TOPICS IN HINDI-URDU PHONOLOGY
Ohala, Manjari
ProQuest Dissertations and Theses; 1972; Dissertations & Theses @ University of California
pg. n/a

INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

A Xerox Education Company

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

72-25,819

OHALA, Manjari, 1943-
TOPICS IN HINDI-URDU PHONOLOGY.

University of California, Los Angeles, Ph.D., 1972
Language and Literature, linguistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

©Copyright by

Manjari Ohala

1972

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UNIVERSITY OF CALIFORNIA

Los Angeles

Topics in Hindi-Urdu Phonology

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Linguistics

by

Manjari Ohala

Doctoral Committee:

Professor William O. Bright, Chairman

Professor Victoria A. Fromkin

Professor Barbara H. Partee

Professor Stanley A. Wolpert

Professor Murray J. Leaf

1972

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The dissertation of Manjari Ohala is approved, and it is acceptable in quality for publication on microfilm.

Stanley A. Wolpert

Murray J. Leaf

Barbara H. Partee

Victoria A. Fromkin

William O. Bright, Committee Chairman

University of California, Los Angeles

1972

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

PLEASE NOTE:

Some pages may have

indistinct print.

Filmed as received.

University Microfilms, A Xerox Education Company

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES iv
ACKNOWLEDGEMENTS vi
VITA AND PUBLICATIONS viii
ABSTRACT ix
INTRODUCTION 1
CHAPTER 1 10
CHAPTER 2 27
CHAPTER 3 170
CHAPTER 4 212
BIBLIOGRAPHY 267
APPENDIX 1 273
APPENDIX 2 282
APPENDIX 3 294

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES AND TABLES

FIGURES

Figure 1 40
Figure 2 71
Figure 3 72
Figure 4 73
Figure 5 73
Figure 6 74
Figure 7 191
Figure 8 192
Figure 9 192
Figure 10 193
Figure 11 194
Figure 12 246
Figure 13 246
Figure 14 246
Figure 15 248
Figure 16 248
Figure 17 275
Figure 18 278
Figure 19 288

TABLES

Table 1 11
Table 2 13
Table 3 18
Table 4 60
Table 5 61
Table 6 63
Table 7 64
Table 8 68
Table 9 70
Table 10 90
Table 11 91
Table 12 92
Table 13 109
Table 14 114
Table 15 115
Table 16 116
Table 17 116
Table 18 117
Table 19 128
Table 20 158
Table 21 159
Table 22 182

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 23 184
Table 24 185
Table 25 189
Table 26 190
Table 27 198
Table 28 212
Table 29 215
Table 30 218
Table 31 219
Table 32 221
Table 33 223
Table 34 225
Table 35 228
Table 36 231
Table 37 232
Table 38 233
Table 39 238
Table 40 240
Table 41 274
Table 42 283
Table 43 286
Table 44 288

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGEMENTS

I would like to thank, first of all, my entire doctoral committee for their patience during the prolonged writing of this dissertation. To my committee chairman, William Bright, I am grateful for his many helpful comments resulting from his careful reading of the preliminary draft of this dissertation. To Victoria Fromkin I owe special thanks for her thorough reading and re-reading of the dissertation and for her lengthy constructive and challenging criticisms of it. But for her efforts this dissertation would be far weaker in both style and content. I am also deeply grateful to her for her constant encouragement. I also give warm thanks to Stanley Wolpert for his generous supply of coffee, sympathy, and help during my tumultuous first years as a foreign student in this country.

In writing this dissertation and in conducting the various experiments reported in it I have also benefited greatly from many discussions and suggestions from my "unofficial committee member", John Ohala. Without his encouragement and help (and prodding) this dissertation might never have been completed. I also thank him for his excellent work in making many of the illustrations for this study.

To my family and friends I am in debt for their unflagging support and understanding during the long years I worked on this dissertation.

I also thank my many informants, both in the United States and in India, for generously giving me their time and cooperation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Much of the experimental or instrumental work reported here was done using the facilities of the Phonology Laboratory, Department of Linguistics, University of California, Berkeley, which is supported in part by a grant from the National Science Foundation. I am especially grateful to Robert Krones for his help and time in getting all the machines to cooperate. The experiment reported in Appendix 1 was conducted in the Research Institute of Logopedics and Phoniatrics, University of Tokyo. I thank Osamu Fujimura for making the facilities of his institute available to me. In particular I thank Dr. Hajime Hirose for making my introduction to physiological experiments quite painless.

Finally I offer my gratitude to Meredith Black Hoffman and to Arax Kizirian for their fine job of translating my manuscript into this typescript.

I am, of course, responsible for any errors contained in this work.

Manjari Ohala

May 1972

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

VITA

September 25, 1943--Born, Murree, West Pakistan

1960--A.B., Banaras Hindu University

1962--M.A., Banaras Hindu University

1965-1967--Teaching Assistant, Department of Classics, Section of Indo-European Studies, University of California, Los Angeles

1967--M.A., University of California, Los Angeles

1967-1969--Teaching Associate, Department of Classics, Section of Indo-European Studies, and Department of Linguistics, University of California, Los Angeles

1968--Language Coordinator, United States Peace Corps (through University of California Extension, Los Angeles)

1970-1971--Teaching Associate, Department of Near Eastern Languages, University of California, Berkeley

PUBLICATIONS

Ohala, M. and Ohala, J. In press. The problem of aspiration in Hindi phonetics. Bhasha. (In Hindi) To appear in English in: The Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, University of Tokyo, and in: Project of Linguistics Analysis Reports, Second Series, No. 16. Phonology Laboratory, University of California, Berkeley.

Ohala, M. In press. Conflicting expectations for the direction of sound change. Project of Linguistics Analysis Reports, Second Series, No. 16. Phonology Laboratory, University of California, Berkeley.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT OF THE DISSERTATION

Topics in Hindi-Urdu Phonology

by

Manjari Ohala

Doctor of Philosophy in Linguistics

University of California, Los Angeles, 1972

Professor William O. Bright, Chairman

This dissertation is a study of selected topics in the phonology of Standard Hindi. The format used is that of generative phonology, with modifications suggested where necessary.

The first chapter discusses the segments that have to be posited at the lexical level, and the features necessary to distinguish these segments and state natural classes. Evidence from a physiological investigation showed that the Chomsky and Halle feature "Heightened Subglottal Pressure" is not appropriate to characterize the Hindi aspirates and breathy voiced obstruents, and instead a feature "Reduced Glottal Resistance after Release" should be used. Also demonstrated is the need for a feature "Distinctive Release" in order to explain why aspirates, breathy voiced stops, and affricates act

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

similarly.

The second chapter deals with the sequential constraints of Hindi. Evidence from an original experiment and from others' experiments suggest that speakers cannot meaningfully make the dichotomous judgment "possible/impossible" with respect to morphemes for their language, but rather only scalar judgments on the continuum: "regular--irregular". It is also argued that morpheme structure conditions should not be restricted to apply only within morphemes; that some of them might be "everywhere" constraints, applying both within and across morpheme boundaries. Thus these are best called simply "sequential constraints".

Experimental evidence is provided for the "psychological reality" of an "abstract" segment in the lexical form of certain morphemes, e.g., that [$g^{h}\tilde{o}sla$] "nest" has the underlying form /$g^{h}\tilde{o}səla$/, the correct phonetic form being derived by the ə-deletion rule.

The initial, medial, and final consonant clusters of Hindi are listed and exemplified in detail and the sequential constraints they manifest are given. Experimental evidence is provided for the psychological reality of some of these constraints.

In chapter three the problem of nasalization in Hindi is discussed. The canonical form of native morphemes containing long nasalized vowels is posited to be different depending on whether voiced or voiceless stop follows, specifically, long nasalized vowel plus stop, if the stop is voiceless, but long nasalized vowel plus homorganic nasal plus stop, if the stop is voiced. It is proposed that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

nasalized vowels have to be admitted at the lexical level.

In chapter four a new formulation of the ə-deletion rule in Hindi is proposed. Whether or not a /ə/ is deleted in a particular form depends not only upon the phonetic environment, e.g., how many consonants precede and follow it and the regularity of consonant clusters that would result upon deletion of the /ə/, but also upon sociolinguistic factors.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Introduction

Hindi-Urdu: The language whose phonology is described in this dissertation is Standard Hindi. The principal informant, myself, is a third generation native speaker of Standard Hindi. Even though the dialect being described is mine, I have made every effort to verify that the description applies to other speakers as well; their responses and judgements have been noted throughout the dissertation.[1] The Hindi that I am referring to is used in everyday casual speech by educated native speakers in cities such as Banaras, Lucknow, Delhi, etc. It is contended that the difference between speakers of these cities is minimal.

The hyphenated term 'Hindi-Urdu' implies that the language described contains what is common to Hindi and Urdu, i.e. it is not the highly Sanskritized Hindi ("high Hindi", or what is called literary style Hindi), nor the highly Perso-Arabicized Urdu ("high Urdu", or what is called the literary style Urdu). In other words it is the dialect spoken by educated urban speakers in casual conversation. My use of the term 'Hindi-Urdu' is thus similar to that of Gumperz and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Naim (1960) and Kelkar (1968). For a detailed account of the different functions of "high Hindi", "high Urdu", and Hindi-Urdu, see Kelkar. Henceforth I will use the terms "Hindi", "Standard Hindi", and "Hindi-Urdu" interchangeably. When I describe characteristics of "high Hindi", "high Urdu", or other dialects of Hindi (e.g. the rural dialects), I explicitly name the dialect.

There are those who doubt that native speakers of Standard Hindi exist, so it is best to be quite clear as to what is meant by a native speaker. By a native speaker of a language is generally meant one who has learned that language as his first language[2]---"at his mother's knee".

Perhaps a century ago it was true that people learned what is called Standard Hindi only as a second language, their first language usually being one of the Hindi vernaculars such as Braj, Bhojpuri, or Awadhi. (For an excellent and detailed account of the functional relationship between the vernacular Hindi and Standard Hindi today, see Gumperz and Naim (1960).) However, today there most certainly are speakers (even third and fourth generation) who use Standard Hindi for their casual everyday conversation. Thus Gumperz and Naim say:

> Native speakers of Hindi-Urdu still represent a minority, although an ever increasing one. They are found primarily among the westernized, urban groups, who have moved away from their home region and have lost connection with their rural past. For a much larger number of individuals the standard continues to function as a second or third speech style, used only in certain social situations... (97)

Standard Hindi is not necessarily to be considered the prestige language to be used in formal situations: although those who are not native speakers of Standard Hindi may use it in this way, the native speakers

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of Standard Hindi would use "high Hindi" or "high Urdu" in formal situations. Thus, for these speakers, the use of the standard does not in itself imply the connotations of "prestige" usually associated with the word "standard".

Aim: This dissertation is entitled "Topics in Hindi-Urdu Phonology". Naturally, the precise form of a phonology of a language will be different depending on the goal. If our aim is simply a descriptive one we are "merely cataloging the results of interacting historical processes" (Ladefoged 1971a:49). In a grammar of this type, as long as the description notes the existing regularities or patterns in the sound system of the language and states these patterns as generally as possible, it will be justified. And of course there might be a number of descriptions which will be equally valid in this kind of treatment of the data. In such a grammar abstract forms could be posited without any further justification than the fact that it is possible to state certain regularities if one assumes them. As Ladefoged has pointed out, this is what Chomsky and Halle's analysis in Sound Patterns of English amounts to. This is true in spite of the modern phonologists' claim that they wish to describe the competence of the native speaker, that is, what he "knows" (tacitly) about the system of his language. In fact, a large number of grammars written in the generative format in the sixties, in that they lack empirical verification, cannot be assumed to represent grammars of competence, but simply descriptive grammars.

On the other hand, instead of description, our goal might be "explanation" of sound patterns. As Ohala (1971c) has pointed out, a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

grammar interested in explanation will have to differentiate between those sound patterns which a native speaker is "aware" of and uses productively and those that are in the language simply due to usage. However it is not simply the goals of the linguist that make the difference between a competence grammar and a purely descriptive grammar because, as has already been pointed out, generativists have always aimed for a competence grammar. The key difference lies in what evidence is provided for the "psychological reality" or lack of it for a given sound pattern (or for a posited 'abstract' segment, see Kiparsky 1968 and 1971). The evidence provided however cannot simply be the fact that a pattern exists and can be formally stated. As Zimmer has remarked, "It is occasionally assumed that, if a regularity can be stated, this alone permits us to infer some kind of psychological reality" (1969:309). Zimmer gives experimental evidence that this assumption is not justified. Nor can the evidence simply be that all the formally stated rules "work" because as Ladefoged has recently pointed out, "Nothing is proved about present day psychological reality by the fact that these unverified guesses all knit together, supporting each other by forming an intricate system of rules" (50). If the aim of our grammar is to account for a speaker's competence, then of course we will have to provide evidence that what we are claiming as part of a speaker's "knowledge" is indeed such. Zimmer and others have indicated appropriate experimental procedures for doing this.

Although construction of a good descriptive grammar is by no means a trivial task, my own interest lies in accounting for the Hindi speakers' competence, i.e. providing evidence for the psychological reality

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of certain sound patterns of Hindi. However if I restrict myself to stating just those patterns for whose psychological reality I had evidence, then I would not be able to say very much about the sound patterns of Hindi (see Ladefoged's similar comment for English (49)).

Therefore this study is a mixture of two things. In a number of cases I have tried to provide evidence for the particular analysis I have chosen. Some of the evidence is from experimental data. Other of the evidence is of a varied sort: from hypercorrection, from children's mistakes, from native speakers' reactions to certain forms, e.g., their regularizing exceptions, their not deleting certain segments that meet the environment of a rule, and their different responses to the same item under different circumstances.

In those cases where I have adopted certain analyses without much evidence, I have generally labelled these as descriptions or as points requiring further testing.

The model used in this dissertation is that of generative phonology, with modifications suggested throughout. Since linguistic theory is still in the stages of finding out relevant facts about languages that need to be incorporated, whenever the current theory was found to be inadequate in some way in accounting for the facts of Hindi in the most explanatory manner, I have chosen to modify the theory.

It has been assumed by generative phonologists that there is a unique grammar of a language, THE one that the child chooses. While I don't think this is the case (I discuss this more in section 2 of chapter 2), certainly the number of such grammars that could be said to be competence grammars will be quite limited since there are undoubt-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

edly severe restrictions on the kind of representation such knowledge can have in the brain. An evaluation procedure or metric is posited by the generativists to measure the complexity of the grammar in order to filter out non-optimal grammars. However none of the proposed evaluation metrics have met with any success or universal endorsement by linguists (see Fromkin 1971c). This is not surprising since little attempt has been undertaken to find out what the functional constraints of the human brain are (see Ohala 1971b). Therefore my grammar is not restricted by any of the proposed evaluation metrics. That is, it is not the case that I have chosen or rejected an analysis solely on the basis of how it would add to the complexity of a grammar. Lacking information on how phonological facts are represented in the brain, it seems to me preferable at this point to try to account for the sound patterns in intuitively the best way and try to secure evidence for the "psychological reality" or productivity of the particular analysis. We can worry about the exact formalism of the productive rule in the speakers' head later.

I have indicated in a number of cases the kind of test or evidence that would be required so that future studies might reveal the inadequacies of my analysis or, hopefully, justify my analysis. In any case I have tried to be explicit about my reasons for adopting the particular analyses that I have chosen.

I have no illusions that this study is the last word on these aspects of Hindi that it touches on. There are a number of open questions on certain points since experimental evidence could not be obtained due to limited resources and time.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I hope that linguists working on linguistic theory and linguists interested in details of Hindi phonology will both find something of interest in this study.

Organization: The first chapter deals with the segments of Hindi that need to be posited at the underlying level, the problem of certain segments that occur only in loans, the features used, including some modifications of the conventional Chomsky and Halle features, and a few segment structure conditions of Hindi. Appendix 1 discusses the Chomsky and Halle feature "heightened subglottal pressure" for Hindi aspirates, and provides experimental evidence indicating its inappropriateness for Hindi.

Chapter 2 deals with the sequential constraints of Hindi. It has three sections:

Section 1 contains a discussion on current morpheme structure condition theory, and some suggested revisions of it.

Section 2 discusses the 'abstractness' issue. I provide some experimental evidence for the "psychological reality" of certain "abstract" segments in Hindi.

Section 3 gives a detailed account of the initial, medial, and final clusters of Hindi, and gives the if-then sequential constraints necessary to account for these.

Appendix 2 reports a test conducted to investigate the psychological reality of certain morpheme-initial consonant cluster constraints in Hindi. Appendix 3 lists examples of the various clusters.

Chapter 3 discusses the problem of nasalization in Hindi, and provides instrumental evidence to show that what is claimed to be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[$\tilde{V}$:C] in Hindi is in a number of cases [$\tilde{V}$:NC] (where N=homorganic nasal), and suggests that this perhaps reflects a sound change that has taken place in Hindi within the last century.

Chapter 4 discusses the ə-deletion rule in Hindi, and proposes a ə-deletion rule that not only takes into account the phonetic environment but sociolinguistic environments as well.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Footnotes

1 I have tried to use only native speakers of Standard Hindi for the most part. I noted down any other languages the informants said they also spoke (these ranged from Marathi and Bengali to Russian and German); but I had no way of finding out how proficient they were in these languages. If an informant also listed one of the rural dialects of Hindi, such as Braj or Bhojpuri, I did not use those responses of his which differed from the responses of other speakers. All my informants also knew English.

2 Sometimes of course, in bilingual situations, a child might learn two languages at the same time, and both would have to be considered his "mother tongue".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 1

1.0. The contrasting phonological segments of Hindi will be described by a set of features chosen from the universal set of distinctive features. For the most part I have used the Chomsky and Halle (1968) feature system not necessarily because it is the best system, but because it is a possible feature system. In certain cases I have made modifications in the features. These will be discussed below.

Table 1 lists the contrasting segments of Hindi at the systematic phonemic level and the features needed to distinguish them. Generally only those feature values have been specified that are necessary for distinguishing the segments. This table is given mainly for descriptive purposes.[1] (That is, so that a reader can interpret the sequential constraints and phonological rules given later; it is not claimed that a comparable chart would form a part of a native speaker's grammar.)

The segments in table 1 also appear at the phonetic level, but so do the segments in table 2, which, however, I do not include in the list of systematic phonemes.[2]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I.P.A. equivalent:	i ĩ ɛ̃ a ɑ̃ ɔ̃ ũ	i ɪ e ɛ æ ɑ ɔ o u ʊ	ɯ j	y ɤ ɐ	h ɾ ɾ	l l̃ ɻ	m ɱ	n ɳ	p p̃	pʰ	b	bʱ	t	tʰ	d	dʱ	ʈ	ʈʰ	ɖ	ɖʱ
sonorant	+	+ + + + + +	+ +	+ + +	+ + +	+ + +	+ + +	+ +	- -	-	-	-	-	-	-	-	-	-	-	-	-
consonantal	-	- - - - - -	- -	- - -	- - -	- - +	+ + +	+ +	- -	-	-	-	-	-	-	-	-	-	-	-	-
syllabic	+	+ + + + + +	+ +	+ + +	- - -	- - -	- - -	- -	+ +	+	+	+	+	+	+	+	+	+	+	+	+
high	+	- - - - - -	+ +	- - -	- - -	+ + +	-														
low	-	- - + + - -	- -	- - +	+ - -	- - -	+														
back	-	- - + + + +	- -	- - -	+ + +	- + +	-														
long	+	+ + - + + -	- +	+ - -	+ - -	- - +	-	+	- -	+	-	+	-	+	-	+	-	+	-	+	-
continuant						- -	+ +	- -	- -	-	-	-	-	-	-	-	-	-	-	-	-
anterior						- + +	-	+ +	+ +	+	+	+	+	+	+	+	+	+	+	+	+
coronal						-		- -	+ +	-	-	-	-	+	+	+	+	+	+	+	+
retroflex																					
nasal	+	+ + + + + +	- -	- - -	- - -	- - -	- -	+	+ +												
voice											- -	-	+	+	-	-	+	+	-	-	+
reduced glottal resistance after release											- -	+	+	- -	+	+	- -	+	+	- -	+
delayed release											- -	-	-	- -	-	-	- -	-	-	- -	-
distinctive release											- -	+	+	- -	+	+	- -	+	+	- -	+

Table 1

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I.P.A. equivalent:	t̪ t̪ʰ tʰ cʰ qʰ	t̪ʲ t̪ʰʲ tʲ tʰʲ qʲ	d̪ d̪ʲ dʲ qʲ	k kʲ kʰ kʰʲ	g gʲ	ɖʰ ɟʰ	ʧ ʧʰ ʧʲ ʧʰʲ ʤʲ ʤʰ ʤʰʲ	ɛ ɛʲ	ʃ ʃʲ	ʂ ʂʲ	ɣ ɣʲ	x xʲ	ʃ
sonorant	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
consonantal	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +
syllabic	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
high				+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +		
low				- -	- -	- -	- -	- -	- -	- -	- -		
back				+ +	+ +	+ +	+ -	- -	- -	- -	- -		
long	- +	- +	- +	- +	- +	- +	- +	- +	- +	- +	- +	- +	- -
continuant	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	+ +	+ +
anterior	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	+ +	+ -
coronal	+ +	+ +	+ +	- -	- -	- -	+ +	+ +	+ +	+ +	+ +	+ +	+ -
retroflex	+ +	+ +	+ +	- -	- -	- -	- -	- -	- -	- -	- -		
nasal													
voice	- -	- +	+ +	- -	- +	+ +	- -	- +	+ +	+ +	+ +	- -	+ -
reduced glottal resistance after release	- -	+ -	- +	- -	+ +	- -	+ +	- -	+ +	- -	+ +		
delayed release	- -	- -	- -	- -	- -	- -	+ +	+ +	+ +	+ +	+ +		
distinctive release	- -	+ -	- +	- -	+ +	- -	+ +	+ +	+ +	+ +	+ +		

Table 1 (continued)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2

Column 1 gives the transcription used in the thesis, column 2 gives the IPA equivalent where necessary, and/or an approximate physiological description.

[ŋ]

[ñ] Lamino-alveolar nasal (Dixit 1963:116)

[ɾ̃] Flapped retroflex nasal[3]

[ɳ] [ɳ̆] (non-flapped retroflex nasal)

[z̃] [ʒ̃]

[ĩ] [ĩ]

[õ]

[ũ] [ɔ̃]

The omission of the segments in table 2 from the list of systematic phonemes in table 1 requires some justification.

1.1. One important question concerns the segments that occur only in loanwords, e.g. [f] [z] etc. Linguists working on languages with a large number of borrowings---such as Hindi and other Indo-Aryan languages--are aware that native speakers may treat some morphemes differently from others; this can be due to such factors as the special nature of certain segments or clusters contained in them. There has been a variety of treatments of loans by structuralist and generative linguists (for a review of this, see Saciuk 1969). I propose that in the lexicon of a grammar of Hindi, morphemes be marked as [± native], and those designated [-native] be either [+Perso-Arabic] (PA), or [+Sanskrit], or [+English].[4] The test reported in Appendix 2 seems to support such a division of the vocabulary in that this division appears to reflect 'competence'.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The use of certain morphemes, or of certain clusters, or certain segments, has, in addition, sociolinguistic significance[5]; it marks the style of speaking, and reveals information on the speaker's background, e.g., whether the speaker draws on PA or Sanskrit loans for his higher vocabulary may indicate whether he is a Hindu or a Muslim.

1.1.1. Of the consonant segments listed in table 1, /f/ /z/ and /š/ occur only in loans--though quite common loans; all other segments occur in what can be called "common core" Hindi-Urdu, i.e. that which is common to all Hindi speakers. /š/ occurs in Sanskrit, English and PA loans, and /f/ and /z/ occur only in English and PA loans.[6] Since most speakers of standard Hindi do have these three segments in their speech, they have been included in table 1. However my calling them 'loan' segments can be questioned. In my view even standard Hindi speakers whose speech contains /f/ /z/ and /š/ are aware that these segments are in some way different from the "common core" segments. This is in part due to their awareness of the fact that many non-standard Hindi speakers do not have some or all of the segments [f] [z] and [š] (and substitute [pʰ] [j] and [s] for them, respectively) whereas they always have the correct "common core" segments. This knowledge can be represented by a sociolinguistic stylistic rule which informally can be stated as (1).

(1) S-rule: $\begin{bmatrix} š \\ f \\ z \end{bmatrix} \rightarrow \begin{bmatrix} s \\ pʰ \\ j \end{bmatrix} \Big/ \begin{Bmatrix} \text{uneducated style;} \\ \text{less elegant style} \end{Bmatrix}$

I have not listed /ž/ and /r̃/. The former occurs in a few English loans like "pleasure", "rouge" in the speech of Westernized, educated, standard Hindi speakers. [r̃] occurs in a few Sanskrit loans like [gur̃] "quality", etc. and is pronounced by most speakers as [n] in informal style, and as [r̃] only in formal style (and only by those speakers who

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

draw on Sanskrit for higher vocabulary).

I have not listed /ŋ ɲ̃ ɳ/ in table 1. This is because I am tentatively adopting the analysis which considers these to be underlying segments of a very special kind---segments that occur only before homorganic segments (see 2.3.5.1. for details). Writers such as Dixit and Kelkar (1968) consider /ŋ/ to be an independent phoneme on the basis of a few Sanskrit loanwords where /ŋ/ occurs before a non-velar segment, e.g. [waŋməy] "literature" and [diŋnad] "reverberating sound". But these loans occur only in 'high' Hindi. Either these few words can be listed as exceptions, or they can be specified as /diNgnad/ etc., i.e. with a following velar stop. In Hindi in certain phonetic environments the g after ŋ is deleted. Thus we have [ʤəŋgəl] "forest" but [ʤəŋli:] "wild" (see the spectographs in chapter 4). We can assume that in the case of exceptions like [waŋməy] the velum remains open in anticipation of the following nasal, and thus the phonetic rule that will be needed to delete the g to obtain [ʤəŋli:] could also delete an underlying g to obtain [waŋməy]. It is possible that this phonetic rule is also required in allegro pronunciation of sequences like -ŋg+nasal or -ŋg#nasal. For example, it is possible that si:ŋg#mẽ "horn+in" ("in the horn") is phonetically [si:ŋmẽ] in running speech.

Another analysis is also possible. Namely, that we derive words such as [waŋməy] from underlying /wak+məy/ via the Sanskrit sandhi rules. I reject this analysis on grounds that there is no evidence for the 'psychological reality' of the Sanskrit sandhi rules in a grammar of a Hindi speaker (for details see 2.3.5.1).

x̲ ɣ̲ and q̲ have not been listed in the list of phonetic segments (table 2), because these occur only in the speech of Urdu speakers, and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

not in the dialect of the informant for this thesis (myself).

ʔ (the so called 'ain'), ɡ̱ ṛ (syllabic ṛ) occur in the writing system, but in the speech of only very few speakers---the former, I think, occurs only in the speech of Maulvis (Muslim 'priests') and the latter two in the speech of Pandits (Hindu 'priests'). Therefore I do not think these should be included in a grammar of Hindi-Urdu.

Among vowels, I have not listed the short nasalized vowels in table 1, since I propose that they are derived by rule from oral vowels. In many morphemes the long nasalized vowels can also be derived by rule, however there are other morphemes where we have to recognize nasalized vowels at the lexical level; therefore I have included long nasalized vowels in table 1 (see chapter 3 for details on nasalization).[7]

Of the other vowels [æ] occurs in my speech only in English loans such as [bæṭ] "bat". All other vowels listed in table 1 occur in native words, e.g.,:

(2)

[i]	[gin]	"count"	Also	in	the	English	loan	[bil]	"bill"
[i:]	[si:kʰ]	"learn"	"	"	"	"	"	[si:l]	"seal"
[e]	[ek]	"one"	"	"	"	"	"	[kek]	"cake"
[ɛ]	[bɛṭʰ]	"sit"	"	"	"	"	"	[čɛk]	"check"
[ə]	[əb]	"now"	"	"	"	"	"	[bəs]	"bus"
[a]	[ap]	"you"	"	"	"	"	"	[fadər]	"father"
[ɔ]	[kɔn]	"who"	"	"	"	"	"	[bɔl]	"ball"
[o]	[nok]	"point"	"	"	"	"	"	[kona]	"kona coffee"
[u:]	[u:n]	"wool"	"	"	"	"	"	[sṭu:l]	"stool"
[u]	[bun]	"knot"	"	"	"	"	"	[buk]	"book"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I am treating 'geminates' as long consonants, i.e. as long segments and not as clusters. Gumperz and Naim (1960) claim that "Aspirates do not occur as geminates"(103); however phonetically it seems to me that what is written as kk^{h} (i.e. k+k^{h}) etc. can simply be treated as a long aspirate. My justification for treating geminates as units is that we thereby avoid all the following disadvantages which would result from treating them as clusters:

(a) Geminates would be clusters of a very peculiar sort: they would be the only clusters that require that a short vowel precede and that a vowel follow.

