheme is a verb. On the other hand, we have much more data for the least favorable following morphological category, which is noun, at 56%. It is clear that a following noun is relatively unfavorable for sandhi in all dialects in our set of data, with values ranging from 44% to 64%. However this percentage is an average of one item at 99% and one item at 97% with two items at 2% and eight items scattered between these extremes. It does not make sense to use an average for such a data set. It may also be worth noting that the category of the preceding morpheme across the dialect areas shows mostly the sorts of gradual transitions of feature values that could be expected in a dialect continuum, but the geographic discontinuity between Dialect Area 1 and Upland is clearly manifested when the data is graphed by the crossing of lines and dramatic changes of value levels.

Before we look at the patterning of the type of internal morpheme boundary across the dialect areas to test Hypothesis H (Sandhi is favored morpheme-internally > inflectional boundaries > derivational boundaries > formal compounds > regular compounds), a few clarifications of the approach taken here may be helpful, since there are several ambiguous cases that do not easily fall into a single category. I have omitted Lexical

Items c., k., m. and s. from these calculations, since Lexical Item c. encompasses both an inflectional boundary (MfoN 23) and a regular compound boundary (MfoN 24), and Lexical Items k., m. and s. are predicted to behave anomalously by Hypothesis J, Hypothesis K and Hypothesis I, respectively. The remaining Lexical Items have been categorized as follows:

<u>Boundary type:</u>	<u>Lexical Items:</u>
Monomorphemic	i p
Inflectional	a d f j n o t u y
Derivational	l r w x z 3
Reduced Compound	e
Regular Compound	b g h q v 1 2 4

I have chosen here to categorize Lexical Item i., the feminine agentive suffix *-erska*, as monomorphemic, although an argument could be made for deriving it from the masculine (or gender-unspecified) agentive suffix *-are*. Most of the inflectional boundaries here precede the genitive suffix *-s/* (which sometimes functions to link the two halves of a compound), except for the superlative suffix *-sta* on Lexical Item o. *största* ‘biggest’ (< *stor* ‘big’) (although this word might also be argued to be effectively monomorphemic, since the umlaut of the vowel renders the stem *stör-* distinct from *stor*, and if those two are not synchronically derivable from each other then it has no other source). I have chosen to regard the *-s* suffix on Lexical Items w. and x. (*här_s* and *tvär_s*, ‘hither’ and ‘thither,’ two adverbs that form a lexicalized pair) as derivational, although it could be argued to be identical to genitive *-s*. I have chosen to regard Lexical Item e. *eftersom* ‘because’ (< *efter* ‘after’ + *som* ‘that’) as a special category which I have labeled ‘reduced compound’, since *Svenska ord* (1994) gives its pronunciation as having two long segments (like regular compounds) without also having Accent 2 (which most regular compounds have). It clearly does not fit into any of the earlier categories, nor can it be considered a regular compound. On the other hand, it also does not belong in the same category as the examples of formal compounds provided by Eliasson (1986: 278),

which took Accent 2 even though the first element of the compound was a minor category word. The first element of Lexical Item z. *försiktig* ‘cautious’ (literally: ‘fore’ + ‘sighted’) is probably etymologically identical to the first elements of those examples of formal compounds, but the combined form does not take Accent 2, so they cannot be equated. Lexical Item 4. *urskog* ‘primeval forest’ has a minor category prefixal first element, and it both contains two long segments and takes Accent 2, making it similar to formal compounds, but both Garlén (1988: 134) and *Svenska ord* (1994: 763) list *ur-* as a productive prefix, so *urskog* also does not properly belong to the category of formal compounds. If it were placed as the sole member in that category, formal compounds would have sandhi at the 92% level, like inflectional suffixes. On the other hand, *urskog* is also not truly a regular compound, although that is where I have placed it. Two of the words that I have classified as regular compounds, Lexical Items h. and v., also contain a linking *-s-* between the nouns, but I have assumed that to be irrelevant. Lexical Item h., the family name *Persson*, might also be placed with reduced compounds, since it is pronounced with Accent 1 instead of Accent 2.

Keeping in mind all these caveats, let us now look at the patterning of the type of internal morpheme boundary across the dialect areas in Table 15 and Figure 23.

Table 15. Internal /rs/ sandhi, by dialect and type of internal boundary

Boundary type	Dialect Areas										Comb.	
	U	1	2	3	4	5	6	7	8	9		10
Monomorphemic	100%	96%	98%	99%	98%	96%	100%	93%	98%	93%	100%	97%
Inflectional	91%	89%	80%	90%	95%	97%	95%	97%	96%	85%	88%	92%
Derivational	100%	96%	93%	98%	95%	95%	93%	93%	92%	86%	89%	94%
Reduced compound	69%	70%	79%	73%	82%	81%	93%	78%	92%	58%	68%	77%
Regular compound	55%	78%	61%	72%	73%	69%	67%	69%	50%	51%	51%	67%
Totals	80%	87%	79%	87%	89%	88%	87%	87%	82%	75%	79%	85%

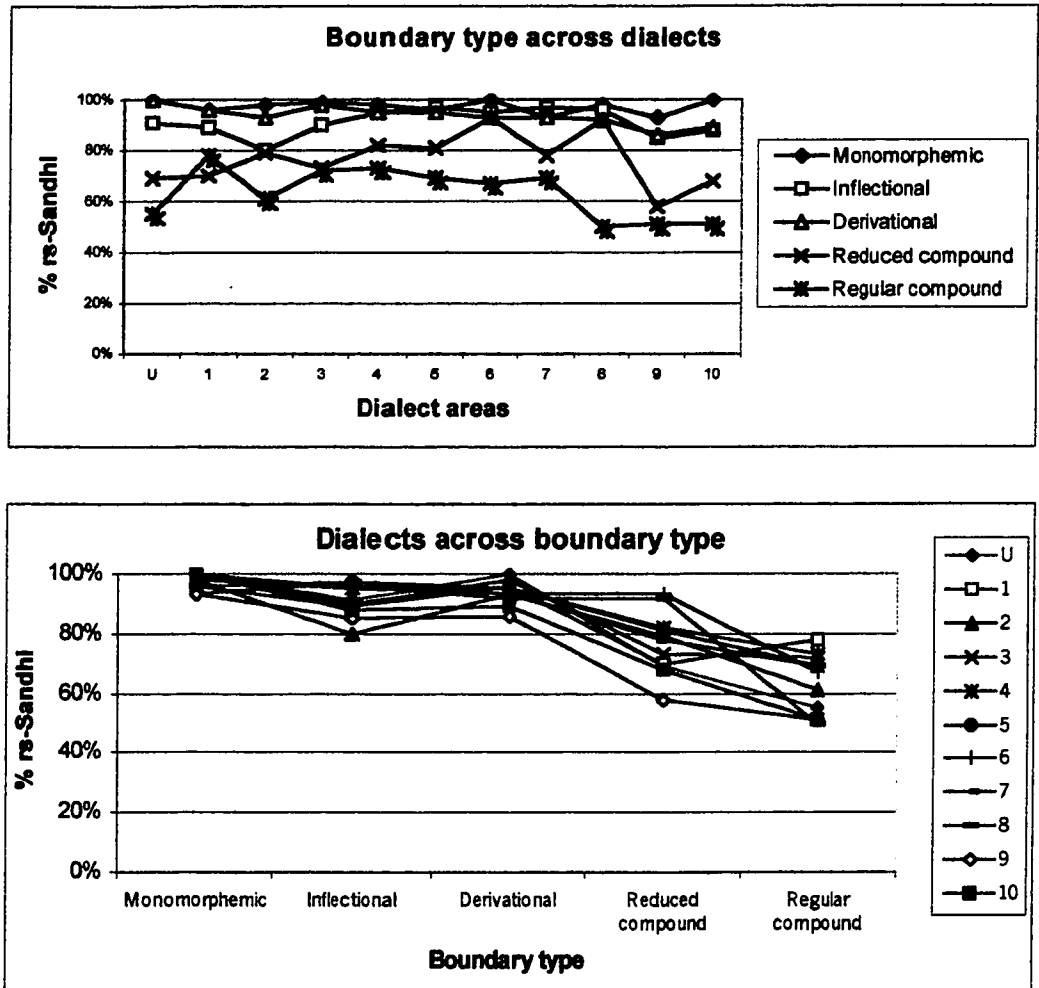


Figure 23. Internal /rs/ sandhi, by dialect and type of internal boundary

Even with all of our caveats and amendments, it seems clear that the general sense of Hypothesis H (Sandhi is favored morpheme-internally > inflectional boundaries > derivational boundaries > formal compounds > regular compounds) is correct. The results for inflectional and derivational boundaries are not clear-cut, but the distinction between inflectional and derivational boundaries is not clear-cut in our data either. It is regrettable that our data did not include a proper example of a formal compound, but the reduced compound *eftersom* may be a suitable stand-in. There is a marked difference in our results at the two ends of the range of categories, monomorphemic and regular

compounds, and the general validity of Hypothesis H seems clear. A closer relationship between the morpheme containing the /r/ and the morpheme containing the /s/ favors /rs/-sandhi, with the monomorphemic case being the closest and most favorable one. It is also interesting to note that the dialects outside the old Swedish speaking region, our Dialect Areas 8-10, stand apart from their nearest neighbors in their treatment of regular compounds.

Now that we are looking at the data for monomorphemic /rs/ sequences, it is time to go back to one of the questions raised in the introduction, and compare the patterning of those monosyllabic sequences to the patterns we found for *sje* in Chapter 3. In terms of prosodic patterns, our two monomorphemic Lexical Items, i. *bagerska* / *skolsköterska* ‘female baker’ / ‘school nurse’ and p. *Helsingfors* ‘Helsinki’, do not have the /rs/ sequence immediately following the vowel with primary stress, which is Lindblad’s (1982:104) Position 1 where NN speakers almost uniformly use [ʂ] to realize /s/. Since they have the /rs/ sequence in Position 2, some variation between [ʂ] and dark-*sje* might be expected to be a possibility in these Lexical Items. Instead we find almost uniform [ʂ] and a small number of cases where sandhi has not been applied to the /rs/ sequence, but no cases where they are pronounced with a dark-*sje*. When we look at the totality of our database for internal /rs/-sandhi, however, we do find three instances where /rs/ has been realized with a dark-*sje*! All three involve the same verb root, Lexical Item u. *försöker* / *försöka* ‘attempts / to attempt’, and all three speakers are male and from northern Dialect Areas. Specifically for MföN speaker #103 (age 49, DA-8) and speaker #122 (age 18, DA-6), and for Msi speaker #67 (age 58, DA-7) all had dark-*sje* in *försöker* / *försöka*. Since /rs/ occurs before the vowel with primary stress, this root clearly fits Position 2a., where variation could be expected. On the other hand, it may be unique to this lexical item. The ages of the three speakers vary from 18 to 58, so it does not appear to be a new innovation, but it also does not appear to be only an old relic form. This lexical item is one that dialectologists and phonologists should keep an eye on for future developments.

It suggests that neutralization between /rs/ and /ʃ/ is an active issue in NN, and not just the theoretical possibility that led me to look at this question originally.

4.5 SUMMARY AND CONCLUSIONS FOR INTERNAL SANDHI

To summarize, let us first look back at the hypotheses that we started with, and review the results that we have gotten for each of them.

Hypothesis G

Sandhi occurs regularly through sequences of dental consonants both within and between morphemes

Sandhi does pass through intervening dental consonants (except /l/) from /r/ to a following /s/ with no apparent diminution of frequency. However, there is so much variation in the rate of application of /rs/-sandhi to the various Lexical Items that describing it with the word “regularly” must be an exaggeration.

Hypothesis H

Sandhi is favored morpheme-internally > inflectional boundaries > derivational boundaries > formal compounds > regular compounds

This hierarchy seems to be supported in rough outline, although uncertainties of classification make it difficult to assess its precise details.

Hypothesis I

Sandhi does not occur following geminate /rr/

This hypothesis was strongly supported for the only example in our data, Lexical Item s. *barrskogen* ‘pine forest’. This can reasonably be ascribed to geminate inalterability or

the linking condition, which have been associated with similar effects in a variety of languages.

Hypothesis J

Sandhi with non-laterals does not occur following /r/

This hypothesis was strongly supported for the only example in our data, Lexical Item k. *pärsocker* ‘crushed loaf sugar’.

Hypothesis K

Sandhi tends to not occur in the words *absurd, urdu, nokturn, Saturnus, Ursula*

The only one of these particular words containing the sequence /rs/ is the woman’s name *Ursula* (Lexical Item m.), which took sandhi in less than half the cases in our database, so there was a certain tendency for sandhi not to occur here. On the other hand, Lexical Items b. *surströmming* and 1. *sommarställe* took sandhi at equally low rates, and they cannot be described as foreign, nor does *sommarställe* contain the vowel /u/, so our data provide only weak and qualified support for Hypothesis K.

Hypothesis L

/rs/-Sandhi tends to occur less frequently in Dialect Area 9 and among males, especially following the prefix *för-*

When we looked at gender vs. dialect area in Table 12 and Figure 20, we saw that the lowest rate of 57% was reached by the males of DA-9. However, when we looked at specific lexical items vs. dialect area, we saw that the three Lexical Items r., z., and 3. do not all behave alike in DA-9, suggesting that there is more than one homographic prefix *för-* in this dialect and that they receive different phonological treatments. Thus there may be a degree of truth in Hypothesis L, but it is not clear precisely what is involved.

When we sought to make a comparison of monomorphemic /rs/ and *sje*, we found no evidence of confusion between the two. However, we did find that three males used a dark-*sje* for the bimorphemic /rs/ sequence in the verb stem of the forms *försöker / försöka*, suggesting that neutralization between /rs/ and /ʃ/ is a real issue for the future of the dialects in NN, although it is still so rare that no-one has commented on it.²⁶ I found evidence for it only because the phonological patterns of the dialects suggested to me that it was a possible tendency, and I collected and listened carefully to a set of data that might include it.

In the next chapter we will look at external /rs/-sandhi, and compare the factors that govern it to those that we have found to influence internal /rs/-sandhi.

²⁶ Per Lindblad (p.c. 1999) remembers hearing pronunciations of this type.

CHAPTER 5: EXTERNAL SANDHI OF /RS/ SEQUENCES

Since no-one has studied /rs/ sequences in North Norrland, there are many unanswered questions about their patterning that are of interest for dialectologists and theoretical phonologists, among others. What governs the occurrence of sandhi in /rs/ sequences where a word boundary intervenes between the two segments? Are there any similarities in the factors that govern the possibility of sandhi between /rs/ sequences that span word boundaries (external sandhi) and the word-internal /rs/ sequences that we looked at in Chapter 4 (internal sandhi), so that we can say that there is a continuum across these two apparently different phonological contexts in NN?

5.0 OVERVIEW OF EXTERNAL SANDHI

In this chapter we will look at examples of the underlying sequence /rs/ that occur on the two sides of the boundary between words, i.e. in what are known as external sandhi contexts. These inter-word contexts will be analyzed by the population factors that we have looked at in previous chapters. They will also be analyzed according to a variety of linguistic factors such as the frequency of the collocation of the words and their closeness, their syntactic categories (parts of speech), the c-command relation between the words, the segments adjacent to the /rs/-sandhi site, and the various possible combinations of stress on the preceding and following syllables in the prosodic environment. We will also look at the individual lexical items separately in various contexts, primarily in order to give us another angle on our other factors so that we may be able to locate individual items that may deviate from the other items within the factor, and thereby be able to focus on the differences between them and gain a clearer perspective on what is involved.

Besides comparing each of these different types of external sandhi to each other, there are also other comparisons to make. The primary one is the comparison between /rs/ sequences that overlap word boundaries and the word-internal /rs/ sequences that we looked at in Chapter 4. For one thing, we can compare the influence of similar stress patterns in the environment. Also, we saw that word-internal /rs/ sequences containing a morpheme boundary encompassed a range of variation in type of boundary, and hence in the degree of distance between their morphemes. After we have looked at the range of variations in external sandhi contexts in this chapter, we will see that the factors affecting internal and external sandhi contexts form a single continuum across all the dialects.

5.1 PREVIOUS STUDIES

In the past fifteen years, the relation between different components of the grammar has been a major field of study in theoretical linguistics. In the present context, the interface that is most relevant is between phonology and syntax. Classics in that area of inquiry include Selkirk 1984, Kaisse 1985, and the collection of papers in Inkelas & Zec 1990. The theories put forward in those works suggest that syntactic influence on phonology may arise from c-command relations, edge relations, or by way of prosody as a mediating influence between syntax and phonology. We will look at each of these for its possible impact on /rs/-sandhi.

As was the case with internal sandhi, no-one has looked specifically at external sandhi of /rs/ sequences in NN. After his long discussion of internal sandhi, Eliasson's (1986: 282) only general statement about external sandhi in Central Standard Swedish is the following: "The situation with respect to external sandhi is distinctly different. Here the assimilation is no longer totally automatic, but depends to a considerable extent on parameters that lie outside the linguistic system in a narrow sense. In general, it can be said that the closer the transition is between two words, the more easily the assimilation will take place. Among the relevant parameters are speech rate including pause and

hesitation patterns, speech style, frequency of word collocations, individual speech habits, etc. Allegro speech and ordinary conversation (as opposed to lento speech) and casual speech (as opposed to a formal and solemn style) regularly favor sandhi.”

From this statement we can extract the generalization that internal sandhi occurs more frequently than external sandhi, as well as three specific hypotheses (Hypothesis O, Hypothesis R, and Hypothesis T, below) that can be tested with my data and methods. I have used my results in Lindblad (1987a) to form one additional hypothesis (Hypothesis S) and those in Lindblad (1987b) to form two additional hypotheses (Hypothesis V and Hypothesis W).

5.2 HYPOTHESES TO BE TESTED RE EXTERNAL SANDHI OF /RS/

Since no-one has studied external /rs/-sandhi in NN previously, we have no basis for any hypothesis regarding the population factors and diachronic trends within NN. However, we will look at them in various compilations of our data to see if the patterns we found with our other data sets also apply here for those factors, and to form a basis for possible future research hypotheses. As regards the linguistic factors that have been suggested to relate to the syntax-phonology interface, no-one has claimed any particular application of them to /rs/-sandhi in Swedish, much less in NN, so only the most general hypotheses can be formed. Kaisse (1985: 186) concisely sums up “the parameters and setting that may be employed to characterize unmarked rules of external sandhi.

1. The c-command condition: One of the words must c-command the other.
2. The edge condition: The sandhi pair (i.e., the words participating in the phonological rule) must lie on the edge of the constituent that contains them.”

These two conditions can be formalized for our purposes as two general hypotheses, Hypothesis M and Hypothesis N.

Hypothesis M

External /rs/-sandhi may be favored by the existence (or direction) of c-command.

Hypothesis N

External /rs/-sandhi may be favored by the words being on an edge of the constituent that contains them.

A phonological approach that somewhat resembles the syntactic concepts of c-command and edge in these hypotheses is found in Eliasson (1986: 282), when he says, “[i]n general, it can be said that the closer the transition is between two words, the more easily the assimilation will take place.” This can be formulated as Hypothesis O.

Hypothesis O

External /rs/-sandhi is favored by a close transition between the two words.

As noted previously, Eliasson’s discussion of internal and external sandhi suggests that degree of separation between elements is a factor that runs across both categories, and that the two form a continuum. This can be formulated as Hypothesis P, which is an extension of Hypothesis H.

Hypothesis P

Sandhi is favored morpheme-internally > inflectional boundaries > derivational boundaries > formal compounds > regular compounds > close word pairs > moderately close word pairs > distant word pairs.

Selkirk (1984), Inkelas & Zec (1990), and others have suggested a variety of prosodic contexts as determinative of sandhi in various languages. Here we will look at the degree of stress on the preceding and following syllables. Since we have no reason to expect

any particular prosodic configuration to be more favorable, we will formalize this also in a very general way as Hypothesis Q.

Hypothesis Q

/rs/-Sandhi may be favored by the stress pattern on preceding and/or following syllables.

Eliasson (1986) makes only very general claims about external sandhi sequences of /r/ and a following (series of) dental(s) in Central Standard Swedish. However, we can also test the possibility of extending certain more specific claims that he made about internal sandhi into the external domain as well. As before, in formulating my hypotheses I have focused primarily on those claims that involve /s/ as the target segment, regardless of other intervening dentals.

One of the broadest claims states that sandhi regularly occurs and recurs both morpheme-internally and across all sorts of morphological and word boundaries. “[P]ostalveolarization is a recursive or iterative rule, i.e., it can in most cases apply and reapply as long as there occur dentals in a direct succession after an /r/ ... (a) morpheme-internally, (b) in internal and (c) in external sandhi” (Eliasson 1986: 279). If this is true, then sandhi is an across-the-board process (a P2 or fast speech rule), and not sensitive to the syntax involved (not a P1 rule)(cf. Kaisse 1985, 1990, for discussion of the distinction between P1 and P2 rules). This can be formulated in the same words as in the previous chapter, as Hypothesis G.

Hypothesis G

Sandhi occurs regularly through sequences of dental consonants both within and between morphemes

Eliasson (1986: 282) also suggested that “frequency of word collocation” could be a factor in sandhi. This can be formulated as in Hypothesis R.

Hypothesis R

Sandhi is favored within frequent word collocations

In Lindblad (1987a), I found that the word pair (*för X*) *år sen* '(X) years ago' always was linked by /rs/-sandhi in all of the speakers in Borås, who used this phrase twelve times in total. This is clearly a frequent collocation of words, and hence is a specific, testable case of Hypothesis R, and can be formulated as in Hypothesis S.

Hypothesis S

Sandhi is required within the phrase (*för X*) *år sen*

Since sandhi is a process that involves direct contact between sounds and creates a linked structure (cf. Figure 8), it is only natural that interrupting the continuity of phonation should inhibit it (if it did not create a linked structure, then pause could be a criterion for determining whether sandhi is a P1 or P2 rule (Kaisse 1990: 131)). The blocking effect of pauses may be formulated as in Hypothesis T.

Hypothesis T

Sandhi does not occur across pauses and hesitations

In Chapter 4 we observed that Eliasson lists three types of exceptions to the generalizations represented by Hypothesis G. The first involves geminate /rr/, which "generally does not cause postalveolarization of the following dentals." (Eliasson 1986: 280) Since he does not specify that this applies only word-internally, we can test it on our external sandhi data in the same formulation as we used there in Hypothesis I.

Hypothesis I

Sandhi does not occur following geminate /rr/

He further observed that the normally free recursion of sandhi (i.e., the spread of the feature [-anterior]) through a string of dentals “does not ... extend to non-laterals that follow a lateral,” (Eliasson 1986: 280), for which he gave only word-internal examples. However, since he does not specify that this applies only word-internally, we can test Hypothesis J too on our external sandhi data.

Hypothesis J

Sandhi with non-laterals does not occur following /r/

Eliasson (1986: 281) notes that certain words “tend not to undergo the process.” He observes that these words are mostly of foreign origin and mostly contain /u/ immediately preceding the sandhi site, and he gives as examples the words *absurd*, *urdu*, *nokturn*, *Saturnus*, and *Ursula*. Although he lists only examples that are word-internal, we can still extract from this and test the hypothesis that the vowel /u/ preceding the sequence /rs/ is an unfavorable environment for external sandhi. We can adapt it by slightly reformulating Hypothesis K in the form of Hypothesis U.

Hypothesis U

Sandhi tends to not occur when the vowel /u/ immediately precedes the site

In Lindblad (1987b), I found a tendency for my informant, who was from a small town west of Gävle in the southern part of Norrland, to produce /rs/-sandhi at a somewhat greater rate when the second word began with the sequence /sm-/. Although I have no good explanation for why that occurred, drawing on those results I hypothesize that the same will be true also in NN, and formulate it as Hypothesis V.

Hypothesis V

/rs/-Sandhi tends to occur more frequently when the following segment is /m/

In the course of gathering data for Lindblad (1987b), my informant expressed the opinion that in natural speech he would produce sandhi between *har* ‘has/have’ and a following verb more often than between the modal *får* ‘may’ and following verbs, and more often following *får* than between copulas and following verbs, and that each was a subset of the previous one. The data that I collected from him seemed to fit this model, but since he only consented to give opinions about sandhi after *får* and the copulas before verbs that had already produced sandhi following *har*, the methodology was highly suspect. Also, we have no reason to believe that speakers have conscious access to the factors that govern their choice to apply sandhi. Nonetheless, his opinion does provide a hypothesis that is testable with our data, which we will formulate as Hypothesis W.

Hypothesis W

/rs/-Sandhi is favored for *har* > modals > copulas, when followed by verbs

In the next section we will discuss briefly the factors that will be used to analyze the data and test these hypotheses.

5.3 FACTORS IN EXTERNAL /RS/-SANDHI

The demographic factors in this chapter are the same as those described in Chapter 2, namely the gender, age, and dialect area of the speakers. The linguistic factors to be considered here are stress, c-command, constituent edge, syntactic category, lexical items, frequency and closeness of word collocations, and the segments adjacent to the sandhi site.

By stress I mean the degree of stress on the syllabic nuclei immediately preceding and following the sandhi site.

By c-command is meant standard branching-node c-command: α c-commands β iff α does not dominate β and the first branching node which dominates α also dominates β .

By constituent edge is meant that the pair of words lie on the edge of the constituent that contains them.²⁷

By preceding and following syntactic category, I mean the grammatical class category of the morphemes containing respectively the /r/ and /s/ that potentially participate in /rs/-sandhi, and that immediately precede and/or follow its site.

By lexical items I mean the words that contain the internal sandhi site, regardless of their syntactic function (for example, *som* 'as, that' is treated as the same lexical item, regardless of whether it functions as a conjunction or a pronoun) . It should be noted that the requirement of the Varbrul program that a factor group have no more than 30 members again played a constraining role in deciding how to code these items, since there were 144 potential sandhi sites, and grouping the recurring words together did not bring the number of distinct forms down to 30. Therefore a number of items have been lumped together into relatively homogenous groups of various sorts, largely based on word-class criteria.

Frequency and closeness of word collocations are somewhat amorphous concepts, but are invoked by Eliasson 1986 as possible factors for sandhi. Obviously you cannot look up word collocations in a frequency dictionary in the same way that you can look up single words in such books for various languages, including Hassler-Göransson (1966) and Allén (1970) for Swedish, so properly analyzing these factors requires the judgement of an insightful native speaker.²⁸

By adjacent segments I mean those immediately preceding and following the potential /rs/-sandhi site.

²⁷ I am grateful to Anders Holmberg for going through my two stories, MfoN and Msi, and rendering his judgements about the c-command and edge relations of each of their possible external /rs/-sandhi contexts.

²⁸ I am grateful to Stig Eliasson for going through my two stories, MfoN and Msi, and rendering his judgements about the frequency and closeness of the word collocations in both stories.