(b) The statement of SQC 2 (cf. 2.3) would be more cumbersome (SQC 2 states that no two stops of the same point of articulation follow each other); we would have to add that stops with the same point of articulation which also have voicing agreement are permissible.

(c) SQC 3 (cf. 2.3) which states that no two [+distinctive release] segments can occur together, would similarly have to be complicated. It would have to allow [c̆c̆] [ɟ̆ɟ̆] etc. but rule out [c̆p^{h}] etc. Moreover it seems that in all other clusters where affricates are involved there is a release of the affricate but we would have to posit a rule for [c̆c̆] etc. saying that in such cases the first affricate is not released.

r y and h do not occur as long consonants. Writers like Arun (1961) and Dixit also list /y:/ and give /nyay:/ "just" and /tay:ari:/ "readiness" as examples. However I think the former is phonetically [nyay] and as far as the latter is concerned it could as well be transcribed as [təiyari:] which would yield the same pronunciation. Therefore I have not included [y:]. I have also not been able to find examples of f and b^{h} as long consonants. In the case of [s̆:] I have not

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

been able to find examples which could clearly be said to be morpheme medial. In examples such as niššiːl "without shame" niššəbd "without sound" it is possible to posit a morpheme boundary (whether we should or not is another question) /niš+šiːl/ /niš+šəbd/. Therefore I have not listed [šː]. Some examples of medial long consonants are given in table 3.

Table 3

[məkʰːiː]	"fly"
[bəgʰːiː]	"carriage"
[əcʰːa]	"good"
[ujʰːit]	"abandoned"
[pəṭʰːe]	"nerves"
[buḍʰːa]	"old"
[pətʰːər]	"stone"
[budʰːiː]	"intellect"
[pʰupʰːus]	"lungs"
[izːət]	"honour"
[pətːa]	"leaf"
[ḍibːa]	"box"
[həlːa]	"noise"

I have listed /ṛ/ as a segment. I think most writers today treat [ṛ] (and ṛʰ) as phonemes rather than as allophones of /ḍ/ (and /ḍʰ/ in the case of ṛʰ), e.g., Arun, Mehrotra (1964), Dixit, Kelkar. The problem is that until certain loanwords like English "soda"[8] came into Hindi, [ṛ] and [ḍ] were in complementary distribution. (For details see the above writers and Misra 1967, Gumperz and Naim.)

I have not listed lh rh mh nh in the list of phonetic segments,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

since I am treating them as clusters (see section 3, chapter 2, for some discussion). I have also not listed ṛh. The case of ṛh is more difficult. In some cases it seems that it should be treated as a cluster, e.g., for the word written गुड़हल {guṛhal}, "hibiscus", one pronunciation is [guṛhəl] (this involves the problem of VhV which I shall not be discussing). In other cases such as in the word [pi:ṛha] "stool" it seems to make no difference whether we treat the -ṛh- as a unit or a cluster. I am tentatively treating all cases of ṛh as a cluster of ṛ+h.

Some writers such as Srivastava (1970) have treated [y] and [w] as allophones of /i/ and /u/. However I am treating them as segments at the lexical level. My reasons for doing this are presented in chapter 2 (2.3.8).

1.2. Features: The following is the list of features I am going to use. Some of these are for the sake of defining the contrasts that exist at the systematic phonemic level, and others are used simply for the sake of more adequate representation of natural classes. The first twelve of these features are fairly straightforward; for their definition see Chomsky and Halle (1968).

1. Sonorant
2. Consonantal
3. Syllabic
4. High
5. Low
6. Back
7. Continuent
8. Anterior
9. Coronal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

10. Nasal

11. Voice

12. Delayed release

The following features need some comment:

13. Long: I have not used the feature 'tense' for long vowels and consonants and have simply used [± long]. There does not seem to be any advantage in using 'tense' over 'long' since the physical correlates of 'tense' are still controversial.

In the case of the vowel pairs [i:] vs. [i], [u:] vs. [u], [ɛ] vs. [e], [ɔ] vs. [o], [ə] vs. [a], the difference is not just one of length but also of quality (Bright 1965). Dixit in his phonetic study on Hindi vowels also comes to the conclusion that "quality is the basic and primary characteristic of Hindi vowels and that length is secondary"(57). However since it is only the feature 'long' that differentiates /i/ from /i:/ and /e/ from /ɛ/ etc. (the problem would be the same if tense had been used) we are forced into making length the primary characteristic and quality the secondary. This is an inadequacy in a feature system which recognizes only three tongue heights for vowels. We will thus need phonetic rules to specify that a qualitative difference is correlated with length in order to obtain the correct phonetic output.

14. Retroflex: Chomsky and Halle use the feature 'distributed' for retroflex segments and define "Distributed sounds [as those which] are produced with a constriction that extends for a considerable distance along the direction of air flow; nondistributed sounds are produced with a constriction that extends only for a short distance."(312). This is, in my view, an ill-defined feature. The difference between retroflex [ʈ] and nonretroflex [t] is not so much the place of articula-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tion or the area of contact, but the 'retroflexion' of the tongue. Thus I use the feature retroflex.

I have distinguished [ɽ] from other liquids by the feature 'retroflex'. Chomsky and Halle do not have any feature proposal regarding this sound. Since the only retroflex liquid in Hindi is [ɽ], my analysis works. The Hindi [ɽ] is, however, very different from the English retroflex [ɹ], therefore at the phonetic level a feature that indicates that the Hindi retroflex is 'flapped' is needed. There can be a low level phonetic rule like (3):

(3) Phonetic rule: $\begin{bmatrix}+\text{liquid}\\+\text{retroflex}\end{bmatrix} \longrightarrow [+\text{flapped}]$

15. Reduced glottal resistance after release: Chomsky and Halle use the feature "heightened subglottal pressure" to characterize aspirates and breathy voiced stops. I did not find any evidence for this feature in Hindi in a test I conducted (for details see Appendix 1). It was found however that there was a drop in pressure immediately after the release of the closure in both the aspirates and breathy voiced stops. Therefore I have used the feature "Reduced glottal resistance after release" (henceforth RGRR). Aspirated and breathy voiced stops have been distinguished from each other by the feature 'voice'.

16. Distinctive release: I also found that affricates, aspirated stops, and breathy voiced stops belong to a natural class which is needed in certain rules (e.g., SQC 3, 28 of chapter 2.). Phonetically what they have in common is the acoustic feature of "distinctive release"; all three require the release of the closure for their identification (unlike other obstruents, where the obstruent can be identified even without the release). The need for this feature reveals the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

inadequacy of the Chomsky and Halle feature system which is based only on articulatory correlates. It seems to me that using only articulatory definitions or only acoustic definitions is simplistic.[9] Sounds may pattern similarly or differently in a given language due to physiological, perceptual-acoustic, or historical reasons. Since Chomsky and Halle's features are based primarily on articulatory correlates they are thus not completely adequate from either a phonological or phonetic point of view.

1.3. Some tentative segment structure conditions: The constraints on permissible segments and permissible sequences in Hindi have been stated in the if-then condition format proposed by Stanley (1967). I will discuss the Stanley model in more detail in chapter 2 where I discuss the Hindi sequential constraints. Before stating the Hindi segmental constraints here, let me briefly mention their function. These conditions state the constraints on what features can cooccur in a segment, and also predict those features that are redundant. Furthermore the segment structure conditions "provide a definition of the set of systematic phonemes"(403).

1. if: [+sonorant]
   ⇓
   then: [+voice]

This condition indicates that there are no voiceless vowels or liquids or nasals in Hindi. As for the glides, Dixit mentions that [h] is voiceless in some environments; however the phonetic (allophonic) rules would take care of this.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. if: [-son. +contin. +ant. -cor.]
⇓
then: [-voice]

This states that Hindi has /f/ but no /v/. Since it is possible to include /ž/, a similar condition has not been stated for /š/.

3. if: {[+son.], [-son. +contin.]}
⇓
then: [-RGRR -distinctive release -delayed release]

The above condition simply states that only stops (including affricates) are aspirated or breathy voiced. Also the features distinctive release, and delayed release are only relevant to stops.

4. if: [+cons.]
⇓
then: [-syll.]

The above condition states that there are no syllabic nasals or liquids in Hindi.

5. if: [-son +back]
⇓
then: [-contin.]

There are no back fricatives in Hindi. This condition would not hold for Urdu speakers who have /x/ and /ɣ/.

6. if: [-syll. -cons.]
⇓
then: [+contin. -nasal -retroflex]

The above condition says that all glides are continuant, non-nasal, and non-retroflex.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7. if: [+cons.; +back]
⇓
then: [+high; -cor.]

The above condition says that the back consonants (i.e. velar stops and /w/) are high and noncoronal. This would rule out q as a segment of Hindi. For Urdu speakers who have q this condition would not hold.

8. if: [-son.; +retro.]
⇓
then: [-ant.; +cor.]

The above condition states that the features anterior and coronal are redundant for retroflex stops.

The features 'anterior' and 'coronal' are not necessary to distinguish glides from each other, but since I need to refer to these features in the case of some SQCs (namely SQC 39 of chapter 2) the following segment structure condition can be stated:

9. if: [-syll.; -cons.; +high; +back]
⇓
then: [+ant.; -cor.]

I.e. /w/ is [+anterior] and [-coronal].

10. if: [+syllabic]
⇓
then: [-retroflex]

I.e. there are no retroflex vowels in Hindi.

11. if: [-sonorant; +contin.]
⇓
then: [-retroflex]

There are no retroflex fricatives in Hindi.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It is also the case that long consonants do not occur initially and finally and /r̥/ does not occur initially, but these are sequential constraints and I will treat them in chapter 2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Footnotes

1 For examples of all the contrasting segments in initial medial and final position see Dixit (1963) or Arun (1961).

2 For a detailed phonetic description of these, and the full range of various positional allophones see Dixit.

3 Most writers have symbolized these two retroflexes by the same symbol, i.e. they have used n for both the flapped and the non-flapped retroflex nasal. It is however usually clear by the context which one they are referring to.

4 This is similar to Srivastava 1969. I will include under 'PA', Persian, Arabic and Turkish loans, and under 'English', English, Portuguese and other European language loans. Even though historically certain loans are English, or Portuguese or Arabic or Turkish, a native speaker of Hindi as far as admissability of certain segments or clusters are concerned, simply treats all European loans alike (as English), and all Persian, Arabic, and Turkish loans alike as 'Urdu' or 'high Urdu', which I denote by 'PA'.

5 I say 'certain morphemes' because in the cases where a loanword is used for something for which no native word exists, the use of the loan morpheme does not seem socially significant, e.g., the use of the English loan "station" in Hindi does not reveal anything about the social status of the speaker.

6 This is not completely accurate. Some words which are historically from Sanskrit, e.g., the words [pʰəl] "fruit" and [ǰi:ra] "cumin", are pronounced as [fəl] and [zi:ra] by a number of speakers. Originally this was due to 'hypercorrection', but now the use of f in the former is recognized as non-standard; while in the case of the latter it is the pronunciation [zi:ra] that has become the standard.

7 Dixit claims that Hindi does not have nasalized e and o i.e., [ẽ] [õ]. I disagree with this. It is generally true that vowels are lowered when nasalized, and so in Hindi also when [e] is nasalized it gets lowered, however I do not think it gets lowered to [ɛ̃]. I have audibly different vowels in [mẽ] "in" and [mɛ̃] "I".

8 The non-English examples that writers such as Dixit list are Sanskrit loans such as [u̯ɖu] "star". The native example g̃ɔ̃ɖa that he gives is phonetically [gɔɳɖa].

9 A similar point has been made by Ladefoged 1971b.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 2

2.0. This chapter is concerned with the sequential constraints of Hindi. The segment structure constraints were discussed in the preceding chapter and will not be considered here. The first part of section 2.1. considers theoretical problems concerning the function of morpheme structure conditions (MSCs) and the latter part their domain. Section 2.2 discusses the level of abstractness appropriate for lexical representation of morphemes and section 2.3 gives the Hindi sequential constraints in detail.

2.1.1. Morpheme structure conditions: their origin and purpose.

Not all possible sequences or types of sounds occur in the morphemes of any single language. These restrictions are called "phonotactic" or "distributional" constraints by structuralists, and morpheme structure (MS) rules (Halle 1959) or MS conditions (Stanley 1967) by the generative phonologists. There is more than just a difference in terminology between the structuralist and the generative school. To the generativists these MS conditions pertain to the "underlying" or "lexical" level and thus to morphemes only, and not to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the phonetic level or to words as such. Also the generativists' treatment would differ from that of the structuralists' with respect to which patterns of segment sequences they would include in the MS conditions. The mere existence of a pattern does not neccessarily mean that an MS condition would be given for it; only very general patterns would be mentioned. E.g., words like "sphinx" [sfiŋks] although listed in most English dictionaries, are treated as exceptions by some generativists, that is, their initial consonant clusters are not considered to be "normal" English clusters (e.g., Schane 1970:100). Moreover these MS conditions are part of a competence grammar, and reflect what a speaker knows about the sound system of his language.

Besides the above tasks, MS conditions perform one other function: they are said to characterize "possible" and "impossible" morphemes for the given language. Thus Stanley says:

> ....we regard a 'possible' morpheme as a form which may or may not occur in the lexicon, but whose phonological structure violates no sequence structure rule (and, of course, no segment structure rule) of the language. (401)

And McCawley (1968):

> ..the least complex grammar of the language will necessarily be one in which /blik/ is treated as possible but */ftik/ as impossible although, in fact, neither item appears in the lexicon. (50)

Similar views are expressed by Chomsky (1964) and Halle (1959, 1964).[1]

I accept some of Stanley's modifications of Halle's version of MS theory----for example, his proposal that both segment structure rules and sequence structure rules be included in the morpheme structure rules; that these redundancy rules be stated as if-then conditions, negative conditions, etc; that they are unordered; that they do not delete or insert segments. However I will point out below certain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

problems with the current theory of MS conditions, which indicate the need for even further modification.

Are exceptions to MSC's "impossible" morphemes?

Generative phonologists recognize that MSC's can have exceptions and Stanley proposes simply that such exceptions be so marked in the lexicon (431). However according to this, exceptions are treated as "impossible" morphemes (i.e. "not among the possible" cf. Stanley 428), which is strange, since as they do exist in the lexicon they must certainly be "possible" in some sense. (By "impossible", I mean a morpheme which is "impossible" for the language in question, not humanly impossible.)

It seems likely that speakers may have some intuition about the degree of closeness to the native system of a particular morpheme but may not be able to make the dichotomous judgement possible/impossible (this would perhaps be especially true in the case of a language like Hindi, which has a great number of borrowed morphemes from sources as unrelated as Sanskrit, Perso-Arabic, and English). The Greenberg and Jenkins experiment (to be reviewed below) seems to support this claim. Therefore I do not think the notion "impossible morpheme" has a place in a grammar of competence. A more realistic view of the functions or capabilities of sequential constraints should be taken. There are certain patterns to be found in languages, and these patterns may have some exceptions.[2] I would call such patterns "tendencies" of the language. No one would deny that such tendencies should be reflected in a grammar, that is, that the speaker may be (tacitly) aware of them. Accordingly the MSC's of the language determine which morphemes are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

regular/irregular, lawful/unlawful (where the possible existence of irregular or unlawful morphemes is not ruled out) but not "impossible" (or "inadmissable") morphemes. The problem lies in the absolutist meaning implied in the word "impossible".

Chomsky and Halle (1968:416) seem to have modified their earlier stance, when they very briefly mention in the epilogue of their work that it seems that lexical items are not simply divided into occurring/accidental gaps/inadmissable but that there may be degrees of admissability. Chomsky and Halle do not provide any evidence as to why they think this modification is necessary. Since linguists (e.g., Schane 1970) still talk about possible/impossible morphemes I think a detailed discussion may be useful.

Psychological reality of MSC's:

Perhaps we can first ask, what evidence is there that speakers "know" (tacitly or otherwise) the sequential constraints of their language? One cannot go simply on the basis of existing morphemes because these would be known by the speaker simply as individual lexical items, the phonological patterns exemplified by them possibly escaping his (tacit) awareness completely. One must examine the speaker's behavior when he is given the chance to operate on new morphemes.

I think there is quite a bit of evidence which shows that speakers are indeed aware of the sequential constraints of their language. Let me briefly review this evidence.

Informal tests conducted by Fromkin and T. Smith (personal communications) showed mixed results. Fromkin spoke certain words to her introductory linguistics class and reported that the students had a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

pretty good idea of which were "possible English words" and which were not. Smith's informal test however, revealed that native speakers differed greatly in whether they would accept some sequences or not, e.g., some subjects accepted "sclerosis" and others didn't.[3] The mixed acceptance that the borrowed cluster skl- received, suggests that the MSC's characteristic of subsets of vocabulary perhaps differ from speaker to speaker.

Anecdotal evidence of the psychological reality of MSC's comes from the fact that speakers often restructure loans to conform to native systems, thus Weinrich, Labov, and Hertzog (1968) report that non-standard English speakers restructure the Yiddish loan [štik] to [stik]. However, not all loan restructuring is motivated by the SQC's because in some cases the loan is restructured in a way that still violates the sequential restrictions of that language. Thus, originally Hindi had no clusters involving retroflex stops initially, though retroflex stops did occur in CV- situations (i.e. in non-cluster situations), and they were the common substitute for alveolar stops in loans from English.[4] and although, in general, initial clusters in Hindi are severely restricted (for details see 2.3), English loanwords "drum", "train", and "station" were borrowed with dr- tr- st- etc. Thus restructuring of the loan yielded a form that violated the original MSC's of Hindi. (I should mention that stop+r, or s+stop clusters, where the stop was not a retroflex, did occur in Sanskrit loans, so perhaps dr- etc. violated the original MSC's in a way that made these MSC's more general.)

Esper's (1925) test, though constructed to reveal the force of analogy in sound change, also reveals rather nicely that speakers are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

aware of the MS conditions of their language.[5] He presented to three groups of subjects 16 figures each having one of four different shapes and one of four different colors and saw how long it would take them to learn the names of the figures. The names of the figures varied for each group:

1. For the first group the names were all bisyllabic and bimorphemic of the sort naslin, nasčaw, nasdeg, naskop etc., where the first syllable nas referred to a particular color and the second syllable referred to a particular shape. The morphemes in this case adhere to English MSC's.

2. The second group learned names that were similar to those given in the first group, in that they were bi-syllabic and bimorphemic, of the sort nulgen, pelgen, nugdet, pegdet etc. except that in this case it was the sequence pe and nu which referred to two of the four colors and the sequence lgen and gdet which were associated with two of the four shapes. Here the second morpheme of the names violated English MSCs.

3. The third group learned names that were completely idiosyncratic: some were monosyllabic and some were bi-syllabic but there was no systematic morphemic relationships between one or the other syllable or sequence of sounds and either the shape or the color of the figure.

Of course Esper did not reveal to any of the subjects the morphemic structure, if any, contained in the names. They had to discover this for themselves. The first group took 28 trials to learn all 16 words right 60% of the time, and the third group 60 trials to reach the same level. This is understandable since the first group's task was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

simplified in that they had to learn essentially only 8 morphemes whereas the third group had to learn all 16 words separately. Interestingly the second group took as long as the third group--approximately 60 trials--indicating that they were not able to take advantage of (that is, recognize) the systematic morphemic make-up of the names since they violated English MSC's, and instead had to learn all 16 names separately, the same as the third group did.

Another piece of evidence that native speakers have some tacit knowledge of the MSCs of their language comes from Fromkin's (1971a) "slip-of-the-tongue" data. In none of the speech errors she gathered were the posited MSC's of English violated.

Among the few rigorous tests performed to specifically investigate the psychological reality of MSCs are those of Greenberg and Jenkins (1964) and Zimmer (1969).

Greenberg and Jenkins gave subjects a list of 24 words, some of which were English words, (e.g., græs), some very close to English (e.g., slʌk) and some quite unlike English (e.g., zbüy). The speakers were asked to indicate the subjective distance of these words from English. The results seem to show that even though the words such as slʌk, the so called "accidental gaps" in the lexicon were not rated as far from English as words such as zbüy, they were still not judged as close to English as words such as græs. The experiment also showed that the words were considered equally far from English whether they violated a sequential constraint or segmental constraint. The test seems to show that speakers have some awareness of the constraints in their language.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Zimmer (1969) gave some native speakers of Turkish pairs of made-up words and asked them to indicate their preference. The words were constructed such that some of them adhered to the major and minor regularities of Turkish morpheme structure conditions and some of them were clear violations. Although speakers gave a clear preference to those forms adhering to the major regularities there was no clear preference to those exemplifying minor regularities.

To see if native speakers of Hindi are aware of the sequential constraints of their language, and also to see if speakers to consider certain morphemes to be "possible" for their language and others "impossible", I conducted a test in which I asked 25 Hindi speakers to indicate whether 30 written words presented to them were (a) familiar existing words, (b) unfamiliar but permissible Hindi words, or (c) unfamiliar and not permissible Hindi words. The words were not presented auditorily since there is always the problem that subjects might misperceive the sounds and thus base their response on a form different from that intended (this proved to be a problem with a few of the words in Greenberg and Jenkin's experiment, cf. their pg. 189). But since the Devanagari orthography can be used phonetically, I chose to present the words in writing rather than auditorily. The subjects were also led to believe (by the instructions) that there might be some Hindi words which they do not know (this was reinforced by including certain rare loans). This was so that they would not just make the distinction of "exists in Hindi" vs. "doesn't exist in Hindi". Twenty-four of the 30 test words were constructed, half of these so that they would violate the posited MS conditions of Hindi (given in 2.3). For further details

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the test see Appendix 2. On the whole the test showed that the speakers overwhelmingly reject what I hypothesized to be impermissible sequences for Hindi. However the reactions to presumed "accidental gaps" in the vocabulary, i.e. supposedly permissible but nonexistent forms, was rather mixed. In a number of cases such words were treated more like the impermissible category. And for reasons that I cannot explain, there was a somewhat higher rejection rate for 'accidental gaps' containing native clusters, e.g., stop+glide, than for 'accidental gaps' containing clusters that occur only in loans, e.g., s+stop. For example, gwek received 17 no's, although gw- exists in native words such as [gwala] "cowherd" but step received only 11 'no's, even though st- occurs only in Sanskrit loans. It is possible that speakers may be more confident of their knowledge in native-like forms (Bright, personal communication).

Thus this test as well provides evidence that Hindi speakers are indeed aware of the sequential constraints of their language. However none of the evidence reviewed gives any clear support to the claim that speakers are able to make the dichotomous judgement possible/impossible morpheme. Thus for English speakers the Hindi word mlan "weak" would violate the sequential constraints and would of course not be in their lexicon either, but do speakers of English feel that it is 'impossible' for their language? The fact that Bret Harte (1937) titled one of his short stories "Mliss" seems to indicate that speakers may not judge a word with the cluster ml- as impossible.

The results of the Hindi test discussed above also lend support to this analysis. Although speakers did respond that some words were 'not

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

possible' words for Hindi, the pattern of their responses reveals that they really were not able to make this judgement in any reasonable way and perhaps it is unreasonable to ask them to, i.e. the task Greenberg and Jenkins gave their subjects seemed more reasonable. Their rejection of /gwek/ when /gwala/ "cowherd" and /gwar/ "a kind of vegetable" exist; likewise the rejection of /ḍyom/ when /ḍyorʰa/ "2-1/2 times" and /ḍyorʰi:/ "porch" exist, suggests that the subjects classed as "not possible" words that certainly are possible. The reason for this may be explained in part by reference to figure 1 given below. I hypothesize there that two 'tests' are administered by the speaker to decide on the acceptability of new forms: the usage test and the sequential constraint test. Failing the usage test led to rejection 55% of the time; but failing both tests led to rejection 77% of the time. Insofar as usage plays a partial role in the judgement, it is clear that "rejection" does not and cannot imply, strictly speaking, that the form is "impossible". Thus there is no need to differentiate between unlawful forms and exceptions, because exceptions <u>are</u> unlawful. Exceptions differ from unlawful, not-yet existent forms only in that usage has given them a place in the lexicon (and one might expect that native speakers will often try to regularize these). One might expect that a native speaker confronted with a morpheme that violates some MS condition might reject it as irregular; however, when he is exposed to it again and again in everyday situation, he may very well come to accept it--as an exception. E.g., some speakers of English might have rejected "shmoo" as irregular at first, and later, after seeing it in the comic strip many times, accepted it.[6]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

On what does a speaker base his reaction to new morphemes?

What determines a speaker's reaction to the phonological shape of new morphemes? It is useful to consider a number of hypotheses:

(a) The existence or non-existence of that morpheme in his stored vocabulary.

(b) The use of pronounceability as a criterion.

(c) The use of some kind of internalized abstract rules.

Hypothesis (a) claims that usage solely determines a speaker's reaction to morphemes. This undoubtedly is a factor in some cases, however it cannot account for many instances of speakers' acceptance of new morphemes, because speakers do learn and do accept new words. Thus we have to reject this as the sole criterion which a speaker uses.

The second hypothesis equates "possible" morpheme with "pronounceable" morpheme and is in fact the position Shibatani (1971) takes.[7] Equating "possible" morpheme with "pronounceable" morpheme seems to reduce it to a question of skill and neuromuscular coordination. At first this hypothesis seems quite reasonable. In my experiment (Appendix 2) informants pronounced (or tried to pronounce) the words given to them before passing judgement on them, even though they had not been asked to. This would seem to support Shibatani's position. However again, even though pronounceability seems to play some part in a speaker's reaction to morphemes, it cannot be a very important factor in his acceptance of them because of the following kinds of evidence. My experiment (appendix 2) shows that "possible" morpheme cannot be equated with "pronounceable" morpheme. Many of the subjects in my test rejected bmi:n, etc., (i.e. words in category III of table 2 of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

test) even though they did pronounce them correctly.

Another example of this is the common case of many American English-speaking children who cannot pronounce the retroflex r and substitute w for it thus saying [θwi:] for "three" or [twək] for "truck", but still recognize the r. Thus if an adult says [θwi:] they insist that he is wrong. Here the child recognizes the r even though he cannot pronounce it.[8]

Cleft-palate speakers also give evidence for this; they cannot produce certain sounds, but they certainly recognize them. "Pronounceability" may have some effect on a speaker's judgement of a given word, but equating "possible" morpheme with "pronounceable" morpheme is too simplistic.

Thus some speakers' behavior in reacting to new morphemes, namely the restructuring of loan words, etc., can be accounted for by assuming they have limitations on their ability to pronounce certain sounds or sequences of sounds. However there still remains much that cannot be explained by this.

So we arrive at the third hypothesis, that speakers have some kind of abstract rules in their head regarding the sound sequence constraints of their language--'abstract'in the sense that they are not necessarily directly equivalent to the neuromuscular constraints that determine what a person can and cannot pronounce, and in the sense that they are applicable to more morphemes than those which already exist in the speaker's active and passive vocabulary. It seems inescapable that besides the previous two factors, namely usage and pronounceability, this factor plays a role in speakers' reaction to new morphemes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 1. shows in flow chart form how these three factors might operate.

This figure requires some comment. First of all from box 3a we go to 4 and not back to 2, because the speaker bases his judgement on the original input, not on the restructured morpheme. The subjects of my test (except one) pronounced the words given to them before rendering their judgement, thus it seems "pronouncing" had some part to play,[9] and it seems this "test" comes first. Next it would seem easier and faster to find out if a word existed or not than to find out if it adheres to the sequential constraints of the language. Thus I hypothesize that the usage test precedes the sequential constraint test. Moreover if the form exists, even though it violates a sequential constraint—i.e. if it is an exception to a sequential constraint, it is accepted, which again indicates that the usage "test" probably comes earlier than the sequential constraint "test". The chart also shows that the pronounceability of a form is independent of its acceptability on other grounds. It should also be emphasized that rejection at box 6 in figure 1 does not imply that the morpheme is impossible, because the morpheme could enter the speakers passive or active vocabulary if he was exposed to it often enough.