5.4 THE EXTERNAL /RS/-SANDHI DATA

A summary of the results for each of the 144 potential external /rs/-sandhi contexts from MfoN and Msi is detailed in Appendix H. The first column of Table 1 gives the preceding word, which contains the /r/. The second column gives the following word, which contains the /s/. The third column gives the Item number of the context, which allows you to locate it in Appendixes A and B (MfoN) or Appendixes C and D (Msi) in order to see its broader context, morphological analysis, translation, etc. The fourth column gives the [ʂ] results, meaning that I heard the phonemic /rs/ sequence pronounced as the phonetic retroflex [ʂ] sandhi product. The fifth column gives the /rs/ results, meaning that I heard the phonemic /rs/ sequence pronounced as a sequence of those two phonemes, without sandhi. The sixth column gives the ‘maybe’ results, meaning that I could not decide whether sandhi had occurred or not.²⁹ The seventh column gives the ‘pause’ results, meaning that I perceived a complete break in the stream of phonation, often marked by audible inhalation. The eighth column gives the ‘omitted’ results, meaning that the speaker did not audibly articulate one or both of the words in question, whether or not they provided a substitute. The ninth column gives the ‘other’ results, meaning a variety of things, such as that other phonetic material was inserted between the /r/ and /s/, the speaker “corrected” the pronunciation, an onlooker provided the “correct” pronunciation, etc. The tenth column gives the ‘total’ results, meaning the number of all the responses from all my readers, including also both of the two responses from those who read the same story a second time in “dialect”, even though I have not used the “dialect” readings that I collected (20 for MfoN, 1 for Msi) for statistical purposes in this study. On the other hand, however, I have used both readings by the two men in NN

²⁹ The reasons for this varied, with the main ones being background noise that was too loud and pronunciations that seemed to be intermediate between a dental and a postalveolar articulation. It could be interesting to do a narrow phonetic analysis of these intermediate pronunciations, but that is beyond the scope of the present study.

whom I encountered a second time with my second story, and who thus read each story once.

In my statistical analysis in Varbrul, I used only the data in the [s] (sandhi occurred) and /rs/ (sandhi did not occur) columns, and all other results have been discarded. Although I have not attempted an analysis of the variety of other results for certain items that this approach does not include, it may be worth noting that my orthographic punctuation marks <,> and <.> did coincide with a natural pause for many speakers for several items, such as MfoN 31 and 55 and Msi 102, 135 and 177. The results for a number of other items were more ambiguous. Perhaps my non-native-speaker tendencies in writing Swedish were ignored by many of the readers in cases such as MfoN 50 and Msi 4, 23, 72, 99, 250, and 255 where the majority apparently did not recognize a need for a pause (or perhaps the pause in such cases was so slight that I did not perceive the complete cessation of phonation, which was my criterion for pause). Eva Strangert (p.c., 1991), a phonetician at Umeå University who has studied pauses extensively, has kindly recorded her own reading of MfoN and then measured the length of pauses at six sites. There is general agreement between the length of pauses in her reading and the number of pauses that I found in my data, though the correlation is not exact, as may be seen in the following summary table.

Table 16. Eva Strangert's measurements of pause length in MfoN

Lexical Item A	Lexical Item B	Item #	# Pauses	Pause length
förorter,	, så	MfoN 31	198	475 ms.
syster,	, som	MfoN 50	33	150 ms.
snart.	. Sara	MfoN 55	239	1320 ms.
tusenkonstnär,	, skulle	MfoN 80	133	115 ms.
Helsingfors,	, Sundsvall	MfoN 101	93	790 ms.
semestern,	, så	MfoN 122	102	415 ms.

Already now, before we even begin analyzing our data for external sandhi with Varbrul, we can make observations that are germane to our Hypotheses, at least in four cases where only one or two Items from the stories are involved.

Item # MfoN 6 was specifically intended to test Hypothesis S (Sandhi is required within the phrase (*för X*) *år_sen* '(X) years ago'), which was found in Lindblad (1987a) to be true in every instance in the casual speech of speakers in Borås. Our unfiltered data in Appendix H (where duplicate readings by the same person have not been removed) show sandhi occurring at a 77% rate for this item, which is far above the 25% average rate for all external sandhi contexts in our data. Since that data gathered in Borås by the Talspråksyntaxprojekt (Spoken language syntax project) consisted entirely of spontaneous speech, while our primary set of data from NN consists entirely of readings, it is plausible to attribute the difference in results to a difference of register between casual speech and reading register, and to assume that there is no difference between the dialects on the item MfoN 6 *år_sen* 'years ago'. On the other hand, it would also be possible to attribute the difference entirely to a difference between these dialects. However, I made an additional observation during my interview of a man in his 60's in Dialect Area 2 that seems relevant here. After reading the story with no sandhi on MfoN 6, he then went on to chat casually with me for a few minutes about various things, and in the process used the phrase *år_sen* 'years ago' three times, producing sandhi every time! That points toward the difference of frequency of sandhi on this item being primarily a matter of register, not dialect. It also suggests a possible metric for the conversion between reading register and casual speech: all of the items where we got results of more than about 75% sandhi in readings might get 100% sandhi in casual speech (e.g., MfoN 28, 46, 85, Msi 123, 191, 276, 280). Unfortunately, our data here are inadequate to test that hypothesis, so it must be left for future research.

Items # Msi 72, Msi 177 are relevant to Hypothesis I (Sandhi does not occur following geminate /rr/). The frequent pauses following Msi 177 render the data from it almost insignificant. However, when combined with the data from Msi 72, the low overall 3% rate tends to support Hypothesis I. Unfortunately the strength of this result is weakened by the fact that both items Msi 72 and Msi 177 contain an orthographic punctuation mark, so it is possible that the low rate of sandhi should be attributed to pauses, even though I

did not perceive pauses for most instances of Msi 72.³⁰ Thus our evidence for Hypothesis I must be regarded as only provisional.³¹

Items Msi 235 and Msi 261 do not lend support to Hypothesis J (Sandhi with non-laterals does not occur following /r/), since the 40% rate of occurrence of sandhi on these items is not only higher than the 25% overall rate for all of our data, but also is higher than the overall 32% rate when *som* is the second lexical item.

Items Msi 99 and Msi 135 support Hypothesis U (Sandhi tends to not occur when the vowel /u/ immediately precedes the site). In fact, the results were so clear-cut that Varbrul signaled this as a Knockout factor, meaning that when /u/ preceded the potential /rs/-sandhi site, the result was always the same, and it was impossible to do statistical calculations. Unfortunately the strength of this result too is weakened by the fact that both items Msi 99 and Msi 135 contain an orthographic punctuation mark, so it is possible that the low rate of sandhi should be attributed to pauses, even though I did not perceive pauses for most instances of Msi 99. Thus our evidence for Hypothesis U must be regarded as only provisional.

5.5 TABLES AND GRAPHS OF RESULTS

Let us now begin to examine our data for external /rs/-sandhi in the light of these hypotheses, starting with the two genders observed across ten-year age groups in Table 17 and Figure 24.

³⁰ Stig Eliasson (p.c. 1999) considers sandhi to be impossible on Msi 72.

³¹ I regret no having anticipated all the possible questions that might arise. For example, I took the internal sandhi prohibition against sandhi after /r/ from Eliasson (1986) and included a test example in MfoN, but at the time I wrote the stories I did not anticipate the question of whether this could also be a prohibition in external sandhi. I regret this lack of foresight, and hope only that it is compensated for in some small part by the possibilities for serendipitous discoveries created by my choice of methodology.

Table 17. External /rs/-sandhi, by gender and age in decades

Age	Male		Female		Both	
< 20	521	33%	399	23%	920	28%
20-29	894	26%	686	23%	1580	25%
30-39	567	28%	385	25%	952	27%
40-49	605	27%	478	29%	1083	28%
50-59	326	20%	260	22%	586	21%
60-69	377	25%	133	22%	510	24%
70-84	144	19%	112	28%	256	22%
Overall Totals	3434	26%	2453	25%	5887	25%

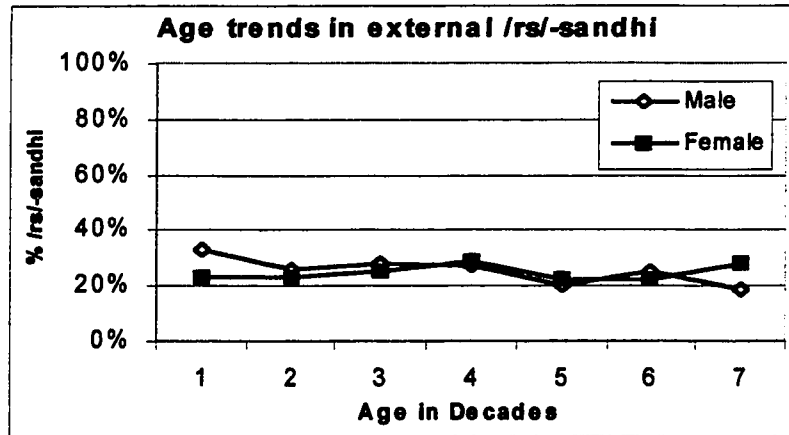


Figure 24. External /rs/-sandhi, by gender and age in decades

It appears from Figure 24 that there is very little difference between males and females at any age. Figure 24 seems to show a slight trend of increased usage of external /rs/-sandhi by younger male speakers and older female speakers, but neither one runs in a straight line and there is no clear overall pattern. Breaking this down into five-year age cohorts shows some differences as large as 11% between male and female speakers, but again there is no clear pattern, so I will not take the space to present the table or figures here. In brief, the most favorable cohort for males was aged 15-19 and for females was aged 40-44, while the least favorable cohort for males was aged 55-59 and for females was aged 65-69.

Next let us look at the patterning of gender across the dialect areas (for a map of the dialect areas, see Figure 9) in Table 18 and Figure 25.

Table 18. External /rs/-sandhi, by gender and dialect area

Dialect Area	External Sandhi				Combined	
	Male		Female			
U	64	14%	74	19%	138	16%
1	438	23%	408	25%	846	24%
2	124	17%	172	20%	296	19%
3	530	27%	308	25%	838	26%
4	885	35%	668	30%	1553	33%
5	426	28%	229	31%	655	29%
6	313	28%	149	26%	462	27%
7	292	28%	218	30%	510	29%
8	218	23%	83	13%	301	19%
9	110	17%	146	16%	256	17%
10	98	13%	72	13%	170	13%
Overall Totals	3498	26%	2527	24%	6025	25%

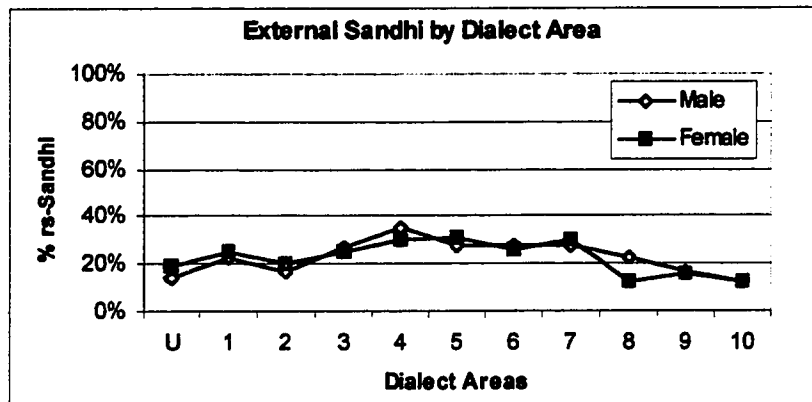


Figure 25. External /rs/-sandhi, by gender and dialect area

Figure 25 shows very little difference between males and females except in DA-8. Usage of external /rs/-sandhi appears to be slightly higher in the center of NN (Dialect Areas 3-7). The general contours of Figure 25 show NN as an apparent dialect continuum, though the differences between the results for Uppland and our DA-1 (or 2) are not especially striking considering the considerable distance between them. I also looked at the

patterning of the ten-year age groups across the dialect areas, but it revealed no clear patterns, and therefore will not be presented here.

We have now concluded our examination of population factors in relation to each other. The majority of our remaining comparisons will pair one of the population factors (most often the two genders) with one of the various linguistic factors. Let us now look at the pronunciations of the preceding lexical items in Table 19 and Figure 26. The items in the table and figure have been arranged by part of speech in the following sequence: noun, pronoun, adverb/preposition, conjunction, verb.

Table 19. External /rs/-sandhi, by first lexical item and gender

First Lexical Item	Gloss	Male		Female		Both	
a. Noun - monosyllabic		175	38%	138	42%	313	39%
b. Noun - multisyllabic		325	34%	248	35%	573	34%
c. Noun - plural		148	21%	110	20%	258	21%
d. Sigvard	(personal name)	276	57%	235	62%	511	59%
e. styvfar	stepfather	64	17%	74	26%	138	21%
f. -r̄ (= karl)	man	34	30%	22	25%	56	28%
g. -rr̄ (=kärr / tyvär̄r)	marsh / too bad	4	5%	1	1%	5	3%
h. vår̄(t)	our	202	33%	144	30%	346	32%
i. stor	large	113	93%	77	86%	190	90%
j. snart / säker	soon / certain	68	57%	45	51%	113	54%
k. här	here	121	56%	80	48%	201	52%
l. där	there	109	37%	79	35%	188	36%
m. när	when	5	2%	7	4%	12	3%
n. ner	down	64	21%	78	32%	142	26%
o. över	over	237	39%	142	31%	379	35%
p. efter	after	70	41%	58	44%	128	42%
q. (där)för	(there)for	103	32%	68	29%	171	31%
r. eller	or	14	13%	5	6%	19	10%
s. Verb - monosyllabic		247	22%	167	19%	414	21%
t. Verb - present tense		703	26%	478	23%	1181	24%
u. gjort	made	93	77%	57	68%	150	73%
v. har [auxiliary]	has / have	130	8%	69	5%	199	7%
w. är [copula]	is / are / am	129	11%	71	8%	200	10%
Overall Totals		3434	26%	2453	25%	5887	25%

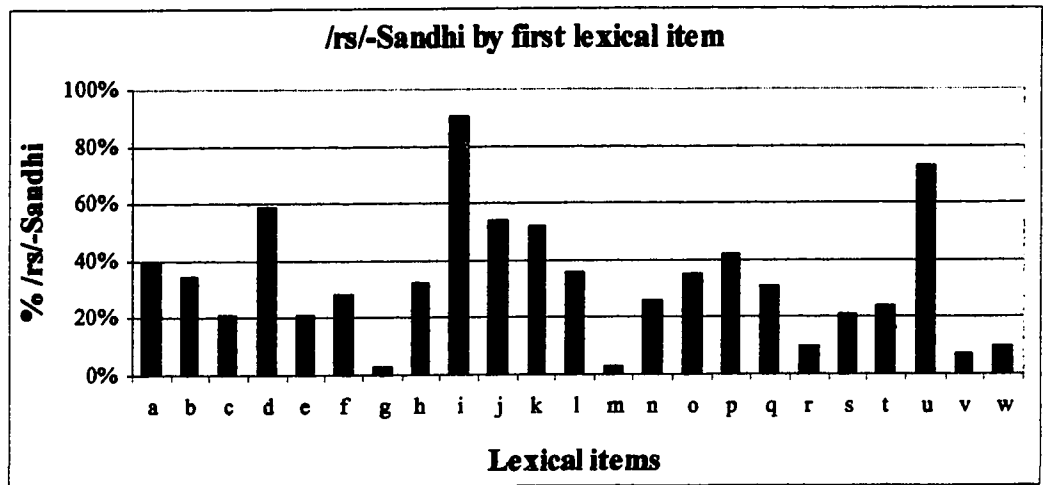


Figure 26. External /rs/-sandhi, by first lexical item

When we look at these results in Table 19 and Figure 26 in more detail, we can observe several points. Since there is very little difference between male and female speech in regard to /rs/-sandhi here, only the combined figures are presented in Figure 26. Looking at cases where another dental consonant is interposed between /r/ and /s/, we see that the sandhi rate for Lexical Item d is 59%, for Lexical Item f is 28%, and for Lexical Item u is 73% (cf. also MfoN 55, 122 Msi 114, 186, 194). Thus the regularity of application postulated by Hypothesis G (Sandhi occurs regularly through sequences of dental consonants both within and between morphemes) is supported to the extent that intervening consonants in general seem to play no role in any irregularity, and certainly do not block /rs/-sandhi. On the other hand, the overall total percentage rate of 25% here makes it apparent that Hypothesis G is not supported as strongly for external /rs/-sandhi as it was in the case of internal sandhi, where the corresponding figure was 79%. Since more than two-thirds of the preceding Lexical Items show external /rs/-sandhi at a rate below 40%, it is clear that sandhi does not apply regularly between all words, and is applied at a high rate only for a minority of Lexical Items.

Repeating observations that we made already directly from the data in Appendix H, when we look in Table 19 and Figure 26 we find confirmation for Hypothesis I (Sandhi does

not occur following geminate /rr/) in our results for Item g. *kärr* ‘marsh’ / *tyvärr* ‘too bad’, with only 3% sandhi overall. Although Item g. *kärr* ‘marsh’ / *tyvärr* ‘too bad’ clearly showed the least frequent usage of sandhi, it was tied for that honor with Item m. *när* ‘when’ (also 3%), and others were nearly as low, including Item v. *har* ‘has/have’ [auxiliary verb] (7%), Item w. *är* ‘is/are/am’ (10%), and Item r. *eller* ‘or’ (10%). On the other hand, we do not find confirmation for Hypothesis J (Sandhi with non-laterals does not occur following /r/) in our results for Item f. *karl* ‘man’, which has 28% sandhi and thus is slightly above the overall rate.

I have also looked at pronunciations of the preceding lexical items as a factor across dialects, but it does not present a clear picture or contribute greatly to our understanding of the factors influencing sandhi, so I will not include it here. I will mention only that the graph of preceding lexical items across dialects makes clear that the lower rate of sandhi in DA-2 as compared to its neighbors applies to many lexical items.

The varied results within the different parts of speech in Figure 26 suggest it not to be an important factor. However, there are questions about the effect of syntactic category that the arrangement of these lexical items cannot answer, so let us now look at the patterning of the preceding syntactic categories in Table 20 and Figure 27.

Table 20. External /rs/-sandhi, by preceding syntactic category and gender

Syntactic category	Male		Female		Both	
	Count	%	Count	%	Count	%
Subject noun	102	18%	85	20%	187	19%
Predicate nominative	16	37%	15	35%	31	36%
Direct object noun	45	43%	36	43%	81	43%
Other noun	859	36%	691	38%	1550	37%
Pronoun	202	33%	144	30%	346	32%
Determiner	91	41%	49	29%	140	35%
Adjective	181	75%	122	69%	303	72%
Adverb	290	28%	247	30%	537	29%
Preposition	330	40%	212	35%	542	38%
Conjunction	15	8%	7	5%	22	6%
Main verb	777	26%	568	25%	1345	25%
Auxiliary <i>har</i> 'has/have'	121	8%	60	5%	181	7%
Modal verb	241	28%	132	20%	373	24%
Copula	164	11%	85	7%	249	9%
Totals	3434	26%	2453	25%	5887	25%

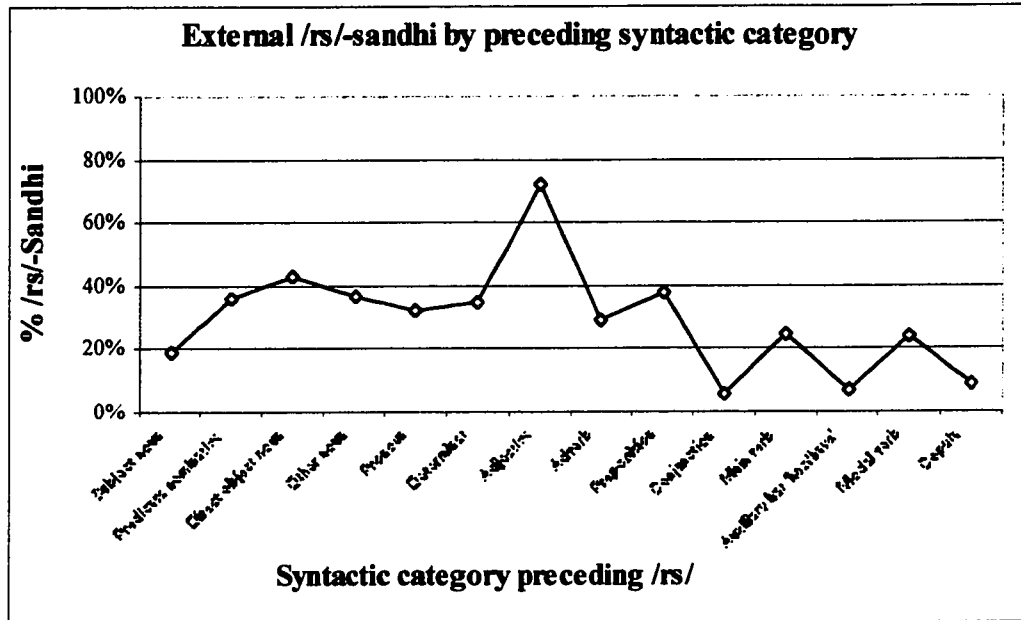


Figure 27. External /rs/-sandhi, by preceding syntactic category

The pattern for both genders in Table 20 is very similar, once again, so only the combined figures have been used in Figure 27. The most striking point on Figure 27 is

the high rate of sandhi for preceding adjectives, for which we seem to have no relevant hypothesis.

The other observation that we should make is the pattern shown by the last three syntactic categories of Table 20 and Figure 27, which are subsets of verbs. These show that our data do not support Hypothesis W (/rs/-Sandhi is favored for *har* > modals > copulas, when followed by verbs). This hypothesis was based on a conjecture about his own behavior by my informant for an earlier paper (Lindblad 1987b), so the failure of our data to support this hypothesis can be interpreted either as a difference between dialects in southern Norrland and NN, or it can be interpreted as showing that speakers do not have conscious access to frequency judgements of their own behavior. In our data as represented in the last three syntactic categories, the hierarchy is instead: modals > copulas > *har*, with very little difference between the last two. However, the syntactic category of copula in Table 20 and Figure 27 is not limited by the category of following words. A closer look at the data for copulas reveals that only one context preceded a verb, MfoN 37, where the rate was 3%. Therefore the more accurate hierarchy for our data is modals > *har* > copulas, although again there is not a great difference between the last two categories, and our quantity of data for copulas preceding verbs is too minimal to be dependable.

Let us now look at the pronunciations as a factor of the following lexical items in Table 21 and Figure 28. These lexical items have been arranged by part of speech in the following order: noun, pronoun, number, adjective, adverb, verb.

Table 21. External /rs/-sandhi, by gender and second lexical item

Second Lexical Item	Gloss	Male		Female		Both	
a. Sara	(personal name)	27	11%	10	6%	37	9%
b. Sigvard	(personal name)	5	2%	2	1%	7	2%
c. Proper name - other		129	27%	59	17%	188	23%
d. Sveriges / svenska / Svensson	Sweden's / ...	63	20%	28	12%	91	17%
e. samerna(s) / samiska	Samis / Samish	15	4%	9	3%	24	4%
f. staket(et)	(the) fence	188	64%	127	59%	315	62%
g. stad / stan	city / the city	170	74%	116	67%	286	71%
h. surströmming	sour herring	68	34%	42	30%	110	32%
i. Noun - other		457	34%	301	29%	758	32%
j. sina	their	32	20%	44	36%	76	26%
k. sej / sig	himself	322	52%	289	60%	611	55%
l. som	that	385	30%	296	30%	681	30%
m. sånt / somliga	such / some	26	12%	23	14%	49	13%
n. Numeral		26	13%	13	9%	39	11%
o. säker	certain	37	9%	33	11%	70	10%
p. Adjective - other		75	11%	32	6%	107	9%
q. så	so	289	24%	250	27%	539	26%
r. sällan	seldom	11	5%	4	2%	15	4%
s. Adverb - other		193	37%	129	34%	322	36%
t. stanna(t)	stayed	34	14%	16	9%	50	12%
u. söker / sökt	seeks / sought	34	16%	20	12%	54	15%
v. sysselsatt / sysslar	employed / works	16	8%	6	4%	22	6%
w. säger / säga(s)	says / to say	122	16%	113	19%	235	17%
x. sa	said	138	44%	118	50%	256	47%
y. sett	seen	28	6%	25	7%	53	6%
z. ska / skulle	will / would	160	43%	127	42%	287	43%
1. Verb - infinitive - other		179	33%	90	21%	269	28%
2. Verb - present tense - other		137	50%	105	50%	242	50%
3. Verb -supine - other		68	12%	26	6%	94	10%
Overall Totals		3434	26%	2453	25%	5887	25%

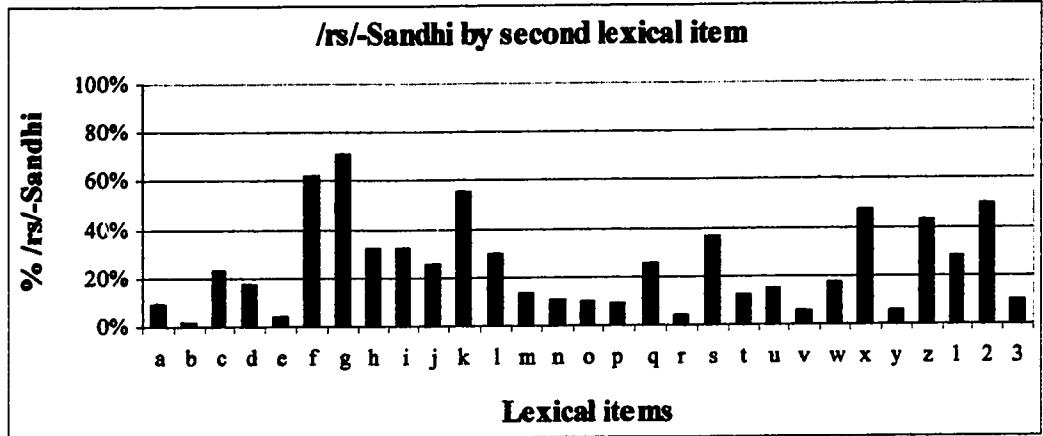


Figure 28. External /rs/-sandhi, by second lexical item

From Table 21 and Figure 28 we can see that nouns (Lexical Items a-i) produced both some of the lowest rates (b, e) and some of the highest rates (f, g). Similarly, pronouns (Lexical Item j-m) and verbs (Lexical Items t-3) produced a variety of results. It is clear that there is a complex set of factors involved in sandhi production. I have also looked at the pronunciations by the dialects as a factor of the following lexical items, but found only the same sorts of variation between dialects that we have noted previously.

Next let us look at the patterning of the following syntactic categories in Table 22 and Figure 29.