Of course a lot of structure is hidden in some of these boxes. For example in box 3a instructions are given to restructure an unpronounceable form so that it is pronounceable, but the nature of these instructions is glossed over. No doubt there are definite rules for accomplishing these, and conceivably some reference is made to the sequential constraints. It is not unlikely that pronounceability

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

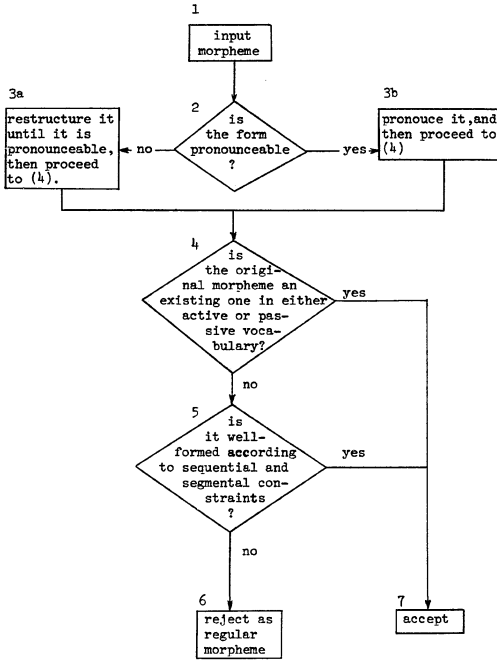


Figure 1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

constraints overlap the sequential constraints to a large extent--since a speaker gets more practice and thus more skill in pronouncing the forms existing in his language, most of which by definition adhere to the sequential constraints. However in many cases there are also alternate ways of doing this restructuring, e.g., the English loan 'station' is restructured by speakers of Hindi dialects with severely restricted initial clusters, as either [isṭeš̆ən] or [səṭeš̆ən] or [ṭeš̆ən]. Or the Vietnamese name "Nga" [ŋa] is restructured by some American English speakers as [na] and by others maybe as [əŋga] or [əŋa].

I have in this section briefly reviewed the evidence that can be cited for the claim that native speakers are aware of the constraints on morphemes of their language, and have speculated on the factors that may underlie a speaker's behavior when he accepts or rejects new morphemes.

2.1.2. The question of the domain of MSC's:

In the previous section I discussed the purpose of MSC's, and certain problems in achieving this. But there are still some controversies in the field as to the domain over which MSC's apply.

Halle (1959), Stanley, and McCawley (1970) notice that certain restrictions hold at the lexical level of morphemes which may not hold at the systematic phonetic level, and, according to McCawley, vice versa:[10]

> ..restrictions on what may occur in pronunciation do not always parallel those on what may occur in basic forms.. (817)

McCawley also notes that in many cases the restrictions on both the dictionary level and the phonological level will be the same:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> The exclusion of [zb] (indeed, of any initial consonant cluster beginning with a voiced spirant) in English can be regarded as either a constraint on possible pronunciations or a constraint on possible basic forms.[11] (817)

Even though the above-mentioned linguists have noted that certain constraints apply to only one level or the other, they do not make a formal distinction between the two kinds of constraints. All the restrictions--redundancy rules--are given at one level, the lexical level, and apply to morphemes only. In the case of initial clusters in English the restrictions are about the same at both the lexical level and the phonetic level so this leads to no problems. Problems develop only when the restrictions appear to be different at the two levels. Some of the restrictions are different only in the domain to which they apply, to the morpheme at the lexical level, but to the word, i.e. across morpheme boundaries, at the phonetic level. These differences are handled by P-rules.

This problem is explicitly recognized and discussed by Brown (1969), Fromkin (1971b), and Shibatani (1971).

Let me briefly review the kind of data that is involved in this controversy:

If we look at languages we find that certain constraints apply only within a morpheme, e.g., within a morpheme in Hindi we cannot get *td as a cluster, but across morpheme boundaries we can, as in pot+dar "someone dealing in beads".[12] Certain other restrictions apply both within and across a morpheme boundary, e.g., Kisseberth's Yawelmani example (to be discussed shortly).

Since certain constraints in languages apply both within a morpheme boundary and across it, the generative phonologist's restriction

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of having MS constraints apply only within a morpheme, has made necessary the writing of certain P-rules which duplicate the constraints already mentioned in the MS conditions. For example Kisseberth (1970) points out that in "...the underlying representation of Tawelmani morphemes, there are no triliteral clusters" (294), and posits a morpheme structure condition to account for this. He goes on to say:

> The existence of this morpheme structure condition is not sufficient to account for the absence of triliteral clusters, for it blocks such clusters only inside the morpheme. Morphological processes of suffixation and special stem formation rules operate so that morphophonemic representations of words do contain triliteral clusters. (294)

And he then mentions that several phonological rules take care of these clusters at the morphophonemic level, so that at the phonetic level these clusters do not occur.[13]

Brown criticizes this approach as being an unnecessary proliferation of rules and questions the restriction of the application of the MS conditions only to within a morpheme. Fromkin seems to agree with the general idea of having MS conditions apply across morpheme boundaries but points out certain difficulties with Brown's proposals due to certain mechanical problems which do not concern us here.

Shibatani approaches the problem by dividing what has usually been called MS conditions into MS constraints (MSC) and surface phonetic constraints (SPC). According to him SPC's "are constraints on possible pronunciation" (18). They apply at the phonetic level, and are not limited to applying only within a morpheme. As a matter of fact, as he mentions in a footnote, they cannot be limited to within a morpheme, since the information on the location of morpheme boundaries is considered to be unavailable at the phonetic level. (I might add that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

this is by current convention, and doesn't necessarily have to be so). Shibatani indicates that information on word boundaries is available at the phonetic level, however, according to McCawley (1968:53) it has been maintained since the Chomsky, Halle, and Lukoff (1956) study that even word boundaries are erased at the phonetic level. Also the SPCs "state inter-morphemic constraints"(7). The MSCs on the other hand state constraints at the "morphophonemic level, and they apply only intra-morphemically" (7).

Thus Shibatani accepts the generative convention of restricting MSC's to applying within morphemes only, and tries to find a solution to this proliferation of rules by claiming that constraints at the systematic phonetic level be formally recognized. Brown and Fromkin on the other hand support an extension of the theory in order to solve this problem: they allow sequential constraints to apply across morpheme boundaries as well as within.

Above I reviewed how Brown, Fromkin, and Shibatani have tried to find a solution to the problem of proliferation of rules. Even though Shibatani's positing of sequential constraints at two different levels solves this problem to a certain extent, it does not solve it completely, and moreover it creates certain other problems, to be discussed below.

1. First it is contradictory to equate SPC's with constraints on pronunceability and then allow exceptions to them as Shibatani does (cf. his footnote 20). Shibatani claims that it is SPC's which a speaker uses to determine whether a given item is an "accidental gap" in his vocabulary or "impossible" morpheme, and elsewhere also claims

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that SPC's "are constraints on possible pronunciation" (18). If speakers judgement of possible/impossible morpheme is based on pronunceability then I don't see how exceptions to SPC's can be permitted--because presumably these constraints amount to limitations on the neuromuscular ability of the speaker. Moreover as discussed earlier the test I conducted shows that it is not solely on the basis of pronunciation that a speaker accepts or rejects a given word.

2. Besides the above-mentioned difficulty there still remains the problem of duplicating rules at various levels.

Shibatani has three kinds of conditions: MSC's, SPC's and ^{m}SPC's. The ^{m}SPC's "..state the constraints that hold at both morphophonemic and phonetic levels" (11). Since his ^{m}SPC's are part of SPC's and since SPC's do not have grammatical information or information on morpheme boundaries available to them, the problem of duplicating information still remains. This is evident in the well-known case of the devoicing of word-final obstruents in German:

Word pairs such as bund [bunt] ~ bundes [bundəs] "union" (sg. and pl.), rad [rat] ~ rades [radəs] "wheel" (sg. and pl.) suggest that word-finally, underlying voiced stops become voiceless. First, Shibatani is of course correct in holding that this devoicing cannot be accounted for by any MSC. This is because in order to derive the correct plural forms it is necessary to posit voiced stops in underlying morpheme final position for these words. The traditional way of generating the surface forms then would be to let a P-rule do the devoicing. However Shibatani suggests this devoicing reflects a constraint on pronunciation and attributes it to one of his SPC's,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

which are supposed to operate on the output of P-rules. This is certainly a possible treatment of the devoicing phenomenon but I see no compelling reason to choose the SPC treatment over the P-rule treatment; on the surface both are equally plausible.

However the devoicing of stops in German is actually more complex than the above examples reveal. Shibatani notes in a footnote that Zimmer pointed out that the devoicing also occurs word-medially at the end of a syllable that is also morpheme final. Thus alongside bund [bunt] we have bündnis [bʏntnis] "union", and alongside rad [rat] there is radfahrer [ratfarər] "to cycle". The word-medial retention of voicing in bundes [bundəs] etc. are presumably explained by claiming that in these cases the voiced stop occurs in syllable initial position and thus escapes the devoicing.[14] For this expanded set of examples of the stop devoicing it is possible to posit a single P-rule something like:

(1) [-continuent] --> [-voice] /____syllable boundary

Condition: syllable boundary must coincide with morpheme boundary.

This handles the word-medial case and the word-final case too, since word-final position is just one particular case of syllable-final and morpheme-final. It is necessary to refer to the morpheme boundary in the above rule, because voiced obstruents in syllable-final position which are not also morpheme-final do not undergo devoicing, e.g., stieglitz [stiglits] "gold finch", adler [adlər] "eagle".

But Shibatani's SPC's cannot refer to morpheme boundaries or any other grammatical markers, thus the word-final devoicing could be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

handled by an SPC but the word medial devoicing would have to be handled by a P-rule. (There are perhaps other cases too where such grammatical information might be needed. As Shibatani notes himself (footnote 10) there are cases in natural languages where his SPC's might require grammatical information.) This is definitely uneconomical. Essentially the same rule would have to be given twice. In this case there is a clear preference to handling all the devoicing phenomena by a single P-rule rather than an SPC plus a P-rule.

There might be some advantage in treating certain phenomena as due to an SPC, where otherwise it would appear that several P-rules are involved in a "conspiracy" (e.g., Kisseberth's Yawelmani example)[15] but there seems to be no advantage in writing an SPC for what can be handled by a single P-rule (however, I will shortly propose that SPC's are not even needed in the "conspiracy" case).

To consider another example: in the case of the English plural s~z alternation, generativists up to now, have handled the voicing agreement by means of a P-rule but would (presumably) handle the highly similar voicing agreement within morphemes, as in the final cluster in apse, adze, etc., by a redundancy rule—i.e. a MS condition. Brown suggests however that the redundancy rules be applied across morpheme boundaries so that the redundancy rule that handles voicing agreement of obstruents within morphemes, also take care of the correct voicing of the plural morpheme in 'cats', 'dogs', etc. She correctly (in my view) points out the arbitrariness of choosing either s or z as the specified form of the plural morpheme when it enters the phonological component, and later changing the voicing by a P-rule. In her

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

system the plural would not be specified for voicing, and the redundancy rule would supply the correct voicing. This still would not take care of the -əz plural in words like judges as Fromkin (1971b) points out. So Brown would have to take care of the -əz plural by a P-rule (since, by convention, redundancy rules can only fill in unspecified features, not insert or change segments). Similarly the plural morpheme shape z after sonorants would also have to be taken care of by a P-rule (since there can't be a redundancy rule for this since both voiced or voiceless consonants can appear after sonorants, e.g., 'lens' [lɛnz], 'dance' [dæns]). If the redundancy rule specifying voicing agreement is allowed to operate across morpheme boundaries, Brown would of course have to list words such as 'width' [widθ], and 'hundredth' [hʌndredθ] as exceptions to this redundancy rule (pointed out to me by J. Ohala).

Shibatani proposes choosing z as the representative shape of the plural, and having the s~z alternation specified by a P-rule and the insertion of the necessary ə in the case of the -əz plural after sibilants, by another P-rule (9):

(2) Ø---> ə / [+strident] _____ [+strident]

(3) z---> s / [-voiced]_____

Thus he too needs two P-rules, however his P-rules are somewhat simpler; since after vowels and after voiced consonants the z would remain unchanged, and only after voiceless consonants would the z be changed to s. However he loses the advantage of a "nonarbitrary" specification of the voicing of the underlying plural morpheme.

Shibatani, however does not stop at just the positing of P-rules

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

since according to him they don't explain enough. He posits two SPCs which more or less repeat the same thing as the P-rules--however according to him the P-rules exist only because of the SPCs (9):

(4) [-sonorant] [-sonorant] #
↓ ↓
[avoiced] [avoiced]

(5) * [+strident] [+strident]

By the way, as stated, his SPC's don't work. SPC (4) states that two [-sonorant] segments must agree in voicing, however as mentioned above, 'width' [widθ] and other words constitute exceptions. SPC (5) states that no two strident consonants can occur together, however words like "sphere", "svelte", "laughs", "wives", etc., are exceptions to this, since sf-, sv-, -fs, -vz, are all [+strident] [+strident] clusters. The exceptions to SPC (5) needn't be exceptions of course, if one modifies (5) to read not simply [+strident] [+strident] but $\begin{bmatrix}+\text{strident}\\+\text{coronal}\end{bmatrix}\ \begin{bmatrix}+\text{strident}\\+\text{coronal}\end{bmatrix}$ (The exceptions to SPC (4) would, as mentioned above, be exceptions for Brown also, but exceptions to SPC (5) need not be exceptions for her if she spelled out the plural shape -əz by a P-rule, rather than just inserting a ə as Shibatani does). However as mentioned earlier exceptions to SPC's are inconsistent if SPC's reflect "possible pronunciation". Thus exceptions to Brown's redundancy rules are not as damaging as exceptions to Shibatani's.[16]

Moreover it is not at all clear how Shibatani accounts for the voicing agreement in apse, adze, etc. He mentions that SPC's can act as linking rules (18): "..when the output of P-rules are found phonetically conflicting, SPCs can link them to well-formed phonetic representations". But he notes that not all types of SPCs can function

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

as linking rules. Since he specifically rules out (4) as operating as a linking rule (18) (since it contains variables and so would not provide a unique output), it presumably could not be used to account for the voicing agreement in apse, adze, etc. It is possible that he would handle these words also by *P*-rules (2) and (3). However would he then claim that apse also has an underlying z /apz/ which the *P*-rule changes to s? This is unlikely. He might account for these words by an MSC, as is usually done. But he has claimed that SPC's (4) and (5) "explain" the existence of *P*-rules (2) and (3); do they also explain the existence of the MSC? If so, this would contradict his previous claim of the independence of SPC's and MSC's.

Perhaps at this point one should ask whether the voicing agreement of the final apse and adze clusters, and the final cats and dogs clusters really are related phenomena or not, i.e. does treating them as the same reflect competence? It is not at all clear that it does because:

(a) Within morphemes there are no exceptions (in the final clusters), but across morpheme boundaries exceptions such as 'width' [widθ] do exist.

(b) We cannot have the MSC's (or SPC's) specify the shape of the plural after sonorants, because as mentioned earlier both voiced and voiceless consonants can occur after sonorants. So at least this part of the plural rule will have to be treated seperately from the MSC's, and no SPC can be posited to 'explain' it. Of course it can be claimed--as Shibatani does--that the plural shape is /z/ and it is changed to [s] or has a [ə] inserted before it, only in the appropriate

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

environments. However Berko's (1958) data suggests that this is not necessarily true. Young children had significantly greater difficulty with the plural after sonorants that after stops, (presumably since both s and z are phonetically possible). If the plural was /z/ then one would have expected that children would have had no more problem with sonorant+z than voiced stop+z in the plural.

(c) Bellugi (1964; cited by McNeill (1970)) found that the s-z alternation is learned in pluralization first, then in the possessive and last in the present tense marker. This suggests that perhaps these alternations are not really treated as one rule. (Of course as Fromkin (personal communication) pointed out, "although the s-z alternation may be learned in stages as related to syntactic or morphophonemic rules, it is still possible that at some stage the generalization is made relating all the different rules into one general rule.")

Thus perhaps the current generative practice of treating the voicing agreement in the plural as seperate from that exhibited in apse, adze, etc., is correct. If, however, further evidence showed that the phenomena were similar then it would still be preferable to follow the Brown-Fromkin approach rather than Shibatani's because although some of Shibatani's SPC's play an active role in altering the output of P-rules, in this case his SPC's repeat what P-rules or MSC's already do. Thus it seems to me the SPC's just duplicate information.

3. Finally some of the evidence Shibatani gives for the need for SPC's requires the prior acceptance of some very questionable points in phonology. He notes for example that Chomsky and Halle (1968:47) posit consonant clusters having identical consonants in the underlying forms

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of Mississippi, Kentucky, etc., namely /mississippi/, /kVntukki/, etc. This is required, according to them, to make the stress assignment correct, and to account for the lack of voicing of /s/ intervocalically, and also to account for the lax vowel in the penultimate syllable of Kentucky. However since identical double consonants are not permitted in English at the phonetic level, Shibatani posits an SPC which rules out clusters of identical consonants. However it is just such highly "imaginary" (to use Crother's term) underlying forms, which have come under some of the most severe attacks in generative phonology (Kiparsky 1968, Crothers 1971, Ladefoged 1971a). The controversy is not settled yet, of course; however the fact of the dispute hardly lends much support to arguments for the need for SPC's.

Thus even though Shibatani's proposal to correct certain deficiencies in the handling of sequential constraints in generative phonology is a very interesting one we find that the theory of SPC's seem to be beset by problems itself. What then is the solution? I propose that an extension of the current theory of MS conditions is needed (along the lines suggested by the brief treatment of this problem by Brown and Fromkin). That is, MS conditions should be permitted to operate across morpheme boundaries too, therefore I will not call them MSC's but sequential constraints (SQC).

The unmarked case will be very general SQC's that apply within and across morpheme boundaries and at the lexical as well as the phonetic level. Those that apply only within a morpheme or only at the systematic phonemic level will be so marked. Morphemes could be recycled back through the sequential constraints and it will probably be just

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

those constraints that apply across morpheme boundaries as well that would be relevant in this recycling. Or rather than recycling the morphemes, these constraints can be "everywhere" constraints applying wherever applicable. There are examples of constraints that apply only at the systematic phonemic level (e.g., the Hungarian example McCawley gives (1970:817) or the Hindi SQC 9 which pertains to homorganic nasals+stop, given in 2.3), and there are examples of constraints that apply to both the systematic phonetic level and the lexical level, e.g., almost all the SQC's of Hindi mentioned at the end of this chapter. However I do not find any clear-cut examples of constraints that apply to just the systematic phonetic level. Most such examples can easily be accounted for by ordinary P-rules. For example, in McCawley's example of those dialects of English which do not permit a syllable-final vowel+r--thus 'saw' and 'soar' both would be [sɔ]--the underlying form for the latter would be /sɔr/ because of forms like 'soaring' [sɔriŋ]. If this is due to a pronunciation constraint, i.e. if they cannot be pronounced, then it gives no evidence for these being a surface phonetic constraint. This is because constraints on "pronunciation" I consider to be constraints at the physical phonetic level, i.e. limitations on the skill or neuromuscular coordination of the speaker, for which constraints there can be no exceptions. There are obviously constraints that hold at this level which do not hold at the lexical level (or systematic phonetic level). If McCawley's example does not reflect a pronunciation constraint, then I do not see any advantage (but there are several disadvantages) in saying that it is a constraint of the systematic phonetic level, i.e. an SPC, rather

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

than treating it as an ordinary *P*-rule.

McCawley refers to Kisseberth's Yawelmani case as evidence that pronunciation effects phonological rules. This is an interesting case of phonological rules "conspiring" to yield forms having similar phonological shape (I have discussed this example in more detail above). This is perhaps one of the strongest points in favour of Shibatani's theory. But we see that *P*-rules are involved in a "conspiracy" in just those cases where they have to duplicate information of MSC's. This "conspiracy" can be quite adequately accounted for by the extensions of the MS theory as I have suggested above. This would permit the restriction in Yawelmani on triliteral clusters that holds within a morpheme boundary, to hold as well across a morpheme boundary. And this is all that Shibatani's SPC theory would be saying by means of the ^{m}SPC. Therefore if one allows MSC's to operate across morpheme boundaries, there is no need for a seperate level of SPC's, nor is a theory of SPC's needed to "explain" the conspiracy. Kisseberth does not say if it is a pronunciation constraint as well or not. If, given new words with triliteral clusters the Yawelmani speakers can correctly pronounce them but reject them as normal Yawelmani words, then the ban on triliteral clusters would be recorded as an SQC, one that applied across morpheme boundaries as well as within. Otherwise it would at least be a pronunciation constraint, possibly also a SQC.

Shibatani also mentions that it is SPC's not MSC's which operate in restructuring loan words. More evidence is needed before one can conclude this. That loanword restructuring exists, is of course, not questioned. That this necessitates positing SPC's, i.e. constraints at

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the systematic phonetic level has not been justified. This is very likely a constraint at the physical phonetic level, that is constraints due to lack of neuromuscular coordination. Moreover I would not rule out the possibility that lexical constraints are relevant in restructuring of loans. Secondly, Shibatani's suggestion that "SPC's determine the phonetic shape of a loanword" (27) should, in my view, read "may determine" since there are a number of cases of loanwords which violate the sequential constraints of the language, and are borrowed unchanged. It is a fact that somehow šm- clusters e.g., shmoo, entered English where they had not existed before). It is also possible that loanwords may be restructured in a way that actually violates the SQCs of the language (cf. the examples cited on pg.31).

At this point one might ask whether it "costs" more to write two rules that apply at separate levels in the derivation, or to add a recycling feature? My answer to this is that firstly the whole theory of "cost" has been under attack recently (Fromkin 1971c, J. Ohala 1971b). The claim that the evaluation measure could indicate the optimal grammar was based on the assumption that a child chooses THE grammar of the language he is learning. However the test discussed in the next section (2.2) shows that it is doubtful that there is only one optimal grammar of a language (see also Gleitman and Gleitman 1970, and Weinrich, Labov, and Herzog 1968). Secondly, rather than trying to evaluate a grammar on the basis of some internal structural properties of the grammar itself, it is more important to establish how well a particular model accounts for all of the existing evidence. In developing the alternative I put forth in the preceding section I have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tried not only to account for the sequential patterns in the intuitively simplest way but also to account for behavioral data.

To briefly summarize: I propose that the notion "impossible" morpheme has no place in a competence grammar, and instead we should talk of "irregular" or "unlawful" morphemes. I also suggest that a number of factors such as internalized rules and usage play a part in a speaker's reaction to morphemes and although pronunciation constraints obviously determine a speaker's pronunciation of morphemes they play no obvious role in his acceptance or lack of it as a permissible morpheme for his language. I further suggest that pronunciation constraints are perhaps best considered at the physical phonetic level, and not at the systematic phonetic level. Regarding the domain of MSC's I suggest that current generative theory should be modified to allow some MSC's to operate across morpheme boundaries as well as within, and therefore I use the term SQC rather than MSC. Some of these SQC's will only pertain to the lexical level and will have to be so marked. Others will pertain to both the lexical level and the systematic phonetic level. Also some of these SQC's will be "everywhere" constraints and will apply wherever applicable.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.2. In this section I attempt to provide evidence for the "psychological reality" (for at least some speakers) of segments which would probably be considered 'imaginary' (to use Crother's (1971) term) in that they (usually) do not show up in any of the phonetic realizations of these morphemes.

For many of the morphemes of a language there is no doubt that the relationship between underlying form and the phonetic form of a morpheme may be quite direct; the two may differ only in small details, e.g. degree of aspiration, vowel length, etc., which may have to be specified by low-level phonetic rules which take into account the presence or absence of accent on the word, location of the word in the sentence, and so on. But for some morphemes the dictionary form may be more "abstract", that is, may differ quite a bit from the phonetic form.

In Hindi, initially and finally the constraints on clusters are quite severe in native morphemes (see 2.3). However medially at the phonetic level, even in native morphemes quite a variety of clusters occur; of course, statistically their incidence is still rather low: in the dictionary (Varma) that I consulted, medial clusters in both native and non-native morphemes were found in only 9% of the total entries. The question arises as to whether these clusters should be posited at the lexical level too? I would like to claim that in a large number of cases these clusters do not exist at the lexical level, but are brought about by the deletion of a /ə/ by a very general rule--the ə-deletion rule (to be discussed in detail in chapter 4). Although in a large number of cases there are alternating forms with [ə] obviating the necessity of positing abstract /ə/s, in a few cases there

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

are no alternating forms, and yet I will propose that the forms contain a /ə/ at the lexical level.

But the desirability of positing highly abstract underlying forms for lexical items has been debated in phonological literature. The so-called abstractness controversy in phonology started with Kiparsky's still unpublished paper "How abstract is phonology?" in which he objects to the practice in generative phonology of positing in certain circumstances highly abstract underlying forms for morphemes--that is, forms that bear little resemblance to their eventual phonetic shape. In answer to Kiparsky's article, Hyman (1970) asks "How concrete is phonology?" and provides data from Nupe that suggests some rather abstract forms are sometimes required. Kisseberth (1969) and Brame (1969) also try to provide evidence favoring "abstractness". Crothers (1971) has provided a good summary of the controversy and reviews the various kinds of evidence brought forth by the "abstractness" advocates. In a recent paper Kiparsky (n.d.) attempts to peek "into the abstract morphological structure of Sanskrit" through the Vedic metrical system and tries to show that the setting up of certain abstract segments for Sanskrit seems to be required because they have metrical function. And then he shows that on phonological grounds, too, these underlying representations are justified. Thus Kiparsky seems to somewhat soften his stance on disallowing absolute neutralization.

However it seems to me that the important question is not whether the underlying forms should be abstract or concrete, but rather, what evidence is there for the "psychological reality" of any posited underlying form whether highly abstract or not?

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.2.1. The kind of data that is involved in the issue to be discussed are the pairs of related Hindi words in table 4. Many more such pairs could be cited. The words with medial clusters at the phonetic level (column 2) have alternating forms with the cluster broken up by a [ə] (column 1). If we assume that the words in column 2 have as their underlying form, forms similar to those words in column 1 plus a suffix or inflectional ending (e.g., both [nəmək] "salt", and [nəmki:n] "salty" have /nəmək/ as an underlying form), then a perfectly general rule can be written[17] to delete the /ə/:

(6) ə ----> Ø / VC____CV

The above rule states that a /ə/ is deleted if it is preceded and followed by consonants, which consonants in turn must be preceded and followed respectively by vowels. (Thus the /ə/ in the first syllable of a morpheme or word would not be deleted.)

Examples of derivations are given below:

/čəmək+a/

yields

[čəmka] by ə-deletion rule.

But /čəmək+ta/

yields

[čəmakta] (ə-deletion rule inapplicable).

I assume thus that a native speaker derives the words in column 2 from those in column 1; i.e. the words in the two columns are considered related not only by common semantic features, but also in a systematic phonological way. An alternate hypothesis, which I reject here, is that speakers do not derive both words from a single underlying form

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 4

1.	[bədəlna]	"to change" (inf.)	[bədlana]	"to cause to change" (inf.)
2.	[čəməkna]	"to shine" (inf.)	[čəmkana]	"to shine" (trans. inf.)
3.	[čipəkna]	"to stick" (intrans. inf.)	[čipka]	"stuck"
4.	[ləṭəkna]	"to hang, suspend" (intrans. inf.)	[ləṭka]	"hung"
5.	[kʰəṭəkna]	"to click, tap" (in trans. inf.)	[kʰəṭka]	"noise made by tapping" (noun)
6.	[hičəkna]	"to hesitate" (intrans. inf.)	[hički:]	"hiccup" (noun)
7.	[əṭəkna]	"to be obstructued" (intrans. inf.)	[əṭka]	"was obstructed"
8.	[sisəkna]	"to sob" (inf.)	[siski:]	"a sob" (noun)
9.	[dʰəṛəkna]	"to palpitate" (inf.)	[dʰəṛkən]	"palpitation of the heart" (noun)
10.	[sənək]	"craze" (noun)	[sənki:]	"cracy" (adj.)
11.	[ləṛəkpən]	"boyhood" (noun)	[ləṛka]	"boy" (noun)
12.	[səṛək]	"road" (noun)	[səṛkẽ]	"roads" (noun)
13.	[upəǰ]	"produce" (noun)	[upǰau:]	"productive" (adj.)
14.	[nəmək]	"salt" (noun)	[nəmki:n]	"salty" (adj.)
15.	[kʰəpət]	"obsession" (noun)	[kʰəpti:]	"obsessed" (adj.)

but have separate entries in their lexicon whose relationship is "known" to them only because of their common semantic features.