Table 22. External /rs/-sandhi, by following syntactic category and gender

Syntactic category	Male		Female		Both	
Subject noun	42	12%	19	7%	61	10%
Predicate nominative	51	9%	23	5%	74	7%
Direct object noun	191	35%	116	30%	307	33%
Other noun	861	36%	547	29%	1408	33%
Pronoun	590	38%	531	45%	1121	41%
Numeral	26	13%	13	9%	39	11%
Adjective	62	7%	30	5%	92	6%
Adverb	396	24%	310	25%	706	25%
Preposition	89	47%	68	50%	157	48%
Conjunction	191	24%	141	24%	332	24%
Main verb	775	21%	528	19%	1303	20%
Modal verb	160	43%	127	42%	287	43%
Totals	3434	26%	2453	25%	5887	25%

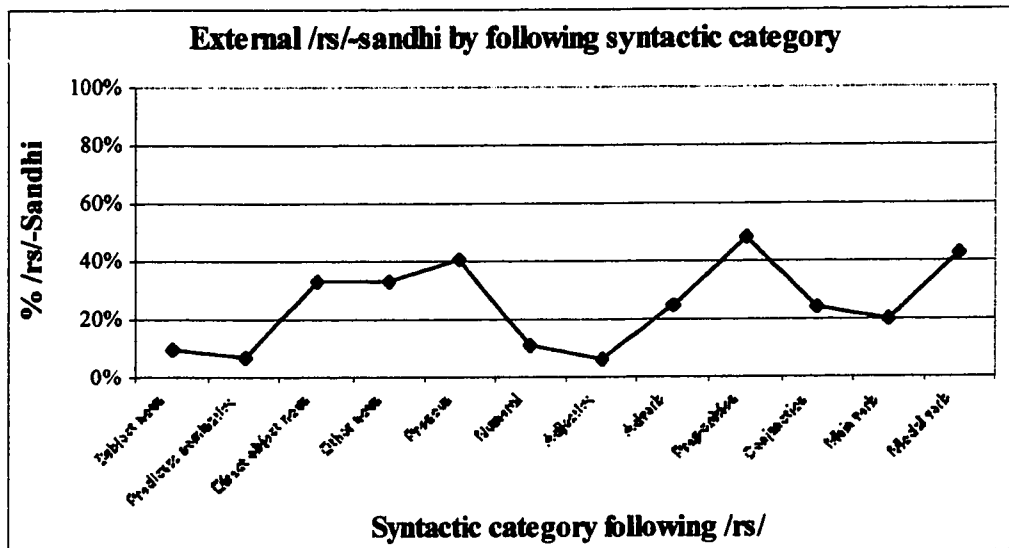


Figure 29. External /rs/-sandhi, by following syntactic category

Although the theoretical linguistic literature on the syntax-phonology interface has not found syntactic category to play a role, we may still be able to get some clues as to what is going on from this data display. The clear jump between subject nouns and object nouns suggests that a factor such as direction of c-command or edge of containing constituent may be in play, and some of the other categories can be interpreted in a similar way, so we will look at that next. It may be noted in passing that Figure 29 shows

an interesting contrast to Figure 27 as regards adjectives, which there preceding the sandhi site were the most favorable environment, but here when they follow the sandhi site go to the opposite extreme and provide the least favorable environment.

Since the direction of c-command and the edge of constituent were the same in almost all cases,³² we will present the results only for the direction of c-command in Table 23 and Figure 30. Those who believe that edge is the true determinant can instead read these results that way with only the smallest of distortions. We will separate our data by dialects to make the individual local situations clearer.

Table 23. External /rs/-sandhi, by direction of c-command across dialects

Dialect	Direction of c-command								Total	
	Left		Right		Both		Neither			
U	52	10%	18	72%	52	26%	16	23%	138	16%
1	382	18%	53	48%	282	36%	129	23%	846	24%
2	122	13%	21	45%	105	30%	48	24%	296	19%
3	370	20%	44	45%	261	38%	163	30%	838	26%
4	783	28%	74	53%	455	44%	241	32%	1553	33%
5	307	23%	33	48%	198	39%	117	30%	655	29%
6	211	20%	31	60%	158	40%	62	34%	462	27%
7	262	26%	26	51%	150	38%	72	25%	510	29%
8	135	14%	19	39%	94	25%	53	28%	301	19%
9	100	11%	10	21%	86	25%	60	26%	256	17%
10	77	11%	5	13%	48	18%	40	16%	170	13%
Totals	2801	20%	334	46%	1889	35%	1001	27%	6025	25%

³² There was a difference in only four items, of which two (MfoN 101, Msi 4) varied between Both and Neither. The other two (MfoN 3, MfoN38) varied between Left/Both and Neither/Left, respectively, so that the very low percentage of sandhi for each produced virtually the same net result in either analysis.

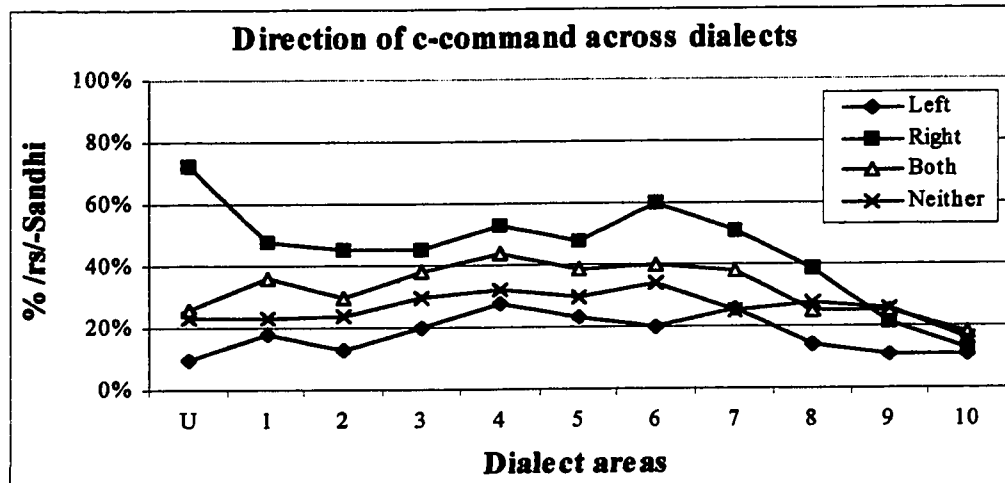


Figure 30. External /rs/-sandhi, by direction of c-command across dialects

The data in Table 23 and Figure 30 shows a pattern of preference for Right c-command over Left c-command in all dialects, with the options of Both and Neither falling in between in most Dialect Areas, except for the far north, outside the old Swedish-speaking region in Dialect Areas 9 and 10, where Both and Neither are slightly preferred options. Although our data do not show the categorical distinction based on direction of c-command (or edge) that are implied by the wording of the parameters and settings proposed by Kaisse (1985: 186), still c-command (or edge) does appear to be an important factor in determining the probabilities of occurrence of /rs/-sandhi in NN (and even more so in Uppland), with a clear preference for Right c-command (or edge) and a clear dispreference for Left c-command (or edge). The only exceptions are in Dialect Areas 9 and 10, where it seems to play very little if any role (and in fact Neither is the slightly preferred option in DA-9). The clear pattern of lines in Figure 30 continues smoothly from dialect to dialect except for the sharp jump between the geographically discontinuous DA-1 and Uppland, which may suggest that this really is a meaningful factor in the competence of these speakers.

Our data support the very general formulations for an influence from c-command and/or edge that were given above in Hypothesis M and Hypothesis N, since c-command and/or

edge does seem to favor sandhi. Based on our data, we may revise and further specify those hypotheses, as in Hypothesis M' and Hypothesis N'.

Hypothesis M'

External /rs/-Sandhi is favored when word *b* c-commands word *a*.

Hypothesis N'

External /rs/-Sandhi is favored by the words being on the right edge of the constituent that contains them.

Although Eliasson (1986: 282) does not formulate a specific definition of 'closeness' as applied to the transition between two words, he kindly agreed to assess the two word external /rs/-sandhi pairs in my stories. The categories he used were 'close', 'somewhat less close', and 'more distant', which I have simplified to 'Close', 'Moderate', and 'Distant'. Although he had no access to my data, his judgements fit the data rather nicely, as may be seen below in Table 24 and Figure 31.

Table 24. External /rs/-sandhi, by 'closeness' and dialect

Dialect	Degree of closeness						Total	
	Close		Moderate		Distant			
U	86	22%	49	12%	3	8%	138	16%
1	415	33%	418	20%	13	8%	846	24%
2	153	24%	139	16%	4	5%	296	19%
3	377	34%	446	24%	15	9%	838	26%
4	739	43%	787	28%	27	11%	1553	33%
5	308	39%	339	25%	8	7%	655	29%
6	231	32%	217	24%	14	18%	462	27%
7	241	40%	262	25%	7	8%	510	29%
8	161	24%	136	16%	4	5%	301	19%
9	117	20%	139	16%	0	0%	256	17%
10	71	18%	98	12%	1	1%	170	13%
Totals	2899	33%	3030	22%	96	8%	6025	25%

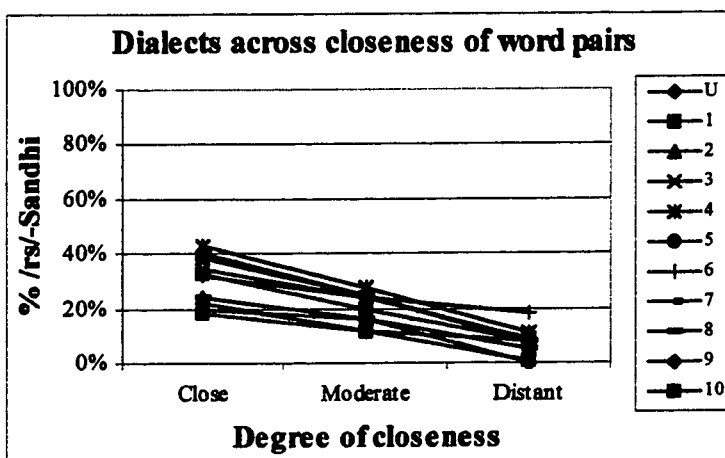
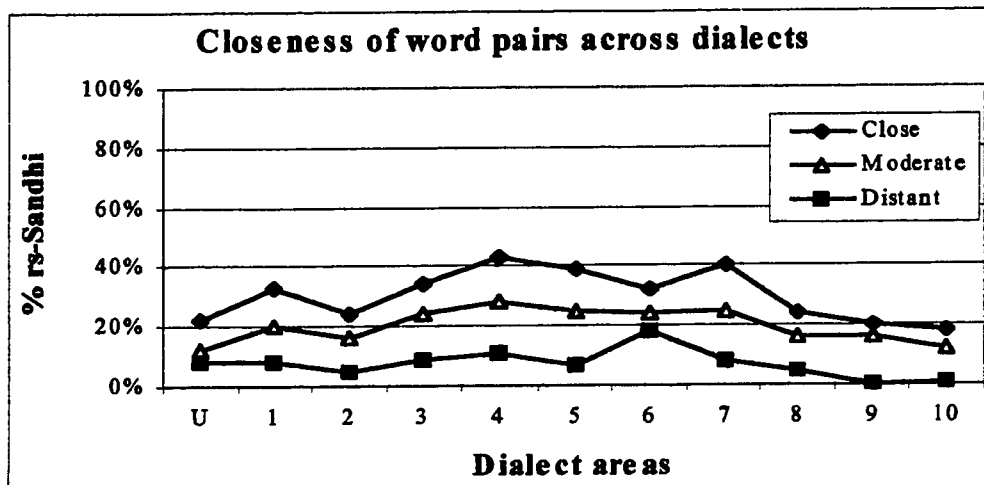


Figure 31. External /rs/-sandhi, by 'closeness' and dialect

Despite its somewhat vague formulation in Hypothesis O (External /rs/-Sandhi is favored by a close transition between the two words.), our data support 'closeness' as a meaningful factor in all of our dialects, although its importance varies between dialects. Many of the word pairs rated by Eliasson as 'distant' were rated by Holmberg as having c-command by Neither, and also were pronounced by many people with an intervening pause. Thus these disparate ways of analyzing the data are not entirely independent of each other, and it is not always possible to decide which is the best and most accurate way to describe the factors involved.

At this point, we may recall that this factor of closeness between words in external /rs/-sandhi is similar to the factor of morphological closeness in internal /rs/-sandhi that we looked at in Chapter 4 (cf. Table 15, Figure 23). Combining that previous set of data for boundaries in internal sandhi with the one immediately above for boundaries in external sandhi allows us to produce Table 25 and Figure 32.

Table 25. /rs/-Sandhi, by dialect and boundary type (internal & external combined)

Boundary type	Dialect Areas											Comb.
	U	1	2	3	4	5	6	7	8	9	10	
Monomorphemic	100%	96%	98%	99%	98%	96%	100%	93%	98%	93%	100%	97%
Inflectional	91%	89%	80%	90%	95%	97%	95%	97%	96%	85%	88%	92%
Derivational	100%	96%	93%	98%	95%	95%	93%	93%	92%	86%	89%	94%
Reduced compound	69%	70%	79%	73%	82%	81%	93%	78%	92%	58%	68%	77%
Regular compound	55%	78%	61%	72%	73%	69%	67%	69%	50%	51%	51%	67%
Close word pairs	22%	33%	24%	34%	43%	39%	32%	40%	24%	20%	18%	33%
Mid word pairs	12%	20%	16%	24%	28%	25%	24%	25%	16%	16%	12%	22%
Distant word pairs	8%	8%	5%	9%	11%	7%	18%	8%	5%	0%	1%	8%

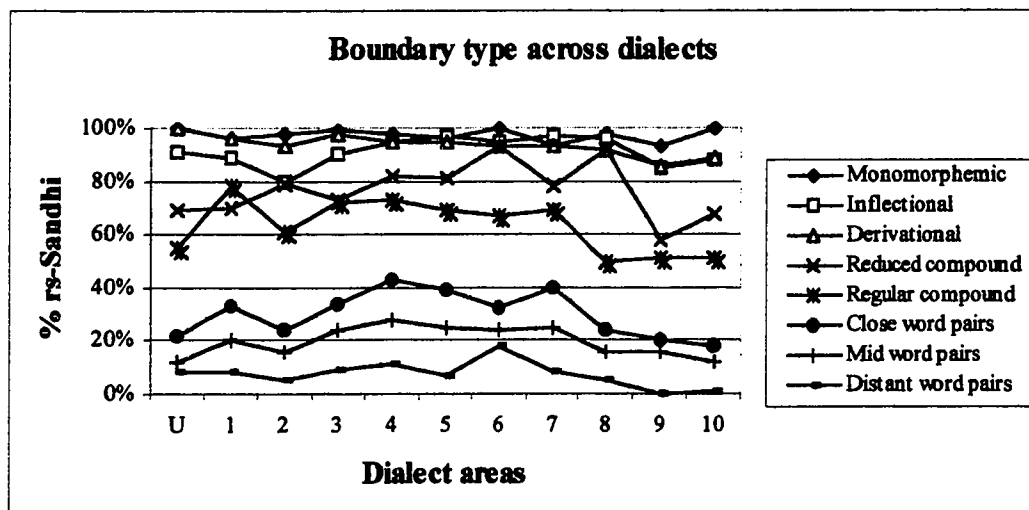


Figure 32. /rs/-Sandhi, by dialect and boundary type (internal & external)

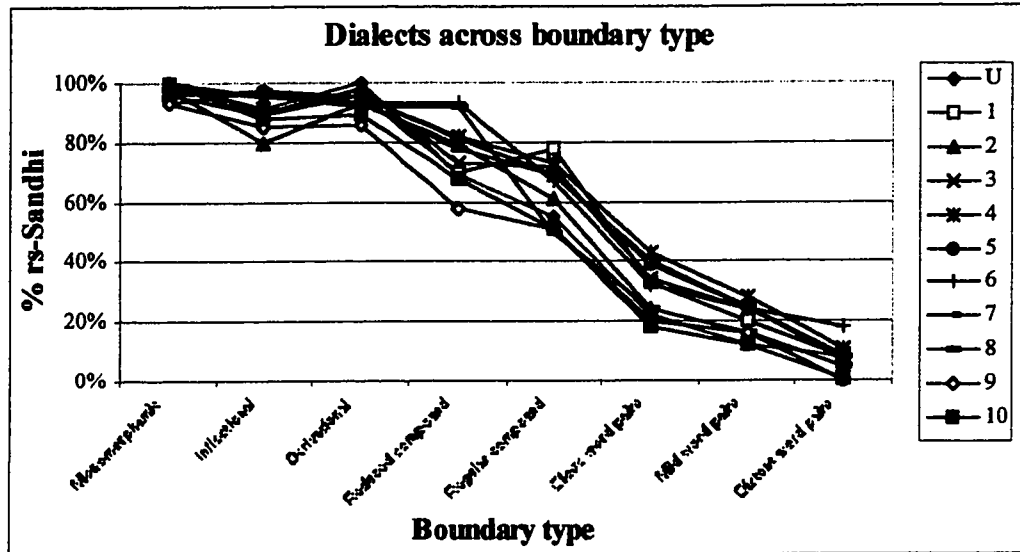


Figure 32. /rs/-Sandhi, by dialect and boundary type (internal & external)(continued)

Putting these two sets of data together produces a nice confirmation for Hypothesis P (Sandhi is favored morpheme-internally > inflectional boundaries > derivational boundaries > formal compounds > regular compounds > close word pairs > moderately close word pairs > distant word pairs). Figure 32, when looking at dialects across boundary types, shows a smooth transition from the categories of internal to external sandhi that supports Hypothesis P. Looking at boundary types across dialects reveals a remarkable parallel between the lines for regular compounds and close word pairs, and this suggests that the two are truly related to each other in a systematic way.

Eliasson (p.c. 1999) has favored me with his judgements of which pairs of words occur more frequently and which are rarer, allowing me to test Hypothesis R (External /rs/-sandhi is favored within frequent word collocations). The result of processing those judgements with Varbrul was that both the frequent and the rare pairs of words came out at exactly 25%, showing that frequency is probably not a meaningful factor and causing me to reject Hypothesis R. In the cases of dark-*sje* in Chapter 3, internal /rs/-sandhi in Chapter 4, as well as in the results I reported in Lindblad (1987b), we found no evidence that frequency was a meaningful factor, so our results here fit with what we knew before,

and lend support to the belief that we can rule out frequency as a factor in determining whether or not sandhi will occur.

Next let us look at the patterning of the vowel preceding /rs/ and gender in Table 26 and Figure 33.

Table 26. External /rs/-sandhi, by gender and the vowel preceding /rs/

Vowel	Male		Female		Both	
	Count	%	Count	%	Count	%
a	974	20%	732	20%	1706	20%
e	1182	31%	855	29%	2037	30%
i	23	29%	12	15%	35	22%
o	244	64%	157	59%	401	62%
y	31	36%	39	58%	70	45%
á	424	33%	282	29%	706	32%
ä	388	19%	256	16%	644	17%
ö	168	32%	120	30%	288	31%
Totals	3434	26%	2453	25%	5887	26%

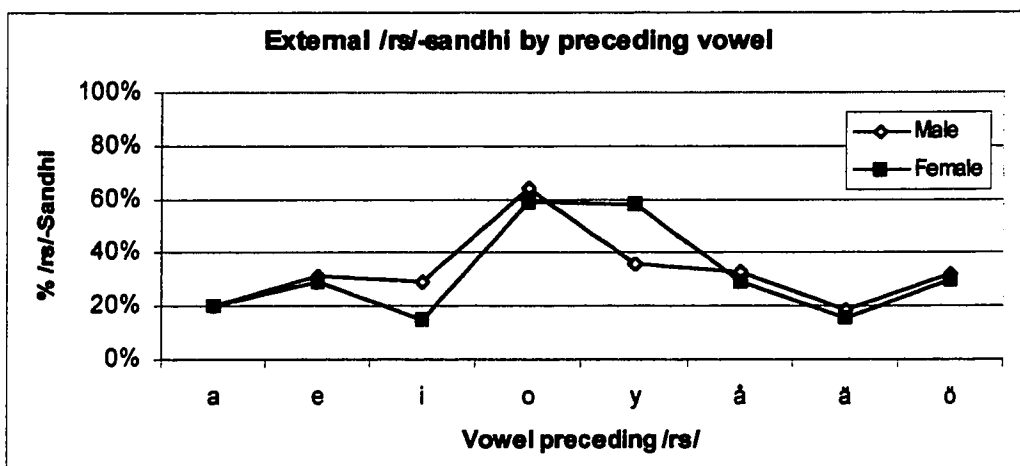


Figure 33. External /rs/-sandhi, by gender and the vowel preceding /rs/

You may observe that the vowel /u/ is not shown as one of the vowels preceding the sandhi site. The reason for this is that the vowel /u/ preceding the potential sandhi site served as a knockout factor for Varbrul because it was never followed by sandhi in our data, and therefore the data for contexts with a preceding /u/ had to be removed from the

database before the program could proceed. This suggests that Hypothesis U (Sandhi tends to not occur when the vowel /u/ immediately precedes the site) was rather strongly supported. Unfortunately, however, the clarity of this result was obscured because both of the two relevant contexts (Msi 99, 135) contained orthographic punctuation marks.

The low rate for /ä/ (17%) can be attributed to the large number (12) of occurrences of the copula *är* in our data, since the rate for /ä/ when they are omitted is 27%, slightly above our average. It is interesting to note that the high front vowels /i, y/ had similar rates for males (29% vs. 36%), but widely disparate rate for females (15% vs. 58%). I have no explanation for this, nor for why /o/ had the highest rate for both males and females.

Next let us look at the patterning of the vowel following /rs/ by gender in Table 27 and Figure 34.

Table 27. External /rs/-sandhi, by gender and the vowel following /rs/

Vowel	Male		Female		Both	
a	1003	32%	632	26%	1635	29%
e	591	28%	444	28%	1035	28%
i	187	25%	127	22%	314	24%
o	545	26%	393	25%	938	25%
u	335	43%	229	39%	564	42%
y	72	9%	44	7%	116	8%
å	418	29%	360	32%	778	30%
ä	223	14%	188	16%	411	15%
ö	60	16%	36	12%	96	14%
Totals	3434	26%	2453	25%	5887	25%

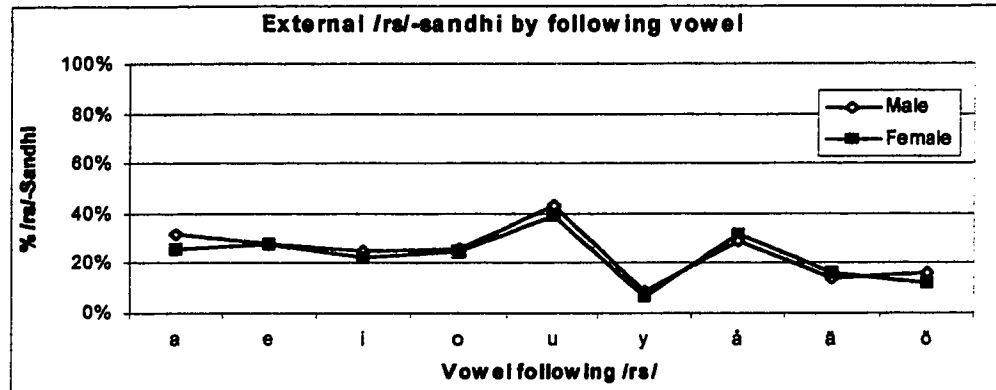


Figure 34. External /rs/-sandhi, by gender and the vowel following /rs/

It is remarkable that the vowel /u/, which was a knockout factor for vowels preceding /rs/ because it was never followed by sandhi, here is at the opposite extreme, providing the most favorable environment when it follows /rs/. Conversely, the vowel /y/ is the least favorable environment here, whereas it was one of the more favorable factors when it preceded /rs/. This may be somehow related to the fact that /u, y/ are both described as high front rounded vowels (see Figure 1), and that they therefore require careful attention to certain other articulatory details in order to maintain their differentiation, especially manner of lip rounding (cf. Wood 1982).

Next let us look at the interaction of the gender of the speakers and the segment following /s-/ in Table 28 and Figure 35.

Table 28. External /rs/-sandhi, by gender and segment following /s-/

Segment	Male		Female		Both	
a	200	18%	149	17%	349	18%
e	499	32%	404	34%	903	33%
i	37	9%	46	14%	83	11%
o	475	30%	352	29%	827	29%
u	94	35%	54	32%	148	34%
y	43	6%	22	4%	65	5%
á	296	23%	265	27%	561	25%
ä	170	12%	150	14%	320	13%
ö	34	16%	20	12%	54	15%
m	68	58%	45	51%	113	55%
v	102	26%	61	21%	163	23%
p	277	41%	173	34%	450	38%
n	59	21%	26	13%	85	17%
l	227	50%	106	27%	333	39%
t	581	32%	380	28%	961	30%
k	272	31%	200	29%	472	30%
Totals	3434	26%	2453	25%	5887	25%

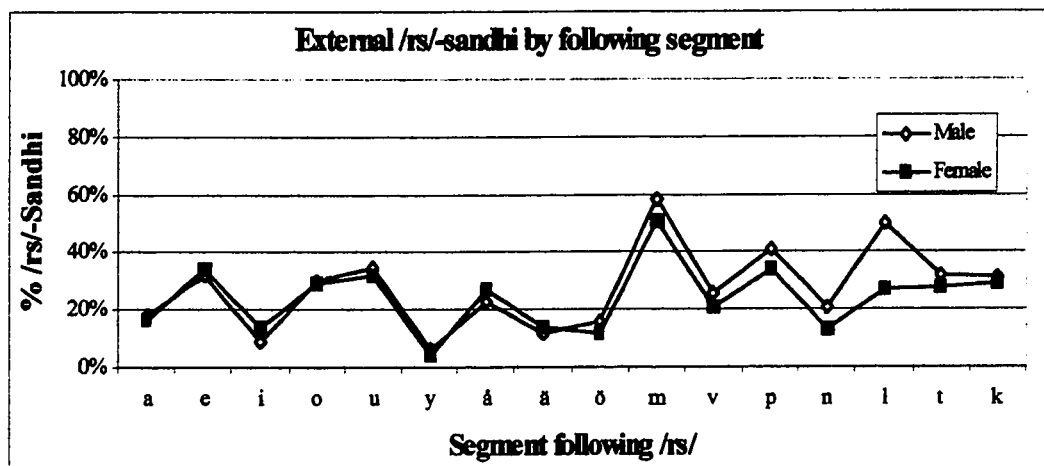


Figure 35. External /rs/-sandhi, by gender and segment following /s-/

Figure 35 shows support for Hypothesis V (/rs/-Sandhi tends to occur more frequently when the following segment is /m/), and at even a higher rate than was found in Lindblad 1987b. However, this result is based solely on one item (MfoN 70), so it may be influenced by other factors, and should not be considered robust. Figure 35 also shows a

considerable difference between male and female rates of /rs/-sandhi preceding /l/, which possibly may be related to a gender difference in production of a back variant of /l/ that occurs in some dialects of Swedish and is popularly known as ‘thick-l’; investigating that possibility is beyond the scope of this study.

Next let us look at the patterning of age in decades across stress on the preceding vowel in Table 29 and Figure 36.

Table 29. External /rs/-sandhi, by age and stress on the preceding vowel

Stress A	Age							Combined
	< 20	20-29	30-39	40-49	50-59	60-69	70-84	
Primary	46%	40%	40%	42%	27%	34%	34%	39%
Secondary	25%	22%	24%	23%	19%	22%	16%	22%
Unstressed	24%	21%	25%	27%	20%	23%	24%	23%
Totals	28%	25%	27%	28%	21%	24%	22%	25%

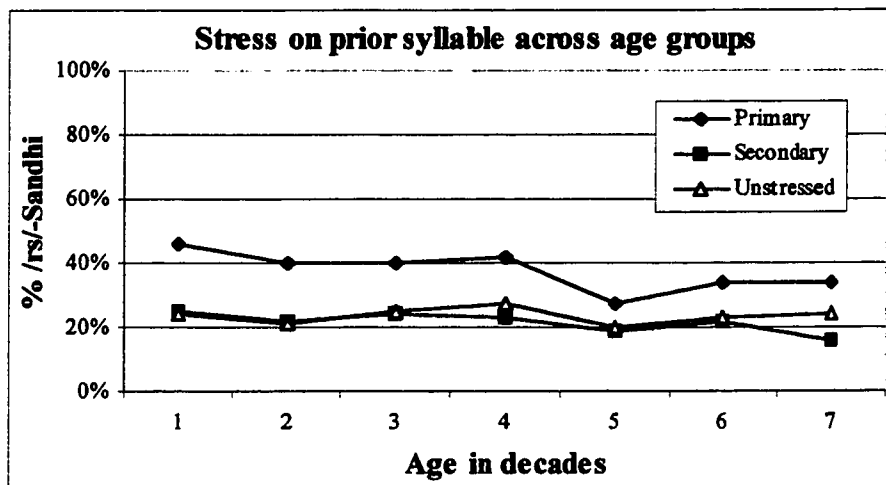


Figure 36. External /rs/-sandhi, by age and stress on the preceding vowel

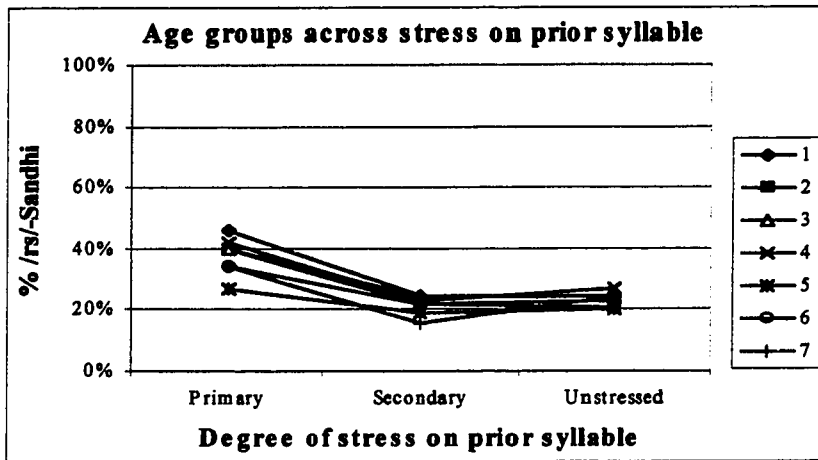


Figure 36. External /rs/-sandhi, by age and stress on the preceding vowel (continued)

We can see from Table 29 and Figure 36 that primary stress on the preceding syllable is the most preferred environment for all age groups, and by a good margin. The difference between secondary stress and unstressed syllables is less marked. Secondary stress is the least favorable environment for most age groups, except for the two youngest ones where sandhi is more favored than no stress by a margin of one percent. The only clear anomaly in the generally smooth trends is a relatively sharp drop in primary stress as a favorable factor for people in their 50's, whom we have referred to before as the WWII generation, and who have broken the pattern of surrounding age groups in other ways that we noted earlier. This instance argues again for their special status as contrarians and innovators.

Next let us look at the patterning of stress on the following vowel across the different age groups in Table 30 and Figure 37.

Table 30. External /rs/-sandhi, by age and stress on the following vowel

Stress B	Age groups							Combined
	< 20	20-29	30-39	40-49	50-59	60-69	70-84	
Primary	22%	18%	21%	21%	17%	20%	19%	20%
Secondary	52%	47%	42%	50%	41%	45%	42%	46%
Unstressed	38%	34%	39%	40%	25%	30%	29%	34%
Totals	28%	25%	27%	28%	21%	24%	22%	25%

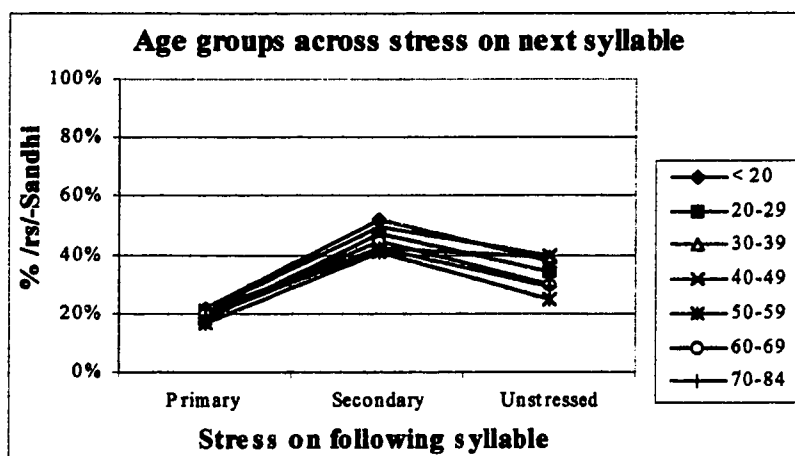
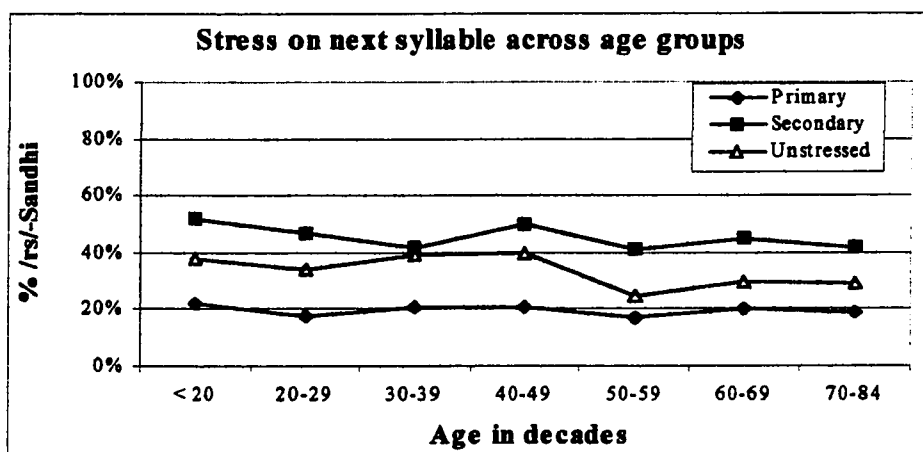


Figure 37. External /rs/-sandhi, by age and stress on the following vowel

From Table 30 and Figure 37 we can see that secondary stress on the following syllable is the most favorable environment for all age groups, and primary stress is the least

favorable. It is interesting that this is the opposite of what we saw for stress on preceding syllables above in Table 29 and Figure 36.

Since the effects of the two stress positions seem to be opposites, let us now look at them in relation to each other in Table 31 and Figure 38, for one final comparison.

Table 31. External /rs/-sandhi, by stress on preceding and following vowels

Stress B	Stress A						Combined	
	Primary		Secondary		Unstressed			
Primary	754	35%	1035	16%	1322	19%	3111	20%
Secondary	48	34%	360	50%	275	46%	683	46%
Unstressed	726	44%	632	31%	735	31%	2093	34%
Totals	1528	39%	2027	22%	2332	23%	5887	25%

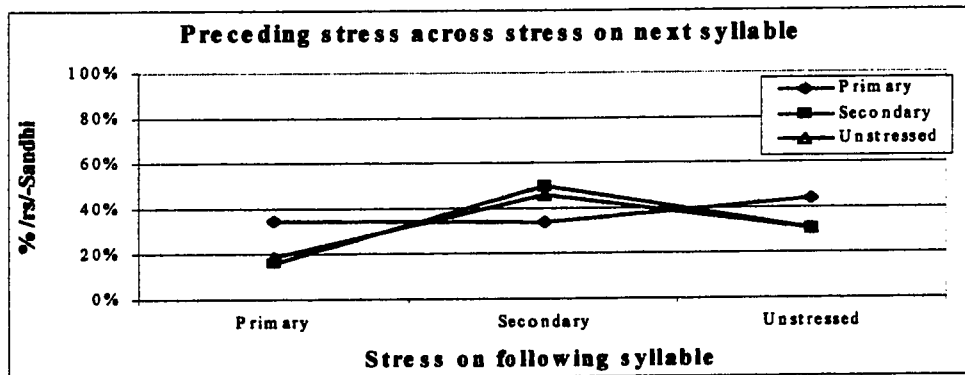
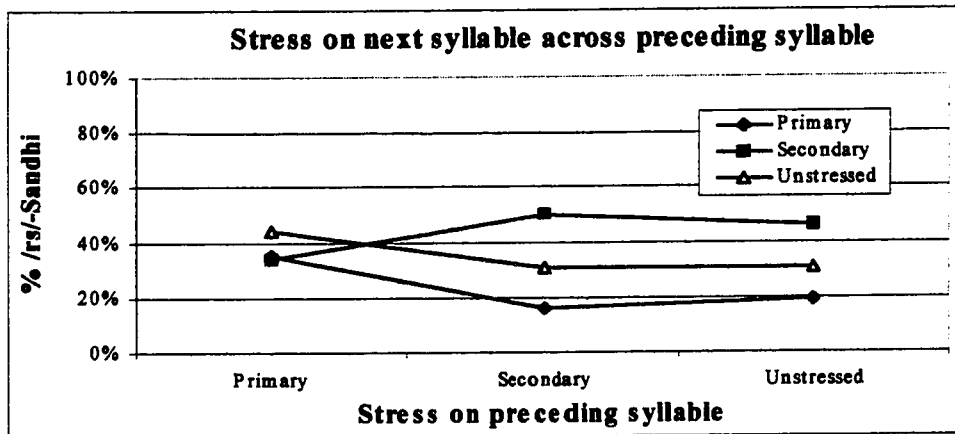


Figure 38. External /rs/-sandhi, by stress on preceding and following vowels

We see from Table 31 and Figure 38 that the relationship between preceding and following stress is complex. Primary stress on the preceding syllable is the most favorable environment if the following syllable has primary stress or is unstressed, but it is the least favorable environment if the following syllable has secondary stress. Secondary stress on the following syllable is the most favorable environment if the preceding syllable has secondary stress or is unstressed, but it is the least favorable environment if the preceding syllable has primary stress. Perhaps these relationships can be expressed by constraints, but to do so is complicated.

5.6 SUMMARY AND CONCLUSIONS FOR EXTERNAL /RS/-SANDHI

In broad terms, we have confirmed Eliasson's generalization that internal sandhi occurs more frequently than external sandhi, with the specific figures for our database being 79% vs. 25% overall. This is a natural result if one factor is the 'syntactic distance' between the two morphological entities that the /r/ and /s/ form a part of.

To summarize our more specific findings regarding external /rs/-sandhi, let us begin by looking back at the hypotheses that we started with, and review our results for each of them.

5.6.1 SUMMARY OF RESULTS FOR HYPOTHESES

Hypothesis M

External /rs/-sandhi may be favored by the existence (or direction) of c-command.

Hypothesis N

External /rs/-sandhi may be favored by the words being on an edge of the constituent that contains them.

Hypothesis M and Hypothesis N are equally supported by our data for all Dialect Areas except 9 and 10, with the preference being for right c-command or the right constituent edge. This led us to reformulate those hypotheses in more specific terms as Hypothesis M' and Hypothesis N'.

Hypothesis M'

External /rs/-Sandhi is favored when word *b* c-commands word *a*.

Hypothesis N'

External /rs/-Sandhi is favored by the words being on the right edge of the constituent that contains them.

Hypothesis O

External /rs/-sandhi is favored by a close transition between the two words.

Our data supported closeness of transition as being a favorable factor for sandhi in an analysis with three degrees of closeness, despite its vagueness.

Hypothesis P

Sandhi is favored morpheme-internally > inflectional boundaries > derivational boundaries > formal compounds > regular compounds > close word pairs > moderately close word pairs > distant word pairs.

Despite some difficulties in categorizing the types of internal boundaries that were discussed in Chapter 4, as well as there being a problem with the vagueness of the phrase 'closeness of transition', our data support this as a generalization across all of our dialects.

Hypothesis Q

/rs/-Sandhi may be favored by the stress pattern on preceding and/or following syllables.

Our data showed that /rs/-sandhi was most favored by primary stress on the preceding syllable and by secondary stress on the following syllable. It was least favored by secondary stress on the preceding syllable and by primary stress on the following syllable.

Hypothesis G

Sandhi occurs regularly through sequences of dental consonants both within and between morphemes

Sandhi effects can pass from /r/ through intervening /t, d, l/ to a following /s/ with no apparent diminution of frequency (there was no appropriate example of -/rn/ in our database). However, there is so much variation in the rate of application of /rs/-sandhi to the various Lexical Items that describing it with the word “regularly” is somewhat of an overstatement.

Hypothesis R

Sandhi is favored within frequent word collocations

Our data do not support Hypothesis R, just as they have not supported other hypotheses involving frequency.

Hypothesis S

Sandhi is required within the phrase (*fōr X*) *ār_{sen}*

There are indications that this is true in casual speech but it is not totally uniform in our data from reading, and therefore perhaps has potential as a good marker of register.

Hypothesis T

Sandhi does not occur across pauses and hesitations

This was true in all cases in my database. Since it would have been a knockout factor for Varbrul, I have discarded all such instances from my data. Although it is not a tested factor in this study, that should not be interpreted as meaning that it is not true. On the contrary, it is a solid generalization.

Hypothesis I

Sandhi does not occur following geminate /rr/

This hypothesis was apparently strongly supported for the only example in our data, Lexical Item g. *kärr* ‘marsh’ / *tyvärr* ‘too bad’. However, this result is weakened by the fact that both contexts contained orthographic punctuation marks between /rr/ and the following /s/.

Hypothesis J

Sandhi with non-laterals does not occur following /r/

This hypothesis was not supported for the only example in our data, Lexical Item f. *karl* ‘man’, which took part in sandhi at a higher than average rate.

Hypothesis U

Sandhi tends to not occur when the vowel /u/ immediately precedes the site

The vowel /u/ preceding the potential sandhi site served as a knockout factor for Varbrul because it was never followed by sandhi in our data. However, the clarity of this result was obscured because both contexts contained orthographic punctuation marks.

Hypothesis V

/rs/-Sandhi tends to occur more frequently when the following segment is /m/

This Hypothesis is confirmed, and at a higher rate than in the previous study (Lindblad 1987a). However, this result is based solely on one item, MfoN 70, so not highly robust.

Hypothesis W

/rs/-Sandhi is favored for *har* > modals > copulas, when followed by verbs

In our data, the hierarchy is instead: modals > *har* > copulas, with very little difference between the last two. However, our data for copulas before verbs is minimal, being based on only one context, and therefore not robust.

5.6.2 SUMMARY OF FAVORABLE AND UNFAVORABLE FACTORS

Now let us summarize the results that we have obtained from inspection of the tables and figures that were generated using the Varbrul statistical program.

<u>The factors:</u>	<u>most favorable</u>	<u>least favorable</u>
male speaker:	aged 15-19,	aged 55-59
female speaker:	aged 40-44	aged 65-69
Dialect Area:	4	10
preceding syntactic category:	adjective	conjunction
following syntactic category:	preposition	adjective
c-command (or edge)	right	left
closeness	close	distant
preceding vowel:	o	u
following vowel:	u	y
following segment:	m	y
preceding stress:	primary	secondary
following stress:	secondary	primary

In this chapter we have looked at external /rs/-sandhi also in terms of comparing the factors that govern it to those that we found to influence internal /rs/-sandhi in the previous chapter. We found that internal sandhi occurs more frequently than external sandhi, with the specific figures for our database being 79% vs. 25% overall, and furthermore that even within the categories of internal and external sandhi, assimilation was generally more likely to occur if the two parts were more closely connected. We found some confirmation that the following hypotheses apply both to internal and to external sandhi contexts: (i) sandhi occurs regularly through sequences of dental consonants both within and between morphemes, (ii) sandhi does not occur following geminate /rr/, and (iii) sandhi tends not to occur when the vowel /u/ immediately precedes the site.

It seems clear that external /rs/-sandhi is constrained by a wide variety of factors, including some that have been proposed by previous researchers, some that we have identified here, and probably some that have not yet been discovered.

In the next chapter we will look another overlapping topic involving the relationship between two words, the question of when the copula *är* 'is/are/am' drops its coda /r/ rather than participating in external sandhi or simply manifesting both of its underlying phonological segments.

CHAPTER 6: PROSODIC FACTORS IN CODA RETENTION BY *ÄR* 'IS/ARE/AM'

This chapter focuses on the question of when the copula *är* is reduced to [e] and when it is pronounced as [er] (unreduced, with an audible [r]) in the dialects of NN. Some surprising data are presented first. (A summary of the complete data for *är* may be found in Appendix I. To clarify their contexts, the reader is referred to Appendixes A - D.) Next these data are analyzed in the light of certain details of Swedish phonology and linguistic universals of prosodic lapses. Then new data are presented that were collected to test this analysis. I conclude that the pronunciation [er] has not been replaced by [e] in underlying representation in these dialects (as might be expected from other dialects), but rather variation between the two pronunciations is dependent upon postlexical stress assignment within the prosodic component of the grammar: prosodic lapse contexts favor coda retention, while lexically stressed syllables adjacent to *är* favor coda loss.

6.0 INTRODUCTION: THE PROBLEMATIC DATA

Most Swedes would probably expect the copula *är* 'is/are/am' to be pronounced without a coda consonant. In fact, what first attracted my attention to this topic was that I had learned (in America, from central and southern Swedish speakers) to pronounce the copula *är* 'is/are/am' as [e]³³. In fact, I thought that pronunciations that included [r] must be overarticulations of the type I will refer to as 'reading pronunciations', in which the reader is influenced away from a more natural pronunciation by the written form of the

33 This vowel is lowered to [ɛ] or [æ] before [r] in most dialects of Swedish. However, I will use [e] to represent all realizations of the vowel of *är* in this paper because when I listened to the tapes my focus was on the consonant [r], and not on the precise degree of height of the vowel. Stig Eliasson (p.c.) observes that this vowel's height could be significant, because a lowered variant, even without an audible [r] segment, could still imply an underlying /r/. After listening to some of my data again, I agree that some of these vowels are lowered, but I still have not looked at all of these data systematically.

word. However, when I analyzed my data from the first 203 readings I collected of MfoN, I found that I got strikingly different data for the six occurrences of the word *är* that I was looking at. Four of these six cases showed approximately 10% to 20% of the pronunciations with [r], and those I still regard as reading pronunciations. What really caught my attention was MfoN-39, where over 96% (almost all) of the readings contained an audible [r]. This was totally unexpected, and it demanded an explanation. I looked first at the immediate segmental context. Since MfoN-39 had *är* before *det*, a word starting with /d/, I looked at the other items, and found that in MfoN-136 *är* also occurred before a word that started with /d/, namely *därför*. Nevertheless, in the context of MfoN-136 I had heard only a normal 17% of pronunciations with audible [r], despite the fact that MfoN-136 occurs in the final sentence of the story, a rhetorical context that produced a wide range of intonation patterns. So what was special about MfoN-39 that induced the full pronunciation [er]? Although it is reasonable to question whether my ear attuned to an American-English phonemic inventory with alveolar /d/ is not missing the Swedish distinction between dental and retroflex 'd', that cannot explain why I heard rather consistent and dramatic differences between MfoN-39 and MfoN-136, so the problem remains. When I looked closer at my data, I realized that another item, MfoN-89, was also unusual in its realizations. It showed approximately one-third apiece for each of three alternatives: no [r], an audible [r], and the retroflex [ʂ] that is the product of /rs/-sandhi. Since /s/ does not freely transform into the retroflex [ʂ] except when following /r/, it is clear that a pronunciation of [ʂ] for /s/ implies that /r/ has not been totally deleted. When the data for MfoN-89 are compared to the much lower retention of /r/ found for MfoN-3 and MfoN-21 the contrast is also dramatic.

Since the immediate segmental context of MfoN-39 (*är* followed by a word starting with /d/) could not explain why I heard an [r] there but not in MfoN-136, I sought a better explanation for it that could also explain the higher percentages found with MfoN-89. The word *är* is a copula, one type of function word, and most of the contexts of interest here contain function words or unstressed syllables of other polysyllabic words adjacent

to it. Consideration of this circumstance eventually led me to conclude that the best explanation for this phenomenon depends on prosodic factors: avoiding a lapse that is a sequence of three unstressed syllables by adding stress to the middle one requires *är* in that position to be pronounced as the full form [er] instead of the reduced form [e]. This may be expressed as in Hypothesis X.

Hypothesis X

The copula *är* retains its coda /r/ when assigned prosodic stress to repair lapses

I have not conducted a thorough spectrographic analysis of my data; but a brief inspection of a few items using a Kay sound spectrograph (Kay CSL 4300B) reveals that, at least for some speakers, the word *är* in MfoN-39 (where I mostly heard [er]) is longer than are the preceding and following syllables, and the *är* in MfoN-39 is also longer than the *är* in MfoN-136 (where I mostly heard [e]). In MfoN-89, on the other hand, *är* is more variable in its realization. Since *är* seems to be longer in MfoN-39, this suggests the correctness of assigning it two moras in that context (cf. Figure 3); the phonetic evidence appears to support this phonological analysis.

In order to make this analysis clear, I will briefly discuss some terminology, symbols, concepts and assumptions.

6.1 EURHYTHMY, AND A PROSODIC LAPSE ANALYSIS OF *ÄR* 'IS/ARE/AM'

It has long been recognized that speech naturally gets broken up into small groups of syllables that are organized by their varying degrees of stress, tone contours, length, etc. This is known as 'eurhythmy' and captures the intuition that it is difficult to produce or listen to long strings of syllables that are pronounced with the same intensity and pitch level without subdividing them. Eurhythmy is reflected in Dauer's (1983: 51) generalization that "a tendency for stresses to recur regularly appears to be a language universal property." It also underlies Nespor & Vogel's (1989: 82) Preference Rule I:

“Prefer maximal alternation.” Eurhythm has long been noted as a factor in the prosody of a wide variety of languages, both internal (within individual words) and external (adjusting the stress levels of a sequence of words). A violation of eurhythm due to stresses coming too closely spaced in a sequence of syllables is referred to as a clash, while a violation of eurhythm due to too many unstressed syllables in sequence is known as a lapse. Selkirk presents the admirably clear, formal definition of (a) clashes and (b) lapses in Figure 39. *The Principle of Rhythmic Alternation* (Selkirk 1984: 52).

- a. Every strong position on a metrical level n should be followed by at least one weak position on that level.
- b. Any weak position on a metrical level n may be preceded by at most one weak position on that level.

Figure 39. *The Principle of Rhythmic Alternation* (Selkirk 1984: 52)

Bruce (1984: 32) has observed that although clashes in Swedish are mostly not resolved by stress shift, lapses in Swedish are frequently repaired by adding extra stresses to break them up, whether word-internally or over a series of underlyingly unstressed words. A representation of the repair of a trisyllabic lapse situated between two stressed syllables using metrical grid notation (cf. Prince 1983) is given in Figure 40.

x	x		x	x	
x	x	=>	x	x	x
x	x	x	x	x	x

Figure 40. Lapse resolution by beat addition

Nespor & Vogel (1989: 83) claim that “[a]ccording to Selkirk’s PRA, seen above in [Figure 39], the maximum number of weak positions allowed together is two. As soon as there are three, a rule will strengthen the middle one to create an alternating rhythmic effect.” However, this generalization fails to observe the distinction between two kinds of lapses exemplified by Selkirk’s example 7.23, involving the sentences in Figure 41.

(Here I have marked lexically stressed syllables with acute accents, enclosed lapse sequences in curly brackets, and underlined the vowels of all function words in the lapse sequences that are not adjacent to stressed syllables and therefore are available to take the added postlexical stress needed to repair the lapses.)

- a. {You could be} sléeping.
- b. Jáne {was for the} Dódgers.
- c. They sánɡ {it in a} fúnny wáy.

Figure 41. Lapse contexts requiring added stress (adapted from Selkirk 1984: 363)

The lapse in Figure 41a, between the beginning of the sentence and the first stress, is a type that Nespor & Vogel apparently did not contemplate when they penned the foregoing statement; Selkirk correctly suggests that it may be satisfactorily repaired by a stress on either the first or second word. This sequence of three unstressed syllables prior to the first lexically stressed syllable of the sentence is parallel to MfoN-89, and the availability of two options for its repair explains the variability of results for it that we see in Appendix I. However, in accord with Nespor & Vogel's interpretation of Figure 39, Selkirk gives only one solution for the type of lapses in Figure 41b and Figure 41c where three unstressed syllables are in the middle of a sentence between two stresses, namely repair by stressing the middle word of the lapse (cf. also Sweet 1876: 11; Hayes 1995: 372). The lack of alternative resolutions for lapses of three syllables between two stresses provides an explanation for why I consistently found unreduced forms with this kind of lapse in MfoN-39. Curiously, although they seem to focus on precisely this type of lapse, Nespor & Vogel (1989: 87, 95) insist that lapse repair is optional. However, my data for MfoN-39 support the contention that lapse repair can be mandatory, not optional, at least in the context of this type of trisyllabic lapse between two stresses in Swedish.

The stress on MfoN-39 is clearly a postlexical phenomenon, since it depends on the specific prosodic context in which the word *är* is located. This context is framed by the

lexically assigned stresses on the bisyllabic words *landet* and *inte*, both of which are stressed on the first syllable. Thus, prosodic lapse repair accounts for the consistent stressing of *är* in MfoN-39, and also accounts for the frequent but more varied stressing of *är* in MfoN-89 that we see in Appendix I.

6.2 NEW DATA, AND ITS RELEVANCE TO THE PROSODIC LAPSE ANALYSIS

After positing the significance of prosodic lapses for pronunciation variants of *är* 'is/are/am', I listened again to the tapes of MfoN for occurrences of *är* in contexts other than only before dental consonants, and this data has also been included in Appendix I. The new item of most particular interest, MfoN-32, has 38% unreduced forms [er], about double the background rate of 'reading pronunciations'. This tends to confirm the hypothesis that strong forms occur in lapse contexts, since the prosodic context here is between two function words (*nu* 'now' and *jag* 'I'), though this lapse context is not identical to the others because it is at the start of a non-initial clause and involves a sequence of four function words.

After my preliminary analysis of the data from MfoN, I wrote the second story, Msi, making a special point of including a wide variety of lapse contexts for *är* 'is/are/am', and again traveled around in NN collecting 155 readings of it. Of the 25 items in Msi involving *är*, seven had more than 55 (=36%) unreduced pronunciations. Those seven items with the highest frequency of unreduced forms all fit neatly into the pattern described, and support the lapse analysis. This does not mean that the entire problem has been solved, however, since the variation between those seven is not yet explained, and some of the items with a lower frequency of unreduced pronunciations are less easily analyzed.

The prosodic context of Msi-226 (*är* as the second of three unstressed syllables in between two lexically stressed syllables) was constructed to mimic the context of MfoN-39 as closely as possible, except for a reversal of the theme/rheme relationship around

*är*³⁴. Since only about half the pronunciations of Msi-226 include an audible [r], this raises new questions about the differences between the two that must be left for further research.

It is also worth noting that not all contexts are easily explainable simply by reference to the category of function words. For example Msi-63 seems at first glance to put *är* in the middle of a sequence of six function words, yet it has one of the lowest scores for unreduced forms. However closer inspection reveals that this sequence “och han är till och med” involves the frequent idiom *till och med* ‘up to and including’ (commonly abbreviated *t o m*), listed in *Svenska ord* (1984) with stress on *till*. It seems that this phrase has been lexicalized, so that even though *till* is a function word, in this phrase it receives what can be considered lexical stress. Thus *är* is not in a lapse context in Msi-63, despite first appearances. Five other items pose similar problems, being in what at least superficially look to be lapse contexts, yet resulting in relatively few forms with audible [r]. In some of these cases *är* is adjacent to adverbs, and adverbs are sometimes considered to be major-category words, so that may provide a partial explanation.