However besides ə-less words with medial clusters such as those in column 2 of table 4 which have alternating forms with /ə/ (those in column 1), there are some native words which are pronounced with clusters at the phonetic level but which have no alternating forms with a /ə/. Examples are given in table 5.

Table 5

1.	[čuṭki:]	"snap of a finger"
2.	[gʰõsla]	"nest"
3.	[ɣʰɔ̃pri:]	"hut"
4.	[kčkra]	"crab"
5.	[sčkra]	"hundred"
6.	[čʰipkəli:]	"lizard"
7.	[čuṭkula]	"anecdote"
8.	[titli:]	"butterfly"
9.	[pəčʰta]	"lament" (imperative)

There are probably a few more words of this type however not very many more.

The question thus arises as to whether these words should be entered in the dictionary with an "imaginary" /ə/ or whether they should be entered with the clusters? The former solution would be the less "costly" alternative according to the present marking conventions in that the words would more closely approximate the "preferred" CVCV... morpheme structure. But the criterion for choosing one representation or another should not be the "cost" according to the marking conven-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tions, but rather the amount of evidence we can provide for one underlying representation or the other being "psychologically real".

To see if a /ə/ is "psychologically real" for these words it is necessary to find some way of making the posited /ə/ manifest itself in a surface form.

2.2.2. Perhaps an environment can be found in which the /ə/ would be free to appear. Although the /ə/ is deleted in the majority of cases where the environment specified by the ə-deletion rule is met, there are some suffixes which, even though they provide this environment, seem to block the application of the ə-deletion rule[18]. One such suffix is -iya,[19] which forms adjectives. Thus although /kesər+õ/ "saffron+plural" is [kesrõ], /kesər+iya/ is [kesəriya] and not *[kesriya].

Since suffixes such as -iya block the application of the ə-deletion rule, addition of this suffix to these ə-less forms might allow the /ə/ to appear if it is there in underlying forms. This was tested. In a pilot test three words were chosen from table 5. The words chosen were nouns in common use in everyday speech, and the addition of -iya to them would give nonexistent but semantically reasonable forms. Two words from table 4 were also used (these two of course already had other alternating forms with a /ə/ but do not have existing forms with the suffix -iya[20]. These five words are given below in table 6.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 6

1. [gʰõsla] "nest"
2. [kẽkṛa] "crab"
3. [ʤʰɔ̃pṛi:] "hut"
4. [čipək] "stickiness"
5. [lətək] "the hanging, or suspension"

Seven college-educated native speakers of Hindi served as subjects. Of these six were males and were students or workers in engineering or science. The single female was a housewife. These speakers were given the following two "warm-up" examples (verbally, by me):

"saffron"	kesər+iya	is	kesəriya	"saffron-colored"
"milk"	du:dʰ+iya	is	du:dʰiya	"milky"

They were asked to form in like manner adjectives with -iya using the words given in table 6 above. The results are given in (7).

(7) 5 speakers responded with forms with the [ə]:

[gʰõsəliya]

[ʤʰɔ̃pəṛiya]

etc.

2 speakers gave the forms with clusters only:

[gʰõsliya]

[ʤʰɔ̃pṛiya]

etc.

Though tentative these results suggest that for five of these speakers the /ə/ which never appears at the phonetic level was "psychologically real", and for their grammar these forms should be entered with a /ə/. However for two speakers who gave only the forms with the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

clusters the /ə/ in the underlying form would not be justified.

2.2.3. As mentioned above this was a pilot test. The results were interesting and certainly pointed to the usefulness of conducting a more elaborate test, since a number of factors had not been controlled for in this test. For example, I did not check the speakers for internal consistency. Also it is possible that the speakers could have been analogizing on the basis of the two "warm-up" examples presented to them. Moreover it was desirable to get a larger number of speakers. Thus the following test was performed as a more rigorous version of the pilot test.

Procedure: The test was given to 27 subjects, most of whom were the same used in the test reported in the appendix of this chapter. The following words were presented orally to the subjects by me:

Table 7

1.	ǰ^{h}ag	"foam"
2.	čipək	"stickiness"
3.	čuṭki:	"snap of a finger"
4.	piǰ^{h}ər	"the lagging"
5.	mɔsəm	"weather"
6.	məṭak	"the sway"
7.	su:rət	"countenance"
8.	čəṭak	"brightness"
9.	kəbu:tər	"pigeon"
10.	əṭək	"obstruction"
11.	p^{h}en	"foam"
12.	t^{h}u:k	"spit" (n.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

13. pani: "water"

14. məst "happy-go-lucky"

15. səfed "white"

16. titli: "butterfly"

17. kisəm "kind" (noun)

18. dəst "diarrhea"

19. $g^{h}\tilde{o}sla$ "nest"

20. $j^{h}\tilde{ɔ}p\underset{\cdot}{r}i:$ "hut"

21. $\underset{\cdot}{j}isəm$ "body"

22. $k\tilde{e}k\underset{\cdot}{r}a$ "crab"

23. $\underset{\cdot}{j}^{h}a\underset{\cdot}{r}ən$ "dusting cloth"

24. čust "tight"

25. kala "black"

26. $d^{h}oti:$ "kind of wearing apparel"

27. pi:la "yellow"

28. $lə\underset{\cdot}{t}ək$ "the suspension"

29. gərəm "hot"

30. supari: "betel nut"

This represents a randomized ordering of these words and is the order in which they were given to the subjects. Of these 30 words, 9 (items no. 1, 11, 12, 13, 15, 25, 26, 27, 30) were not of interest to this test, but were put in so that the subjects would not answer automatically according to a certain pattern. The 21 words of interest to the test were of the following types:

(a) words like [$lə\underset{\cdot}{t}ək$] which have a /ə/ in the base, and have existing alternating forms with $\underline{\emptyset}$ as shown in table 4 (the alternation

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is due to either derivation or inflection). Henceforth I will call words of this type 'ə-base words'.

(b) words like [gʰõsla] (i.e. words like those in table 5), which don't have a /ə/ in the base, and don't have existing alternating forms where a [ə] shows up. I will call words of this type 'Ø -base words'.

(c) words like [kisəm][21] which have a [ə] for speakers not well-versed in Urdu (Urdu speakers would instead have [kism]), and which have alternating forms with Ø (even for non-Urdu speakers) when inflectional (or derivational) endings are added, e.g., [kismõ] "kinds" (oblique, plural). (Here the environment of the ə-deletion rule is met thus: kisəm+õ --> kismõ).

(d) words like [dəst] which even though they are historically borrowed words, have to be considered [+native] in a synchronic grammar of Hindi. In words of this type the final cluster is retained by all speakers (all involve the --st cluster, which, along with homorganic nasal+stop, is the only cluster non-high-Hindi and non-high-Urdu both tolerate word finally).

(I might mention here that there are 5 or 6 suffixes which all have the shape -iya but which have different functions; e.g., -iya forms dimunitives: [ḍib:a] "small box", [ḍibiya] "very small box"; it forms occupational nouns: [kəbaṛ] "junk", [kəbaṛiya] "junk collector"; it forms nouns from adjectives: [pi:la] "yellow", [pi:liya] "jaundice" etc. And as mentioned above one function of -iya is to form adjectives. In the 21 words of interest I have tried to use this last adjective-forming -iya.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The subjects were asked to add -iya to the words in table 7. They were told that this might result in words that don't exist in Hindi today, but nevertheless make sense. No 'warm-up' examples were given to the subjects as had been done in the pilot test. Occasionally, throughout the test I would give the meaning that would result from adding -iya to the word, so that the subject would not add -iya automatically as to a nonsense word, but would react to the word as a word that made sense, e.g., "if one wanted to say that a girl's hair is nestlike and we say this by adding -iya to [gʰõsla], what do we get?" etc.

After giving the subjects the 30 words in table 7 I then gave the subjects another task to do (the results of which I wasn't really interested in for this test) namely, to add the suffix -ən to 5 words which I presented to them.

After this I pretended that I had mismarked a few items in the first test and had forgotten to ask for responses to a few words, and so asked for a few more responses. My assumption was that the subject would not remember his previous response because of this added task; however this is an assumption which I can not justify. I have no way of knowing whether the subjects remembered their previous response or not.

This time I asked them to again add -iya to the following words, in the order given below:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 8

1.	du:d^{h}	"milk"
2.	kə̌kṛa	"crab"
3.	ləṭək	"suspension"
4.	ʃ̌h^{h}ɔ̌pṛi:	"hut"
5.	čipək	"stickiness"
6.	kesər	"saffron"
7.	g^{h}ə̌sla	"nest"
8.	čuṭki:	"snap of a finger"
9.	kisəm	"kind"
10.	mɔsəm	"weather"
11.	čəṭək	"brightness"
12.	titli:	"butterfly"

Ten of these words are from the previous list, and 2 are new, namely, [du:d^{h}] and [kesər]. Both of these words have existing -iya forms, e.g., [du:d^{h}iya] and [kesəriya]. These 2 were used as "cue" words to see if speakers were influenced by them, i.e. to see whether speakers just analogized on the basis of patterns exhibited by them. Since [du:d^{h}] was a form which did not involve a [ə], this was the best form I could get to see if the lack of a [ə] influenced the subjects to change their answer (for the four words that follow [du:d^{h}] from an original [ə] to a $\underline{\emptyset}$. Similarly [kesər] was the 'cue' word to see if they would change their original answer to one with a [ə].

In recording the subjects' responses, I recorded whether the subject added or retained the [ə] in giving the base+iya ([ə] was recorded [ə], and if they did not add or retain a [ə] the response was recorded

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 9:

1. In the table only the 21 words of interest are shown.

2. "Other responses" have been given as OR, not the exact response that was given.

3. The 'cue' word [du:d^h] has been left out. Except for the "other responses", all informants gave [du:d^hiya]; since the response does not involve [ə] or ∅, I have left it out of the table. It would have appeared at the top of the list of words in the table.

4. The order of the words does not represent the order given the first time; however, the first 11 words are shown in the order that they were presented to the subjects in the repetition.

5. For the first 11 words there are two columns for each subject. The first column indicates the response given the first time, the second column the response given during the repetition.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Subj. no:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
kẽkra	ər ər	ØØ	ər ər	ØØ	əə	əə	ØØ	əə	ØØ	əə	ØØ	ØØ	ØØ	ØØ	əə	əə	əə	ə	ə	ə	əə	əə	əə	əə	ØØ	ØØ	əə
leṭək	ə	əə	əə	əə	əə	əə	əə	əə	ØØ	əə	ØØ	ØØ	ØØ	ØØ	əə	əə	ØØ	əə	ə	əə	əə	ə	ə	ØØ	ØØ	ØØ	əə
jʰɔ̃pri:	ØØ	ØØ	ØØ	ØØ	ØØ	ə	əə	ə	ØØ	ə	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	əə	ə	ə	əə	ØØ	əə	əə	əə	ØØ	ØØ	ØØ
ṭipək	əə	əə	əə	əə	əə	əə	əə	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ə	əə	ə	ə	ə	əə	əə	ØØ	ØØ	ØØ	ØØ	əə	əə
kesər	ə	ə	ə	ə	ə	ə	ə	Ø	ə	Ø	Ø	Ø	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə
gʰõsla	ər	ØØ	ØØ	ə	ØØ	ə	ə	ØØ	əə	əə	ØØ	ØØ	ØØ	ØØ	ØØ	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə
ʧuṭki:	ØØ	ə	ØØ	ØØ	ØØ	ə	ə	ØØ	ØØ	əə	ØØ	ØØ	ØØ	ØØ	ə	ə	ə	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ
kisəm	ə	əə	əə	əə	əə	əə	əə	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	əə	əə	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ə
məsəm	ə	əə	əə	əə	əə	əə	əə	əə	ØØ	əə	ØØ	ØØ	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə
ʧətək	ə	əə	əə	əə	əə	əə	əə	əə	ØØ	əə	ØØ	ØØ	ØØ	ØØ	əə	əə	əə	əə	əə	əə	əə	əə	əə	əə	ØØ	əə	əə
titli:	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ə	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ	ə	ØØ	ØØ	ə	ØØ	ØØ	ØØ	ØØ	ØØ	ØØ
piʰɔ̃r	ə	ər	ə	ə	ə	ə	ə	ə	ə	ə	Ø	ə	ə	ər	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə
məṭək	Ø	ə	ə	ə	ə	ə	ə	Ø	ə	ə	Ø	Ø	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	Ø	Ø	ər	Ø	ə
su:rət	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	Ø	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə
kabu:tar	Ø	ə	ə	ə	ə	ə	ə	ə	ə	ə	Ø	Ø	ə	ə	Ø	ə	ə	Ø	ə	ə	ə	ə	ə	Ø	ə	Ø	ə
ətək	ə	Ø	ə	ə	ə	ə	ə	ə	ə	ə	Ø	ə	ə	ə	ə	ə	Ø	ə	Ø	ə	ə	ə	Ø	ə	ə	Ø	ə
ɟisəm	Ø	ə	ə	ə	ə	ə	ə	ə	ə	ə	Ø	ə	ə	Ø	ə	Ø	Ø	Ø	Ø	ə	Ø	ə	Ø	ə	ə	Ø	Ø
ɟʰərən	ə	ə	ə	ə	ə	ə	ə	Ø	ə	ə	Ø	Ø	ə	ə	ə	ə	ər	ə	ə	ə	ə	ə	ə	ə	ə	ə	ə
gərəm	Ø	Ø	ə	ə	ə	ə	ə	ə	ə	ə	Ø	Ø	ə	ə	Ø	ə	Ø	Ø	Ø	ə	Ø	Ø	Ø	Ø	Ø	Ø	Ø
məst	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
dəst	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
ʧust	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø

Table 9

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

as ø). I also noted any other response (OR).

Results: The raw data from this test are presented in table 9. Figures 2-6 present an analysis of the responses to the 10 words that were repeated. Some of the percentages in the following figures may not add up to 100% because either the "other responses" have not been included, or there was some rounding off of the figures. Figure 2 gives the breakdown of the responses to the 10 words as to whether they were consistent or inconsistent, and how they were consistent or inconsistent, i.e. whether the consistent response was with a [ə] or a ø̲. Subjects were consistent 77.5% of the time.

Figure 2

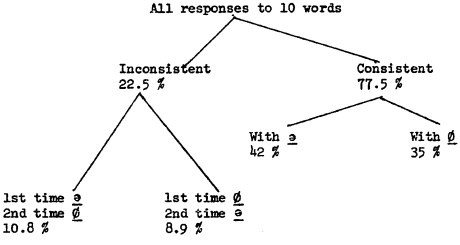
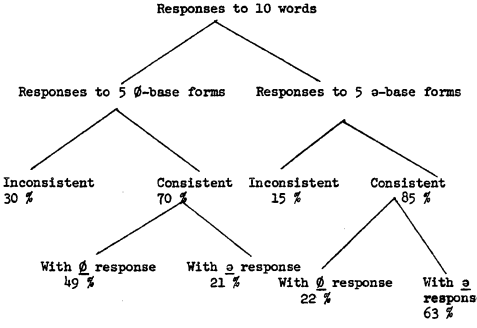


Figure 2 shows that there were more consistent responses with a [ə] than with a ø̲. In the inconsistent responses there was more tendency to first give a [ə] then change it to a ø̲ rather than the other way around.

Figure 3 breaks down these 10 words into whether they were ø-base or ə-base. This same figure also shows the percentage of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 3



consistent responses with [ə] and the percentage of consistent responses with Ø for both the ə-base and the Ø-base words. For the ə-base words, 63 % of the time the [ə] was retained and 22 % of the time a Ø was given consistently. If we look at the consistent responses to the Ø-base words, we find that 49 % of the time the Ø was retained, 21 % of the time a [ə] was given consistently.

To determine if the "cue" word influenced the response the second time, we want to know if there is any change in the incidence of a given type of response, i.e. percentage of [ə] responses and of Ø responses, from the first presentation to the second presentation which was preceded by the given "cue" word. We can analyze the responses in two ways: first we can look at the responses regardless of whether they were of the Ø-base or ə-base type; this is done in figure 4. Second, we can analyze the responses taking into consideration whether

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

they were of the Ø-base or ə-base forms, this is done in figures 5 and 6.

Figure 4

	"cue" word		% response 1st time	% response 2nd time	Net change
All 10 words	Ø	ə	59 %	52.5 %	-6.5 %
		Ø	41 %	48.5 %	+6.5 %
	ə	ə	51 %	52.5 %	+1.5 %
		Ø	49 %	47.5 %	-1.5 %

Figure 4 shows that there might have been an influence in the direction of the 'cue' word, but it was very small. When the 'cue' word with Ø was presented, there were 6.5 % more responses with Ø than there were upon the initial presentation of these words. When the 'cue' word with [ə] was presented, there was a 1.5 % increase in responses with [ə]. On the whole this is not much of a change.

Figures 5 and 6 indicate the percentages of responses as a function of the cue word and as a function of whether they were Ø-base or ə-base, respectively.

Figure 5

	"cue" word		responses 1st time	responses 2nd time	Net change
Ø-base words	Ø	ə	46 %	38.5 %	-7.5 %
		Ø	54 %	61.5 %	+7.5 %
	ə	ə	25 %	36.5 %	+11.5 %
		Ø	75 %	63.5 %	-11.5 %

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 6

	"cue" word	responses 1st time	responses 2nd time	Net change
ə-base words	ø	ə	72 %	66 %	-6 %
		ø	28 %	34 %	+6 %
	ə	ə	77 %	68.5 %	-8.5 %
		ø	23 %	31.5 %	+8.5 %

In the case of ø-base words there is a slightly larger influence of the "cue" word. In the case of the ə-base words, not only is the influence of the "cue" word very little, but the second time the [ə] response went down from 77 % to 68.5 % even after a ə-cue word: i.e., the change in responses was opposite to what would be expected if the cue word was influencing the responses. The slight increase in the responses in the direction of the "cue" word, represented in figure 4, could be either because there was some influence of the 'cue' word or because of chance, i.e. it was random. The -8.5 % net change in the case of the ə-base words, when presented with a ə-cue word, suggests that the slight influence of the "cue" word as shown in figure 4 is due to chance.

Discussion: The results show that subjects were more consistent than they were inconsistent, i.e. their responses were rule governed[22] and not completely random.

We also find that there were more consistent answers in forms with a [ə] base (85%) than in forms with a ø base (70%). There is also greater percentage of adherence to the base in the ə-base forms (63%) than in the ø-base forms (49%), but greater than chance adherence to the base in both, i.e. it seems that most subjects preferred to keep

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the base unchanged.

A question that could be asked is whether or not speakers have pre-existing rules to which they subject new forms, or whether they just analogize on the basis of existing forms, making up rules on the spur of the moment. If the latter was the case, then we might expect subjects to be greatly influenced by "cue" words. However, this does not seem to be the case. Of the subjects that were inconsistent, 6 were inconsistent in a way that showed no influence due to the "cue" word. There were no speakers who changed all the responses the second time around to conform to the "cue" word. However, as mentioned above the Ø-base words seemed to be somewhat more influenced by the "cue" words than ə-base words.

We also find that one subject gave a [ə] no matter what the form (she did this for these 10 words, and for all the remaining 21 words of interest except for those ending in -st); I will call her dialect the ə-dialect. One subject gave a consistent Ø response no matter what the word, including the "cue" word [kesəriya]; I will call her dialect the ə-less dialect. Another subject gave a consistent Ø response to all 21 words and the "cue" words except for one word--[ətək] "obstruction". One other subject gave a consistent Ø response to all the 10 words that were repeated, but kept the [ə] in the "cue" word [kesəriya]. Although these two subjects retained a [ə] in at least one other word, they are perhaps closer to the ə-less dialect than the ə-dialect.

For the subjects that consistently gave a Ø response, the assumption that the -iya suffix blocks the ə-deletion rule was wrong. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ə-deletion rule's environment still seems to be undergoing change, and speakers of this ə-less dialect possibly represent a dialect where the change has been completed. Perhaps we can say that this dialect is a more innovating dialect, since it applies the ə-deletion rule in a more general environment.[23] I do not know whether or not any of the other blocking suffixes (e.g., -izm, thus /sənatən+izm/ "a sect+the English suffix -izm" is usually rendered [sənatənizm] not *[sənatnizm] are blocking suffixes for speakers of this dialect. For the speakers of this ə-less dialect, this test provides no justification for setting up a /ə/ in the underlying forms of Ø-base words; i.e. for speakers of this dialect we have no evidence for or against underlying /ə/'s in words like [ghõsla] because for them the -iya suffix no longer provides an environment for this [ə] to show up. This dialect, however, does not seem to represent the majority of the Hindi speakers (only 1 subject was a clear case of this dialect, and 2 subjects were possible ones).

For speakers of the ə-dialect, i.e. speakers who gave a [ə] in all 10 words, as well as the other words of the test (but not the dəst-type words) the setting up of underlying /ə/s in the Ø-base words is, I think, justified. This dialect does not represent the majority of the speakers either. Only one subject out of 27 was a clear case of this dialect. I have not used myself as a subject, but I, too, use this dialect. Five of the 7 subjects in the pilot test who gave [ə] responses to the items in table 2 may belong to this dialect, --though it can't be said with certainty since they were not checked for consistency. For the speakers of this dialect the -iya suffix does indeed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

provide an environment which blocks the ə-deletion rule, and the /ə/'s which do not show up at the phonetic level for existing words do indeed show up in these "new" words.

One subject always gave the response which kept the base unchanged. For words of the [ghõsIa] type, this subject in this test gave no evidence for underlying /ə/'s. Before setting up (or not setting up) underlying /ə/'s in these words for her grammar, further tests would have to be done. For example, she has the alternations described in table 4. However all these are existing words. If we were to test her with nonsense words and the same suffixes, would she still keep the base unchanged, or would she then change it? E.g., she has [lə̤ṭak] and [lə̤ṭka]; given a word *[bə̤ṭak], would she give *[bə̤ṭka] or *[bə̤ṭaka]?

In my view the spreading environment in which the ə-deletion rule applies represents an on-going sound change. The rule has persisted since Old Hindi, and is still in the process of change. Since the change has not been completed, one might expect a certain amount of variation in subjects' responses, which is what we did find.

For the subjects who gave a mixed response to the words in the test (and this seems to represent the majority of speakers) the -iya suffix sometimes blocks the ə-deletion rule and other times does not. In the existing word [kesəriya], however, the [ə] is <u>always</u> retained. For these subjects we do not have clear-cut evidence for either setting up a /ə/ or not setting up a /ə/ in the Ø-base words, assuming we want the grammar to reflect "psychological reality". There are two possibilities:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) We claim that the underlying forms are the same as the phonetic ones for these words and set up no underlying /ə/'s in them; i.e. we have [ghõsla] at the phonetic level and /g^{h}õsla/ at the lexical level.

(b) We set up underlying forms with a /ə/ for these words; i.e. we have phonetic [g^{h}õsla] but underlying /g^{h}õsəla/.

I will choose analysis (b) here, and set up an underlying /ə/ for Ø-base words, for speakers of the mixed dialect also, accounting for the appearance or nonappearance of the [ə] in all the words of the test (except of course dəst-type words), by saying that this was due to the fact that the -iya suffix is in the process of being removed as a blocking suffix. (This might be best formulated as a sort of "variable rule" (cf. Weinreich, Labov, and Herzog (1968)).

I reject analysis (a) because I can explain the deletion of a /ə/ but I cannot explain the insertion of a /ə/. One might say that the [ə] was sometimes inserted on the basis of analogy with other forms having the [ə], e.g., [kesəriya] etc., that is sometimes speakers analogized on [kesəriya] type words, and other times on the [du:d^{h}iya] type words. First, as mentioned earlier, the test did not show that speakers were greatly influenced by "cue" words, i.e. that they analogized. Second, and more important, subjects never inserted a [ə] in words of the dəst type producing *[dəsətiya]. If analogy played a role then this should have been a possible response. This seems most easily accounted for by forgetting about analogy and simply saying that words like [dəst] do not have an underlying /ə/ therefore a [ə] never shows up in the derivation of these words. Words of the [g^{h}õsla] type do

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

have a /ə/ and therefore it sometimes did show up--the variation between ə~Ø in these words being attributed to the variability of ongoing sound change.

A question that could be asked at this point is whether my subjects were influenced by the writing system when they gave a [ə] response to words like [g^hõsla], that is, although the words were presented to them orally, is it possible that the conventional spelling of these words influenced speakers to put in a [ə] where there was none previously? First, as far as Hindi orthography is concerned, the rendering of words in Devanagari is not so conventionalized. In Devanagari these words are usually written as if they did not have a cluster even though they are pronounced with a cluster. (The same is true of words such as [čipka] etc., i.e. words pronounced with clusters but which, unlike the g^hõsla-type words, do have alternating forms with the cluster broken up by a [ə] (see table 4 col. 2)). But people can and frequently do write these words as if they did contain a cluster.

Furthermore three of my subjects did not know Devanagari well enough to be influenced by it. Their responses (informants no. 23, 24, 25, in table 9) were no different from the responses of other subjects.

However, to find out for sure whether orthography might have partially influenced speakers' response, it would be necessary to add another feature to the test, namely that of giving the test to preliterate children also, or to adults who are not fluent in Devanagari (which is common among scientists and engineers).

One other question (raised by Zimmer, personal communication) is how did the children who heard words of the [g^hõsla]-type always with a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cluster deduce an abstract form with a /ə/? I cannot provide any conclusive answer to this question. However, it is well known that children produce words they have not heard before, e.g., a child hears poise, assumes it is a plural form and asks "what is a poy?". I can speculate that upon hearing existing forms such as [čipka]-[čipək] (cf. table 4) the child forms underlying forms with a /ə/ for these words, which /ə/ he deletes with the ə-deletion rule. Similarly, hearing a form such as [g^{h}ɔ̃sla], he hypothesizes that it too had a /ə/ which has been deleted by the ə-deletion rule, and thus posits underlying /g^{h}ɔ̃səla/. Of course not all children would arrive at the same conclusion, and it is entirely possible that some children have only /g^{h}ošla/ as the underlying form. Some children's treatment of the Hindi word for "lizard" [č^{h}ipkəli:] (to be discussed in 2.2.5) provides additional evidence that children may somehow formulate "abstract" forms quite different from phonetic forms.

To summarize briefly, for the speakers of the ə-dialect (including speakers who gave a mixed response), I propose that words such as [g^{h}ɔ̃sla] "nest", which do not have existing alternating forms with a [ə], are represented in the lexicon with an underlying /ə/, e.g., /g^{h}ɔ̃səla/. The ə-deletion rule applies to forms such as these to yield the correct phonetic output, e.g., [g^{h}ɔ̃sla]. However, suffixes such as -iya block the application of the ə-deletion rule such that /g^{h}ɔ̃səla+iya/ yields [g^{h}ɔ̃səliya]. I cannot propose underlying /ə/s in such words in the lexicon of speakers of the ə-less dialect. Since for these speakers suffixes such as -iya do not block the application of the ə-deletion rule and therefore there is as yet no evidence for any

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

underlying /ə/.

2.2.4. The results also suggest another interesting point: namely that speakers who produce forms which are identical phonetically may still have quite different grammars. (See also Zimmer (1969:320, and Gleitman and Gleitman). Such speakers would usually be considered to be speakers of the same dialect. Current generative phonology would assume that they have THE ideal grammar of their language, which as children learning the language, they had to choose. But if what I have said above is correct, then this is not the case. Given the same phonetic data, e.g., words of the [g^{h}ōsla] type, one child might form an abstract form with an underlying /ə/ and another an underlying form with no /ə/, i.e. with the cluster. Or one child might on the basis of [kesəriya] hypothesize a grammar where -iya blocks the ə-deletion rule, whereas another child might hypothesize a grammar where -iya does not block the ə-deletion rule, simply marking the existing [kesəriya] as an exception. Thus:

Speaker A	Speaker B
/g^{h}ōsəla/	/g^{h}ōsla/
↓ (by the ə-deletion rule)	↓
[g^{h}ōsla]	[g^{h}ōsla]

And:

/kesər+iya/	/kesər+iya/
↓ (-iya blocks ə-deletion rule)	↓ (exception to ə-deletion rule)
[kesəriya]	[kesəriya]

The difference in their grammar would only become apparent when they attempted to formulate new words (under 'grammar', I include the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the lexicon as well as the rules). Thus I think this is evidence for the view that there is more than one possible grammar that children may select. I have not seen any empirical evidence which supports the counter-assumption: that there is a single ideal grammar of the language which a child must select. I am not proposing that we write a separate grammar for every speaker; we can probably narrow the choice down to a small number of grammars. But it seems unlikely that we can settle for just one maximally "optimal" grammar.