In the following section we will look at the patterning of the data for unreduced *är* with the aid of tables and graphs. The population factors that we will consider are age, gender and dialect, and the linguistic factors that we will consider are the word class and stress of the following word.

6.3 TABLES AND GRAPHS OF RESULTS

The data from Appendix I for the presence vs. absence of coda-[r] on the copula *är* ‘is/are/am’ are displayed in a bar graph in Figure 42.

³⁴I am indebted to Staffan Hellberg for suggesting the hypothesis that the theme/rheme structure surrounding MfoN-39 might play a role in its abnormal phonological realization.

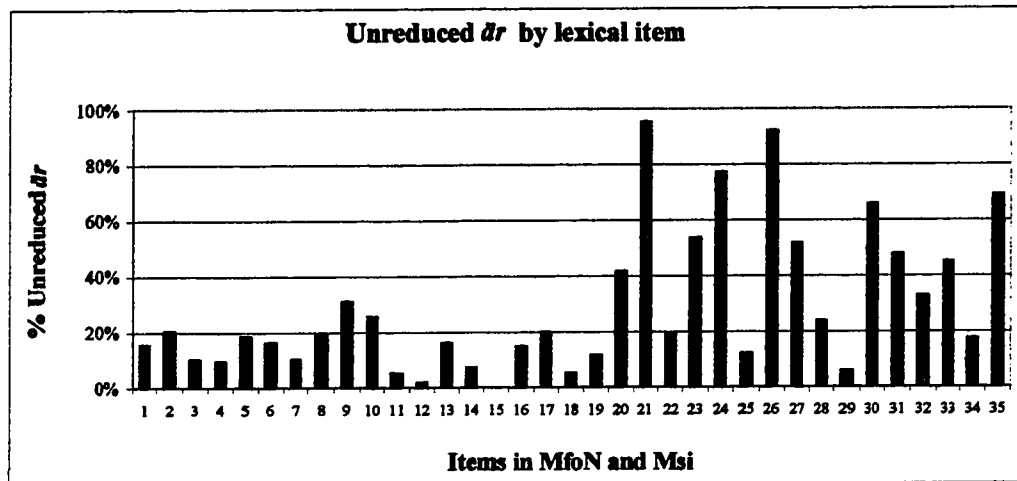


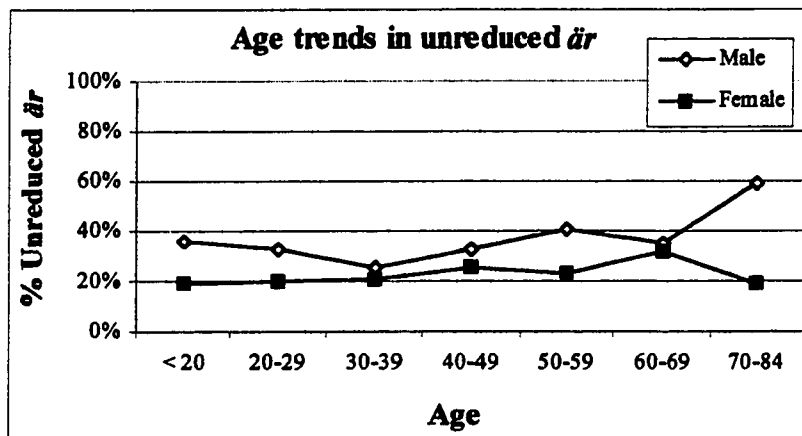
Figure 42. Unreduced *är* 'is/are/am' by lexical item

The data from MfoN and Msi have been intermixed here (and in Appendix I) based on the stress on the following syllable, so that the first nineteen items in Figure 42 occur before stressed syllables. It is evident in this set of data that all of the contexts in Msi that result in large percentages of full forms of *är* occur in prosodic contexts between unstressed syllables. This is confirmation of Hypothesis X, that the explanation for the abnormal pattern of consistent [er] pronunciations noted for MfoN-39 lies in prosody, and more specifically in lapse repair.

Let us begin the further analysis of this data for unreduced *är* by looking at the population factors, beginning with age and gender in Table 32 and Figure 43.

Table 32. Unreduced *är*, by age and gender

Age	Unreduced <i>är</i> pronunciations					
	Male		Female		Both	
< 20	140	36%	83	19%	223	27%
20-29	323	33%	169	20%	492	27%
30-39	127	26%	72	21%	199	24%
40-49	179	33%	109	26%	288	30%
50-59	195	41%	63	23%	258	34%
60-69	132	35%	56	32%	188	34%
70-84	103	59%	16	19%	119	46%
Totals	1199	35%	568	22%	1767	30%

Figure 43. Unreduced *är*, by age and gender

We can see from these displays that unreduced *är* is more common in male speech at all ages. Since there is approximately the same amount of usage across the generations, it seems possible that the current pattern of usage has been in place for a long time.

In Lindblad 1998 I speculated that the pattern of unreduced coda-/r/ on the copula *är* 'is/are/am' might be unique to NN. I based this on several personal observations about central standard Swedish. In the materials that Lund University's Talspråksyntaxprojekt (Syntax of spoken language project) collected from Borås, *är* was consistently transcribed as <e>, and I never heard it pronounced as [er] on the tapes. I did not recall hearing unreduced pronunciations of *är* in central Sweden. Also, other linguists in

Sweden to whom I described my findings from NN did not seem to recognize the pattern. A major motivation for my decision to collect data with MfoN in Uppland in 1996 was a desire to determine whether or not this pattern extended beyond NN. Since gender also seems to play a role, let us now look at dialects and gender in Table 33 and Figure 44.

Table 33. Unreduced *är*, by dialect and gender

Dialect Area	Unreduced forms of <i>är</i>				Combined	
	Male		Female			
U	29	41%	11	19%	40	31%
1	175	30%	55	14%	230	24%
2	69	35%	33	24%	102	30%
3	197	38%	87	23%	284	32%
4	250	40%	129	22%	379	31%
5	133	32%	70	29%	203	31%
6	78	33%	25	21%	103	29%
7	106	39%	64	30%	170	35%
8	77	34%	19	19%	96	30%
9	51	40%	47	21%	98	27%
10	63	30%	39	25%	102	28%
Overall Totals	1228	35%	579	22%	1807	30%

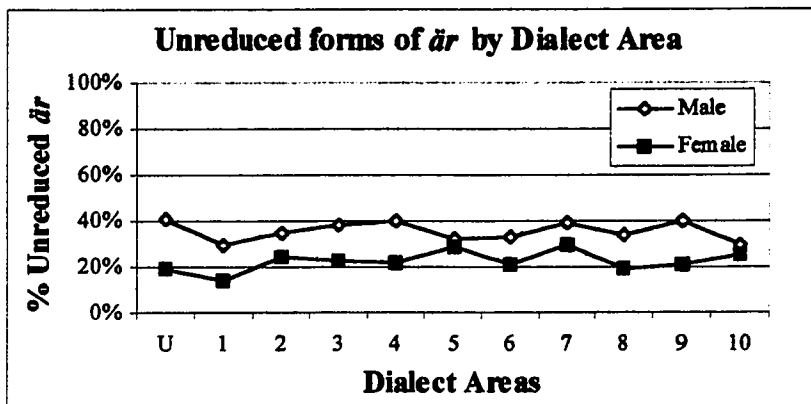


Figure 44. Unreduced *är*, by dialect and gender

We see here in Table 33 and Figure 44 that males use unreduced *är* at a greater rate in all dialect areas, just as they did across all age groups in Table 32 and Figure 43 above. Comparing the various dialect areas in Table 33 to each other, we find that the highest

rate of unreduced *är* among males occurs in Uppland. This runs contrary to our surmised that unreduced *är* does not occur in central standard Swedish, and raises several new questions. How many dialects in central or southern Sweden have unreduced *är* as a regular fixture in their repertoire? Do all of these dialects of Swedish use unreduced *är* in lapse contexts? Does the use of unreduced *är* in lapse contexts occur systematically in casual speech in the same way that it does in reading register? Unfortunately our data do not allow us to investigate the situation in other dialects of Swedish. All that we can surmise from the data in Figure 44 is that a very similar situation regarding unreduced *är* seems to hold in all the dialect areas of NN and in Uppland.

Now let us look one more time at gender, this time as it interacts with the syntactic category of the following word in Table 34 and Figure 45.

Table 34. Unreduced *är*, by gender and following word class

Word class	Male		Female		Both	
Noun	96	25%	35	12%	131	19%
Adverb	196	21%	66	9%	262	16%
Adjective	155	20%	37	6%	192	14%
Pronoun	602	59%	352	46%	954	54%
Determiner	71	42%	36	27%	107	35%
Conjunction	79	64%	42	46%	121	57%
Totals	1199	35%	568	22%	1767	30%

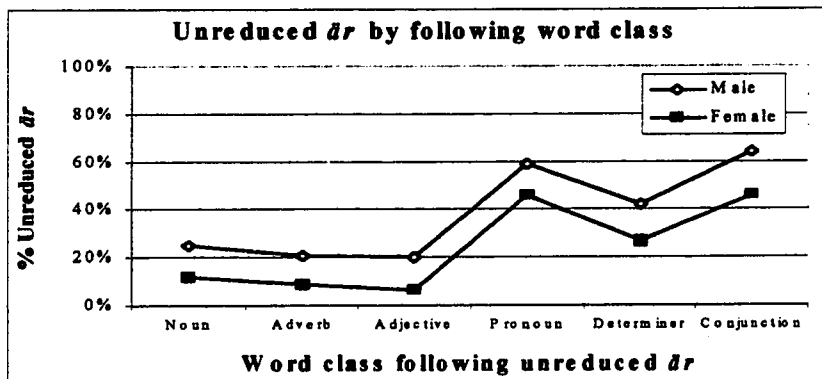


Figure 45. Unreduced *är*, by gender and following word class

Once again we find that males used unreduced *är* at a higher rate in all categories. In Figure 45, the first three word classes are major category words, and the last three are function words. Since we have assumed that all major category words are inherently stressed in the lexicon, while monosyllabic function words get stress postlexically only if they are in a lapse position, these data suggest that stress on the following word might be the factor that is actually involved, and not word class. Inspection of the data reveal that the correlation is quite thoroughgoing, with only a few exceptions, including the adjective *försiktig* 'cautious' (Msi-83), which is unstressed on its first syllable, and a few function words that are stressed for various reasons.

Before we go on to look directly at stress, let us look one more time at the syntactic category of the following word, this time as it relates to age groups in Table 35 and Figure 46.

Table 35. Unreduced *är*, by age and following word class

Second word class	Age							Combined
	< 20	20-29	30-39	40-49	50-59	60-69	70-84	
Noun	15%	16%	13%	26%	24%	24%	30%	19%
Adverb	14%	12%	9%	18%	25%	20%	31%	16%
Adjective	14%	12%	9%	7%	21%	19%	40%	14%
Pronoun	50%	57%	47%	52%	54%	57%	63%	54%
Determiner	28%	30%	36%	43%	36%	44%	63%	35%
Conjunction	48%	52%	47%	67%	63%	74%	53%	57%
Totals	27%	27%	24%	30%	34%	34%	46%	30%

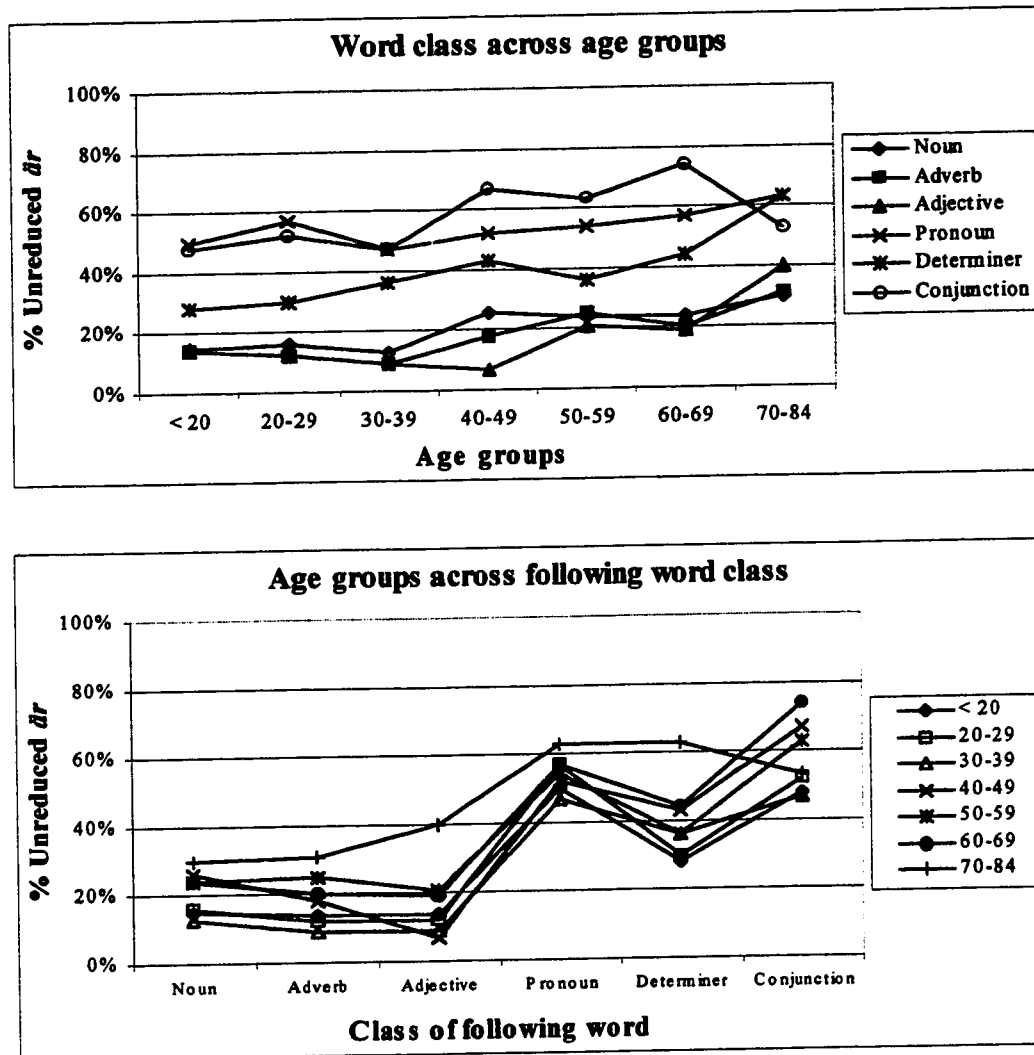


Figure 46. Unreduced *är*, by age and following word class

We see an interesting pattern in Figure 46 in the graph of word class across age groups, where the major category words generally pattern together, but a step down occurs between the 50-59 group and the 30-39 group, and the 40-49 transitional group treats the three categories Noun, Adjective and Adverb differently. The graph of age groups across following word class in Figure 46 resembles that seen above in Figure 45 for all age groups except the oldest, those people aged 70-84, who also utilize unreduced *är* more frequently before major category words, but whose unique trajectory across the word

classes suggests that there may have been a different pattern of usage of unreduced *är* in NN in an earlier era, and that its last traces are now disappearing from the scene.

Now let us look at age groups again, this time in direct relationship to the degree of stress on the following syllable in Table 36 and Figure 47.

Table 36. Unreduced *är*, by age and following stress

Age	Following stress				Combined	
	Primary		Unstressed			
< 20	57	13%	166	45%	223	27%
20-29	107	11%	385	47%	492	27%
30-39	36	8%	163	45%	199	24%
40-49	71	13%	217	50%	288	30%
50-59	83	20%	175	51%	258	34%
60-69	58	19%	130	53%	188	34%
70-84	45	31%	74	67%	119	46%
Totals	457	14%	1310	49%	1767	30%

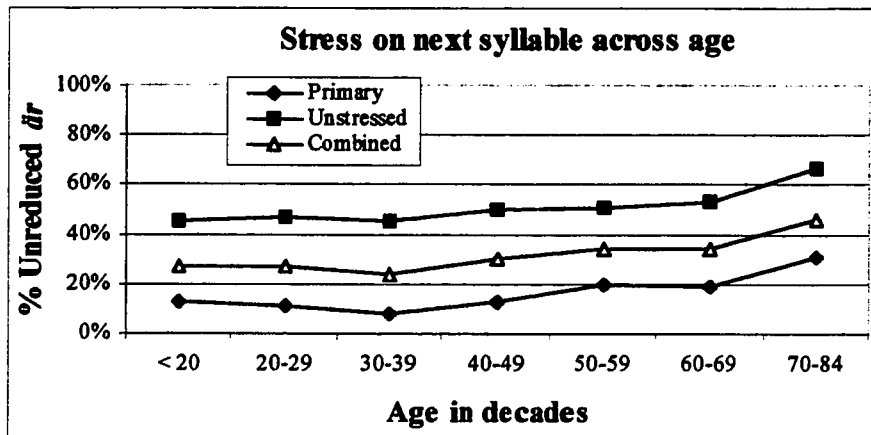


Figure 47. Unreduced *är*, by age and following stress

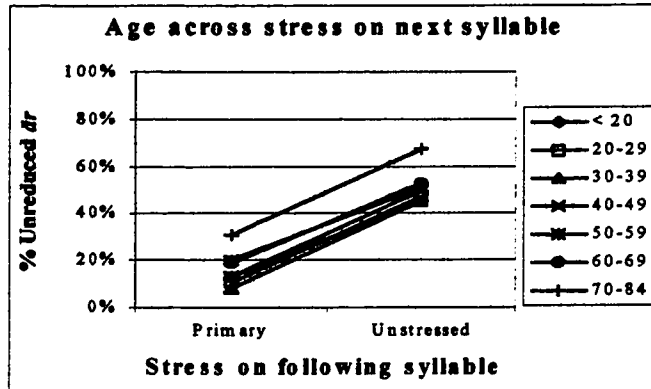


Figure 47. Unreduced *är*, by age and following stress (continued)

In Table 36 and Figure 47 we can see that the difference between primary stress and unstressed following syllables is more than 30% for all age groups. This clear and consistent difference suggests strongly that stress on the following syllable is the most significant factor that we have been able to isolate. In the context of Hypothesis X, this naturally raises the question of what effect is produced by the degree of stress on the preceding syllable. Unfortunately, our data regarding the preceding syllable are not able to produce a clear result because we have no examples of a syllable with primary stress preceding the copula *är*. The results for secondary stress and unstressed preceding syllables do not reveal a clear pattern, and will not be shown here. Although this does not produce the clearest possible confirmation of our hypothesis, the lapse contexts with unstressed syllables both preceding and following the copula *är* that are contained in our data provide good evidence in its support. It will have to be left to future research to confirm that unstressed syllables and syllables with primary stress preceding the copula *är* produce the same clear lapse-based pattern when the following syllable is unstressed.

Finally, let us look one more time at our most important factor, stress on the following syllable, and this time relate it to the dialect areas, where we originally thought that Uppland would lie outside the region where unreduced *är* in lapse contexts are found, as seen in Table 37 and Figure 48.

Table 37. Unreduced *är*, by dialect and following stress

Dialect	Type of stress on the vowel following <i>är</i>				Total	
	Primary		Unstressed			
U	12	16%	28	54%	40	31%
1	59	11%	171	39%	230	24%
2	30	16%	72	50%	102	30%
3	77	16%	207	51%	284	32%
4	113	17%	266	48%	379	31%
5	51	14%	152	51%	203	31%
6	29	14%	74	49%	103	29%
7	47	18%	123	56%	170	35%
8	21	11%	75	54%	96	30%
9	16	8%	82	51%	98	27%
10	14	7%	88	52%	102	28%
Totals	469	14%	1338	49%	1807	30%

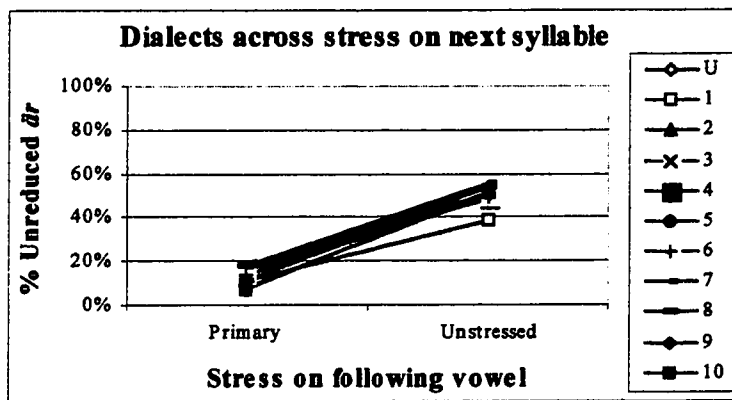
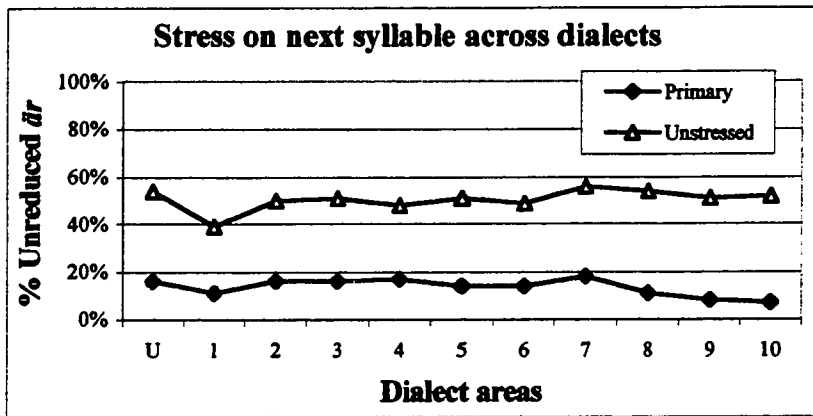


Figure 48. Unreduced *är*, by dialect and following stress

In Table 37 and Figure 48 we see that the differential between primary stress and unstressed following syllables is greatest (45%) in DA-10 in the far north, and lowest (28%) in DA-1 in the southern end of NN. The pattern for the intermediary dialects between DA-10 and DA-1 that can be seen in Figure 48 shows smooth transitions, but there is a sharp turn of the trends between the discontinuous areas DA-1 and Uppland that can only leave us wondering what the situation regarding unreduced *är* may be in the large region of southern Norrland between the two. Since our data cannot provide an answer for those dialects, this too must be left for future research.

In sum, our data have shown that unreduced *är* is most common in lapse contexts, and is strongly disfavored by primary stress on the following syllable. The superficial evidence for influence from the syntactic category of the following word is best interpreted as being a reflection of the difference in lexical stress between major category words and function words. Males produce unreduced *är* at a somewhat higher rate than females. There is some sketchy evidence for an age-based change toward less usage of unreduced *är*, with speakers aged 40-49 seeming to be the group that changed most rapidly, and speakers aged 70-84 showing some evidence of an earlier pattern that was quite different. The dialects of NN are not very different from each other, and even as far south as Uppland the same general pattern for unreduced *är* prevails.

6.4 LAPSES WITH OTHER FUNCTION WORDS

When listening to the Msi tapes, I transcribed data not only for *är* but also for all the other function words with a distinct weak form. We will not present a detailed analysis of these data here, but it is clear that the other function words do not pattern exactly like *är*. The median rate of production for non-reduced forms of the 76 items of *jag* 'I', *sig* 'self', *och* 'and', *det* 'it, the (neuter)' and *med* 'with' is 4%, while the median for the non-reduced form of the 25 items with *är* is 24%. All eleven occurrences of *mig* /*mig*/ 'myself' and *sig* /*sig*/ 'his-/her-/itself' show less than 6% strong forms, even in prosodic

lapse contexts; it seems probable that these words have been relexicalized as /mej/ and /sej/ for most speakers, and that those few strong forms that I recorded are reading pronunciations. The pattern of data for *jag* 'I' does not fit the prosodic lapse model, but rather shows strong forms most frequently in the two (of sixteen) sites where it is the second member of a complex subject (*X och jag...*, 'X and I...'). The data for *och* 'and', *med* 'with' and *det* 'it, the (neuter)' are generally compatible with the prosodic lapse model, with some reservations. Strong forms of *och* 'and' are somewhat more frequent than strong forms of other function words (except for *är*), possibly due to influence from a partially homophonous adverb meaning 'also' that has no weak form, or possibly due to syllable geometry. The four (of eleven) sites of *med* 'with' that occur most frequently in their strong form are located later in the sentence than the verb, and could be partially influenced by the regular pattern of stress for lexicalized verb-adverb pairs in Swedish that puts stress on the adverb. To summarize these data, coda loss does occur in other function words, and partially follows the same pattern as *är*, with unreduced forms in lapse contexts. Other factors seem also to play a role, but cannot be properly analyzed here and will be left for further research.

6.5 SUMMARY AND CONCLUSIONS

I have argued that:

1. the pronunciation [er] has not been replaced by [e] in underlying representation in the dialects of northern Norrland (as might be expected from other more southerly dialects); but rather variation between different pronunciations is dependent upon postlexical stress assignment within the prosodic component of the grammar, with prosodic lapse contexts generally favoring coda retention and lexically stressed syllables adjacent to *är* favoring coda loss,
2. prosodic lapses of three syllables between two stressed syllables can mandate an added stress on the middle syllable,

3. at least *är* and probably also other monosyllabic function words when stressed are phonologically bimoraic and phonetically longer than unstressed function words,
4. unstressed monosyllabic function words have only one mora in Swedish, like other unstressed syllables and unlike all major-category words,
5. postlexical stress can influence the segmental make-up of the surface form of function words,
6. stress on a function word in a lexicalized phrase like *till och med* can create a non-lapse context as effectively as a major-category word's lexical stress,
7. function words are visible to phrasal stress rules, and
8. the patterns of secondary stress in polysyllabic words can vary depending on prosodic context.

The data I collected in Uppland with MfoN following my presentation of a paper based on this material at Nordic Prosody VII (Lindblad 1998) did not show the deletion of coda-/r/ from *är*. This means that the pattern described here may extend not only throughout NN but also through much of central Sweden. However, our data cannot reveal how extensive the geographic spread of this pattern may be, so that question will have to be left for future research.

In the Chapter 7 we will review the results from this and the previous chapters.

CHAPTER 7: SUMMARY AND CONCLUSIONS

The general purpose of this research project was to investigate the phonological realm of /rs/-sandhi in a region that approximates the northernmost 40% of the geographical area of Sweden. This region, which we have referred to as North Norrland (NN), is of particular phonological interest because it is the locus of a potential neutralization between the phonetic product of /rs/-sandhi, [ʂ], and the phoneme of Swedish that is known as *sje* or /ʃ/, which in this region is also pronounced as [ʂ] by most people most of the time.

Previous studies have looked at the phonology of /rs/-sandhi in Central Standard Swedish, at the phonology and phonetics of *sje* in various dialects of Swedish, and at the phonetics of *sje* in six towns in NN in some detail. None have attempted to look at the phonology of *sje*, of /rs/-sandhi, and of the copula *är* in NN in detail, and to examine them in conjunction with each other.

In the dissertation I have used new data collected in NN with two stories, MfoN and Msi, to look at these questions. We have examined them using three population factors and a variety of linguistic factors. The three population factors are the age, gender and dialect area of the speakers. The linguistic factors varied for the different chapters, but included the position and types of adjacent stress, the identity of the lexical items, the word-class categories of the words, and the c-command relationship between the words.

In Chapter 3 we looked at the variation in pronunciation of underlying *sje* between [ʂ] and the dark-*sje* variants such as [ʌ, ɸ], and found effects from the population factors of age, gender and dialect. We found considerable variation between lexical items that can be attributed in large part to their stress patterns. It appeared that stress is a major factor, and that dark-*sje* was found for the most part only before vowels with primary or

secondary stress in the same syllable, but we did not have enough variety in our data to determine precisely how to describe all the components of this environment. There was no evidence that frequency of occurrence of the words was a meaningful factor in the choice of pronunciation variants for *sje*.

In Chapter 4 we looked at /rs/ sequences within words, i.e. internal sandhi. There was only a small amount of variation for each of the population factors, so they did not appear to play a major role, although we did note a general trend toward increased usage by younger speakers. We found extreme variation between lexical items, ranging from 2% to 99% sandhi. The strongest factors blocking sandhi appeared to be the preceding segments; we found some evidence that sandhi is disfavored by preceding /u/ or /r/, and that it does not follow /r/ sequences. With that one exception, we found that sandhi (the feature [-anterior]) spreads freely from /r/ through following uninterrupted sequences of dental consonants. Much of the remaining variation can be attributed to stress patterns, with sandhi favored by preceding secondary stress and following primary stress. We also compared /rs/ sequences within words to the results we had obtained in Chapter 3 for underlying /ʃ/ to see if there is evidence for or against neutralization of the two as the phonetic segment [ʃ], either now or in the past, and we found some tentative evidence that this is the case, with three instances of dark-*sje* being substituted for bimorphemic /rs/ sequences by men of varying ages from northern dialect areas. There was no evidence that frequency of occurrence of the words was a meaningful factor in the production of sandhi.

In Chapter 5 we looked at external /rs/-sandhi, and as with internal /rs/-sandhi, the population factors played a role that was relatively minor, though noticeable. We found nearly as wide a range of variation among our lexical items as with internal /rs/-sandhi, with values ranging from 3% to 90% for the preceding word, and values ranging from 2% to 71% for the following word. In broad terms, we confirmed Eliasson's generalization that internal sandhi occurs more frequently than external sandhi, with the specific figures

for our database being 79% vs. 25% overall, and a more detailed examination showed that greater closeness between the elements containing the /r/ and /s/ favors sandhi across a whole spectrum of morphological categories. For external sandhi environments, we found that sandhi (the feature [-anterior]) spreads freely from /r/ through following uninterrupted sequences of dental consonants, even including the /rɫ/ sequence of the word *karl* ‘man’. We found some evidence that specific phonological segments either preceding or following the sandhi site can influence its probability of occurrence. Some variation can be attributed to stress patterns, with sandhi favored by preceding primary stress and following secondary stress, which rather curiously is the opposite of the stress patterns that favored internal sandhi. C-command (or edge) appeared to play a role as well, with c-command from the right being favorable for sandhi to occur. The considerable variation we found in terms of preceding and following syntactic categories may better be explained by other factors such as c-command. There was no evidence that frequency of occurrence of the word pairs was a meaningful factor in the production of sandhi.

In Chapter 6 we looked at the copula *är* ‘is/are/am’ in regard to the contexts in which it does or does not drop the coda segment /r/. Males produce unreduced *är* at a somewhat higher rate than females, and there is a slight trend toward less usage of unreduced *är* by younger speakers. The dialects of NN are not very different from each other, and even as far south as Uppland the same general pattern for unreduced *är* prevails. The prosodic context of a lapse is a strong factor in the maintenance of unreduced *är*, and its maintenance is strongly disfavored by primary stress on the following syllable. Certain other function words also tend to preserve their unreduced forms in lapse contexts, but for others the evidence is somewhat mixed.

Overall, we have found a wide range of variation in NN between the various lexical items for each of our four areas of interest, the choice of variants of *sje* (Chapter 3), the occurrence of internal /rs/-sandhi (Chapter 4), the occurrence of external /rs/-sandhi (Chapter 5), and the production of the copula *är* ‘is/are/am’ with an unreduced coda-/r/

(Chapter 6). The population factors of age, gender and dialect played a strong role in regard to the choice of variants of *sje*, a somewhat smaller role in regard to the production of the copula *är* 'is/are/am' with an unreduced coda-/r/, and seemed to have little influence on internal or external /rs/-sandhi. Both internal and external /rs/-sandhi seemed to be influenced by adjacent segments, although the number of different words with each of the specific segments in question was so small that we cannot be highly confident in reaching those conclusions when other explanations might be possible. For each of the four sets of data, stress was a major factor. Primary stress on the following vowel had a range of influences; it facilitated the choice of dark-*sje* variants, favored internal /rs/-sandhi, disfavored external /rs/-sandhi, and disfavored the maintenance of coda-/r/ on the copula *är* 'is/are/am'. There was no evidence that frequency of occurrence of the words was a meaningful factor in the choice of variants for any of the topics that we looked at.

We can safely say that each of the phonological processes that we have looked at in this study is governed and constrained by a variety of factors, including some that have been proposed by previous researchers, some that we have identified here for the first time, and others that have not yet been discovered by the scholarly community. Therefore I have attempted throughout this dissertation to use the work of previous researchers to help in the formulation of hypotheses, to assess the importance of all the factors that I have looked at in my data, and to point up those questions that have arisen and cannot be answered by my data but that could be resolved by well-designed research in the future.

BIBLIOGRAPHY

- Allén, Sture. 1970. *Nusvensk frekvensordbok baserad på tidningstext. Vol. 1. Graford, Homografkomponenter. Frequency dictionary of present-day Swedish based on newspaper material. Vol. 1. Graphic Words, Homograph Components.* Stockholm: Almqvist & Wiksell.
- Allén, Sture, and Staffan Wåhlin. 1979. *Förnamnsboken.* Stockholm: Almqvist & Wiksell.
- Andersson, Ingvar. 1970. *A History of Sweden.* Translation: Carolyn Hannay and Alan Blair. Stockholm: Natur och Kultur.
- Björkhagen, Im. 1962. *Modern Swedish Grammar.* Ninth Edition (Revised). Stockholm: Norstedts.
- Blom, Jan-Petter, and John J. Gumperz. 1972. Social Meaning in Linguistic Structure: Code-Switching in Norway. In John J. Gumperz & Dell Hymes (eds.) *Directions in Sociolinguistics: The Ethnography of Communication.* New York: Holt, Rinehart and Winston. 407-434.
- Bruce, Gösta. 1977. *Swedish Word Accents in Sentence Perspective.* PhD dissertation, Lund University. Lund: Gleerup.
- Bruce, Gösta. 1984. Rhythmic alternation in Swedish. In C.-C. Elert, I. Johansson, & E. Strangert (eds.) *Nordic Prosody III: Papers from a Symposium.* Stockholm: University of Umeå. 31-41.
- Bruce, Gösta, and Eva Gårding. 1978. A Prosodic Typology for Swedish Dialects. In Eva Gårding, Gösta Bruce, & Robert Bannert (eds.) *Nordic Prosody: Papers from a Symposium.* Lund: Lund University. 219-228.

- Castensson, Reinhold (ed.). 1992. *The Infrastructure: Administration, Energy, Communications*. The National Atlas of Sweden. Stockholm: SNA Publishing.
- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Clements, George N., and Samuel Jay Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge and London: The MIT Press.
- Dahlstedt, Karl-Hampus and Per-Uno Ågren. 1954. *Övre Norrlands Bygdemål: Berättelser på Bygdemål med Förklaringar och en Dialektöversikt*. Skrifter Utgivna av Vetenskapliga Biblioteket i Umeå. Umeå: Almqvist & Wiksell.
- Dauer, R. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11. 51-62.
- Davis, Lawrence M. 1990. *Statistics in Dialectology*. Tuscaloosa and London: University of Alabama Press.
- Elert, Claes-Christian. 1964. *Phonologic Studies of Quantity in Swedish*. Uppsala: Almqvist & Wiksells.
- Elert, Claes-Christian. 1970. *Ljud och ord i svenskan*. Stockholm: Almqvist & Wiksell.
- Elert, Claes-Christian. 1981. *Ljud och ord i svenskan 2*. Umeå Studies in the Humanities 40. Umeå: Almqvist & Wiksell.
- Elert, Claes-Christian, and Sigurd Fries (eds.). 1982. *Nordsvenska: Språkdrag i övre Norrlands tätorter*. Umeå Studies in the Humanities 49. Umeå: Almqvist & Wiksell.
- Eliasson, Stig Sune Evald. 1970. *Studies in Swedish Phonology*. PhD dissertation, Harvard University.
- Eliasson, Stig. 1986. Sandhi in Peninsular Scandinavian. In Henning Andersen (ed.) *Sandhi Phenomena in the Languages of Europe*. Berlin, New York, Amsterdam: Mouton de Gruyter. 271-300.

- Garlén, Claes. 1988. *Svenskans fonologi*. Lund: Studentlitteratur.
- Halle, Morris, and Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge and London: The MIT Press.
- Hammermo, O. 1976. Sje-ljudet i Eskilstunaspråket. FUMS Report No. 46. Uppsala: FUMS.
- Hassler-Göransson, Carita. 1966. *Ordfrekvenser i nusvenskt skriftspråk*. Lund: Skriptor.
- Hayes, Bruce. 1984. The Phonology of Rhythm in English. *Linguistic Inquiry* 15. 33-74.
- Hayes, Bruce. 1988. Metrics and phonological theory. In Frederick J. Newmeyer (ed.) *Linguistics: The Cambridge Survey. Volume II, Linguistic Theory: Extensions and Implications*. Cambridge, New York, New Rochelle, Melbourne, Sydney: Cambridge University Press. 220-249.
- Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago and London: The University of Chicago Press.
- Hedelin, Per, Anders Jonsson, and Per Lindblad. 1989. *Svenskt uttalslexikon*. Göteborg: Chalmers University of Technology, Department of Information Theory.
- Hellberg, Staffan. 1974. *Graphonomic Rules in Phonology. Studies in the Expression Component of Swedish*. Nordistica Gothoburgensia 7. Göteborg: Gothenburg University.
- Hovdhaugen, Even. 1969. *Transformasjonell generativ grammatikk. En inføring*. Oslo: Universitetsforlaget.
- Inkelas, Sharon, and Draga Zec (eds.). 1990. *The Phonology-Syntax Connection*. Chicago and London: University of Chicago Press.
- Jespersen, Otto. 1949. *A Modern English Grammar on Historical Principles. Part I: Sounds and Spellings*. Copenhagen: Einar Munksgaard.

- Kaisse, Ellen M. 1985. *Connected Speech: The Interaction of Syntax and Phonology*. Orlando: Academic Press.
- Kaisse, Ellen M. 1990. Toward a Typology of Postlexical Rules. In Sharon Inkelas & Draga Zec (eds.) *The Phonology-Syntax Connection*. Chicago and London: University of Chicago Press. 127-143.
- Kaisse, Ellen M., and Patricia A. Shaw. 1985. On the theory of Lexical Phonology. *Phonology Yearbook* 2. 1-30.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford UK & Cambridge USA: Blackwell.
- Larsen, Karen. 1948. *A History of Norway*. Princeton: Princeton University Press.
- Lindblad, Per. 1980. *Svenskans sje- och tje-ljud i ett allmänfonetiskt perspektiv*. Travaux de L'institut de linguistique de Lund XVI. Lund.
- Lindblad, Per. 1982. Sje- och tje-ludet i sex nordnorrländska stadsmål. In Claes-Christian Elert and Sigurd Fries (eds.) *Nordsvenska: Språkdrag i övre Norrlands tätorter*. Umeå Studies in the Humanities 49. Umeå: Almqvist & Wiksell. 99-122.
- Lindblad, Vern M. 1987a. Sandhi between r and s in Swedish. Ms, University of Washington.
- Lindblad, Vern M. 1987b. Sandhi of /-r/ and /s-/ Between Verbs in the Speech of an Informant from Central Sweden. Ms, University of Washington.
- Lindblad, Vern M. 1989. An autosegmental approach to Swedish stressed syllables. Ms, University of Washington.

- Lindblad, Vern M. 1998. Prosodic factors in coda loss in Norrland Swedish. In Stefan Werner (ed.) *Nordic Prosody: Proceedings of the VIIth Conference, Joensuu 1996*. Frankfurt am Main: Peter Lang. 159-170.
- Lundholm, Kjell O., Östen J. Groth, and Rolf Y. Petersson. 1996. *North Scandinavian History*. Translation: Paul Fisher. Luleå: I-Tryck/Grafiska Huset.
- Mohanan, K.P. 1986. *The Theory of Lexical Phonology*. Dordrecht: D. Reidel Publishing Company.
- Nespor, Marina, and Irene Vogel. 1989. On clashes and lapses. *Phonology* 6. 69-116.
- Nordberg, Karin. 1995. Från byklubbar till radiodistrikt: En västerbottensk radiohistoria 1923-1961. In Britta M Lundgren (ed.) *västerbotten 1-95*. Umeå: Västerbottens museum. 2-18.
- Öberg, Sture, and Peter Springfeldt (eds.). 1991. *The Population*. The National Atlas of Sweden. Stockholm: SNA Publishing.
- Ordbok över Svenska Språket*. Lund: The Swedish Academy.
- Pamp, Bengt. 1978. *Svenska dialekter*. Stockholm: Natur och Kultur.
- Prince, Alan S. 1983. Relating to the Grid. *Linguistic Inquiry* 14. 19-100.
- Modern svensk engelsk ordbok*. 1980. Stockholm: Bokförlaget Prisma.
- Riad, Tomas. 1988. Tracing the foot. A metrical analysis of change in Nordic languages. *Arkiv för nordisk filologi* 103. 1-35.
- Riad, Tomas. 1992. *Structures in Germanic prosody. A diachronic study with special reference to the Nordic Languages*. PhD dissertation, Stockholm University. Stockholm: Department of Scandinavian languages.
- S.A.O.B.* (= *Svenska Akademiens Ordbok*) = *Ordbok över Svenska Språket*.
- Scott, Franklin D. 1977. *Sweden: The Nation's History*. Minneapolis: University of Minnesota Press.

- Selkirk, Elisabeth. O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge and London: The MIT Press.
- Sellberg, B., Berg, T., and M. Bolander. 1975. Hur talar Umestudenter? Report No. 11, Department of Nordic Languages. Umeå: Umeå University.
- Sjöstedt, Gösta. 1936. *Studier över r-ljuden i sydkandinaviska mål*. Skrifter utgivna genom landsmålsarkivet i Lund 4. Lund: Gleerup.
- Svenska Akademiens Ordbok* -- see *Ordbok över Svenska Språket*.
- Svenska ord*. 1984. Stockholm: Norstedts.
- Svenska ord*. Second edition. 1994. Stockholm: Norstedts.
- Sweet, Henry. 1876. Words, Logic, and Grammar. In Henry Sweet, 1913, *Collected Papers of Henry Sweet*. Oxford: Oxford University Press. 1-33.
- Trudgill, Peter. 1974. *Sociolinguistics: An Introduction*. Harmondsworth, Middlesex, England: Penguin Books.
- Wood, Sidney. 1982. *X-Ray and Model Studies of Vowel Articulation*. PhD dissertation, Lund University. Lund: Department of Linguistics.
- Wuorinen, John H. 1965. *A History of Finland*. New York and London: Columbia University Press.

APPENDIX A: MIN FAMILJ OCH NORRLAND (NUMBERED)

Min familj och Norrland

1

Jag heter Svensson. Jag är stolt över att vara infödd norränning. Det var
 2 3 4
 för sextio år sen som mina föräldrar flyttade till Norrland, och jag har sällan tänkt på att
 5 6 7 8
 flytta härifrån. Man får säga att jag har stark hemortskänsla. Vi äter surströmming
 9 10 11 12 13 14
 varje helg, trots att utlänningar säger sånt som inte bör sägas om vår surströmming.
 15 16 17 18 19
 Jag talar svenska och lite finska. Det är sant att finskan inte är något världs-språk,
 20 21 22 23 24
 men jag tycker så mycket om det, eftersom det var min morfars språk.
 25 26 27 28
 Nuförtiden bor jag i en av Luleås förorter, så nu är jag storstadsbo. Men när jag
 29 30 31 32 33
 var liten, bodde vi ute på landet. Min far sa alltid att barrskogen lockade honom.
 34 35 36
 Jag kommer bäst ihåg honom när han var sysselsatt med att snickra eller stämpla
 37 38
 träd. Så tyst som på landet är det inte i den här stan. Men man sitter där så ofta
 39 40 41 42
 utan lönearbete. Det verkar som om folk som tänker skaffa sej pengar gör det
 43 44 45
 i någon stor stad, där man får slita och släpa, och blir slav under klockan.
 46 47 48 49

Min äldsta syster, som heter Sara Persson, säger att hon tänker sluta arbeta
 50 51 52 53
snart. Sara jobbar och strävar som bagerska i vår systems restaurang. Numera
 54 55 56 57 58 59
 använder Sara pärlsocker sparsamt, för hon är ganska tjock. Min andra syster Ursula
 60 61 62 63 64 65 66
 bestämmer över spisen. Hon arbetar snabbt, och har säker smak. Då hon
 67 68 69 70

kokar stuvning på potatis och renkött, säger somliga som smakar den att den är
 71 72 73 74 75
världens bästa.
 76

Själ är jag en sorts tusenkonstnär, skulle man kunna säga. Jag har sökt och haft
 77 78 79 80 81
 så många olika jobb. Jag har skott hästar, och jag har skottat snö. Jag har timrat hus
 82 83 84
 och gjort staket. Jag har spelat clown, och jag har cyklat och jonglerat på enhjuling.
 85 86 87 88

Det är som vägdirektör som jag har kört runt under de senaste åren, och som jag
 89 90 91 92 93 94
har sett Nordens största städer. Jag har stannat några dagar i Stockholm, Oslo,
 95 96 97 98 99
 Helsingfors, Sundsvall, o s v [och så vidare].
 100 101

Jag har sällan åkt till Medelhavsländerna. Turister söker sej dit, och jag vill inte
 102 103 104 105 106
 vara med. Ofta ligger soldyrkarna på stränderna som sardiner stekta i sololja.
 107 108 109 110 111
 I januari -68 [sextio åtta] reste vi på vintersemester till Jugoslavien. På hemvägen
 112 113 114
 reste vi över Spanien till Portugal. Då vi flög över Spanien blev jag flygsjuk. Om
 115 116 117 118
 jag får sammanfatta den där semester, så säger jag att jag aldrig har sett så många
 119 120 121 122 123
fattiga. Här slipper jag se dem, så jag försöker stanna hemma här i höga Norden.
 124 125 126 127
 I Norrlands städer går säkerligen många arbetslösa, men det finns ändå många
 128 129 130 131
fördelar som uppväger nackdelarna. Det är därför som jag tycker bäst om mitt
 132 133 134 135 136 137
Norrland.
 138

APPENDIX B: MIN FAMILJ OCH NORRLAND (INTERLINEAR TRANSLATION)

Min familj och Norrland
 min familj och norr land
my family and north land
 Pro N Conj N N

My family and Norrland

Jag heter Svensson. Jag är stolt över att vara
 jag het -r Sven -s son jag är stolt över att vara
I call -s Sven -'s son I is proud over that to_be
 Pro V -Pres N -Gen N Pro Cop A P Part Cop -Inf

My name is Svensson. I am proud to be

infödd norrlänning. Det var för sextio år sen som
 infödd norr länning det var för sextio år sen som
native north lander it was for sixty year since that
 A N N Pro Cop -Past P Num N Adv Pro

a native Norrlander. It was sixty years ago that

mina föräldrar flyttade till Norrland, och jag har
 min föräldrar flytt -de till norr land och jag ha -r
my parents move -ed to north land and I have -s
 Pro N -Pl V -Past P N N Conj Pro V -Pres

my parents moved to Norrland, and I have

sällan tänkt på att flytta härifrån. Man
 sällan tänk -t på att flytt härifrån man
seldom think -ed on that move from_here one
 Adv V -Sup P Part V Adv Pro

seldom thought about moving away from here. You may

får säga att jag har stark
 få -r säg -a att jag ha -r stark
may -s say -Inf that I have -s strong
 V -Pres V -Inf Part Pro V -Pres A

say that I have strong feelings for my home locality.

hemortskänsla. Vi äter surströmming varje helg,
 hem ort -s känsla vi ät -r sur strömming varje helg
home place -'s feeling we eat -s sour herring each holiday
 N N -Gen N Pro V -Pres A N A N

We eat fermented herring every holiday,

trots att utlänningar säger sånt som inte bör
 trots att ut länning -r säg -r sånt som inte bör
despite that out lander -Pl say -s such as not should
 Conj Part P N -Pl V -Pres Pro Conj Adv V -Pres

even though foreigners say things that shouldn't

sågas om vår surströmming. Jag talar
 säg -a -s om vår sur strömming jag tala -r
say -Inf -Pass about our sour herring I speak -s
 V -Inf -Pass P Pro A N Pro V -Pres

be said about our fermented herring. I speak

svenska och lite finska. Det är sant att finskan inte
 svenska och lite finska det är sant att finskan inte
Swedish and little Finnish it is true that Finnish not
 N Conj A N Pro Cop A Part N Adv

Swedish and a little Finnish. It is true that Finnish

är något världsspråk, men jag tycker så mycket
 är något värld -s språk men jag tyck -r så mycket
is some world -'s language but I think -s so much
 Cop A N -Gen N Conj Pro V -Pres Adv A

is not a world-language, but I like it so much,

om det, eftersom det var min morfars språk.
 om det efter som det var min mor far -s språk
about it after as it was my mother father -'s language
 Adv Pro P Conj Pro Cop -Past Pro N N -Gen N

because it was my grandfather's language.

Nuförtiden bor jag i en av Luleås
 nu för tid -en bo -r jag i en av Luleå -s
now for time -the dwell -s I in one of Luleå -'s
 Adv P N -Det V -Pres Pro P Num P N -Gen

Now I live in one of Luleå's

förorter, så nu är jag storstadsbo. Men
 för ort -r så nu är jag stor stad -s bo men
fore place -Pl so now is I large city -'s dweller but
 Adv N -Pl Conj Adv Cop Pro A N -Ger.N Conj

suburbs, so now I'm a city-dweller. But

när jag var liten, bodde vi ute på landet.
 när jag var liten bo -de vi ute på land -et
when I was little dwell -ed we out on land -the
 Conj Pro Cop -Past A V -Past Pro Adv P N -Det

when I was little, we lived out in the country.

Min far sa alltid att barrskogen lockade
 min far sa alltid att barr skog -en locka -de
my father said always that needle forest -the entice -ed
Pro N V-Past Adv Part N N -Det V -Past

My father always said that the pine-forest drew him.

honom. Jag kommer bäst ihåg honom när han var
 honom jag komm -r bäst i håg honom när han var
him I come -s best in mind him when he was
Pro Pro V -Pres A P N Pro Conj Pro Cop -Past

I remember him best when he was

sysselsatt med att snickra eller stämpla träd. Så tyst
 sysselsatt med att snickra eller stämpla träd så tyst
occupied with that do_carpentry or stamp tree so quiet
A P Part V-Inf Conj V-Inf N Adv A

employed doing carpentry or marking trees. As quiet

som på landet är det inte i den här stan. Men
 som på land -et är det inte i den här stad -en men
as on land -the is it not in the here city -the but
Conj P N -Det Cop Pro Adv P Det Adv N -Det Conj

as in the countryside it isn't in this here city. But

man sitter där så ofta utan lönearbete. Det
 man sitt -r där så ofta utan lönearbete det
one sit -s there so often without paid_work it
Pro V -Pres Adv Adv Adv P N Pro

so often there you have no paying job. It

verkar som om folk som tänker skaffa sej pengar
 verka -r som om folk som tänk -r skaffa sig pengar
seem -s as if people that think -s obtain self money
V -Pres Conj Conj N Pro V -Pres V-Inf Refl N

seems like people who intend to obtain money

gör det i någon stor stad, där man får
 gör det i någon stor stad där man få -r
do it in some large city there one may -s
V-Pres Pro P A A N Conj Pro V -Pres

do it in some big city, where you can

slita och släpa, och blir slav under klockan.
 slita och släpa och blir -r slav under klocka -n
drudge and drag and become -s slave under clock -the
V-Inf Conj V-Inf Conj Cop -Pres N P N -Det

drudge and drag, and become a slave to the clock.

Min äldsta syster, som heter Sara Persson, säger att
 min äldsta syster som het -r Sara Per -s son säg -r att
my eldest sister that call -s Sara Per -'s son say -s that
Pro A -Super N Pro V -Pres N N -Gen N V -Pres Part

My eldest sister, who is named Sara Person, says that

hon tänker sluta arbeta snart. Sara jobbar och
 hon tänk -r slut -a arbeta snar -t Sara jobba -r och
she think -s stop -Inf work prompt -ly Sara work -s and
Pro V -Pres V -Inf V -Inf A -Adv N V -Pres Conj

she intends to quit work soon. Sara works and

strävar som bagerska i vår systers restaurang. Numera
sträva -r som bagare -ska i vår syster -s restaurang numera
strive -s as baker -ess in our sister -'s restaurant nowadays
V -Pres Conj N -FemN P Pro N -Gen N Adv

strives as a baker in our sister's restaurant. Nowadays

använder Sara pärlsocker sparsamt, för hon är ganska
använd -r Sara pärl socker spar -sam -t för hon är ganska
use -s Sara pearl sugar spare -some -ly for she is rather
V -Pres N N N V -A -Adv Conj Pro Cop Adv

Sara uses pearl-sugar sparingly, for she is rather

tjock. Min andra syster Ursula bestämmer över spisen.
tjock min andra syster Ursula bestämm -r över spis -en
stout my second sister Ursula decide -s over stove -the
A Pro A N N V -Pres P N -Det

stout. My second sister Ursula is in charge of the stove.

Hon arbetar snabbt, och har säker smak. Då hon
hon arbeta -r snabb -t och ha -r säker smak då hon
she work -s quick -ly and have -s certain taste then she
Pro V -Pres A -Adv Conj V -Pres A N Adv Pro

She works fast, and has a sure sense of taste. When she

kokar stuvning på potatis och renkött, säger
koka -r stuvning på potatis och ren kött säg -r
cook -s stew on potato and reindeer meat say -s
V -Pres N P N Conj N N V -Pres

makes stew from potatoes and reindeer meat, some people

somliga som smakar den att den är världens bästa.
 somliga som smaka -r den att den är värld -en -s bäst
some that taste -s it that it is world -the -'s best
 Pro Pro V -Pres Pro Part Pro Cop N -Det -Gen A

who taste it say that it is the world's best.