This bears crucially on the problem of what kind of evidence is required to verify a particular phonological analysis of a language. In particular it undermines the claim of Kiparsky (1971) that (586):

> We have to find external evidence which shows that the abstract analyses posited in certain cases do have psychological reality. Of course, this cannot and need not be done in every single example. Once we have provided external justification in a few clear ones, we can provisionally adjust phonological theory so that it will require the abstract analyses in these cases. The resulting theory will then lead to specific predictions about what the correct analysis is in many other cases, where external empirical evidence may or may not be available to test the theory.

But the evidence cited above suggests that speaker A may have abstract forms, and speaker B may not. If we followed Kiparsky we would set up abstract forms for speaker B, too, on the evidence of their existence for speaker A.

2.2.5. Certain other aspects of Hindi phonology can also be better accounted for if we posit an underlying /ə/.

1. The word for lizard in Hindi is [čʰipkəli:]. The word has a medial -pk- cluster at the phonetic level. Does it have this cluster at the abstract level also? There are no alternating forms for this word where a [ə] shows up. However children are reported to say

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[č^{h}ipəkli:]. Normally the ə-deletion rule applies from right to left (see chapter 4 for details). If we posit the underlying form to be /č^{h}ipəkəli:/, with a feature which marks it as being an exception to the right-to-left application of the ə-deletion rule, then it explains the child's behavior. The child treats it as a normal word and applies the ə-deletion rule from right to left getting [č^{h}ipəkli:]. Children's "mistakes" are usually in the direction of regularizing words that are exceptions. What he has to learn, in order to copy the adult model, is that this word is an exception; the ə-deletion rule seems to apply from left to right here. Thus it seems that there is an underlying /ə/ in the word. (Another analysis, of course, is that the child metathesizes the adult's kə to ək in this word, but then why is this word only such a popular candidate for metathesis?)

2. Additional evidence for abstract /ə/'s is the following: There is an overwhelming tendency for Hindi to have only homorganic nasals before stops (for details see section 2.3.5.1 of this chapter). This is true mostly of [+native] words considered at the lexical level, since obviously the ə-deletion rule produces non-homorganic nasals before stops at the phonetic level. However, there are some exceptions to the generalization that nasals are homorganic before stops:

(8) tinka "bit of dry grass"
mənka "a kind of bead"
kənk^{h}i: "corner of the eye"
kunba "family"
čingari: "spark"
inkar "denial"[24]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the case of [inkar], [čingari:], and [kunba], I think these should perhaps just be listed as exceptions because sometimes speakers do regularize them, pronouncing them as [iŋkar] [čiŋgari:] etc. If we recognize them as exceptions to begin with we can understand why speakers would change them in this way--that is, regularize them. If we treat them as underlyingly regular as has been proposed by some writers (Narang and Becker (1971), see 2.3 for details). we cannot account for this change. The other words in the list are never regularized in my experience. These are best accounted for by positing a /ə/ between the cluster, i.e. /kənək^{h}i:/, etc., and then deleting the /ə/ by the ə-deletion rule. I will discuss the case of this word and the topic of homorganic nasals in general in the next section.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.3. Much valuable work on the sequential constraints of Hindi consonant clusters has already been done by various writers (Bhatia 1964, Gupta 1966, Arun 1961, Dixit 1963, Gumperz and Naim 1960, etc.). However there are various aspects of the problem that some or all of these writers fail to treat.

Bhatia, Arun, Gupta, and Dixit treat medial clusters, but unfortunately do not distinguish intermorphemic vs. intramorphemic clusters. They also do not give separate treatment to those clusters that occur in loanwords, as opposed to those in native words.

Furthermore, Dixit does not consider Hindi to have any final clusters; he posits a [ə] after all clusters that are considered to be final by other authors. It is true that there is quite an audible release after most final clusters. But in my view, phonologically, this release should not be considered to be a final vowel for the following reasons: certain types of clusters occurring in final position in loan words are simplified by most Hindi speakers, e.g., the Sanskrit loan for "ashes" b^{h}əsm is rendered in informal style by most Hindi speakers as [b^{h}əsəm] (see also discussion in 2.3.4.). Medially, however, clusters are not simplified, and when the word for "ashes" appears with an inflectional ending the cluster is maintained: [b^{h}əsmõ] "ashes" (oblique). Dixit would have no difficulty in accounting for the form [b^{h}əsmõ]; but since he would transcribe b^{h}əsm as [b^{h}əsm^{ə}], i.e. with a medial, not a final cluster, it is not at all clear how he would explain why b^{h}əsmə is often simplified to [b^{h}əsəm] but b^{h}əsmõ is never simplified to *[b^{h}əsəmõ], nor is the sm cluster simplified in words like [kismət] "fate", which do not involve

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

inflectional endings. Presumably he would have to mark specially all those words where the cluster is simplified. However these marked words would be precisely the class of words which other writers have considered to contain final clusters.

Of the above-mentioned writers, Gumperz and Naim are the only ones who differentiate between loan and native clusters. However they only deal with initial and final clusters, not medial. They also do not give any examples to illustrate the clusters they list in their charts.

Due to the above-mentioned limitations in descriptions of clusters in Hindi, I decided to gather the data anew for this study.

Before undertaking the study, however, the following question had to be faced: in a grammar of competence of Hindi, should these clusters be differentiated as to whether they occur in loanwords versus native words? My answer is 'yes'. As mentioned in chapter 1, speakers are aware of some words as being [+foreign], and one of the ways in which they recognize certain items as being loans is by the clusters they contain. For example in the test reported in appendix 2 (involving native Hindi speakers' response to nonsense words containing a variety of consonant clusters) some of the subjects reported that certain words could be either "Urdu" or "English". As the words were constructed by me, they obviously did not exist in my informants' vocabulary; the informants were presumably basing their judgment on the initial clusters contained in these words.

Moreover the retention or simplification of clusters has, among other things, social significance in communication. Gumperz and Naim

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

note this when they say (116): "Degree of formality and informality is signalled by .... consonant clusters". Depending on whether a speaker is knowledgeable in High Hindi or High Urdu, and depending on whether he is 'westernized' or not, i.e. knows English or not, he will or will not use certain clusters. The use of certain clusters also depends on his level of education, whether the person he is speaking to is an Urdu/Hindi speaker, and the source from which most of his vocabulary is drawn (e.g., Sanskrit or Perso-Arabic, etc.) with the "uneducated low prestige speech [showing] the smallest number of clusters" (112). However if certain clusters don't appear in a given style of speech, it does not necessarily mean that the speaker cannot pronounce them. (As mentioned earlier in this chapter, pronunciation constraints need not be the same as the sequential constraints specific to the language). The facts are much more complicated. We find that a "high Hindi" speaker may correctly pronounce a certain cluster in a Sanskrit loan, but simplify the same cluster in a PA loan. I will discuss this in more detail with an example shortly.

2.3.1. The initial, medial, and final clusters of Hindi are given in the tables, below. I have indicated whether a cluster occurs in native words or only in loans, and the source of such loans, i.e. Sanskrit or English or Perso-Arabic. Although the most common clusters occurring in loans are listed, the table does not list all possible loan clusters. I have tried to be as complete as possible in the case of native clusters, but it is possible that I have missed a few. Perhaps, since the informant (myself) draws more on Sanskrit than Perso-Arabic for learned vocabulary, the listing of the Sanskrit loan

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

clusters might be more complete than the Urdu ones.

In trying to list as complete a list of clusters as possible, my aim was mainly a descriptive one. I am not claiming that all these clusters occur in a grammar of competence of any one native speaker, though it is not impossible that there is such a speaker. Examples of words containing these clusters are given in Appendix 3. They are represented there in a broad phonetic transcription. A number of low level phonetic rules will be needed to give the final phonetic output. For example, in word-final position these rules will have to state that there is quite an audible release after clusters. There are different degrees of this release depending on the nature of the clusters; when the final member of the cluster is [y] the release seems to be most distinct, e.g., [nrity] "dance" in a more narrow phonetic transcription would be [nrity^{ə}]. This final vowel, however, is much more attenuated than the regular [ə]. Another phonetic detail which I have not taken into account in listing the clusters and the examples (mainly due to lack of data) is that of deaspiration of final stops (and in a number of cases where the aspirate is the first member of a cluster). In allegro style word-final aspirates seem to be deaspirated. I should also mention that a number of the clusters only exist due to the application of the ə-deletion rule. E.g., [-wṭ-] occurs only in words such as [bənawṭi:] "artificial" which must have an underlying /ə/ breaking up the cluster since there are alternating forms such as [bənawəṭ] "make" (noun). Thus at the systematic phonemic level such clusters would not exist (cf. also the discussion in 2.2). All other clusters exist at both the systematic

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

phonemic level and the systematic phonetic level.

In the following tables native clusters are represented by an asterisk '*' and clusters occurring only in loans by an x. The source of the loan is shown at the top of the square, where PA=Perso-Arabic, S=Sanskrit, and E=English.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

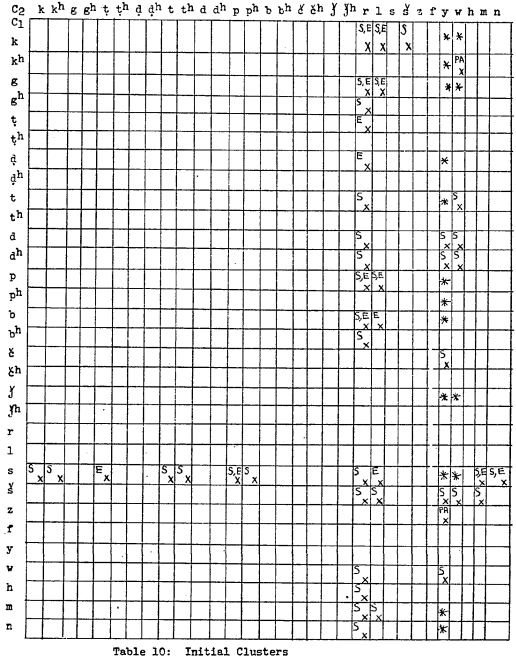


Table 10: Initial Clusters

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

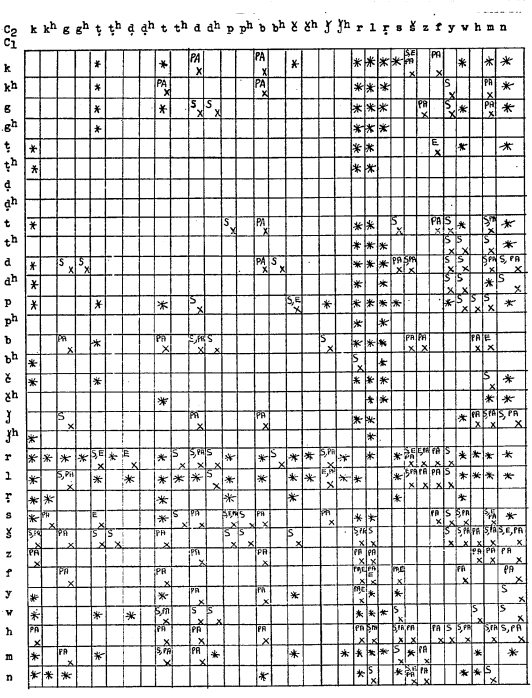


Table 11: Medial Clusters

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

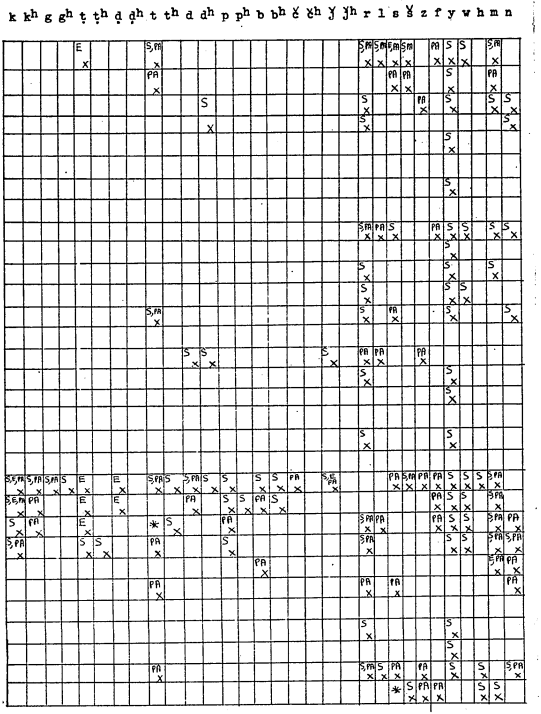


Table 12: Final Clusters

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

For ease of reference I also list these clusters in groups according to the classification below. Also listed are three-consonant clusters and homorganic nasal+stop clusters which are not shown in the preceding tables. Examples of all of these clusters are given in Appendix 3.

Initially: kr-, kl-, kw-, ky-, kš-; k^hy-, k^hw-; gr-, gl-, gy-, gw-; g^hr-; čy-; jy-, jw-; ṭr-; ḍr-; tr-, ty-, tw-; dr-, dy-, dw-; d^hr-, d^hy-, d^hw-; pr-, pl-, py-; p^hy-; br-, bl-, by-; b^hr-; nr-, ny-; mr-, ml-, my-; wr-, wy-; hr-; šr-, šl-, šw-, šy-, šm-; sr-, sl-, sw-, sy-, sk-, sk^h-, sṭ-, st-, st^h-, sp-, sp^h-, sn-, sm-; zy-; str-, spr-, spl-, skr-, smr-.

Of these those occurring only in native words: ky-, kw-; k^hy-; gy-, gw-; jy-, jw-; ḍy-; ty-; py-; p^hy-; by-; ny-, my-; sy-, sw-; only in PA loans: k^hw-; zy-; only in English loans: ṭr-, ḍr-; bl-; sṭ-, sl-; spl-, skr-; only in Sanskrit loans: kš-; g^hr-; čy-; tr-, tw-; dr-, dy-, dw-; d^hr-, d^hy-, d^hw-; b^hr-; nr-, mr-, ml-; wr-, wy-; hr-; šm-, šl-, šw-, šr-; sr-, sk^h-, st-, st^h-, sp^h-, sk-; str-, smr-; šy-; in more than one type of loan: kr-, kl-, gr-, gl-, pr-, pl-, br-; sp-, sm-, sn-; spr-.

Finally: -kṭ, -kt, -kr, -kl, -ks, -kš, -kf, -ky, -kw, -km; $-k^h$t, $-k^h$s, $-k^h$š, $-k^h$y, $-k^h$m; $-gd^h$, -gr, -gz, -gy, -gm, -gn, $-g^h$r, $-g^h$n; -ṭy; -ḍy; -tr, -tl, -ts, -tf, -ty, -tw, -tm, -tn; $-t^h$y; -dr, -dy, -dm; $-d^h$r, $-d^h$y, $-d^h$w; -pt, -pr, -py, -pn, -ps; -bd, $-bd^h$, -bj, -br, -bl, -bz; $-b^h$r, $-b^h$y; -čy; -jr, -jy; -rk, $-rk^h$, -rg, $-rg^h$, -rṭ, -rḍ, -rt, $-rt^h$, -rd, $-rd^h$, -rp, -rb, $-rb^h$, -rč, -rj, -rs, -rš, -rz, -rf, -ry, -rw, -rh,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-rm; -lk, -lk^{h}, -lṭ, -lḍ, -ld, -lp, -lp^{h}, -lb, -lb^{h}, -lf, -ly, -lw, -lm; -sk, -sk^{h}, -sṭ, -st, -st^{h}, -sp, -sr, -sl, -sf, -sy, -sw, -sm, -sn; -šk, -šṭ, -šṭ^{h}, -št, -šp, -šr, -šy, -šw, -šm, -šn; -zb, -zm, -zn; -ft, -fr, -fs, -fn; -wr, -wy; -hy; -mt, -mr, -ml, -ms, -mz, -my, -mh, -mn; -ns, -nš, -nf, -nz, -nh, -nm; all homorganic nasal+stop clusters.

Those occurring in native words: -st, -ns, and all homorganic nasal+stop clusters; only PA loans: -kf; -k^{h}t, -k^{h}s, -k^{h}š, -k^{h}m; -gz; -ps, -br, -bl, -bz; -tl, -tf; -rč, -rs, -rz, -rf; -lk^{h}, -ld, -lb, -lf; -sk^{h}, -sp, -sl, -sf; -št; -zb, -zn; -ft, -fr, -fs, -fn; -mt, -ms, -mz; -nz, -nf; only Sanskrit loans: -ky, -kw; -k^{h}y; -gr, -gy, -gm, -gn, -gd^{h}; -g^{h}n, -g^{h}r; -ṭy; -ḍy; -ts, -ty, -tw, -tm, -tn; -t^{h}y; -dr, -dy, -dm; -d^{h}y, -d^{h}r, -d^{h}w; -pr, -py, -pn; -bd, -bd^{h}, -bj; -b^{h}r, -b^{h}y; -čy; -jr, -jy; -rg^{h}, -rt^{h}, -rd^{h}, -rp, -rb, -rb^{h}, -ry, -rw, -rh; -lp, -lp^{h}, -lb^{h}, -ly, -lw; -sk, -st^{h}, -sy, -sw; -šṭ, -šṭ^{h}, -šp, -šy, -šw; -wr, -wy, -hy; -ml, -my, -mh; -nš, -nh, -nm; only in English loans: -kṭ, -sṭ, -rḍ, -rṭ, -lṭ, -lḍ; in more than one kind of loan: -kt, -kr, -kl, -ks, -kš, -km; -tr; -rk, -rg, -rk^{h}, -rt, -rd, -rj, -rš, -rm; -lk, -lm; -sr, -sm, -sn; -šk, -šr, -šm, -šn; -zm; -mr, -mn; -pt.

Final three-consonant clusters: -kšṇ, -kšy, -kšm; -gd^{h}y, -try, -tsy, -štr, -šṭ^{h}y, -str, -st^{h}y, -sty, -rky, -rk^{h}y, -rg^{h}y, -rjy, -rty, -rtm, -rt^{h}y, -rdr, -rd^{h}w, -ršw, -rhy, -rṇy, -rmy, -lky, -ṇḍy, -ṇdr, -ŋk^{h}y, -nty, -ntr, -ndr, -ndw, -ndy, -nd^{h}y. All of the three-consonant clusters occur in Sanskrit loans only.

Final four-consonant clusters: -rtsy. Occurs only in one Sanskrit loan.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Medially:

stop+stop: -kṭ-, -kt-, -kd-, -kb-, -kč-; -k^{h}ṭ-, -k^{h}t-, -k^{h}b-; -gṭ-, -gt-, -gd-, -gd^{h}-; -g^{h}ṭ-; -ṭk-; -ṭ^{h}k-; -tk-, -tp-, -tb-; -dk-, -dg-, -dg^{h}-, -db-, -db^{h}-; -d^{h}k-; -pk-, -pṭ-, -pt-, -pd-, -pč-, -pj-; -bg-, -bṭ-, -bt-, -bd-, -bd^{h}-, -bj-; -b^{h}k-; -čk-, -čṭ-; -č^{h}t-; -jg-, -jd-, -jb-; -j^{h}k-;

stop+liquid: -kr-, -kl-, -kṛ-; -k^{h}r-, -k^{h}l-, -k^{h}ṛ-; -gr-, -gl-, -gṛ-; -g^{h}r-, -g^{h}l-, -g^{h}ṛ-; -ṭr-, -ṭl-; -ṭ^{h}r-, -ṭ^{h}l-; -tr-, -tl-; -t^{h}r-, -t^{h}l-, -t^{h}ṛ-; -dr-, -dl-, -dṛ-; -d^{h}r-, -d^{h}ṛ-; -pr-, -pl-, -pṛ-; -p^{h}ṛ-, -p^{h}r-; -br-, -bl-, -bṛ-; -b^{h}r-, -b^{h}ṛ-; -čr-, -čl-, -čṛ-; -č^{h}l-, -č^{h}ṛ-; -jr-, -jl-; -j^{h}l-;

stop+fricative: -ks-, -kš-, -kf-; -gz-; -ṭf-; -ts-, -tf-; -ds-, -dš-; -ps-; -bš-, -bz-;

stop+glide: -kw-; -k^{h}y-; -gy-, -gw-; -ṭw-; -ty-, -tw-; -t^{h}y-, -t^{h}w-; -dy-, -dw-; -d^{h}y-, -d^{h}w-; -py-, -pw-, -ph-; -bh-; -jw-, -jh-;

stop+nasal: -km-, -kn-; -k^{h}m-, -k^{h}n-; -gm-, -gn-; -ṭn-; -tm-, -tn-; -t^{h}m-, -t^{h}n-; -dm-, -dn-; -d^{h}m-, -d^{h}n-; -pm-, -pn-; -bm-; -čm-, -čn-; -č^{h}m-, -č^{h}n-; -jm-, -jn-;

fricative+stop: -sk-, -sk^{h}-, -sṭ-, -st-, -st^{h}-, -sd-, -sp-, -sp^{h}-, -sb-, -sǰ-; -šk-, -šg-, -šṭ-, -šṭ^{h}-, -št-, -šp^{h}-, -šp-, -šč-; -zk-, -zd-, -zb-; -fg-, -ft-;

fricative+liquid: -sr-, -sl-; -šr-, -šl-; -zr-, -zl-; -fr-, -fl-;

fricative+glide: -sy-, -sw-; -šy-, -šw-, -šh-; -zh-; -fw-;

fricative+fricative: -fs-, -sf-;

fricative+nasal: -sm-, -sn-; -šm-, -šn-; -zm-, -zn-; -fn-;

glide+stop: -yk-, -yt-, -yd-, -yb-, -yč-; -wk-, -wṭ-, -wḍ-, -wt-,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-wd-, -wd^{h}-; -hk-, -ht-, -hd-, -hb-;

glide+liquid: -yr-, -yl-; -wr-, -wl-, -wṛ-; -hr-, -hl-;

glide+fricative: -ys-; -ws-; -hs-, -hš-, -hf-;

glide+glide: -wh-; -hy-, -hw-;

glide+nasal: -yn-; -wn-; -hn-, -hm-;

liquid+stop: -rk-, -rk^{h}-, -rg-, -rg^{h}-, -rṭ-, -rṭ^{h}-, -rḍ-, -rt-, -rt^{h}-, -rd-, -rd^{h}-, -rp-, -rb-, -rb^{h}-, -rč-, -rč^{h}-, -rj-, -rj^{h}-; -lk-, -lg-, -lṭ-, -lḍ-, -lt-, -lt^{h}-, -ld-, -ld^{h}-, -lp-, -lb-, -lč-, -lč^{h}-, -lj-, -lj^{h}-; -ṛk-, -ṛk^{h}-, -ṛṭ-, -ṛp-, -ṛč-;

liquid+fricative: -rs-, -rš-, -rz-, -rf-; -ls-, -lš-, -lz-, -lf-; -ṛs-;

liquid+liquid: -rl-; -lr-;

liquid+glide: -ry-, -rw-, -rh-; -ly-, -lw-, -lh-; -ṛw-, -ṛh-;

liquid+nasal: -rm-, -rn-, -rṇ-; -lm-, -ln-;

nasal+nasal: -mn-, -mṇ-; -nm-; nasal+homorganic stop, all possibilities;

nasal+non-homorganic stop: -mk-, -mg-, -mṭ-, -mt-, -md-, -md^{h}-, -mč-, -mj^{h}-; -nk-, -nk^{h}-, -ng-, -nb-;

nasal+fricative: -ms-, -mš-, -mz-; -ns-, -nš-, -nz-;

nasal+liquid: -mr-, -ml-, -mṛ-; -nr-, -nl-;

nasal+glide: -mh-; -nw-, -nh-.

Of these the following occur in native words: -kṭ-, -kt-, -kč-, -kr-, -kl-, -kṛ-, -ks-, -kw-, -km-, -kn-; -k^{h}ṭ-, -k^{h}r-, -k^{h}l-, -k^{h}ṛ-, -k^{h}n-; -gṭ-, -gt-, -gr-, -gl-, -gṛ-, -gw-, -gn-; -g^{h}ṭ-, -g^{h}r-, -g^{h}l-, -g^{h}ṛ-; -ṭk-, -ṭr-, -ṭl-, -ṭw-, -ṭn-; -ṭ^{h}k-, -ṭ^{h}r-, -ṭ^{h}l-; -tk-, -tr-, -tl-, -tw-, -tn-; -t^{h}r-, -t^{h}l-, -t^{h}ṛ-, -t^{h}n-; -dk-, -dr-, -dl-, -dṛ-;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-dʰk-, -dʰr-, -dʰṛ-, -dʰm-; -pṭ-, -pk-, -pj-, -pr-, -pt-, -pl-, -pṛ-,
-ps-, -py-, -pn-; -pʰr-, -pʰṛ-; -bṭ-, -br-, -bl-, -bṛ-; -bʰk-, -bʰṛ-;
-čṭ-, -čk-, -čr-, -čl-, -čṛ-, -čn-; -čʰt-, -čʰl-, -čʰṛ-, -čʰm-, -čʰn-;
-jr-, -jl-, -jw-; -jʰk-, -jʰl-; -rk-, -rkʰ-, -rg-, -rgʰ-, -rṭʰ-, -rt-,
-rp-, -rb-, -rč-, -rčʰ-, -rjʰ-, -rl-, -rs-, -rw-, -rh-, -rm-, -rn-;
-lk-, -lṭ-, -ld-, -lt-, -ltʰ-, -ld-, -lp-, -lb-, -lč-, -lčʰ-, -ljʰ-,
-lr-, -ls-, -lw-, -lh-, -lm-, -ln-; -ṛk-, -ṛkʰ-, -ṛt-, -ṛp-, -ṛč-, -ṛs-,
-ṛw-; -sk-, -st-, -sr-, -sl-, -sn-; -yk-, -yt-, -yč-, -yl-, -ys-; -wk-,
-wṭ-, -wḍ-, -wr-, -wl-, -wṛ-; -mk-, -mṭ-, -mdʰ-, -mč-, -mjʰ-, -mš-,
-mr-, -ml-, -mṛ-, -mh-, -mn-; -nk-, -nkʰ-, -ng-, -nb-, -nr-, -nw-,
-nh-, -ns-; only in English loans: -ṭf-, -bm-, -sṭ-, -rḍ-; only in PA
loans: -kb-, -kd-, -kf-; -kʰt-, -kʰb-, -kʰm-; -gz-, -gm-; -tb-, -tf-;
-db-, -ds-; -bt-, -bg-, -bz-, -bš-, -bh-; -jb-, -jd-, -jh-; -zb-, -zd-,
-zk-, -zl-, -zr-, -zh-, -zn-, -zm-; -ft-, -fg-, -fw-, -fn-; -sb-, -skʰ-,
-sd-, -sj-, -sf-; -št-, -šg-, -šh-; -yd-, -yb-; -hf-, -hš-, -hr-, -hd-,
-hb-, -ht-, -hk-; -rf-; -lz-, -lf-; -md-, -mg-, -mz-; -nz-; only in
Sanskrit loans: -gd-, -gdʰ-, -gy-; -kʰy-; -tp-, -ts-, -ty-; -tʰy-,
-tʰw-, -tʰm-; -dg-, -dgʰ-, -dbʰ-, -dy-, -dw-; -dʰy-, -dʰw-, -dʰn-;
-pd-, -pw-, -ph-, -pm-; -bj-, -bdʰ-; -bʰr-; -čm-; -jg-; -šṭ-, -šč-,
-šṭʰ-, -šp-, -šl-, -šy-, -špʰ-; -stʰ-, -spʰ-, -sy-; -rdʰ-, -rbʰ-, -rtʰ-,
-ry-; -ldʰ-, -ly-; -wd-; -wdʰ-; -wn-, -ws-, -wh-; -yn-; -hy-; -nm-,
-nl-; -ms-; in more than one type of loan: -kš-; -tm-; -dn-, -dm-,
-dš-; -pč-; -bd-; -jm-, -jn-; -fl-, -fs-, -fr-; -šk-, -šr-, -šw-, -šm-,
-šn-; -sp-, -sm-, -sw-; -wt-; -hw-, -hn-, -hl-, -hs-; -rṭ-, -rd-, -rj-,
-rš-, -rz-; -lg-, -lš-, -lj-; -yr-; -mt-; -nš-.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Medial three-consonant clusters: English: -rtg-, -mpr-, -ndr-, -nfl-; PA loans: -stm-, -str-; Sanskrit loans: -krw-, -kšm-, -kšw-, -tkr-, -tsn-, -tsy-, -tpr-, -rtr-, -rtʰn-, -ŋkʰy-, -ŋgy-, -ŋkr-, -ndʰy-; those occurring in native words: -mbr-, -mbʰl-, -ŋgṛ-, -ŋgr-, -ŋgl-, -ŋgʰr-, -ñčm-, -ñjʰr-, -ñjʰl-, -ñjṛ-, -ñjr-, -ndʰl-, -ntr-, -ndr-, -ndʰr-.

2.3.2. The sequential constraints have not been stated in terms of syllables; my reasons for this are discussed in 2.3.6. Before stating some general points on these clusters, let me give some details on why I have excluded or included certain clusters.

Initial consonant clusters: I have not included lh- as in lhasa "capital of Tibet". The word is written with an lh- cluster, but I believe speakers pronounce it simply as [lasa].

Similarly I have omitted hw- which Dixit lists, as in hwel "whale". I think most speakers pronounce it [wel].

Dixit also lists pʰr- and pʰl- which occur only in the mispronunciation of English loans like frame, flute, etc., by speakers of dialects with no [f]. It is my impression that these dialect speakers say [pʰəluːt] and [pʰərem], i.e. they don't have the cluster pʰr- and pʰl-; therefore they are not included.

I am treating [kʰy-] as native even though in only occurs in words that are historically loans, e.g., [kʰyal] "thought" (PA). However the loan originally began with x which has been replaced by [kʰ-] by non-Urdu speakers, and for these speakers I think [kʰyal] behaves like a native word (I might mention that it is a very commonly used word).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I have listed [zy-] and [šy-] as [-native] clusters because [z] and [š] are [-native] segments, but it could be argued that they should be treated as native, since those speakers who do have the [z] or [ś] sound wouldn't regard this cluster as different from any native clusters. [šy-] occurs in common names like [šyam] or [šyama] which were originally Sanskrit loans. (Actually it is possible that at the phonetic level in casual speech these are actually [šam] and [šama] -- there doesn't seem to be much of a difference between [š] and [šy] phonetically.)

[p^hy-] occurs in my dialect in the word [p^hyas] "dandruff", so I have listed it as a native cluster. However this is the only word I could find with this cluster, though it does occur frequently in mispronunciations of English loans with [fy-] as in 'fuel' [p^hyu:l] by speakers who do not have [f]. It is possible that [p^hy-] should not be treated as a native cluster, because in the test reported in the appendix many of the subjects treated nonsense words with this cluster as being possibly English, but not Hindi.

Final clusters: I have listed [-rč] as occurring only in PA loans. The dictionary lists it as occurring in the native word for "chili" which it lists as mirč, however most speakers I know simplify the cluster to [mirič], i.e. they seem to treat it as [-native]. In a variant of the word where the cluster becomes medial it is retained, e.g., [mirči:].[25]

I have treated [-ns] as a native cluster, even though the examples containing it could be considered to be loans, e.g., [həns] "swan". [-ns] seems to be treated as a native cluster by native

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

speakers (see chapter 3 for more discussion on [hɛns]). Moreover it could be considered to be a nasal homorganic with the following [s] (homorganic nasal+stop clusters are native clusters).

[-rh] has been listed on the chart. It is possible that it should have been omitted, because I am not certain if speakers do actually pronounce it as [-rh] or simply [-r]. In any case it only occurs in rare Sanskrit loans.

Similarly although I have listed [-st^h] and [-sk^h], it is possible that they are deaspirated to [-st] and [-sk] at the phonetic level, and should have been omitted.

Among final clusters for Urdu, Gumperz and Naim list -zd, -ǰr, -lt, -rb, -ǰz, -wz, -ǰd, and -md. I have omitted these because I was unable to find examples containing them.

Gumperz and Naim also list -xr and -xl which I have not included. Non-high-Urdu speakers do not have the sound [x]; they replace it by [k^h]. Moreover I do not even list a [-k^{h}r] or a [-k^{h}l], because words that I found with -xr and -xl would not be in the vocabulary of a non-high-Urdu speaker, and a high Urdu speaker would of course have the correct -xr and -xl.

Gumperz and Naim also list Urdu -pt and -ps, for which I have not been able to find examples in the dictionary. However the dictionary lists words with -bt and -bs, e.g., zəbt "seized", həbs "humid", which I would pronounce as [zəpt] and either [həbz] or [həps] , i.e. with voicing agreement in the final cluster. It is likely that Gumperz and Naim had such examples in mind, therefore I have listed [-pt] and [-ps] but not -bt and -bs. Similarly the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

dictionary lists the word for "word" as ləfz, but I would render this as [ləfs]; thus I have not included the cluster -fz. (Gumperz and Naim don't list this cluster either.) For those Hindi speakers (if any) who do pronounce the above clusters in these Urdu words without the voicing agreement (i.e. as [-bt], [-bs], etc.), these would have to be listed as exceptions to the sequential constraints.

Medial clusters: I have treated [-gd-] and [-ǰg-] as non-native clusters occurring only in Sanskrit loans. Arun gives lugdi: "a kind of cake" as an example of the former cluster. This word is a Desi word (a term applied to words of unknown origin), and is the only non-Sanskrit example of this cluster that I could find. It is not a very commonly-used word. In the case of [-ǰg-], Arun provides gəǰga "ornament for elephant", as an example of a native word containing it. I think this word should be analyzed as /gəǰ+ga/ and thus is not an example of a morpheme-medial cluster.

[-bg-] has been treated as a cluster occurring in PA words. It is possible that the one example that I have found, namely [i:səbgol] "a kind of medicine", should be treated as native, but more evidence is needed. Native words like [dəbgər] "a leather cone maker" (actually the word is a Desi word), I consider to be /dəb+gər/, [-gər] occurring in various words like [kari:gər] "artisan", [ǰadu:gər] "magician", etc.

-p^{h}l- has been left out. The examples the dictionary gives for this cluster, e.g., trip^{h}la, dəp^{h}li:, etc. I render with an [f], i.e. [trifla], [ḍəfli:], etc. (ḍəfli: may also be rendered [ḍ^{h}əpli:]).[26]

I have included [-p^{h}r-] which occurs in very few words. It is possible that Hindi speakers have [-fr-] in the examples containing

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[-pʰr-]. I have not personally heard [-pʰr-] pronounced.

I have listed [-pʰṛ-] as in [pʰepʰṛa] "lungs", and [-ṭw-] as in [pəṭwari:] "village accountant". The dictionary lists these words also as pʰupʰus+ṛa and paṭ+wa, i.e. with a morpheme boundary intervening between the cluster. However I doubt that this morpheme boundary has any 'psychological reality' in a synchronic grammar of Hindi.

Among PA clusters, I have left out -kʰš-, -kʰf-, -zg-, -gb-, and -bkʰ-. -kʰf-, -gb-, and -bkʰ- occur only in a few high Urdu words (e.g., rəɣbət "desire"), and speakers who have these words in their vocabulary would probably have the correct [-ɣb-], [-xf-], and [-bx-]. -kʰɣš- occurs in a commonly used PA loan, the word for "tip, alms" but it is pronounced [bəkši:š] and not [bəkʰšiš]. High Urdu speakers would of course have [-xš-]. Thus [-kʰɣš-] as a cluster would not exist. The only example of -zg- that I know of involves a morpheme boundary medially, e.g., rozgar "daily wages, salary" which I think should be analyzed as /roz+gar/ "daily"+suffix.

A number of other clusters which have not been included are possible in a fast tempo of speaking, e.g., -ky-, -čʰw-, -ǰy-, -tʰy-, -by-, -zt- which occur in a fast delivery of words like wakəya (PA) "in reality", kəčʰua (N) "turtle", wiǰəya (S) "a proper name", həthiyar (N) "weapons", təbiyət (PA) "health", izazət (PA) "permission".

-ščʰ- which I have left out occurs in a few high Hindi words like nišč̌ʰəl "guileless", but it is possible that speakers who have this word are also aware that it is /niš+čʰəl/.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[-fl-] has been listed as occurring only in loans. However an alternate analysis could be to consider it as 'native' since some words of native origin have this cluster by hypercorrection, e.g., the example referred to earlier [trifla] "a medicine".

[-wṭ-] has been included. It occurs in some native words due to the ə-deletion rule in words like [bənawṭi:] "artificial" which is related to [bənawət] "make". However [-wṭ-] is not a cluster for all speakers in this word. A non-rigorous test I did (to be discussed in more detail in chapter 4) seemed to indicate that a number of speakers did not delete the /ə/ in such words, and for them [-wṭ-] would not be a cluster.

[-lḍ-] has been treated as a native cluster on the basis of one word [ḍalḍa] "vegetable shortening". My dictionary does not list this word and I do not know of its origin. It is in common use and is a brand name (i.e. it is similar to English "Kodak").

[-mč-] is listed as a native cluster and the example given is [čəmča] "spoon". This is from PA [čəm:əč] "spoon". I think in a synchronic grammar of Hindi there is no reason to treat either [čəmča] or [čəm:əč] as loans.

[-mš-] occurs in [šəmšan] "graveyard" which historically is from Sanskrit <u>šməšan</u>.

Words which are sometimes listed with <u>-ṽv-</u>, e.g., Sanskrit <u>səṃvad</u> "conversation", are pronounced as [səmwad] by most speakers and not as [sə̃ṽvad] or [sə̃w̃wad] and thus I have not listed [-ṽv-] or [-w̃w-].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I have not included -yp^{h}- and -yǰ- which occur in examples like [ǰayp^{h}əl] "nutmeg", and [payǰeb] "anklet"; I think a morpheme boundary can be posited in these words: /ǰay+p^{h}əl/, /pay+ǰeb/, etc. I am not sure if this analysis is best for a competence grammar of Hindi; more evidence is needed.

I have not included -wm-. This would be possible in Sanskrit loans such as /nəwəm+i:/ "nine"+suffix to which the ə-deletion rule would apply. However, in my experience this is rendered as [nɔmi:].

-wb^{h}- has also been omitted. It only occurs in one Sanskrit example b^{h}əwb^{h}u:ti: "name of a writer". This word can be analyzed morphemically into /b^{h}əw+b^{h}u:ti:/ "world+wealth", but I am not sure whether this analysis is justified in a synchronic grammar.

[-yn-] has been listed as a cluster occurring in Sanskrit loans. It is possible that it should have been listed as a native cluster since words such as the word for "mirror" can be transcribed as either [aena] or [ayna].

2.3.3. General statements about the clusters:

Initially, Perso-Arabic words do not seem to have clusters other than C+glide. Gumperz and Naim seem to allow certain other clusters too, for they say ".... Urdu ... [has] only a limited number of initial sequences of the type consonant plus l, r or y" (107). Since they don't give examples, it is difficult for me to tell what sort of words they had in mind for the C+liquid clusters. Since they mention that their informants were educated, in fact, they were Indians already in the U.S.A. (Gumperz, personal communication), it is possible that they were treating common English loans like [glas]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

"glass", etc. as examples of C+liquid.

Initially, native words are also highly constrained; only clusters of C+liquid or glide seem to be tolerated.

As I mentioned in chapter 1, I am treating all so-called geminate clusters like rr, mm, pp, etc., as long consonants; therefore they have not been listed in these Tables.

Medially as well as finally, what is listed by some writers as ny is phonetically [ñy]. E.g. [kəñya] "girl". Therefore I have treated this cluster as homorganic-nasal+y.

Medially and finally, the various clusters usually listed as h+consonant by writers, are not considered here to be clusters in Standard Hindi. Thus Gumperz and Naim list -hr, -hm, -hw, etc., for Urdu; examples are šəhr "city", wəhm "superstition", səhw "mistake",[27] etc. Srivastava (n.d.) lists the word for city as [šəhr]. However, in my view, there is quite a pronounced vocalic release after the [h]. The word for 'city' is a very common word in the vocabulary of Hindi speakers, and the Standard Hindi speakers that I know of render this [šəhɚ]. In Devanagari also this loan is written without a cluster: शहर šəhər.[28] I doubt that a non-high Urdu speaker even realizes that in Urdu this word has a final cluster. Even many Urdu speakers pronounce these words without a cluster, at least in casual speech.

A number of the h+consonant clusters that are listed in the chart for medial clusters, are only clusters for Urdu speakers (e.g., -hk-, -ht-, -hb-, -hd-, -hr-, etc.); Hindi speakers pronounce these without the clusters.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Both medially and finally, what is written in Sanskrit loans as hm and hn (e.g., brəhma "a god", čihn "sign") are pronounced as [mh], [nh], etc., and thus I have treated them as such.

I have not included medial -gh- and medial and final ṛh. These clusters would be possible in Hindi words like that for a particular month, written अगहन əgəhən, and the name of a place मगहि məgəhi: "pertaining to məgəh", and the name for the hibiscus flower written गुड़हल gurəhəl. I mentioned in chapter 1 that in the case of ṛh in some cases (as in the word for "hibiscus") it should perhaps be treated as a cluster, and in other cases (as in [pi:ṛha] "stool") it does not seem to matter whether it is treated as a cluster or a unit. The discussion of these clusters would also involve the interesting problem of Vh in Hindi which is beyond the scope of this study. It is necessary to gather more data on what the variations in the pronunciation of these words are. However, lacking this information, I have chosen not to include them. Medially, if ṛh had been included as a cluster, then the three-consonant cluster -ṛhk- would also have to be listed (e.g., [luṛhka] "toppled"). This is the only three-consonant cluster with ṛ+h as first and second consonants. In final position, after ṛ+h, no other consonant occurs. I am also not certain at this point whether or not in running speech we get [-ṛh] morpheme-finally at all or whether deaspiration takes place giving just [-ṛ] in words such as गढ़ gəṛh "fort".

I have treated [rh] [lh] [nh] [mh] as medial and final clusters but an alternative is to treat these as units, i.e. [r^h], [l^h], [n^h], [m^h], etc. There are pros and cons to both analyses, and the issue is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

still an open one. If we treat them as units these would be segments with a rather limited distribution, since they do not occur initially, and it is not clear whether some of them are pronounced with the aspiration or not when final (i.e. whether we have [-r^{h}] or [-r]). In the case of aspirate or breathy voiced stops, Narang and Becker (1971) suggest that from the point of view of the ə-deletion rule (I discuss this rule in detail in chapter 4) they are regarded as units (in that it deletes the [ə]in words such as [bɛṭ^{h}ək] "meeting" when the appropriate ending is added) and therefore should be treated as units. Narang and Becker do not discuss [nh], [lh], etc. In the case of [lh] we have the word [dulhən] "bride", to which, if we add the inflectional ending [õ] "plural(oblique)", we get [dulhənõ] and not *[dulhnõ], i.e. the ə-deletion rule does not apply, which might suggest that [lh], etc., should be treated as clusters. However, the use of the ə-deletion rule to decide whether certain segments should be treated as units or clusters has certain problems: as I show in chapter 4 in a number of cases it applies even if there is a cluster preceding the [ə] , i.e. in the environment VCC___CV. Therefore the fact that it does not apply to the [ə] in [dulhən] (but does apply to the [ə] in [bɛṭ^{h}ək]) is not a clear criterion by which to decide whether or not [lh] (or any aspirate) is a cluster or a unit. However, if a cluster follows, the [ə] is never deleted. Therefore when the [ə] is deleted in /səməj^{h}+a/ "understood" yielding [səmj^{h}a], this could perhaps be taken as support for the view that aspirates are units. The matter deserves further investigation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I have not listed clusters with $\widetilde{\underline{r}}$ in the cluster charts. Initially $\widetilde{\underline{r}}$ does not occur. Medially and finally it occurs only in very specialized speech -- in Sanskrit loans used by high Hindi speakers in formal style of speech. The restrictions on clusters with $\widetilde{\underline{r}}$ are as follows: Medially after $\widetilde{\underline{r}}$, only vowels occur, and before $\widetilde{\underline{r}}$ only $\underline{r}$ and $\underline{m}$. Finally after $\widetilde{\underline{r}}$, only glides occur; and before $\widetilde{\underline{r}}$, only $\underline{g}$, $\underline{r}$, and $\underline{\text{ɣ}}$.

The first consonant in clusters like [-čṭ-, -č^{h}t-, -b^{h}k-, -ǰg-, -d^{h}k-, -ṭ^{h}k-, -g^{h}ṭ-, -k^{h}ṭ-, -ǰ^{h}k-, -d^{h}ṛ-, -b^{h}ṛ-, -t^{h}l-, -k^{h}r-, -k^{h}l-, -k^{h}ṛ-], etc., has quite an audible release (that is, if deaspiration has not taken place). The common factor in these clusters is that the first consonant is an aspirate, breathy voiced stop, or an affricate. Treating the above as clusters can be questioned. I think that before we can conclusively decide one way or the other, more data from the spoken language is needed. All the above clusters involve very few examples (most have just one or two), but these are very commonly used words.

As mentioned earlier none of the authors who discuss medial clusters distinguish between morpheme-medial clusters and those that occur only across morpheme boundaries. The following are some of the clusters which are generally listed as "medial" which I found to occur only across morpheme boundaries: -gč-, -kp-, -kǰ^{h}-, -čp-, -ṭk^{h}-, -ṭp-, -t^{h}k-, -d^{h}k^{h}-, -ṭs-, -hk^{h}-, -rč^{h}-, -ṛl-, -lṛ-, -ŋg^{h}n-, -r^{h}n-, -nṛ-, -ṭm-, -ṛg-, -ǰp-, -tp^{h}-, -td-, -rč-, -my-, -rb-, etc. This by no means exhausts all clusters occurring across morpheme boundaries. These are the ones usually listed by others as medial

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

clusters (e.g., Arun). Some examples of these are given below in table 13.

Table 13

-kp-	[čəkpəka]	čək+pək+a	"startled"
-kǰ^{h}-	[ǰ^{h}əkǰ^{h}or]	ǰ^{h}ək+ǰ^{h}or	"to shake something violently"
-čp-	[pəčpən]	pəč+pən	"fifty five"
-ṭk^{h}-	[k^{h}əṭk^{h}əṭa]	k^{h}əṭ+k^{h}əṭ+a	"knock"
-ṭp-	[k^{h}əṭpəṭ]	k^{h}əṭ+pəṭ	"noise"
-t^{h}k-	[həṭ^{h}kəṛi:]	həṭ^{h}+kəṛi:	"handcuffs"
-d^{h}k^{h}-	[əd^{h}k^{h}ila]	əd^{h}+k^{h}ila	"half-opened (flower)"
-ṛl-	[laṛla]	laṛ+la	"darling"
-ŋg^{h}n-	[suŋg^{h}ni:]	su:ŋg^{h}+ni:	"little box containing perfume for smelling"
-ṛ^{h}n-	[oṛ^{h}ni:]	oṛ^{h}+ni:	"scarf"
-ṭm-	[k^{h}əṭməl]	k^{h}aṭ+məl	"bed bugs"
-td-	[potdar]	pot+dar	"person dealing in small beads"
-ǰp-	[raǰpu:t]	raǰ+pu:t	"a warrior class"
-ṛg-	[gəṛgəṛ]	gəṛ+gəṛ	"glug glug"
-ṛč-	[g^{h}uṛčəṛ^{h}i:]	g^{h}uṛ+čəṛ^{h}i:	"horse climbing ceremony"
-ṛb-	[gəṛbəṛ]	gəṛ+bəṛ	"confusion"

In final position we find that native words permit very few clusters. The most common is homorganic nasal+stop, but I also consider [-st] to be a native cluster (as mentioned in section 2 of this chapter). Historically most words involving [-st] are loans,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

but these words have been in common use for quite some time and the cluster is not simplified even by uneducated speakers; therefore marking words with [-st] as loans would not be justified. In this respect my analysis seems to agree with Gumperz and Naim's, who say that in uneducated speech "... finally there are only clusters of nasal plus homorganic stop or sibilant plus stop" (112). My analysis differs slightly in that I am only allowing s+voiceless stop; I rule out s+voiced stop. Although I am allowing š+voiceless stop, it is possible that this should be ruled out; a test should be performed to see if speakers simplify this final cluster or not. (The test reported in the previous section does provide evidence for treating [-st] as native. Thus although my subjects treated common loan words like kism "kind" as [kisəm], and some of them seemed to give evidence of having underlying /kisəm/, for words like dəst "diarrhea" they had [dəst] and gave no evidence of having a /ə/ in the underlying form, i.e. */dəsət/.)

In medial position the three-consonant clusters occurring in native words involve a homorganic nasal as the first member of the cluster. In final position three-consonant clusters occur only in Sanskrit loans. Medial and final four-consonant clusters should be ruled out -- there is only one example of each: [wərtsy] "alveolar" and [b^{h}ərtsna] "censure", both extremely rare Sanskrit loans.

2.3.4. Let me now return to a point mentioned earlier. A particular cluster may be pronounced correctly by a speaker in a Sanskrit loan, but the same cluster simplified by that speaker in a PA loan in terms of the native or "common core" system, or vice-versa. For example,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

native words do not have [-sm]. A person who draws on Sanskrit for his learned vocabulary might correctly pronounce the Sanskrit loan for "ashes" as [b^həsm] but simplify the PA loan kism "kinds" to [kisəm]. On the other hand a speaker who draws on PA vocabulary will have [b^həsəm], i.e. break up the final cluster, but will render the word for "kinds" as [kism], i.e. retain the cluster. Morpheme-medially, they will both retain the cluster (if we add an inflectional ending or a suffix). Thus, if the environment of the ə-deletion rule mentioned in section 2 is met, we get [b^həsmõ] and [kismõ], i.e. morpheme-medially the [sm] cluster is permitted. These facts were noticed by Kelkar (1968 pg. 29).

How should these facts be represented in a grammar? There are a number of alternatives:

1. One possibility is to enter words like kism, etc. as /kisəm/ and add the feature +PA, then have a rule that deletes the /ə/ in the environment of [+formal][+high Urdu] style or something (whatever the sociolinguistic features should be). This deletion rule is added by the speaker only when he is exposed to high Urdu, of course. Since the environment of this ə-deletion rule would be socially conditioned, I will mark it as [ə-deletion $rule_{Socio}$] to distinguish it from the ə-deletion rule discussed earlier in this chapter. However, since there are also PA loans like [kəsəm] "swear" where the [ə] is not deleted (even by high Urdu speakers), the morphemes where the [ə] is deleted in high Urdu, e.g., kism, etc., would have to be marked as [+ə-deletion $rule_{Socio}$]; otherwise we would incorrectly get *[kəsm] "swear".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. Another alternative is to say that the child (of a non-high Urdu-speaking family) first learns /kisəm/ and later, when (and if) he is exposed to high Urdu, he enters along with /kisəm/ the variant form /kism/, which shares all the semantic and grammatical features of /kisəm/ but also has the feature [+formal][+high Urdu] style, or such.

3. A third alternative is that the entry in the lexicon should be /kism/ and then a ə-insertion rule would insert the /ə/ in the environment of $\overset{+PA}{}$-high-Urdu, etc., to get [kisəm] in the non-high Urdu speakers case. This is, in fact, what Srivastava (1969:918) suggests. [kəsəm] could be entered simply as /kəsəm/, so of course the ə-insertion rule would not apply.

4. There is also a fourth alternative, namely that the speaker first enters /kisəm/ in his lexicon, and later when he is exposed to the prestigious [kism] he replaces /kisəm/ by /kism/ and then optionally applies the ə-insertion rule.

From a descriptive point of view all of the above analyses are adequate. However, if we are interested in a grammar that reflects competence, we will have to reject analysis 3 because it makes a rather strange claim, namely that a child (whose parents draw on Sanskrit for high vocabulary and not on Perso-Arabic), who hears those around him pronounce the word for "kinds" as [kisəm], somehow hypothesizes an underlying form /kism/ to which he applies a ə-insertion rule for most of his childhood, that is, until he is exposed to the prestigious [kism], and from then on he adds the feature -high Urdu to his ə-insertion rule. This seems highly unlikely.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ultimately, of course, which of these analyses is chosen depends on what evidence can be brought forth to support its "psychological reality". Data from hypercorrection seems to rule out analyses 3 and 4, and seems to provide more evidence for 1 than for 2, however it is not incompatible with 2. Thus I have an open mind to both 1 and 2. The data are as follows:

There are forms of PA which retain the /ə/ even in prestigious speech, e.g., [kəsəm]; hence, when a non-high Urdu speaker learns that the prestigious pronunciation of [kisəm] is [kism], he has to learn which forms delete the /ə/ and which don't. Occasionally, therefore, he "hypercorrects" and deletes the /ə/ of a form where the /ə/ should be retained, i.e. producing [kəsm] "swear" instead of the correct [kəsəm]. This seems to suggest that he is using a rule productively, and I propose that this is the ə-deletion rule$_{Socio}$; however, he hasn't tagged his morphemes 'correctly' as to whether the rule applies to them or not. Of course he learns this rule only after he is exposed to the prestigious [kism] pronunciation; before that he would have no need for this rule -- in fact, some speakers may never acquire it.

For analysis 2 we would have to say that he incorrectly adds to morphemes like /kəsəm/ an alternate /kəsm/ by hypercorrection. But the only way he could generate the alternate /kəsm/ which he has never heard before is by using a rule -- the ə-deletion rule. Thus this analysis seems functionally equivalent to analysis 1.

2.3.5.1. As mentioned earlier, homorganic nasal+stops are among the very few types of clusters that native words tolerate in medial and final position; moreover, these clusters are quite common. They have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

not been listed on the charts, but examples have been given above.

At this point some discussion of the problems and various treatments of homorganic nasals in Hindi is in order.

On the basis of words such as those given in table 14 (and many more could be added), we could make the following generalization at the phonetic level (I will shortly modify this): Generalization A: in Hindi we only get homorganic nasals before stops.[29] This generalization is also made by Narang and Becker (1971).[30]

Table 14[31]

[pʰəŋki:]	"handful"
[kənd^{h}a]	"shoulder"
[gend]	"ball"
[tamba]	"copper"
[ɟʰəɳ̣da]	"flag"
[gəñɟa]	"bald"
[gəŋga]	"Ganges"
[si:ŋg]	"horn"
[gu:ñɟ]	"sounded, echoed"

In final position this generalization holds at both the systematic phonemic level and at the systematic phonetic level. There are a few Perso-Arabic loans with non-homorganic nasals, e.g., [simt] "direction", however these are rare and can be treated as exceptions. Moreover, these words would occur only in the vocabulary of high Urdu speakers, and it is possible that generalization A should not be made in a grammar of these speakers.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In medial position, however, generalization A seems to hold only for native vocabulary. There are a number of examples of loanwords with non-homorganic nasals before stops, and many of these are used commonly by native speakers of standard Hindi:

Table 15

[nəmda]	(P)	"a rug"
[bəramda]	(P)	"veranda"
[umda]	(P)	"good, best"
[gomti:]	(S)	"name of a river"
[imtihan]	(A)	"examination"
[mumkin]	(A)	"possible"
[təmga]	(Turkish)	"medal"
[nankin]	(Chinese)	"Nankin"
[menka]	(S)	"a name"
[tənk^{h}a]	(P)	"pay, salary"
[inkar]	(A)	"denial"
[inkəm]	(E)	"income"
[munkir]	(A)	"one who denies"

(Even though [inkar] is historically a loanword, in a synchronic grammar of Hindi it should be treated as native, thus it belongs in table 18.)

It also seems to be the case that, morpheme-medially, this generalization does not hold at the phonetic level since we find examples such as those in table 16:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 16

[čimṭa]	"tongs"
[səmɟ^{h}a]	"understood"
[d^{h}əmkana]	"to threaten"
[čəmki:la]	"bright"
[sənki:]	"whimsical"
[č^{h}ənka]	"clanked"
[ɟ^{h}umka]	"a kind of earring"
[k^{h}ənkana]	"to clang"

The nonhomorganic nasals in these words have been brought about by the application of the ə-deletion rule referred to in section 2 of this chapter. Thus we have:

Table 17

[čimṭa]	čimət+a
[səmɟ^{h}a]	səməɟ^{h}+a
[sənki:]	sənək+i:
[čəmki:la]	čəmək+i:la

To the list of words in table 17 some non-native words can also be added which show an alternation with /ə/ -- at least in the pronunciation of non-high Urdu speakers. E.g.,

(9)	[k^{h}ušamdi:]	(P)	k^{h}ušaməd+i:	"flatterer"
	[amdəni:]	(P)	aməd+əni:	"income"
	[ki:mti:]	(A)	ki:mət+i:	"expensive"

There are also some native words (though not very many) where we find a nonhomorganic nasal at the phonetic level, where no

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

alternating forms exist with a /ə/ which could be said to be deleted by the ə-deletion rule:

Table 18

[tinka]	"bit of dry grass"
[mənka]	"bead"
[kənk^{h}i:]	"sideways glance"
[kunba]	"family"
[čingari:]	"spark"

Before giving my proposals on the treatment of exceptions listed under table 18 let me first briefly review the analysis of a few other writers.