Själv är jag en sorts tusenkonstnär, skulle man
 själv är jag en sort -s tusen konstnär skulle man
self is I a sort -'s thousand artist should one
 Pro Cop Pro Det N -Gen Num N V Pro

Myself, I'm a sort of jack-of-all-trades, you

kunna säga. Jag har sökt och haft
 kan -a säg -a jag ha -r sök -t och hav -t
can -Inf say -Inf I have -s seek -en and have -en
 V -Inf V -Inf Pro V -Pres V -Sup Conj V -Sup

could say. I have sought and had

så många olika jobb. Jag har skott hästar,
 så många olika jobb jag ha -r sko -t häst -r
so many different job I have -s shoe -en horse -Pl
 Adv A A N Pro V -Pres V -Sup N -Pl

so many different jobs. I have shoed horses,

och jag har skottat snö. Jag har timrat
 och jag ha -r skotta -t snö jag ha -r timra -t
and I have -s shovel -en snow I have -s timber -en
 Conj Pro V -Pres V -Sup N Pro V -Pres V -Sup

and I have shoveled snow. I have framed

hus och gjort staket. Jag har spelat clown,
 hus och gjor -t staket jag ha -r spela -t clown
house and make -en fence I have -s play -en clown
 N Conj V -Sup N Pro V -Pres V -Sup N

houses and built fences. I have played clown,

och jag har cyklat och jonglerat på enhjuling.
 och jag ha -r cykla -t och jonglera -t på enhjuling
and I have -s cycle -en and juggle -en on unicycle
 Conj Pro V -Pres V -Sup Conj V -Sup P N

and I have ridden and juggled on a unicycle.

Det är som vägdirektör som jag har kört
 det är som väg direktör som jag ha -r kör -t
it is as way director that I have -s drive -en
 Pro Cop Conj N N Conj Pro V -Pres V -Sup

It is as Superintendent of Roads that I have driven

runt under de senaste åren, och som jag har
 runt under de senaste år -en och som jag ha -r
round under the latest year -the and that I have -s
 Adv P Det A -Super N -Det Conj Conj Pro V -Pres

around in recent years, and that I have

sett Nordens största städer. Jag har
 se -t Norden -s stor -sta stad -r jag ha -r
see -en Scandinavia -'s large -est city -Pl I have -s
 V -Sup N -Gen A -Super N -Pl Pro V -Pres

seen Scandinavia's largest cities. I have

stannat några dagar i Stockholm, Oslo, Helsingfors,
stanna -t några dag -r i Stockholm Oslo Helsingfors
stay -en some day -PI in Stockholm Oslo Helsinki
V -Sup A N -PI P N N N

stayed a few days in Stockholm, Oslo, Helsinki,

Sundsvall, o s v.
Sundsvall o s v
Sundsvall etc.
N Pro

Sundsvall, etc.

Jag har sällan åkt till Medelhavsländerna.
jag ha -r sällan åk -t till Medelhav -s land -r -na
I have -s seldom travel -en to Mediterranean -'s land -PI -the
Pro V -Pres Adv V -Sup P N -Gen N -PI -Det

I have seldom traveled to the Mediterranean lands.

Turister söker sej dit, och jag vill inte vara
turist -r sök -r sig dit och jag vill inte vara
tourist -PI seek -s self to_there and I want not to_be
N -PI V -Pres Refl Adv Conj Pro V Adv Cop -Inf

Tourists head there, and I don't want to be

med. Ofta ligger soldyrkarna på stränderna som
med ofta ligg -r soldyrkar -na på strand -r -na som
with often lie -s sunbather -the on beach -PI -the like
Adv Adv V -Pres N -Det P N -PI -Det Conj

with them. Often the sunbathers lie on the beaches like

sardiner stekta i sololja. I januari -68 reste vi
sardin -r stekta i sol oljai januari -68 res -de vi
sardine -PI fried in sun oil in January 1968 travel -ed we
N -PI A P N N P N Num V -Past Pro

sardines fried in suntan lotion. In January '68 we went

på vintersemester till Jugoslavien. På hemvägen
på vinter semester till Jugoslavien på hem väg -en
on winter vacation to Yugoslavia on home way -the
P N N P N P Adv N -Det

on a winter vacation to Yugoslavia. On the way home

reste vi över Spanien till Portugal. Då vi flög
res -de vi över Spanien till Portugal då vi flög
travel -ed we over Spain to Portugal then we flew
V -Past Pro P N P N Conj Pro V -Past

we traveled across Spain to Portugal. While we flew

över Spanien blev jag flygsjuk. Om jag får
över Spanien blev jag flyg sjuk om jag få -r
over Spain became I flying sick if I may -s
P N Cop -Past Pro N A Conj Pro V -Pres

over Spain I became airsick. If I were

sammanfatta den där semestern, så säger jag att
sammanfatt -a den där semester -n så säg -r jag att
summarize -Inf the there vacation -the so say -s I that
V -Inf Det Adv N -Det Conj V -Pres Pro Part

to summarize that there vacation, I'd say that

jag aldrig har sett så många fattiga. Här
 jag aldrig ha -r se -t så många fattiga här
I never have -s see -en so many poor here
 Pro Adv V -Pres V -Sup Adv A A Adv

I never have seen so many poor people. Here

slipper jag se dem, så jag försöker stanna
 slipp -r jag se dem så jag för sök -r stanna
escape -s I see them so I for seek -s stay
 V -Pres Pro V -Inf Pro Conj Pro Adv V -Pres V -Inf

I avoid seeing them, so I try to stay

hemma här i höga Norden. I Norrlands städer
 hemma här i höga Norden i norr land -s stad -r
at_home here in high Scandinavia in north land -'s city -PI
 N -Loc Adv P A N P N N -Gen N -PI

at home here in upper Scandinavia. In Norrland's cities

går säkerligen många arbetslösa, men det finns ändå
 gå -r säker -ligen många arbetslösa men det finn -s ändå
go -s certain -ly many jobless but it find -Pass still
 V -Pres A -Adv A A Conj Pro V -Pass Adv

there certainly are many jobless people, but still there

många fördelar som uppväger nackdelarna.
 många för del -r som upp väg -r nack del -r -na
many fore part -PI that up weigh -s back part -PI -the
 A Adv N -PI Pro Adv V -Pres Adv N -PI -Det

are many advantages that outweigh the disadvantages.

Det är därför som jag tycker bäst om mitt Norrland.
 det är därför som jag tyck -r bäst om min norr land
it is therefore that I think -s best about my north land
Pro Cop Adv Conj Pro V -Pres A Adv Pro N N

That's why I like my Norrland best.

APPENDIX C: MIN STYVFARS IDOL (NUMBERED)

Min styvfars idol

1

Jag bor här med min mamma, syster, styvfar, och min hund som heter Sylvester.
 2 3 4 5 6
 Min mamma flyttade mycket när hon var ung, men hon har slagit ner sina bopålar här
 7 8 9
 efter det att jag hade blivit född. Mamma gifte om sig efter det att pappa hade dött,
 10 11 12 13 14 15
 men hon säger att hon aldrig gifter sig med en gruvarbetare igen. Min styvfar sysslar
 16 17 18 19 20 21
med lite av varje. När fabriken lades ner, så började han köra skolbussen. Mamma
 22 23 24
är skolsköterska i vår skola, så jag måste gå i skolan varje dag. Det är inte något tvivel
 25 26 27 28 29 30 31 32 33
 om att min lärare skulle tala om för henne om jag skolkade.
 34 35

Min syster och jag tycker om vår styvfar. Han berättar så många roliga historier.
 36 37 38 39
 Men mamma säger att min styvfar inte alltid håller sig till sanningen när han
 40 41 42
 berättar Sveriges historia för mig, och att jag inte bör tro på allt han säger.
 43 44 45 46 47 48

Min styvfar säger att vår största hjälte någonsin är Sigvard Hjortronsson
 49 50 51 52 53 54 55 56
 i Jokkmokk. Sigvard är större än Ingemar som var världsmästare i slalom, och
 57 58 59 60 61 62
 han är till och med större än den Ingemar som blev världsmästare i tungviktsboxning.
 63 64 65 66 67
 Mamma säger att hon aldrig hört talas om Sigvard.
 68 69 70

--Tyvärr, säger min styvfar, så är det så att myndigheterna inte vill att
 71 72 73 74 75 76 77
människor ska ta efter Sigvard, för han betalar aldrig skatt. Men Sigvard är försiktig,
 78 79 80 81 82 83 84
och ingen kan bevisa att han har några inkomster, så han går fri. Kanske är hans
 85 86 87 88 89
inkomster så små att han trots allt inte är skyldig att betala skatt.
 90 91 92

Sigvard är en speciell sorts människa. Han är stark som en tjur, snabb som en pil
 93 94 95 96 97 98 99
och lång som en flaggstång. Han trivs med utmaningar. Som pojke tröttnade han på
 100 101 102
skolan redan när han gick i sexan. Just den våren, när samerna drev förbi sina renar
 103 104 105
på väg till sommarbetet i fjällen, då följde han med och hjälpte dem. Det var ju ovanligt,
 106 107 108 109
men han trivdes ändå. Sedan dess sägs det att han talar samiska bättre än många av
 110 111 112 113
samebarnen som är tvungna att gå i svenska skolor i flera år. Efter Tjernobyl,
 114 115 116 117
när samerna blev tvungna att slakta ner sina renar, sa min styvfar att han hoppades att
 118 119 120 121 122
Sigvard skulle hjälpa sina vänner då, för det hade blivit svårt för dem att klara sig själva.
 123 124 125 126 127 128 129 130

Samerna lärde honom också att jaga alla sorters djur. Så nu är det ingen risk att
 131 132 133 134 135 136 137
Sigvard svälter. Han är så skicklig med geväret, att han är lika styv som de allra bästa
 138 139 140 141 142 143 144
skarpskyttarna, och han äter därför så mycket kött som han vill. Skogvaktaren bryr sig
 145 146 147 148 149 150 151 152 153 154
inte om Sigvard, eftersom Sigvard inte tar mer än han behöver, och dagsböter
 155 156 157 158
betyder så litet för honom. Och dessutom är det så besvärligt att få tag på honom
 159 160 161 162 163
att det inte är mödan värt att försöka.
 164 165 166 167

I vintras var det mycket halt, och min styvfar körde ner sig i ett kärr. Sen sa min
 168 169 170 171 172 173 174 175 176 177
 styvfar att det var synd att Sigvard inte var med, för han är så stark att vi inte skulle
 178 179 180 181 182
 behövt någon vinsch att dra upp bilen med. Runt vår gård står ett tvåmetersstaket,
 183 184 185 186 187
 men jag kan ändå se över staketet från mitt rum och titta in i skogen. Sigvard skulle
 188 189 190 191
 lätt kunna hoppa över staketet, men min hund Sylvester kan inte göra det.
 192 193

Vårt sommarställe ligger vid Luleälven. Det är snart dags att åka dit. Då
 194 195 196 197 198 199
 blir det en äventyrens tid. När jag plockar så stora hinkar med bär, så har min styvfar
 200 201 202 203 204 205 206
 för vana att säga att det är så gott som i klass med Sigvard. Vår sommarstuga ligger sex
 207 208 209 210 211 212 213 214
 mil från Jokkmokk, men min styvfar säger att Sigvard säkert kommer dit någon gång
 215 216 217 218 219 220
 när han vandrar härs och tvärs över samernas land. Där ute i skogen är det inte svårt
 221 222 223 224 225 226 227 228
 att tänka sig att Sigvard kan stå bakom vilket träd som helst. I fjol när Sylvester och jag
 229 230 231 232 233 234
 strövade omkring i skogen såg vi en jättestor karl som bar ett djur över axlarna. Jag
 235 236 237
 förstod genast att det var Sigvard, och jag sprang hem och berättade det för mamma.
 238 239 240 241 242 243 244 245
 Hon hade svårt att tro mig:
 246 247

--Jag kan inte tro att du har sett Sigvard, sa mamma. Det är bara inbillning.
 248 249 250 251 252

--Jo då, jag har sett Sigvard, sa jag. Jag är säker på att det var han. Han är den
 253 254 255 256 257 258 259 260
 längsta karl som jag någonsin har sett.
 261 262 263 264

Min styvfar säger att Sigvard övervintrar i en koja i urskogen. Så snyggt som
 265 266 267 268
 hemma hos oss är det ju inte i Sigvards koja. Det finns inte någon mamma där som
 269 270 271 272 273 274
ser till att han ständigt städar efter sig. Min styvfar brukar säga att mitt rum ter sig som
 275 276 277 278 279 280 281
Sigvards koja.
 282

APPENDIX D: MIN STYVFARS IDOL (INTERLINEAR TRANSLATION)

Min styvfars idol
 min styvfar -s idol
my stepfather -'s idol
Pro N -Gen N

My stepfather's idol

Jag bor här med min mamma, syster, styvfar, och
 jag bo -r här med min mamma syster styvfar och
I dwell -s here with my mother sister stepfather and
Pro V -Pres Adv P Pro N N N Conj

I live here with my mother, sister, stepfather, and

min hund som heter Sylvester. Min mamma flyttade mycket
 min hund som het -r Sylvester min mamma flytt -de mycket
my dog that call -s Sylvester my mother move -ed much
Pro N Pro V -Pres N Pro N V -Past Adv

my dog named Sylvester. My mother moved a lot

när hon var ung, men hon har slagit
 när hon var ung men hon ha -r slag -t
when she was young but she have -s beat -en
Conj Pro Cop -Past A Conj Pro V -Pres V -Sup

when she was young, but she has

ner sina bopålar här efter det att jag
 ner sin bo påle -r här efter det att jag
down her dwelling pole -PI here after it that I
Adv Pro N N -PI Adv P Pro Part Pro

settled down here since I

hade **blivit** **född. Mamma gifte** **om** **sig**
 ha -de bliv -t född mamma gift -de om sig
have -ed *become* -en *born* *mother* *marry* -ed *again* *self*
 V -Past Cop -Sup A N V -Past Adv Refl

was born. Mother remarried

efter det att pappa hade **dött,** **men hon säger**
 efter det att pappa ha -de dö -t men hon såg -r
after it that papa have -ed *die* -en *but she say* -s
 P Pro Part N V -Past V -Sup Conj Pro V -Pres

after daddy died, but she says

att hon aldrig gifter **sig med en gruvarbetare igen.**
 att hon aldrig gift -r sig med en gruv arbetare igen
that she never marry -s *self with a mine worker again*
 Part Pro Adv V -Pres Refl P Det N N Adv

that she will never marry a mineworker again.

Min styvfar **sysslar** **med lite av varje. När fabriken**
 min styvfar syssla -r med lite av varje när fabrik -en
my stepfather work -s *with little of each when factory* -the
 Pro N V -Pres P A P A Conj N -Det

My stepfather does all sorts of jobs. When the factory

lades **ner, så började** **han köra**
 la -de -s ner så börja -de han kör -a
lay -ed -Pass *down so begin* -ed *he drive* -Inf
 V -Past -Pass Adv Conj V -Past Pro V -Inf

shut down, he started to drive

skolbussen. Mamma är skolsköterska i vår
 skola buss -en mamma är skola sköt -are -ska i vår
school bus -the mother is school nurse -er -ess in our
 N N -Det N Cop N V -Agent -FemN P Pro

the schoolbus. Mother is the school nurse in our

skola, så jag måste gå i skolan varje dag. Det är
 skola så jag måste gå i skola -n varje dag det är
school so I must go in school -the each day it is
 N Conj Pro V V P N -Det A N Pro Cop

school, so I must go to school every day. There is

inte något tvivel om att min lärare skulle tala om
 inte något tvivel om att min lärare skulle tala om
not some doubt about that my teacher should speak about
 Adv A N P Part Pro N V V Adv

no doubt that my teacher would tell

för henne om jag skolkade.
 för henne om jag skolka -de
for her if I shirk -ed
 P Pro Conj Pro V -Past

her about it if I skipped school.

Min syster och jag tycker om vår styvfar. Han
 min syster och jag tyck -r om vår styvfar han
my sister and I think -s about our stepfather he
 Pro N Conj Pro V -Pres Adv Pro N Pro

My sister and I like our stepfather. He

berättar så många roliga historier. Men mamma säger
 berätta -r så många roliga historier men mamma säg -r
tell -s so many funny stories but mother say -s
 V -Pres Adv A A N-Pl Conj N V -Pres

tells so many funny stories. But mother says

att min styvfar inte alltid håller sig till sanningen
 att min styvfar inte alltid håll -r sig till sanning -en
that my stepfather not always hold -s self to truth -the
 Part Pro N Adv Adv V -Pres Refl P N -Det

that my stepfather doesn't always stick to the truth

när han berättar Sveriges historia för mig, och
 när han berätta -r Sverige -s historia för mig och
when he tell -s Sweden -'s history for me and
 Conj Pro V -Pres N -Gen N P Pro Conj

when he tells me the history of Sweden, and

att jag inte bör tro på allt han säger.
 att jag inte bör tro på allt han säg -r
that I not should believe on all he say -s
 Part Pro Adv V-Pres V P Pro Pro V -Pres

that I shouldn't believe everything he says.

Min styvfar säger att vår största hjälte någonsin är
 min styvfar säg -r att vår stor -sta hjälte någonsin är
my stepfather say -s that our large -est hero ever is
 Pro N V -Pres Part Pro A -Super N Adv Cop

My stepfather says that our greatest hero ever is

Sigvard Hjortronsson i Jokkmokk. Sigvard är större
 Sigvard hjortron -s son i Jokkmokk Sigvard är större
Sigvard cloudberry -'s son in Jokkmokk Sigvard is greater
 N N -Gen N P N N Cop A

Sigvard Hjortronsson in Jokkmokk. Sigvard is greater

än Ingemar som var världsmästare i slalom, och
 än Ingemar som var värld -s mästare i slalom och
than Ingemar that was world -'s master in slalom and
 Conj N Pro Cop -Past N -Gen N P N Conj

than Ingemar who was world champion in slalom, and

han är till och med större än den Ingemar som blev
 han är till och med större än den Ingemar som blev
he is to and with greater than the Ingemar that became
 Pro Cop Adv Conj Adv A Conj Det N Pro Cop -Past

he is even greater than that Ingemar who became

världsmästare i tungviktsboxning. Mamma säger att
 värld -s mästare i tungviktsboxning mamma säg -r att
world -'s master in heavyweight_boxing mother say -s that
 N -Gen N P N N V -Pres Part

world champion in heavyweight boxing. Mother says that

hon aldrig hört talas om Sigvard.
 hon aldrig hör -t tala -s om Sigvard
she never hear -en speak -Pass about Sigvard
 Pro Adv V -Sup V -Pass P N

she has never heard of Sigvard.

Tyvärr, säger min styvfar, så är det så att
 tyvärr säg -r min styvfar så är det så att
unfortunately say -s my stepfather so is it so that
 Adv V -Pres Pro N Conj Cop Pro Pro Part

Too bad, says my stepfather, but it's true that

myndigheterna inte vill att människor ska ta efter
 myndighet -r -na inte vill att människa -r ska ta efter
authority -Pl -the not want that person -Pl will take after
 N -Pl -Det Adv V Part N -Pl V V Adv

the authorities don't want people to follow Sigvard's

Sigvard, för han betalar aldrig skatt. Men Sigvard är
 Sigvard för han betala -r aldrig skatt men Sigvard är
Sigvard for he pay -s never tax but Sigvard is
 N Conj Pro V -Pres Adv N Conj N Cop

example, because he never pays taxes. But Sigvard is

försiktig, och ingen kan bevisa att han har
 för siktig och ingen kan bevis -a att han ha -r
fore sighted and nobody can prove -Inf that he have -s
 Adv A Conj Pro V V -Inf Part Pro V -Pres

foresighted, and nobody can prove that he has

några inkomster, så han går fri. Kanske
 några inkomst -r så han gå -r fri kan ske
some income -Pl so he go -s free can happen
 A N -Pl Conj Pro V -Pres A V V

any income, so he goes free. Maybe

är hans inkomster så små att han trots allt
 är han -s inkomst -r så små att han trots allt
is he -'s income -Pl so small that he despite all
 Cop Pro -Gen N -Pl Adv A Part Pro P Pro

his income is so small that he in any case

inte är skyldig att betala skatt.
 inte är skyldig att betala skatt
not is obliged that pay tax
 Adv Cop A Part V N

isn't required to pay taxes.

Sigvard är en speciell sorts människa. Han är stark
 Sigvard är en speciell sort -s människa han är stark
Sigvard is a special sort -'s person he is strong
 N Cop Det A N -Gen N Pro Cop A

Sigvard is a special sort of person. He is strong

som en tjur, snabb som en pil och lång som en
 som en tjur snabb som en pil och lång som en
as a bull quick as a arrow and long as a
 Conj Det N A Conj Det N Conj A Conj Det

as a bull, fast as an arrow and tall as a

flaggstång. Han trivs med utmaningar. Som pojke
 flaggstång han triv -s med utmaning -r som pojke
flagpole he thrive -Pass with challenge -Pl as boy
 N Pro V -Pass P N -Pl Conj N

flagpole. He thrives on challenges. As a boy,

tröttnade han på skolan redan när han gick i
 tröttna -de han på skola -n redan när han gick i
tire -ed he on school -the already when he went in
 V -Past Pro P N -Det Adv Conj Pro V -Past P

he got tired of school already when he was in

sexan. Just den våren, när samerna
 sexa -n just den vår -en när same -r -na
sixth_grade -the just the spring -the when Sami -PI -the
 N -Det Adv Det N -Det Conj N -PI -Det

the sixth grade. Just that spring, when the Sami

drev förbi sina renar på väg till sommarbetet
 drev förbi sin ren -r på väg till sommar bet -et
drove past their reindeer -PI on way to summer grazing -the
 V -Past Adv Pro N -PI P N P N N -Det

drove their reindeer past on the way to the summer pastures

i fjällen, då följde han med och hjälpte
 i fjäll -en då följ -de han med och hjälp -de
in mountains -the then follow -ed he with and help -ed
 P N -Det Adv V -Past Pro Adv Conj V -Past

in the mountains, he followed along and helped

dem. Det var ju ovanligt, men han trivdes
 dem det var ju ovanligt men han triv -de -s
them it was certainly unusual but he thrive -ed -Pass
 Pro Pro Cop -Past Adv A Conj Pro V -Past -Pass

them. It certainly was unusual, but he thrived

ändå. Sedan dess sägs det att han talar samiska
 ändå sedan dess säg -s det att han tala -r samiska
still since its say -Pass it that he speak -s Sami
 Adv Adv Pro V -Pass Pro Part Pro V -Pres N

anyway. Since then it is said that he speaks Sami

bättre än många av samebarnen som är tvungna
 bättre än många av same barn -en som är tvungna
better than many of Sami child -the that is compelled
 A Conj A P N N -Det Pro Cop A

better than many of the Sami children who are forced

att gå i svenska skolor i flera år. Efter
 att gå i svenska skola -r i flera år efter
that go in Swedish school -Pl in several year after
 Part V P A N -Pl P A N P

to go to Swedish schools for several years. After

Tjernoby, när samerna blev tvungna att
 Tjernoby när same -r -na blev tvungna att
Chernobyl when Sami -Pl -the became compelled that
 N Conj N -Pl -Det Cop -Past A Part

Chernobyl, when the Sami were forced to

slakta ner sina renar, sa min styvfar
 slakt -a ner sin ren -r sa min styvfar
slaughter -Inf down their reindeer -Pl said my stepfather
 V -Inf Adv Pro N -Pl V -Past Pro N

slaughter their reindeer, my stepfather said

att han hoppades att Sigvard skulle hjälpa sina
 att han hoppa -de -s att Sigvard skulle hjälp -a sin
that he hope -ed -Pass that Sigvard should help -Inf his
Part Pro V -Past -Pass Part N V V -Inf Pro

that he hoped that Sigvard would help his

vänner då, för det hade blivit svårt
 vän -r då för det ha -de bliv -t svår -t
friend -Pl then for it have -ed become -en difficult -Neut
N -Pl Adv Conj Pro V -Past Cop -Sup A -Neut

friends then, because it had become difficult

för dem att klara sig själva.
 för dem att klar -a sig själv
for them that manage -Inf self self
P Pro Part V -Inf Refl Pro

for them to fend for themselves.

Samerna lärde honom också att jaga alla
 same -r -na lär -de honom också att jag -a alla
Sami -Pl -the learn -ed him also that hunt -Inf all
N -Pl -Det V -Past Pro Adv Part V -Inf A

The Sami also taught him to hunt all

sorters djur. Så nu är det ingen risk att
 sort -r -s djur så nu är det ingen risk att
sort -Pl -'s animal so now is it no risk that
N -Pl -Gen N Conj Adv Cop Pro A N Part

sorts of animals. So now there is no risk that

Sigvard svälter. Han är så skicklig med geväret,
 Sigvard svält -r han är så skicklig med gevär -et
Sigvard starve -s he is so skillful with rifle -the
 N N -Pres Pro Cop Adv A P N -Det

Sigvard will starve. He is so skillful with his rifle

att han är lika styv som de allra bästa
 att han är lika styv som de allra bäst
that he is equally skillful as the very best
 Part Pro Cop Adv A Conj Det Adv A

that he is as good as the very best

skarpskyttarna, och han äter därför så mycket
 skarp skytt -r -na och han ät -r därför så mycket
sharp shooter -PI -the and he eat -s therefore so much
 A N -PI -Det Conj Pro V -Pres Adv Adv A

sharpshooters, and therefore he eats as much

kött som han vill. Skogvaktaren bryr
 kött som han vill skog vakt -are -n bry -r
meat as he want forest watch -er -the concern -s
 N Conj Pro V N V -Agent -Det V -Pres

meat as he wants to. The forest ranger doesn't concern

sig inte om Sigvard, eftersom Sigvard inte tar
 sig inte om Sigvard efter som Sigvard inte ta -r
self not about Sigvard after as Sigvard not take -s
 Refl Adv P N Conj Conj N Adv V -Pres

himself with Sigvard, since Sigvard doesn't take

mer än han behöver, och dagsböter betyder
 mer än han behov -r och dag -s böt -r betyd -r
more than he need -s and day -'s fine -Pl mean -s
Adv Conj Pro V -Pres Conj N -Gen N -Pl V -Pres

more than he needs, and wage-based fines mean

så litet för honom. Och dessutom är det så besvärligt
 så lite för honom och dessutom är det så besvärlig -t
so little for him and besides is it so difficult -Neut
Adv A P Pro Conj Adv Cop Pro Adv A -Neut

so little to him. And besides, it is so difficult

att få tag på honom att det inte är mödan
 att få tag på honom att det inte är möda -n
that get hold on him that it not is trouble -the
Part V N P Pro Part Pro Adv Cop N -Det

to get hold of him that it isn't

värt att försöka.
 värd -t att för sök -a
worth -Neut that for seek -Inf
A -Neut Part Adv V -Inf

worth the trouble to try.

I vintras var det mycket halt, och min styvfar
 I vintras var det mycket halt och min styvfar
last_winter was it very slick and my stepfather
Adv Cop -Past Pro Adv A Conj Pro N

Last winter it was very slick, and my stepfather

körde **ner** **sig** **i** **ett** **kärr.** **Sen** **sa** **min**
 kör -de ner sig i en kärr sen sa min
drive -ed down self in a marsh since said my
 V -Past Adv Refl P Det N Adv V-Past Pro

drove into a marsh. Then my stepfather said

styvfar **att** **det** **var** **synd** **att** **Sigvard** **inte**
 styvfar att det var synd att Sigvard inte
stepfather that it was pity that Sigvard not
 N Part Pro Cop -Past N Part N Adv

that it was a pity that Sigvard

var **med,** **för** **han** **är** **så** **stark** **att** **vi** **inte**
 var med för han är så stark att vi inte
was with for he is so strong that we not
 Cop -Past Adv Conj Pro Cop Adv A Part Pro Adv

wasn't there, because he is so strong that we wouldn't

skulle **behövt** **någon** **vinsch** **att** **dra** **upp** **bilen** **med.**
 skulle behov -t någon vinsch att dra upp bil -en med
should need -en some winch that draw up car -the with
 V V -Sup A N Part V-Inf Adv N -Det Adv

have needed a winch to pull the car out with.