Kelkar (1968) gives considerable data on this problem. For a discussion of his treatment see Srivastava (1969).

Srivastava suggests that the generalization about homorganic nasals be incorporated in a P-rule. He suggests that the P-rule also applies across morpheme boundaries and that the rule functions both as a MS-rule (when it operates within morphemes) and as a P-rule (when it operates across morpheme boundaries). The only example of the latter that he gives is [əheŋkar] "pride", which he says is /əhəm+kar/.

For words such as those listed in table 16, Srivastava also points out that alternating forms with /ə/ exist where a /ə/ breaks up the non-homorganic nasal+stop cluster. He does not restrict the generalization to homorganic nasals before stops, and accounts for apparent exceptions like the Sanskrit loans [činməy] "form of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

intellect", [waŋməy] "literature", [diŋnad] "sound reverberating in all directions", etc., by claiming that they are really /wak+məy/, /čit+məy/, etc., to which Sanskrit Sandhi rules apply to give [waŋməy] and so on.

Narang and Becker treat what I have called generalization A by a P-rule that is restricted to applying within morphemes. They give no reason for treating this phenomenon by a P-rule rather than by a redundancy rule. They make the above generalization not only for native vocabulary, but presumably for the entire vocabulary.

They note that the words given above in table 16 are not exceptions for the reasons given above. [waŋməy], [diŋnad], etc., discussed by Srivastava, are not relevant to their generalization because they restrict their generalization to applying only before non-uvular stops. Of the exceptions I have listed, they only list [imtihan] "examination" and [simt] "direction" as exceptions. They do not consider [inkar] or [munkir] (given in table 15) as exceptions because they posit underlying forms for these such as /nAkArA/, etc. (663), i.e. having a vowel between the nasal and the consonant.

They mention that [mənka] and [tinka] (given in table 18) are also not exceptions for them. They posit a morpheme boundary between the nonhomorganic nasal and the stop, i.e. /mən+ka/ and /tin+ka/ (659).

They do not consider [kunba] (in table 18) to be a real exception because they claim that the regular form is [kumba], and only Urdu speakers pronounce it [kunba] because they confuse the first part kun with the Arabic kun from a line of the Koran; and the last part ba with the Persian prefix ba as in ba+ədəb "with politeness"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(658). They do not mention the other words I have listed in tables 15 and 18.

I disagree for the most part with both Srivastava's and Narang and Becker's treatment. Both Srivastava and Narang and Becker fail to realize that, first, a difference has to be made in the treatment of morpheme-medial and morpheme-final homorganic nasals. Second, morpheme-medially the generalization, that is, that nasals are homorganic with following stops, applies only to the lexical level, not the phonetic.

Narang and Becker state the generalization as "within a morpheme, a nasal consonant must agree in point of articulation with a following non-uvular stop" (654) and indicate that this applies "... at the level of systematic phonetics" (653). They capture this generalization by a P-rule (their P-rule 3). They mention that words such as those in table 16 ([čiṃta] etc.) are not really exceptions, because "the non-homorganic nasal clusters of these words do not exist at the systematic-phonemic level of representation, but arise only through the operation of the rule which we have labelled P1 [the ə-deletion rule]" (662). This is true, but I fail to see what this has to do with eliminating exceptions to their generalization which is supposed to be true at the phonetic (not systematic-phonemic) level. Since at the systematic phonetic level these non-homorganic nasal+stop clusters do exist in vast numbers, violating the authors' generalization (above), how can they still claim that this generalization holds at the systematic phonetic level? The fact is that the above generalization can only be stated for the systematic

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

phonemic level, not the systematic phonetic.

Further, I don't agree with Srivastava that the rule applies across morpheme boundaries. Words such as [əhəŋkar] "egoism, pride" are Sanskrit loans that are not commonly used by native speakers of Hindi. It is doubtful whether native speakers "know" [əhəŋkar] to be derived from /əhəm+kar/ any more than most native English speakers realize that the name "Thompson" is derived from "Thom(as)+son". Furthermore, there are many native words where non-homorganic nasals occur across morpheme boundaries, e.g., [ʄəmta] from /ʄəm+ta/ "freezing", [somwar] from /som+war/ "Monday", etc.

I also disagree with Srivastava's treatment of the Sanskrit loans [waŋməy] etc. These occur only in learned vocabulary, and Srivastava gives no independent evidence for the reality of such abstract underlying forms as /wak+məy/ for [waŋməy] nor for the reality in modern Hindi of the Sanskrit Sandhi rules which would be required in order to derive the correct phonetic output. One might equally well suggest that Hindi speakers who have borrowed electric and electricity from English "know" the (proposed) velar-softening rule. This is highly unlikely.

I think Narang and Becker's treatment suffers in general from their not recognizing that Hindi vocabulary has to be divided into [±native] for generalizations to hold and for the rules to make sense. They say "... the validity of these rules would be open to serious doubt if the number of exceptions were very large" (661) but later "Luckily, the exceptions to our rules are quite limited". However as they have stated it, the exceptions to their P-3 would be quite

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

numerous, because there are quite a few [-native] words with non-homorganic nasals, a small sample of which is given in table 15.

Their treatment of [inkar] and [munkir], etc., is also very suspect. The underlying forms /nAkArA/ and so on, that the authors posit reflect the Arabic origin of the roots. Since [inkar] is a commonly used word among Hindi speakers who do not know "high-brow" Urdu and would thus not know of its Arabic derivation, the underlying form given by the authors has no 'psychological reality' for these speakers. Similarly many Urdu speakers who are not learned in Arabic would perhaps not 'know' of the Arabic history of this word either.

Further, positing a morpheme boundary in words like [tinka] (/tin+ka/) is a diacritic use of the boundary, since there is no justification for setting up a morpheme boundary -- as the authors recognize themselves (cf. their footnote 14). If there is any principled reason for using the boundary marker as a diacritic feature, instead of simply listing these "uncooperative" forms as exceptions, the authors have not revealed it.

For [kunba] "family", Narang and Becker, as mentioned earlier, claim that the word is actually [kumba] but Urdu speakers pronounce it [kunba], because they confuse the first part kun with the Arabic word kun and the last syllable ba with the Persian prefix ba; thus, since these Urdu speakers have /kun+ba/, the nasal is not assimilated to the following stop due to the morpheme boundary.

Firstly it would seem rather unlikely that speakers would take the prefix ba and use it as a suffix. Secondly, the Arabic kun has no semantic features in common with [kunba] "family". The dictionary

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Varma 1958) lists this word as kunba and gives the etymology that it is derived from Sanskrit kuṭumb. Now perhaps most speakers are unaware of this word's true origin, but, on the other hand, there is no reason to expect any non-Urdu speaker to make up this fanciful PA etymology for this word -- and it is not only Urdu speakers who say [kunba], but many Hindi speakers too (after all, this is the pronunciation given by the dictionary). Moreover native speakers of Hindi give evidence of "knowing" that [kunba] is an exception, because a number of speakers sometimes regularize it to [kumba].

It is not clear how Narang and Becker would treat the words I list in tables 15 and 18. They could either treat these as exceptions or they could claim they are not exceptions by either positing a morpheme boundary in them or by positing abstract underlying forms which reflect the historical source of the words. As mentioned above, they posit a morpheme boundary in the case of [tinka] and [mənka]; however, in the case of [kənk^{h}i:] "side-ways glance" this treatment will not work, as I will shortly show.

Moreover using diacritic features to make certain morphemes appear to be non-exceptions hides important facts. Speakers do sometimes regularize some of the above-mentioned words, thus giving evidence of their exceptional status. Thus I have heard speakers say [kumba] [iŋkar] and [čiŋgari:].

How then should the exceptions under table 18 be treated? There are three possibilities:

(a) List them all as exceptions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(b) Treat them all as if they had an underlying /ə/ between the non-homorganic nasal and the stop and then delete the /ə/ by the ə-deletion rule.

(c) List some of them as exceptions, and give the others the underlying /ə/ treatment.

I will choose alternative (c) here. Both alternative (a) and (b) would work descriptively, i.e. they would account for the existing forms; however, behavioral evidence supports (c).

Since speakers do sometimes render [inkar] "deny", [kunba] "family", and [čingari:] "spark" with homorganic nasals, I will list these as exceptions. As mentioned above, I take the fact of occasional regularization as evidence of their exceptional nature.

Since [mənka] [tinka] and [kənk^{h}i:] are never rendered with homorganic nasals (at least as far as I know), I will treat these differently. I will posit an underlying /ə/ for them at the systematic phonemic level, which /ə/ will later be deleted by the ə-deletion rule to give the correct phonetic forms. Let us consider the case of [kənk^{h}i:] "sideways glance" in detail. Here, positing a morpheme boundary will not work. If a morpheme boundary is to be included it would have to be as /kən+ək^{h}i:/ -- the word is made up of two morphemes meaning "ear+eye", i.e. literally "eyes to ears" ("sideways glance"). A morpheme break /kən+k^{h}i:/ would not be justified since there is no morpheme k^{h}i:. The dictionary lists the word as derived from Sanskrit kər̥̄+əksi, which, through historical changes would give Hindi /kən+ək^{h}i:/. If it was /kən+ək^{h}i:/ and if this morpheme boundary had some reality, the [ə] of ək^{h}i: would be protected from

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the ə-deletion rule because it is morpheme-initial, and thus we should get [kənək^{h}i:] (i.e. without the non-homorganic nasal+stop cluster). However the fact that phonetically we have [kənk^{h}i:] shows that in modern Hindi the morpheme boundary has been lost, yielding, from the underlying /kənək^{h}i:/, the phonetic [kənk^{h}i:] by the ə-deletion rule (this analysis also lends support to the discussion on "abstractness" under 2.2.). One might ask, however, how a child forms an underlying form with a /ə/, if there are no alternating forms for this word which contain a [ə]. I discussed this question earlier in 2.2., where I concluded that the evidence seems to show that children do indeed hypothesize lexical forms they have never heard before. Moreover the fact that the nasal in [kənk^{h}i:] is never rendered as homorganic provides evidence that speakers do somehow hypothesize an underlying /ə/ between the nasal and the following stop, whereas other words having non-homorganic nasal+stop clusters do have the nasal made homorganic, e.g., [inkar]~[iŋkar] "denial". Also even though there are no alternating forms with a [ə] for words such as [kənk^{h}i:], there are alternants with a [ə] in the case of words such as [čimṭa] (those forms in table 16). It seems reasonable that a child could use this as a basis for positing underlying forms such as /kənək^{h}i:/.

How should generalization A be revealed in a grammar? There are at least the following possibilities:

(a) We capture it by a P-rule. As mentioned above this is what Narang and Becker propose. They of course claim that the generalization holds at the phonetic level; however, as I have already pointed out, the existence of words such as those in table 16 ([čimṭa]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

"tongs", etc.) argues against this. Thus if we wanted to reveal the generalization by a P-rule we would have to modify the generalization to: nasals are homorganic with following stops unless there was a /ə/ between the nasal and the stop at some point in the derivational history of the morpheme in question. Current generative phonology of course assumes that morphemes do not retain their derivational history. Thus the fact that at the phonetic level we do get non-homorganic nasal+stop due to the application of the ə-deletion rule, and the fact that the P-rule would have to be restricted to within morphemes only argues against this analysis. On the plus side of this analysis is the fact that if we do not insist that morphemes enter the phonological component fully specified, i.e. if we allow archi-segments (this would mean a modification in the Stanley model of MSC theory), then we need not set up /ŋ/ /ñ/ and /ɳ/ as segments at the underlying level; they can be derived by the P-rule.

(b) We state generalization A as applying only to the lexical level, and reveal it by a SQC. In that case we have to set up /ŋ ñ ɳ/ as underlying segments of a special sort -- they are segments that occur only before homorganic consonants -- since SQC's under current MS theory do not derive anything, they simply state constraints. The only time these nasals occur independently is in reciting the Devanagari alphabet -- the alphabet, following Sanskrit, has separate symbols for all five nasals: [ŋ] [m] [n] [ñ] and [ɳ]. Since it is a syllabary, in recitation one says "kə, k^hə, gə, g^hə, ŋə," etc. The fact that both [ñ] and [m] appear before palatal fricatives (see 2.3.5.2.) might also provide support for treating [ñ]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

as some sort of underlying segment.

I am tentatively adopting analysis (b) and treating the generalization by SQC 9 (to be given shortly under 2.3.7.).

A third analysis, which is an intriguing one, is to allow SQC's to be active in derivation, i.e. to allow them to fill in unspecified features. We could then have an archi-segment /N/ with SQC 9 specifying that the point of articulation for the nasal is the same as the following stop. If we also stated that the output of the SQC's is not necessarily systematic phonemes, then we need not consider [ŋ] [ñ] and [n̪] to be segments at the underlying level. The ramifications of this proposal need to be investigated more. This seems to be closer in spirit with Halle's morpheme structure rule theory than Stanley's morpheme structure condition theory.

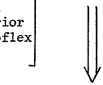
Since I am tentatively adopting analysis (b), two more SQC's need to be posited to reveal the special nature of the nasals [ŋ] [ñ] and [n̪]:

(10) if: $\begin{bmatrix} -\text{syll} \\ +\text{nasal} \\ +\text{high} \\ -\text{low} \\ -\text{back} \end{bmatrix}$ $[+\text{segment}]$

then: $\begin{bmatrix} -\text{syll} \\ +\text{high} \\ -\text{low} \\ -\text{back} \end{bmatrix}$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(11) if: $\begin{bmatrix} -\text{syll} \\ +\text{nasal} \\ -\text{anterior} \\ \alpha\text{retroflex} \\ \beta\text{back} \end{bmatrix}$ [+segment]



then: $\begin{bmatrix} -\text{sonorant} \\ -\text{continuant} \\ \alpha\text{retroflex} \\ \beta\text{back} \end{bmatrix}$

The first SQC states that after a palatal nasal the following segment must be any palatal consonant. The second SQC states that after [ɲ] or [ɲ̣] the following segment must be a homorganic stop.

To summarize: I propose treating generalization A by a sequential constraint (SQC 9 given under 2.3.7.) rather than a P-rule. The fact that both homorganic and non-homorganic nasal+stop clusters appear at the phonetic level is strong justification for not treating the phenomenon by a P-rule. I will also limit the SQC to within morpheme boundaries. Moreover, morpheme-medially I restrict the SQC to applying only at the lexical level and not the phonetic. I also posit two additional SQC's which state that velar and retroflex nasals occur only before homorganic stops, and the palatal nasal only before palatal consonants.

As for the exceptions listed in table 18, I am proposing that some of them (e.g., [inkar] "denial") be listed as exceptions (i.e. entered as /inkar/, etc.) and others (e.g., [kənk^{h}i:]) be listed with an underlying /ə/ (i.e. as /kənək^{h}i:/) to which the ə-deletion rule would apply to yield the correct phonetic output (i.e. [kənk^{h}i:]). 2.3.5.2. I have discussed above the issue of homorganic nasals before stops. However, there still remains the problem of nasal+y. This

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

occurs in Sanskrit loans such as kənya "girl", sənyog "chance", sənyukta "a proper name", ənу "other", etc., which are phonetically rendered with palatal nasals, e.g., [kəñya] and so on. Similarly before š, even though the orthography gives n, phonetically there is the palatal nasal, e.g., [əñš] "portion", [məñša] "desire", etc. SQC 9 (referred to above) will not apply to this palatal nasal since it applies only to nasals before stops. But it is not possible to write another SQC that would state that all nasals before palatal consonants are palatal nasals because of the examples in table 19:

Table 19

[sɔmy]	(S)	"calm"
[sɔmya]	(S)	"a proper name"
[rəmy]	(S)	"pretty"
[šəmšer]	(PA)	"a proper name"
[rəmzan]	(PA)	"a Muslim month"
[šəmšan]	(N)	"graveyard"

We thus need a SQC which states that before a palatal consonant we can either get a palatal nasal or a bilabial but not a dental retroflex or velar:

(12) if: $\begin{bmatrix}-\text{syll}\\+\text{nasal}\end{bmatrix}$ $\begin{bmatrix}-\text{syll}\\+\text{high}\\-\text{back}\end{bmatrix}$

then: $\left\{\begin{bmatrix}-\text{anterior}\\-\text{retroflex}\\+\text{coronal}\end{bmatrix}, \begin{bmatrix}+\text{anterior}\\-\text{coronal}\end{bmatrix}\right\}$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

My analysis of nasal+fricative disagrees with that of Narang and Beckers'. They state that: "... within a morpheme, the only nasal consonant which can occur before ... any fricative is [n]." (654) As examples from table 19 show we do find bilabial nasals before fricatives, therefore their statement is not correct.

They also do not indicate how they will treat nasal+y in words such as [kəñya], but this has been accomplished in the analysis given above.

2.3.6. Let me now state some of the sequential constraints of Hindi. The test reported in Appendix 2 seems to provide evidence of the psychological reality for a number of the initial constraints. Similar tests should be conducted for the medial and final constraints also. All the SQC's except for no.9 apply at both the lexical level and the systematic phonetic level; SQC 9 applies to only the lexical level when it applies morpheme medially (in final position it applies to both levels). I have done no investigation to see if any of the SQC's apply intermorphemically also, therefore at this point I will restrict them all to applying intramorphemically. With the exception of SQC 9 which also applies to three-consonant clusters, only two-consonant cluster constraints are given. SQC's pertaining to three-consonant clusters need further study.

I have stated the SQC's in terms of morpheme-initial, medial and final, rather than in terms of syllables, for the following reasons:

First, the problem of defining a syllable has always plagued linguistic theory. While it is true that native speakers can in most cases consistently say how many syllables a given utterance has,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

demarcating the syllable is quite another problem. Native speakers differ quite a bit in this respect. Kelkar (1968), Mehrotra (1959), and Upraiti (1964) have, among others, tried to demarcate the syllable in Hindi. Although their analyses are similar in some respects, they are different in others. Some writers have suggested that Hindi [ašrəm] "hermitage" be divided in a-šrəm, but it seems to me the break aš-rəm is equally possible.

Secondly, it is not obvious to me that by stating the Hindi initial and final constraints as syllable constraints any significant saving would be achieved in the statement of morpheme-medial constraints. Almost all the consonants occur initially and finally, but medially there are obviously many co-occurrence restrictions on which consonants can occur in clusters.

On the basis of monosyllabic morphemes the structure of the (native) syllable in Hindi can be said to be:

(13) $C_o^2VC_o^2$ Condition: $\sim$ CCVCC

Examples:

V	[a]	"come"
CVC	[kəl]	"tomorrow, yesterday"
VC	[ag]	"fire"
CV	[lɔ]	"flame"
CVCC	[b^haŋg]	"hemp"
CCVC	[pyas]	"thirst"
CCV	[kyõ]	"why"
VCC	[əst]	"set"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

However if loan words are included (mainly Sanskrit loans), the above structure has to be modified to:

(14) $C_o^{\ 3}VC_o^{\ 3}$ Condition: $\sim$CCCVCCC

Examples:[32]

CCVCC	[skənd^h]	(S)	"shoulder"
VCCC	[əstr]	(S)	"weapon"
CCCV	[stri:]	(S)	"woman"
CVCCC	[šəštr]	(S)	"weapon"
CCCVC	[str$\tilde{\text{ɛ}}$]	(S)	"feminine"
CCVCCC	[swast^{h}y]	(S)	"health"
CCCVCC	[spriṣ̌t]	(S)	"touched"

Four-consonant clusters occur in only one example, Sanskrit [wertsy] "alveolar", and in my view should be ruled out.

The various consonant combinations that are possible in sequences have already been given in tables 10, 11, and 12.

The majority of Hindi morphemes consist of up to three syllables, e.g., [a] "come", [sona] "gold", [pəheli:] "puzzle", etc. In morphemes like [kəčɛhɛri:] "court" it seems that four syllables are possible, but it is also possible that phonetically this word is [kəc^{h}ɛri:] (at least for some speakers); this again involves the problem of Vh (referred to earlier) which I am not treating. Five syllables occur only in derived or inflected words.

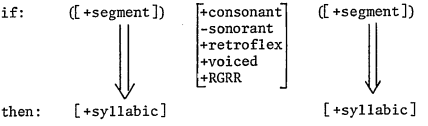
2.3.7. Sequential constraints of consonants:

Stanley states that there are three kinds of constraints: if-then, negative, and positive (for details see Stanley 1967). The following are some tentative general constraints that can be stated

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

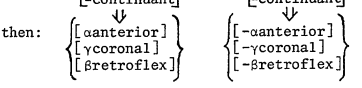
for consonants:

1. [ḍʰ] does not form any cluster, i.e. it occurs as neither C_1 nor C_2 in clusters initially, medially or finally.



The general convention that at least one of the two parentheses has to be chosen will apply here.

2. Initially, medially, and finally, two stops of the same point of articulation do not follow each other.



(The above notation of enclosing the features in square brackets in braces simply means the adjacent segments must differ in either anteriority, or coronality or retroflexion, but not necessarily in any two or all three features.)

Stops with the same point of articulation that have similar voicing, e.g., tt, pp, etc. (the so-called geminates) I am treating as long consonants, therefore this condition will not be relevant to them. It is this condition that gives [pod:ar] instead of [potdar] as the last name (see discussion in footnote 12). In my view it is because of this condition that the [ə] is not deleted in words like [adətĕ].

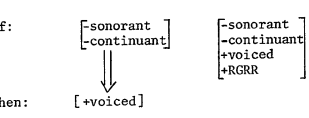
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. No two aspirates or breathy voiced stops occur in a sequence, nor does an affricate, aspirate, or breathy voiced stop occur before or after an affricate. It is because of conditions like this that, as mentioned in chapter 1, I have used the feature "distinctive release".

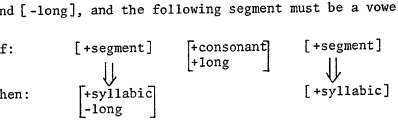
What is usually written as $\underline{pp^h}$, etc. I consider to be long consonants and similarly $\underline{\check{c}\check{c}^h}$, etc.

NC: $\sim$ [+distinctive release] [+distinctive release]

4. In stop+stop clusters, if C_2 is breathy-voiced, C_1 must be voiced.



5. If a consonant is [+long], the preceding segment must be a vowel and [-long], and the following segment must be a vowel.



Long consonants are the so-called geminates. This condition also rules out long consonants initially and finally. Although in final position, the orthography gives long consonants, e.g., $\underline{\text{nirləǰǰ}}$ "shameless", $\underline{\text{rəbb}}$ "God", $\underline{\text{mlečč}^{\text{h}}}$ "barbarian", in my experience, words such as these are not pronounced with long consonants. There are perceptual reasons for the constraint that a vowel must precede a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

long consonant. Unless a vowel precedes the long consonant there would be no way of telling whether the consonant is long or not. The constraint that the preceding vowel be short is common to other languages, e.g., Norwegian and Swedish (cf. Lehiste 1970:42).

I now discuss a notational difficulty encountered in the next few constraints:

In Hindi we find initial and final constraints such as the following:

(a) Initially and finally, if C_1 is either a stop or a fricative and C_2 is either a stop or a fricative, they must agree in voicing.

(b) Initially and finally after glides, no stop, nasals, or fricatives occur.

Now in the above constraints, under the present morpheme structure condition format, two conditions each would be needed, one for when the environment is provided by the initial morpheme boundary and one in which the final morpheme boundary provides the environment. It seems to me that writing two conditions each for the above misses a generalization. Bach (1968) and Langacker (1969), noticing situations similar to these in syntax and phonology, proposed a new convention in the writing of rules. Bach called it the "neighborhood convention" and Langacker the "mirror image" convention. The actual notational conventions proposed by Bach and Langacker are different but both express the same idea.

At first glance it seems that such a convention should be adopted even in the lexicon to handle constraints such as the ones mentioned above. In the case of constraint (a), this mirror image convention

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

would work by writing if-then conditions of the following sort, using Langacker's asterisk (*) convention:

(a) if: *+ [-sonorant, +consonant] [-sonorant, +consonant]

then: [αvoice] [αvoice]

However in the case of constraint (b), neither Bach's nor Langacker's proposal will work. We don't want a mirror image of the entire condition. Let me illustrate: if we stated the condition as:

if: *+ [-consonant, -syllabic] [-syllabic]

then: [+sonorant, -nasal]

the mirror image convention would wrongly yield:

if: [-syllabic] [-consonant, -syllabic] +

then: [+sonorant, -nasal]

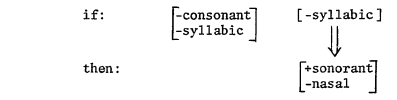
I.e., what we want is the morpheme boundary to be reversed but the order of the other segments to be kept the same. Bach says (130): "Given an abbreviated rule 'c--> d/$a_1a_2....a_n$ should this be defined as the abbreviation of rule A or B?

A. c--> d/{$a_1a_2....a_n$ ____ / ____ $a_1a_2....a_n$} B. c--> d/{$a_1a_2....a_n$ ____ / ____ $a_n....a_2a_1$}"

and concludes that B is correct. Further he says "I know of no rules involving a sequence of several segments where a particular specification is made after or before that sequence in the same order (which

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

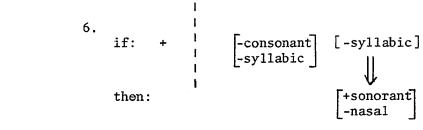
seems inherently implausible)" (130). Constraint (b) given above is a counter-example to Bach's claim. With certain modifications constraint (b) can be expressed in terms of Bach's neighborhood convention:



So far the "double environment" has not been used in the lexicon, but we would have to do so if we were to use Bach's notational convention as above. Thus we would have to modify the current method of writing morpheme structure conditions and would have to restrict Bach's convention to be used with the understanding that mirror imaging is not implied. But rather than redefining its use in this way, and to avoid confusion, I propose using the following convention.

The left-right exchange convention: Simultaneously everything to the left of the dotted line goes to the right and vice-versa.

Thus in the case of constraint (b) we have the following condition:



Since the segments on the right of the dotted line keep the same order, the desired results are obtained. If in phonology and syntax similar situations arise where some segments of a rule reverse but

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

others retain their order, then it would suggest that the "left to right exchange convention" should be used rather than the "mirror image" rule convention even in other parts of the grammar.

In fact, strictly speaking, there is no definitive evidence that mirror image rules play any role in phonology. Most of the examples given by Bach are of the sort where the environment of a rule can be either ____X or X____, but this, of course, could as well be an example of left-right exchange. There is no telling which it is since there are no component elements in the interchanged segments whose order could be reversed. In the few cases where the order of the units within the reversed parts can be determined, there is again no evidence of strict mirror imaging. Bach considers the case of vowel harmony wherein a given vowel's features assimilate to that of the vowel in the environment of either

$$____X \left[\begin{matrix} V \\ \alpha\text{feature} \end{matrix}\right]$$

or

$$\left[\begin{matrix} V \\ \alpha\text{feature} \end{matrix}\right] \quad X____$$

where X can be any number of segments. But strictly speaking, in order for this to be considered a case of mirror imaging, whenever there are two or more segments in X it must be the case that the order of the segments in X, $S_1\ S_2 \ldots\ldots S_n$ be exactly reversed in the mirror image environment. Now it is conceivable that there might be a few cases in which a harmony rule will apply in both the environments $____S_1S_2V$ and $VS_2S_1____$, but it is more likely that we will also find cases where a rule applies in the first environment but there is no

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

occurrence of the second environment. Of course, it could be claimed that the second environment is a proper environment for the application of the rule, but it simply doesn't exist in the language. What is needed to demonstrate the existence of true mirror image environments in vowel harmony are cases where the environments

(a) $____S_1S_2V$

and

(b) $VS_2S_1____$

do trigger the vowel assimilation but the environments:

(c) $VS_1S_2____$

and

(d) $____S_2S_1V$

do not.

(Of course, if what we found in vowel harmony languages was that the vowel harmony rule applied in environments (a) and (c), then there would be no reason to call these environments "mirror image" environments, since strictly speaking they are not).

I know of no such cases; and lacking this crucial evidence, one could simply explain vowel harmony as assimilation across segments and not involve mirror image environments at all. This interpretation seems reasonable, given the findings of Öhman (1966,1967) on vowel coarticulation.[33]

If this is true, then it would appear that perhaps the human brain cannot recognize the similarity of mirror images of strings of segments or create mirror images of strings. More investigation on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

this point is needed; but I would claim that, at least in the lexicon, mirror image rules should be ruled out.