Runt **vår** **gård** **står** **ett** **tvåmetersstaket,** **men** **jag**
 runt vår gård stå -r en två meter -s staket men jag
round our yard stand -s a two meter -'s fence but I
 P Pro N V -Pres Det Num N -Gen N Conj Pro

There is a two meter high fence around our yard, but I

kan ändå se över staketet från mitt rum och titta
 kan ändå se över staket -et från min rum och titta
can still see over fence -the from my room and look
 V Adv V P N -Det P Pro N Conj V

can still see over the fence from my room and look

in i skogen. Sigvard skulle lätt kunna hoppa över
 in i skog -en Sigvard skulle lätt kan -a hoppa över
In in forest -the Sigvard should easily can -Inf jump over
 Adv P N -Det N V Adv V -Inf V -Inf P

out in the woods. Sigvard could easily jump over

staketet, men min hund Sylvester kan inte göra det.
 staket -et men min hund Sylvester kan inte gör -a det
fence -the but my dog Sylvester can not do -Inf It
 N -Det Conj Pro N N V Adv V -Inf Pro

the fence, but my dog Sylvester can't do it.

Vårt sommarställe ligger vid Luleälven. Det är
 vår -t sommar ställe ligg -r vid Luleå älv -en det är
our -Neut summer place lie -s near Luleå river -the it is
 Pro -Neut N N V -Pres P N N -Det Pro Cop

Our summer place is near the Luleå River. It will

snart dags att åka dit. Då blir
 snar -t dag -s att åk -a dit då blir -r
prompt -ly day -'s that travel -Inf to_there then become -s
 A -Adv N -Gen Part V -Inf Adv Adv Cop -Pres

soon be time to go there. Then it will be

det en äventyrens tid. När jag plockar så stora
 det en äventyr -en -s tid när jag plocka -r så stor
it a adventure -the -'s time when I pluck -s so large
 Pro Det N -Det -Gen N Conj Pro V -Pres Adv A

a time for adventure. When I pick such big

hinkar med bär, så har min styvfar för vana
 hink -r med bär så ha -r min styvfar för vana
pail -PI with berry so have -s my stepfather for habit
 N -PI P N Conj V -Pres Pro N P N

pails of berries, my stepfather likes to

att säga att det är så gott som i klass med Sigvard.
 att säg -a att det är så gott som i klass med Sigvard
that say -Inf that it is so good as in class with Sigvard
 Part V -Inf Part Pro Cop Adv A Conj P N P N

say that it's so good that it's in a class with Sigvard.

Vår sommarstuga ligger sex mil från Jokkmokk, men
 vår sommar stuga ligg -r sex mil från Jokkmokk men
our summer cottage lie -s six 10 km from Jokkmokk but
 Pro N N V -Pres Num N P N Conj

Our summer cabin is 60 kilometers from Jokkmokk, but

min styvfar säger att Sigvard säkert kommer
 min styvfar säg -r att Sigvard säker -t komm -r
my stepfather say -s that Sigvard certain -ly come -s
 Pro N V -Pres Part N A -Adv V -Pres

my stepfather says that Sigvard is sure to come

dit	någon gång när	han vandrar	härs	och
dit	någon gång när	han vandra -r	här -s	och
to_there	some time when	he wander -s	here -'s	and
Adv	A N Conj	Pro V	-Pres Adv -Gen	Conj

there some time when he is wandering hither and

tvärs	över samernas	land. Där	ute i
tvär -s	över same -r -na -s	land där	ute i
across -'s	over Sami -PI -the -'s	land there	out in
N -Gen	P N -PI -Det -Gen	N Adv	Adv P

thither across the land of the Sami. Out there in

skogen	är det inte svårt	att tänka sig
skog -en	är det inte svår -t	att tänk -a sig
forest -the	is it not difficult -Neut	that think -Inf self
N -Det	Cop Pro Adv A -Neut	Part V -Inf Refl

the forest it isn't hard to imagine

att Sigvard kan stå	bakom vilket träd	som helst.
att Sigvard kan stå	bakom vilket träd	som helst
that Sigvard can stand	behind which tree	as preferably
Part N V V	P Pro N	Adv Adv

that Sigvard can be standing behind any of the trees.

I fjol	när Sylvester och jag	strövade omkring i
i fjol	när Sylvester och jag	ströva -de omkring i
in last_year	when Sylvester and I	roam -ed around in
P N	Conj N Conj Pro V	-Past Adv P

Last year when Sylvester and I were roaming around in

skogen såg vi en jättestor karl som bar ett djur
 skog -en såg vi en jätte stor karl som bar en djur
forest -the saw we a giant large man that bore a animal
 N -Det V -Past Pro Det N A N Pro V -Past Det N

the woods, we saw a really big man that was carrying an

över axlarna. Jag förstod genast att det
 över axel -r -na jag för stod genast att det
over shoulder -Pl -the I for stood immediately that it
 P N -Pl -Det Pro Adv V -Past Adv Part Pro

animal on his shoulders. I understood immediately that it

var Sigvard, och jag sprang hem och berättade det
 var Sigvard och jag sprang hem och berätta -de det
was Sigvard and I ran home and tell -ed it
 Cop -Past N Conj Pro V -Past Adv Conj V -Past Pro

was Sigvard, and I ran home and told

för mamma. Hon hade svårt att tro mig:
 för mamma hon ha -de svår -t att tro mig
for mother she have -ed difficult -Neut that believe me
 P N Pro V -Past A -Neut Part V -Inf Pro

mother about it. She had difficulty believing me.

Jag kan inte tro att du har sett
 jag kan inte tro att du ha -r se -t
I can not believe that you have -s see -en
 Pro V Adv V Part Pro V -Pres V -Sup

I can't believe that you have seen

Sigvard, sa mamma. Det är bara inbillning.
 Sigvard sa mamma det är bara inbillning
Sigvard said mother it is only imagination
 N V-Past N Pro Cop Adv N

Sigvard, said mother. It's just your imagination.

Jo då, jag har sett Sigvard, sa jag.
 jo då jag ha -r se -t Sigvard sa jag
yes then I have -s see -en Sigvard said I
 Part Adv Pro V -Pres V -Sup N V-Past Pro

Oh yes, I have seen Sigvard, I said.

Jag är säker på att det var han. Han är den
 jag är säker på att det var han han är den
I is certain on that it was he he is the
 Pro Cop A P Part Pro Cop -Past Pro Pro Cop Det

I'm sure that it was him. He is the

längsta karl som jag någonsin har sett.
 läng -sta karl som jag någonsin ha -r se -t
long -est man that I ever have -s see -en
 A -Super N Pro Pro Adv V -Pres V -Sup

tallest man that I've ever seen.

Min styvfar säger att Sigvard övervintrar i en
 min styvfar säg -r att Sigvard över vintra -r i en
my stepfather say -s that Sigvard over winter -s in a
 Pro N V -Pres Part N Adv V -Pres P Det

My stepfather says that Sigvard spends the winters in a

koja i urskogen. Så snyggt som hemma hos oss
 koja i ur- skog -en så snyggt som hemma hos oss
hut in original- forest -the so tidy as at_home at us
 N P A- N -Det Adv A Conj Adv P Pro

hut in the old-growth forest. As tidy as at our place

är det ju inte i Sigvards koja. Det finns inte
 är det ju inte i Sigvard -s koja det finn -s inte
is it certainly not in Sigvard -'s hut it find -Pass not
 Cop Pro Adv Adv P N -Gen N Pro V -Pass Adv

it certainly isn't in Sigvard's hut. There isn't

någon mamma där som ser till att han ständigt
 någon mamma där som se -r till att han ständig -t
some mother there that see -s to that he constant -ly
 A N Adv Pro V -Pres Adv Part Pro A -Adv

any mother there that checks to make sure he always

städar efter sig. Min styvfar brukar säga
 städa -r efter sig min styvfar bruka -r säg -a
clean -s after self my stepfather have_habit -s say -Inf
 V -Pres P Refl Pro N V -Pres V -Inf

cleans up after himself. My stepfather always says

att mitt rum ter sig som Sigvards koja.
 att min rum te -r sig som Sigvard -s koja
that my room appear -s self as Sigvard -'s hut
 Part Pro N V -Pres Refl Conj N -Gen N

that my room looks like Sigvard's hut.

APPENDIX E: POPULATION DATA FOR MFON AND MSI

Population data: MfoN

Age	Gender	Uppland	Dialect Areas										All areas
			1	2	3	4	5	6	7	8	9	10	
<20	m	0	3	0	2	2	3	2	1	3	1	0	17
	f	0	2	2	1	2	0	2	1	2	4	0	16
20-29	m	4	3	0	4	6	1	7	3	1	2	1	32
	f	3	4	3	1	5	0	3	2	0	1	0	22
30-39	m	0	2	1	3	4	3	1	2	1	2	2	21
	f	3	2	1	2	0	3	1	1	4	1	2	20
40-49	m	0	2	0	3	7	4	3	1	3	2	0	25
	f	0	3	4	0	5	0	0	0	2	1	0	15
50-59	m	1	0	1	1	2	2	1	2	1	0	1	12
	f	0	1	2	1	4	1	1	0	1	1	1	13
60-69	m	2	1	3	3	3	1	1	1	1	1	0	17
	f	0	0	2	0	2	0	0	0	1	0	0	5
70-84	m	0	3	0	4	2	0	1	1	0	0	0	11
	f	0	5	0	1	0	0	0	0	0	0	0	6
Total	m	7	14	5	20	26	14	16	11	10	8	4	135
	f	6	17	14	6	18	4	7	4	10	8	3	97

Population data: Msi

Age	Gender	Dialect Areas										All areas
		1	2	3	4	5	6	7	8	9	10	
<20	m	2	1	0	2	2	0	1	0	0	1	9
	f	2	0	2	3	1	0	2	0	1	0	11
20-29	m	7	0	4	8	2	1	4	0	0	2	28
	f	3	0	5	4	3	1	4	0	2	4	26
30-39	m	4	1	2	0	0	1	0	1	0	2	11
	f	1	0	1	1	2	0	0	0	1	1	7
40-49	m	3	0	5	0	1	0	0	2	1	0	12
	f	0	0	3	4	2	0	1	0	1	0	11
50-59	m	1	2	1	3	2	1	2	1	0	2	15
	f	2	0	1	3	0	0	0	0	0	0	6
60-69	m	1	2	0	1	3	0	0	1	1	0	9
	f	1	0	1	1	0	1	0	0	1	0	5
70-84	m	0	0	1	1	1	0	0	0	0	0	3
	f	0	0	0	1	0	0	0	0	0	0	1
Total	m	18	6	13	15	11	3	7	5	2	7	87
	f	9	0	13	17	8	2	7	0	6	5	67

APPENDIX F: LEXICAL ITEMS FOR DARK-SJE

Lexical Item:	Gloss	Item #	[š]	Dark-sje	Maybe	Omit	Other	Total
<u>s</u> jälv	self	MfoN 77	190	46	7	8	1	252
flyg <u>s</u> juk	airsick	MfoN 118	145	88	8	10	1	252
skolsk <u>ö</u> terska	school nurse	Msi 26	136	19				155
människ <u>o</u> r	people	Msi 78	152	3				155
kans <u>k</u> e	maybe	Msi 88	154	1				155
sk <u>y</u> ldig	responsible	Msi 92	133	21		1		155
människ <u>a</u>	person	Msi 96	151	3	1			155
<u>s</u> jälva	selves	Msi 130	115	38		2		155
skickl <u>i</u> g	skilled	Msi 140	134	20	1			155
skarpsk <u>y</u> ttarna	the sharpshooters	Msi 145	122	32			1	155
vins <u>c</u> h	winch	Msi 184	154	1				155

APPENDIX G: LEXICAL ITEMS FOR INTERNAL /RS/-SANDHI

Lexical Item:	Gloss:	Item #	[s]	[rs]	Maybe	Omit	Other	Total
a. <u>hemortskänsla</u>	home place's feeling	MfoN 11	189	19	21	3		232
b. <u>surströmming</u>	sour herring	MfoN 14	78	135	19			232
b. <u>surströmming</u>	sour herring	MfoN 19	76	138	16	2		232
c. <u>världsspråk</u>	world language	MfoN 23	208	18	6			232
c. <u>världsspråk</u>	world language	MfoN 24	196	33	1		2	232
d. <u>världsmästare</u>	world champion	Msi 60	150	3			1	154
d. <u>världsmästare</u>	world champion	Msi 67	146	8				154
e. <u>eftersom</u>	because	MfoN 26	200	25	5	1	1	232
e. <u>eftersom</u>	because	Msi 156	93	61				154
f. <u>morfärs</u>	mother's father's	MfoN 27	219	10	2	1		232
g. <u>storstadsbo</u>	big city resident	MfoN 32	220	8		4		232
h. <u>Persson</u>	Persson	MfoN 51	226	2	2	2		232
i. <u>bagerska</u>	female baker	MfoN 56	228	3		1		232
i. <u>skolsköterska</u>	school nurse	Msi 27	139	14	1			154
j. <u>systers</u>	syster's	MfoN 58	196	29	3	4		232
k. <u>pärlosocker</u>	crushed loaf sugar	MfoN 60	5	222	3	2		232
l. <u>sparsamt</u>	sparingly	MfoN 62	183	47	1	1		232
m. <u>Ursula</u>	Ursula	MfoN 63	86	143	2	1		232
n. <u>sorts</u>	of a sort	MfoN 73	197	24	7	3	1	232
n. <u>sorts</u>	of a sort	Msi 95	150	2		2		154
o. <u>största</u>	largest	MfoN 88	230	1		1		232
o. <u>största</u>	largest	Msi 52	153	1				154
p. <u>Helsingfors</u>	Helsinki	MfoN 91	228	1	1	2		232
q. <u>vintersemester</u>	winter vacation	MfoN 103	110	115	4	2	1	232
r. <u>försöker</u>	attempts	MfoN 116	223	3	1	3	2	232
r. <u>försöka</u>	to attempt	Msi 167	151	2			1	154
s. <u>barrskogen</u>	the pine forest	MfoN 35	4	228				232
t. <u>styvfärs</u>	stepfather's	Msi 1	121	6		27		154
u. <u>sorters</u>	of sorts	Msi 134	86	66		1	1	154
v. <u>tvåmetersstaket</u>	two meter fence	Msi 187	138	15	1			154
w. <u>härs</u>	hither	Msi 221	151	3				154
x. <u>tvärs</u>	thither	Msi 223	150	4				154
y. <u>Sigvar_ds</u>	Sigvard's	Msi 271	141	11	1	1		154
y. <u>Sigvar_ds</u>	Sigvard's	Msi 282	144	9		1		154
z. <u>försiktig</u>	careful	Msi 84	144	9		1		154
1. <u>sommarställe</u>	summer place	Msi 195	57	97				154
2. <u>sommarstuga</u>	summer cabin	Msi 213	109	44	1			154
3. <u>förstod</u>	understood	Msi 238	151	3				154
4. <u>urskog</u>	old growth forest	Msi 268	138	12		4		154
Overall Totals			5814	1574	98	70	10	7566

APPENDIX H: LEXICAL ITEMS FOR EXTERNAL /RS/-SANDHI

Lexical Item A	Lexical Item B	Item #	[ʂ]	/rs/	Maybe	Pause	Omit	Other	Total
heter	Svensson	MfoN 2	51	174	3	2	1	21	252
är	stolt	MfoN 3	12	235			5		252
för	sextio	MfoN 5	45	193	2	5	1	6	252
år	sen	MfoN 6	184	56	3	7	2		252
har	sällan	MfoN 8	10	226	1	8	4	3	252
får	säga	MfoN 9	35	204	4	3	1	5	252
har	stark	MfoN 10	6	186		11	3	46	252
äter	surströmming	MfoN 13	42	145	43	5	1	16	252
utläningar	säger	MfoN 15	26	195	2	15	2	12	252
säger	sånt	MfoN 16	24	164	2	27	11	24	252
bör	sågas	MfoN 17	50	186	5	7	3	1	252
vår	surströmming	MfoN 18	77	117	33	9	3	13	252
talar	svenska	MfoN 20	42	196		9	4	1	252
är	sant	MfoN 21	9	237	2	3	1		252
tycker	så	MfoN 25	77	110	2	31	14	18	252
morfars	språk	MfoN 28	172	36	2	36	4	2	252
förorter,	, så	MfoN 31	5	43		198	4	2	252
far	sa	MfoN 35	91	137	1	7	3	13	252
var	sysselsatt	MfoN 37	6	219	4	12	6	5	252
eller	stämpla	MfoN 38	19	200	2	9	3	19	252
här	stan	MfoN 40	107	104	6	2	5	28	252
där	så	MfoN 42	90	111	4	22	19	6	252
verkar	som	MfoN 43	45	187	1	8	4	7	252
tänker	skaffa	MfoN 44	86	133	6	4	6	17	252
stor	stad	MfoN 46	216	23	2	2	6	3	252
får	slita	MfoN 47	74	134	32	5	3	4	252
blir	slav	MfoN 49	39	142	54	5	4	8	252
syster,	, som	MfoN 50	89	104	1	33	22	3	252
heter	Sara	MfoN 51	12	216	1	1	6	16	252
tänker	sluta	MfoN 53	79	118	31	3	6	15	252
snart.	. Sara	MfoN 55	0	2	0	239	6	5	252
strävar	som	MfoN 56	89	127	2	22	7	5	252
vår	systems	MfoN 58	28	170	11	26	8	9	252
använder	Sara	MfoN 60	26	201	3	10	7	5	252
pärlosocker	sparsamt	MfoN 62	91	118	5	25	10	3	252
över	spisen	MfoN 67	66	154		25	6	1	252
arbetar	snabbt	MfoN 68	79	148	2	1	12	10	252
har	säker	MfoN 69	13	215	3	8	9	4	252
säker	smak	MfoN 70	130	102	4	8	6	2	252
kokar	stuvning	MfoN 71	81	139	2	9	6	15	252
säger	somliga	MfoN 73	28	199	6	3	9	7	252

Lexical Item A	Lexical Item B	Item #	[s]	/rs/	Maybe	Pause	Omit	Other	Total
tusenkonstnär,	, skulle	MfoN 80	35	67		133	10	7	252
har	sökt	MfoN 81	6	226	2	6	10	2	252
har	skott	MfoN 82	15	213	1	12	9	2	252
har	skottat	MfoN 83	10	220		5	9	8	252
gjort	staket	MfoN 85	166	62	7	8	9		252
har	spelat	MfoN 86	13	222	2	3	9	3	252
har	cyklat	MfoN 87	5	226	1	4	9	7	252
är	som	MfoN 89	70	170	1	1	10		252
vägdirektör	som	MfoN 90	80	102	1	49	16	4	252
har	sett	MfoN 95	20	212	1	7	8	4	252
har	stannat	MfoN 98	13	223		7	8	1	252
Helsingfors,	, Sundsvall	MfoN 101	42	72	4	93	8	33	252
har	sällan	MfoN 102	6	232	1	1	9	3	252
Turister	söker	MfoN 105	50	132	10	20	12	28	252
söker	sej	MfoN 106	122	107	3	5	12	3	252
ligger	soldyrkarna	MfoN 107	9	205	3	15	9	11	252
sardiner	stekta	MfoN 111	28	180	1	27	9	7	252
över	Spanien	MfoN 115	78	152		6	11	5	252
över	Spanien	MfoN 117	69	163	3	5	11	1	252
får	sammanfatta	MfoN 119	24	199		16	10	3	252
där	semestern	MfoN 120	47	183	3	8	9	2	252
semestern,	, så	MfoN 122	19	117	3	102	11		252
har	sett	MfoN 123	11	198		4	9	30	252
Här	slipper	MfoN 124	114	101	24	2	9	2	252
försöker	stanna	MfoN 126	37	188	1	1	10	15	252
går	säkerligen	MfoN 129	19	204	1	15	10	3	252
fördelar	som	MfoN 133	35	170	1	31	10	5	252
därför	som	MfoN 137	111	113	2	12	14		252
syster,	, styvfar	Msi 4	18	116	1	6	1	13	155
heter	Sylvester	Msi 6	3	150		1		1	155
har	slagit	Msi 8	52	82	6	5		10	155
ner	sina	Msi 9	28	118		3	1	5	155
gifter	sig	Msi 17	74	74	1	1		5	155
styvfar	sysslar	Msi 20	16	132	1	2	1	3	155
ner,	, så	Msi 23	18	110	1	23	3		155
är	skolsköterska	Msi 25	7	148					155
vår	skola	Msi 28	69	68		3		15	155
lärare	skulle	Msi 34	1	148		5		1	155
vår	styvfar	Msi 38	33	116		2		4	155
berättar	så	Msi 39	42	104		2	2	5	155
håller	sig	Msi 41	69	82		2		2	155
berättar	Sveriges	Msi 43	7	124		16	1	7	155
styvfar	säger	Msi 49	37	110	1	5		2	155
vår	största	Msi 51	33	118		2		2	155
är	Sigvard	Msi 54	2	149	1		3		155
är	större	Msi 58	10	144				1	155

Lexical Item A	Lexical Item B	Item #	[s]	/rs/	Maybe	Pause	Omit	Other	Total
Ingemar	som	Msi 59	19	102		33	1		155
Ingemar	som	Msi 66	56	88		11			155
Tyvär,	, säger	Msi 72	5	137		12		1	155
styvfar,	, så	Msi 74	4	71		77		3	155
människor	ska	Msi 79	23	111	1	10		10	155
efter	Sigvard	Msi 80	4	146	1	2		2	155
inkomster,	, så	Msi 87	18	79		58			155
inkomster	så	Msi 90	74	79				2	155
är	stark	Msi 97	14	140	1				155
tjur,	, snabb	Msi 99		130		23		2	155
utmaningar.	. Som	Msi 102		43		111		1	155
när	samerna	Msi 104	5	146		1	1	2	155
talat	samiska	Msi 113	5	146		2	2		155
samebarnen	som	Msi 114	1	116		33	2	3	155
när	samerna	Msi 118	4	146		3	1	1	155
ner	sina	Msi 121	48	93		3		11	155
renar,	, sa	Msi 122	2	60		90		3	155
Sigvard	skulle	Msi 123	120	29		6			155
djur.	. Så	Msi 135		38		116	1		155
Sigvard	svälter	Msi 138	73	79		3			155
är	så	Msi 139	9	144			2		155
därför	så	Msi 149	32	116		4	1	2	155
bryr	sig	Msi 153	70	85					155
betyder	så	Msi 159	56	94		1		4	155
ner	sig	Msi 174	48	93		10	3	1	155
kärr.	. Sen	Msi 177		13		139	2	1	155
var	synd	Msi 179	7	146		1		1	155
är	så	Msi 182	51	98	1		4	1	155
gård	står	Msi 186	65	72	1	14	2	1	155
över	staketet	Msi 189	91	63		1			155
Sigvard	skulle	Msi 191	114	38		2		1	155
över	staketet	Msi 192	76	78		1			155
Vårt	sommarställe	Msi 194	90	60	1	2		2	155
är	snart	Msi 198	9	146					155
plockar	så	Msi 204	32	111		3	6	3	155
bär,	, så	Msi 206	7	67		81			155
är	så	Msi 209	24	127			2	2	155
Vår	sommarstuga	Msi 212	27	124				4	155
ligger	sex	Msi 214		148		1		6	155
styvfar	säger	Msi 215	43	108		2		2	155
Sigvard	säkert	Msi 217	39	99	1	12	2	2	155
över	samernas	Msi 224	10	141	3	1			155
när	Sylvester	Msi 232	4	149		2			155
karl	som	Msi 235	24	81		49		1	155
var	Sigvard	Msi 240	2	152				1	155
har	sett	Msi 249	1	153				1	155

Lexical Item A	Lexical Item B	Item #	[ʂ]	/rs/	Maybe	Pause	Omit	Other	Total
Sigvard,	, sa	Msi 250	83	55		17			155
har	sett	Msi 254	11	142		2			155
Sigvard,	, sa	Msi 255	88	55	1	7	1	3	155
är	säker	Msi 258		153	1		1		155
karl	som	Msi 261	34	63		9	43	6	155
har	sett	Msi 264	10	137				8	155
styvfar	säger	Msi 265	40	111		2		2	155
där	som	Msi 274	64	73	1	10	2	5	155
efter	sig	Msi 276	124	29				2	155
brukar	säga	Msi 278	13	137	1	1		3	155
ter	sig	Msi 280	113	39		1	2		155

APPENDIX I: LEXICAL ITEMS FOR ÄR 'IS/ARE/AM/

Item #	Item:	[e]	%[e]	[r],[ʂ]	%[r]	Other	Total
MfoN 3	är stolt	208	83%	39	15%	5	252
MfoN 21	är sant	195	77%	51	20%	6	252
MfoN 22	är något	220	87%	26	10%	6	252
MfoN 64	är ganska	220	87%	24	10%	8	252
MfoN 75	är världens	192	76%	47	19%	13	252
MfoN 136	är därför	199	79%	41	16%	12	252
Msi 25	är skolsköterska	139	90%	16	10%		155
Msi 32	är inte	124	80%	30	19%	1	155
Msi 54	är Sigvard	103	66%	48	31%	4	155
Msi 58	är större	114	74%	40	26%	1	155
Msi 63	är till	145	94%	8	5%	2	155
Msi 91	är skyldig	151	97%	3	2%	1	155
Msi 97	är stark	129	83%	25	16%	1	155
Msi 115	är tvungna	143	92%	11	7%	1	155
Msi 143	är lika	154	99%	0	0%	1	155
Msi 165	är mödan	131	85%	23	15%	1	155
Msi 198	är snart	124	80%	31	20%		155
Msi 252	är bara	147	95%	8	5%		155
Msi 258	är säker	135	87%	18	12%	2	155
MfoN 32	är jag	143	57%	105	42%	4	252
MfoN 39	är det	4	2%	241	96%	7	252
MfoN 78	är jag	196	78%	49	19%	7	252
MfoN 89	är som	105	42%	135	54%	12	252
Msi 75	är det	33	21%	120	77%	2	155
Msi 83	är försiktig	134	86%	19	12%	2	155
Msi 89	är hans	10	6%	143	92%	2	155
Msi 94	är en	74	48%	80	52%	1	155
Msi 136	är det	113	73%	37	24%	5	155
Msi 139	är så	144	93%	9	6%	2	155
Msi 161	är det	52	34%	102	66%	1	155
Msi 182	är så	75	48%	74	48%	6	155
Msi 209	är så	100	65%	51	33%	4	155
Msi 226	är det	85	55%	70	45%		155
Msi 260	är den	124	80%	27	17%	4	155
Msi 269	är det	46	30%	107	69%	2	155

215

VITA

VERN MYRON LINDBLAD

University of Washington

1999

1967 B.A. ENGLISH LITERATURE REED COLLEGE

Thesis: Pieces and Colors (creative writing, short stories)

1979 B.A. FOREIGN LANGUAGES PORTLAND STATE UNIVERSITY

1990 M.A. LINGUISTICS UNIVERSITY OF WASHINGTON

Thesis: Neutralization in Uyghur (phonology of an eastern Turkic language)