I will state condition (a) with the left-right exchange convention:

7. if: + | [-sonorant, +consonant] [-sonorant, +consonant]
then: [αvoice] [αvoice]

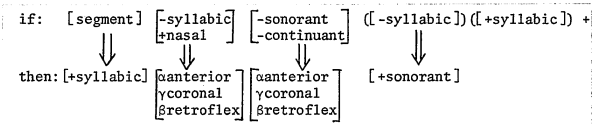
Initially, of course, there are no stop+stop clusters, but this is accounted for by other constraints to be given shortly; thus the present constraint can be stated in a more general form. The high Urdu word fəsd "to take blood out" (noun) would have to be listed as an exception, but in any case this word would not occur in the vocabulary of non-high Urdu speakers.

8. Medially and finally fricatives do not occur before breathy voiced stops.

if: [+syllabic] [-sonorant] [-sonorant, -continuant, +voiced, +RGRR] ([+syllabic]) +
then: [-continuant]

9. As mentioned earlier, morpheme-medially and morpheme-finally, the normal case in nasal+stop clusters is for the nasal to be homorganic before the stop. In morpheme-medial case this is true only at the lexical level (for details on SQCs applying to a particular level see 2.1.2.):

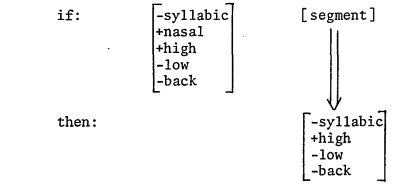
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Condition: Morpheme-medially, this only applies at the lexical level; and only to [+native] morphemes.

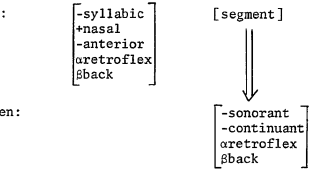
If the last two optional segments are not chosen, we will get a morpheme-final homorganic nasal; if the optional last segment is left out, we will get a morpheme-final three-consonant cluster where the C_1 is a homorganic nasal. If all optional elements are chosen, we will have a morpheme-medial three-consonant cluster with a nasal as C_1; and if the second-to-last optional element is not chosen, a morpheme-medial two-consonant cluster with C_1 as nasal. Not all consonants can occur as the C_3. Stops and fricatives are ruled out both medially and finally. The segments that can occur as C_3 are further restricted, but these restrictions have not been mentioned here.

10. After a palatal nasal the following segment is a palatal consonant (this condition was given and discussed earlier under (10):

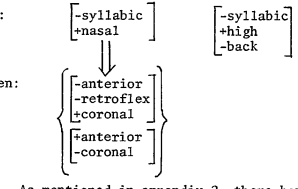


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

11. After [ŋ] or [n̩] the following segment must be a homorganic stop (this condition was given earlier under (11)):



12. If the segment before a palatal consonant is a nasal, it must be a bilabial or palatal nasal (this condition was discussed earlier under (12)):



13. As mentioned in appendix 2, there have been no studies involving the co-occurrence restrictions on vowels after consonant clusters. I have not done such a study either. However the results of tests that I conducted (see the appendix) seemed to suggest that at least kyi: was not acceptable to speakers. On looking at my data I found that I have no examples of initial or medial stop+y+front vowel clusters. On the basis of speakers' rejection of kyi:, I am treating this non-occurrence of stop+y+front vowel as a constraint in the language and not an accidental gap. This is, of course, tentative,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and more investigation should be done in this area of Hindi phonology.

if: [-sonorant, -continuant] [-syllabic, -consonant, +high, -back] [+syllabic]

then: [+back]

This will permit the back vowels including [ɐ] but rule out the front vowels after C+y.

14. Initially and finally, [ṛ] does not occur in any cluster

NC: ~ + ([-syllabic]) [+consonant, +sonorant, +retroflex] ([-syllabic])

(If [-ṛh] was treated as a cluster, this condition would have to be modified to permit -ṛ+h. See discussion in 2.3.3.).

15. Initially and finally after [y], only vowels occur.

if: + [-consonant, -syllabic, -back, +high] [segment]

then: [+syllabic]

A few other conditions could be stated, but I am not sure if what these constraints would rule out are accidental gaps or whether they really are impermissible for Hindi. Therefore I have not stated them formally as conditions.

(a) Initially, medially, and finally, aspirates do not occur as C_2 in stop+stop clusters (pp^h, etc., are treated as long consonants).

(b) Initially, medially, and finally, in stop+stop clusters, if C_2 is retroflex it must be [ṭ].

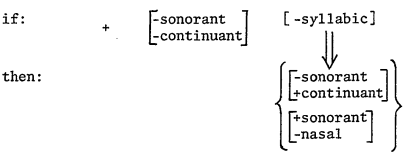
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(c) Medially and finally, if C_2 is [t^h], then the only consonant that occurs before it is [š].

(d) Initially and finally after breathy voiced stops, if C_2 is a liquid it must be [r].

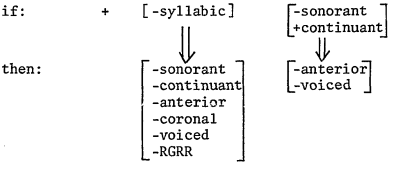
Initial consonant constraints:

16. If C_1 is a stop (here and in what follows 'stop' includes affricates), C_2 cannot be a stop or a nasal.



The rejection of *bmi:n, *t^hgen, and *pdi:l in the test reported in the appendix 2 lends support to the above condition being "psychologically real".

17. In a two-consonant cluster, if C_2 is a fricative, it must be [š] and the preceding consonant [k].

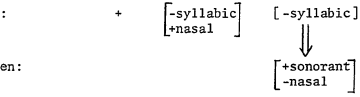


It might seem at first glance that the above condition is too specific and that the non-occurrence of other stop+fricative clusters might be an accidental gap, but in my test the majority of the subjects

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

rejected morphemes like *kfal *čsu:l *zfi:k *msok (the last two words violate more than one condition; *zfi:k violates 7, 17, 23, 21, and *msok violates 17 and 18). This, however, needs further testing, since I did not include kš- in the accidental gap category of words but only in the existing word category.

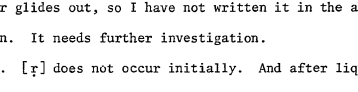
18. If C_1 is a nasal, C_2 cannot be a stop, a fricative, or a nasal.



Again the rejection of *msok and *mkol by the majority of the subjects in the test lends support to the above condition.

There is a further constraint possible, in that if C_2 is a glide it must be [y]; but I am not sure if subjects would indeed rule the other glides out, so I have not written it in the above condition. It needs further investigation.

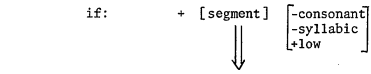
19. [r̥] does not occur initially. And after liquids only vowels occur.



The only word violating this constraint included in the test was *lj^{h}am and the majority of the subjects did reject it.

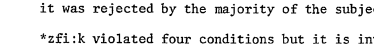
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

20. Before [h] only a vowel can occur.

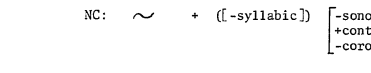


If one were treating the aspirates ([b^h], [p^h], etc.) as clusters of stop+h instead of units, as I do, then of course this condition would not hold.

21. Before or after [f], only a vowel can occur. The word *zfi:k in the test violated this constraint and as mentioned earlier it was rejected by the majority of the subjects. As mentioned earlier *zfi:k violated four conditions but it is interesting to note that the rejection of *zfi:k was not greater than that of *bmi:n which violated only one constraint.



22. If C_1 is a fricative and C_2 is a stop, the fricative must be [s].



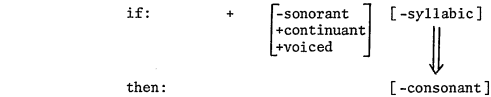
Of the [+anterior] fricatives [f] is already ruled out by condition 21. It is possible that [š] should not be ruled out but treated as an accidental gap. The only word violating this condition in the test was *sb^{h}al. However this was not rejected by the majority of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

subjects perhaps (as I mentioned in appendix 2) due to its similarity in fast speech with [səmb^{h}al] "take care of". More investigation is needed.

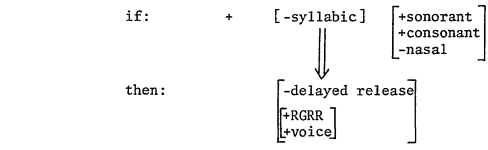
The stop that follows the [s] also has to be voiceless, but this is taken care of by condition 7.

23. If C_1 is [z], C_2 can only be [y].



The test item *zfi:k which violates this condition was rejected by the subjects. Of the glides [h] is ruled out by condition 20. I haven't ruled [w] out since I am not sure at this point whether the absence of [zw-] is not just an accidental gap.

24. If C_2 is a liquid, C_1 cannot be an affricate or an aspirate, i.e. breathy voiced stops are permitted but not aspirates.

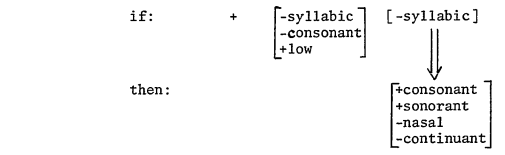


At first it might appear that since breathy voiced stops occur as C_1 the non-occurrence of aspirates is an accidental gap. However in my test there were 20 out of 25 no's to *t^{h}ru:g, but in the case of b^{h}rak only 10 out of 25 no's ([b^{h}r-] occurs in Sanskrit loans like [b^{h}rəm] "confusion").

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

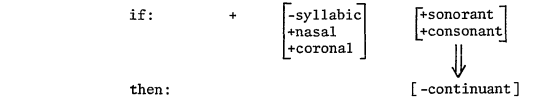
Further, from the tables listing the clusters, it appears that if we have a stop before the liquid [l] the stop must be [-breathy voiced], and must be labial, i.e. before [l] even breathy voiced stops are not tolerated. [bl-] occurs only in a few English loans. The only example relevant to this restriction in my test was dleš, which was not overwhelmingly rejected. However (as mentioned in Appendix 2), this could have been due to its similarity with an existing word. Therefore more investigation is needed on this point.

25. If C_1 is [h] C_2 must be [r].



We don't need to mention [-long] to rule out [r:] because 5 takes care of this. We also don't need to mention [-retroflex] because of condition 14.

26. If C_1 is [n], C_2 cannot be [l].



Nasals are also [+sonorant, +consonant] but they will be ruled out by condition 18.

A few more conditions could perhaps be stated, however more investigation is needed to determine if they might be accidental gaps:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) If C_2 is a glide and C_1 is an aspirated stop, the C_1 must either be bilabial or velar; or if the C_1 is a breathy voiced stop, then it can only be [d^h].

(b) If C_2 is [w], C_1 cannot be a bilabial or retroflex stop.

(c) As mentioned in condition 6, no stops, fricatives or nasals occur after glides. However even the glides and liquids that occur after glides are restricted. Of the liquids only [r] occurs. If the C_1 glide is [y], condition 15 will make sure that only a vowel occurs after it. And if the C_1 glide is [h], only [r] occurs after it by condition 25. If the C_1 glide is [w], only [wy-] occurs. [wh] is ruled out by condition 20, and [ww] is a long consonant which would be ruled out in initial position by condition 5).

Medial consonant cluster constraints:

27. Glides do not occur before breathy voiced stops or as aspirates.

if: + [segment]$_o$ [+syll][-syll] $\begin{bmatrix}\text{-sonorant}\\ \text{-continuant}\\ \text{+RGRR}\end{bmatrix}$ [+syll][segment]$_o$ +

then: [+consonant]

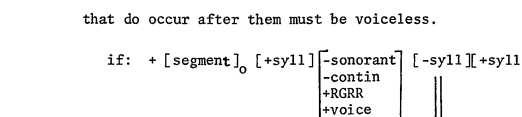
28. Fricatives do not occur after breathy voiced stops, aspirates or affricates.

if: + [segment]$_o$ [+syll][+dist rel][-sonorant][+syll][segment]$_o$ +

then: [-continuant]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

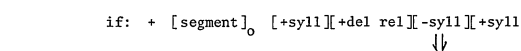
29. Nasals do not occur after breathy voiced stops, and stops that do occur after them must be voiceless.



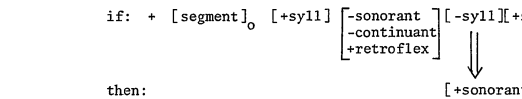
Since fricatives have been ruled out after breathy voiced stops by condition 28, voiceless [-sonorant] will not include fricatives.

There are further restrictions that could be stated for the above consonant clusters, namely that if the breathy voiced stop is velar the following stop must be [t̤], and if the breathy voiced stop is non-velar the following stop must be [k]. However, further investigation is needed to see if this really is a constraint or an accidental gap.

30. Glides do not occur after affricates.



31. If, in a two-consonant cluster, C_1 is [d̤], C_2 cannot be a fricative or a stop.



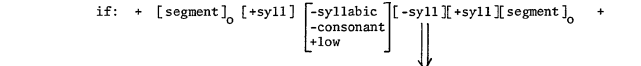
Actually there might be further restrictions. The only glide in C_2 position is [w] and the only nasal is [m]; both of these occur in rare

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

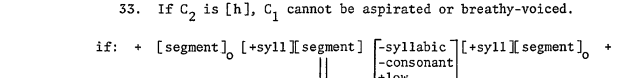
Sanskrit loans (they have not been listed in the cluster charts). Even liquids occur after [ḍ] only in three-consonant clusters where C_1 is a homorganic nasal. However more evidence is needed before these restrictions are included.

Above the feature [-RGRR] has not been used to rule [ḍʰ] out because [ḍʰ] is already ruled out by condition 1.

32. Retroflex segments or affricates do not occur after [h].

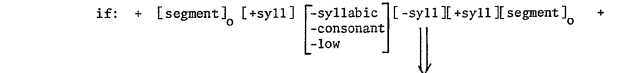


33. If C_2 is [h], C_1 cannot be aspirated or breathy-voiced.



There are further restrictions that could be stated: if C_1 is a stop, it can only be palatal (i.e. an affricate) or bilabial; and moreover, if it is voiceless, it can only be bilabial. However at this point I do not have enough evidence to state these observations as constraints.

34. If C_1 is a [-low] glide, i.e. [y] or [w], then C_2 cannot be a glide.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Due to this there are no glides occurring before [h].

35. [y] does not occur before or after retroflex segments.

NC:

$$\sim + [\text{segment}]_0[+\text{syll}]([+\text{retro}])\begin{bmatrix}-\text{syll}\\-\text{conson}\\-\text{back}\\+\text{high}\end{bmatrix}([+\text{retro}])[+\text{syll}][\text{segment}]_0 +$$

36. The voiced fricative [z] occurs only after voiced segments.

$$\text{if: } + [\text{segment}]_0 [+\text{syll}][\text{segment}]\begin{bmatrix}-\text{sonorant}\\+\text{continuant}\\+\text{voice}\end{bmatrix}[+\text{syll}][\text{segment}]_0 +$$

$$\text{then: } \Downarrow\ [+\text{voiced}]$$

37. Retroflex segments, voiced stops, liquids, and nasals do not occur after [ṛ].

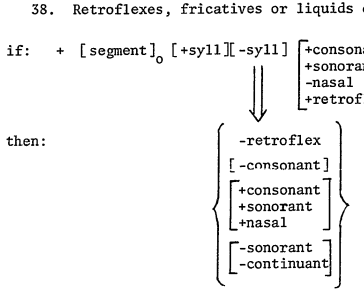
$$\text{if: } + [\text{segment}]_0 [+\text{syll}]\begin{bmatrix}+\text{sonorant}\\+\text{consonant}\\-\text{nasal}\\+\text{retroflex}\end{bmatrix}[-\text{syll}][+\text{syll}][\text{segment}]_0 +$$

$$\text{then: } \Downarrow\ \left\{\begin{array}{l}-\text{retroflex}\\[-\text{consonant}]\\\begin{bmatrix}+\text{consonant}\\-\text{sonorant}\\-\text{voiced}\end{bmatrix}\end{array}\right\}$$

There is a further constraint that if the C_2 is a fricative it must be [s], but I am not sure if the non-occurrence of [š] and [f] in this position is an accidental gap or not; [z] is ruled out by the [-voice]. Of the glides, [y] does not occur after [ṛ], but this is taken care of by condition 35.

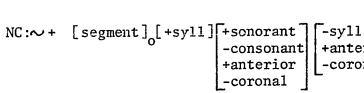
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

38. Retroflexes, fricatives or liquids do not occur before [ɻ].



Of the glides, [y] doesn't occur, but this is ruled out by 35. [w] occurs, and the non-occurrence of [h] is here treated as an accidental gap.

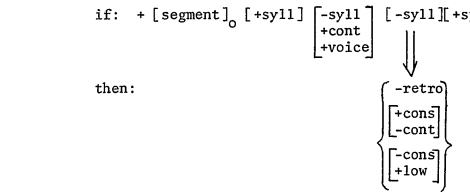
39. [w] does not occur before bilabials (stops, fricatives, nasals, or glides). (What other writers give as the sequence ww I treat as a long consonant.)



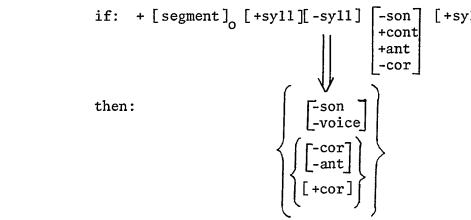
The features anterior and coronal are not needed to distinguish the three glides [y] [w] and [h] from each other, but are relevant to this rule to show that the above constraint exemplifies dissimilation.

40. Fricatives and retroflexes do not occur after [z], and if the consonant following [z] is a glide it must be [h].

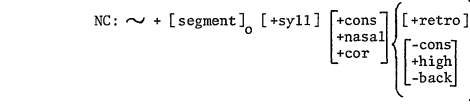
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



41. Bilabials do not occur before [f]. (Again, what others call ff I treat as a long consonant.) wf would be ruled out by condition 39. Also, if the segment before [f] is a stop, it must be voiceless (the voiced fricative [z] would be ruled out by 40).



42. Retroflexes or [y] do not occur after [n].



The following conditions could also be stated for medial clusters, however it is not clear to me at this point whether they should be stated formally:

(a) If C_2 is an affricate and C_1 a stop, the affricate must be [-RGRR] (i.e. unaspirated and non-breathy voiced) and the C_1 must be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

bilabial or velar. (Actually the only combinations we do find are -kč-, -pǰ-, and -bǰ-.)

(b) In stop+stop clusters, if C_1 is an aspirate it cannot be a dental or a bilabial, and C_2 cannot be [p].

(c) In stop+stop clusters, if C_1 is retroflex it must be voiceless, and the C_2 must be [k].

(d) In stop+stop clusters, if C_1 is voiceless and C_2 is voiced, C_2 cannot be a velar. If C_1 is voiced and C_2 is voiceless, C_2 cannot be bilabial.

(e) If C_2 is a glide and C_1 a breathy voiced stop, C_1 must be [d^h].

(f) After [š], if a stop is voiced it must be [g], and if it is aspirated it must not be dental or velar.

(g) After [s], voiceless affricates do not occur, and neither do [g] or [h].

Final consonant constraints:

43. The aspirated and breathy voiced affricates [$ȡ^h$] and [$ʝ^h$] do not form a cluster, neither as C_1 nor as C_2.

$$\text{if: } ([+\text{segment}]) \begin{bmatrix} +\text{del rel} \\ +\text{RGRR} \end{bmatrix} ([+\text{segment}]) \; +$$

$$\text{then: } [+\text{syll}] \qquad\qquad [+\text{syll}]$$

44. [p^h] does not occur as C_1 in any final consonant cluster.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

if: [+cons, -son, +RGRR, -voice, +ant, -cor] [segment] +

then: [+syll]

45. In stop+stop clusters, velars and bilabials do not occur as C_2.

if: [-son, -cont] [-son, -cont] +

then: [+cor]

46. After breathy voiced stops, no fricatives or stops occur.

if: [+RGRR, +voice] [segment] +

then: [+son]

There are further restrictions. If the C_2 is a liquid it must be [r]; however, as stated under (d) on pg. 143. I am not sure whether this is not just an accidental gap. The fact that C_2 cannot be [h] will be ruled out by condition 50. Of the nasals, I haven't found examples of [m] as C_2, but I am treating this as an accidental gap.

47. If C_1 is retroflex, C_2 can only be [y].

if: [+retro] [-syll] +

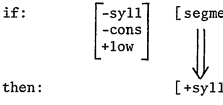
then: [-cons, +high, -back]

After [ɽ] not even [y] occurs, but this is taken care of by condition 14.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

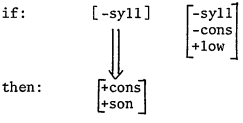
48. After [f] glides do not occur.

49. [h] does not occur as C_1 in any final consonant cluster.



There are a few Sanskrit loans that would occur in high Hindi which contain -hy; these would have to be listed as exceptions. The above condition would not hold for Urdu speakers who are supposed to have the clusters -hr, -hw, etc.

50. Only liquids or nasals occur before [h].



I am not sure if speakers do pronounce the [h] in clusters like -mh, -nh, etc. If they don't, then the above condition would have to be modified to say that before [h] only vowels occur (i.e. it could be combined with no. 49). On the cluster chart -lh has not been listed. The only example I found listed in the dictionary is listed as selh ~ sel "spear".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

51. Liquids do not occur after [n].

if: $\begin{bmatrix} -\text{syll} \\ +\text{nasal} \\ +\text{cor} \end{bmatrix} \quad \begin{bmatrix} +\text{cons} \\ +\text{son} \end{bmatrix} +$

then: $\Downarrow$ [+nasal]

The following conditions could also be stated, but more evidence is needed:

(a) Stop+affricate clusters are rather restricted. Only [ǰ] occurs as C_2, and only bilabials as C_1. (Actually only [b] occurs, since [b^h], [p^h], and [p] would be ruled out by conditions already stated.)

(b) In stop+stop clusters, if C_1 is an aspirate, it must be [k^h] and C_2 must be [-retroflex]. (Again we only get [t] as C_2, since other stops would be ruled out by conditions already stated).

2.3.8. Constraints on vowel sequences:

It has not been possible for me to do a detailed analysis on the sequential constraints on vowels. Therefore the following observations are highly tentative.

Writers such as Arun and Dixit, who have listed vowel sequences, have not listed any for morpheme-initial position. The only initial vowel sequences which I have been able to find examples of are:

(15) ai: e.g., [ai:na] "mirror"

au e.g., [aus] "a kind of grain"

Even these occur in very few words. Moreover, it is debatable whether ai: does exist as an initial vowel sequence, because most speakers render the word for "mirror" as [ayna] or [aena]. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

dictionary also lists a word u:abai: "babble". This might lead one to think that u:a is also a possible initial cluster; however, I think this can be analysed as /u:+a+bai:/, the meaningless noises u: plus a plus the morpheme bai: "lady". This would be the only example of this cluster initially, and it does not even occur morpheme medially or finally.

Initially we do not find diphthongs -- that is, the sequences əi: and əu: listed as diphthongs[34] by writers such as Arun (1961) in their list of phonemes. At the phonetic level of course these are vowel sequences. When writers such as Arun list these as phonemes they are claiming that phonologically these somehow act as a unit. I do not see any evidence of this. As a matter of fact if they are treated as sequences of vowels then the statement of the sequential constraints is simplified (cf. discussion in SQC 56 below). Therefore I propose to treat these as sequences of two vowels at both the systematic phonemic and the systematic phonetic level.

In non-initial position I have found examples of the following clusters:

Table 20

əi:	[kəi:]	"many"
ui:	[sui:]	"needle"
ai:	[nai:]	"barber"
oi:	[loi:]	"small ball of flour"
ei:	[lei:]	"paste made of flour"
au	[blauz]	"blouse"
əu:	[gəu:]	"cow"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ua	[ɣua]	"dice"
oa	[k^hoa]	"dry milk"
uə	[suər]	"pig"
ae	[čae]	"tea" (ae could alternatively be simply treated as [ay]).
ao	[čao]	"desire" (again could be treated as [aw]).
au:	[tau:]	"uncle"

Of the above, [au] and [uə] occur only medially, not finally. I have done no test to see if any of the rest are restricted to just medial or final position, or if their non-occurrence in one or the other position is merely an accidental gap.

Arun and Dixit also list the following sequences:

Table 21

ie	ia	io
ee	ea	oo
əe		
oe		uo

The only examples of these that I have been able to find involve a morpheme boundary, e.g., /ro+e/ "cry+subjunctive" (3rd.pers.). This is not to say that there are no examples of these clusters morpheme-medially, just that in my tentative observations I haven't found any.

Moreover a number of the clusters in table 21 are not really clusters because a glide intervenes, e.g., /dek^h+ie/ "see" (polite imper.) is phonetically [dek^hiye]. In such sequences there always seems to be an intervening glide; for others some speakers have a glide

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and others do not, e.g., /k^{h}a+e/ "eat (subjunctive)" can be [k^{h}ae] or [k^{h}aye] depending on the speaker. In the case of /e+e/ I think most speakers insert a glottal stop or [w] and not the palatal glide [y], i.e. instead of /k^{h}e+e/ "row" (subj.) being rendered *[k^{h}eye], it is rendered [k^{h}e?e] or [k^{h}ewe].

The following conditions for vowels can be tentatively proposed. The first two conditions pertain to nasalized vowels. These will be discussed in detail in chapter 3.

52. if: ([-nasal]) $\begin{bmatrix}+\text{syll}\\-\text{long}\end{bmatrix}$ ([-nasal])

then: $\Downarrow$ [-nasal]

53. if: $\begin{bmatrix}+\text{syll}\\+\text{long}\\+\text{nasal}\end{bmatrix}$ $\begin{bmatrix}+\text{cons}\\-\text{son}\end{bmatrix}$

then: $\Downarrow$ $\left\{\begin{bmatrix}+\text{cont}\end{bmatrix}, \begin{bmatrix}-\text{cont}\\-\text{voice}\end{bmatrix}\right\}$

54. In morpheme-final position, short vowels do not occur.

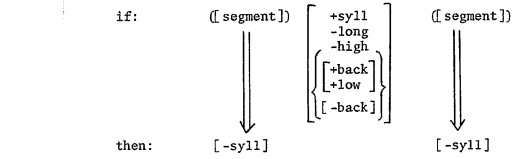
if: [+syll] +

then: $\Downarrow$ [-long]

This also accounts for the fact that in final position in vowel sequences, the V_2 is never short.

55. [ɛ] and [ɔ] do not occur in any vowel clusters, neither as V_1 nor as V_2, initially, medially, or finally.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



56. Morpheme-initially after a short vowel, no other vowel occurs, i.e. the following segment must be a consonant.

The only exceptions I could find are interjections like [əi:], [ui:] "ouch".

As mentioned earlier diphthongs do not occur morpheme-initially. I also proposed earlier that they should be treated as sequences of two vowels -- əi: and əu:. If we treat them as vowel sequences then we don't need a separate condition to rule diphthongs out initially since condition 56 will not permit initial əi: and əu:.

Certain other generalizations could be made about two-vowel sequences:

(a) [i:] and [u:] are never V_1

(b) [i] is never V_2

(c) Initially [e] and [o] are not found as V_1.

However all of these require more investigation before being stated formally.

Three-vowel clusters do not seem to be permitted in Hindi,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

either within morphemes, or across morpheme boundaries. A glide always seems to intervene. Thus /b^{h}ai:+ð/ "brother+plural" (oblique) is [b^{h}aiyð].

57. Three-vowel clusters are not permitted.

NC: $\sim$ [+syll] [+syll] [+syll]

If the above condition is found to be true even on more investigation then this constraint should be treated as a general "everywhere" constraint and will thus apply both morpheme medially and across morpheme boundaries.

I should mention at this point that since I consider [y] and [w] to be segments at the lexical level, my analysis differs from that of writers such as Srivastava (1970) who do not. Srivastava says "... it is not necessary in Hindi to posit any underlying semivowel phonemes like y and w. In cases where y and w appear in final output, i and u vowel phonemes .... can be said to exist in the phonological representation in the lexicon" (134). Thus he considers both [aya] "maid" and [aya] "came" to be without glides in the underlying representation, i.e. /aia/ and /a+a/ respectively, and has the glide introduced by P-rules. I agree with him in the case of the verb (i.e. [aya] "came") but not in the case of the noun, where I would claim that the underlying form is /aya/. I do not see any advantage in this case in positing the abstract underlying form /aia/ for [aya] (or /aa/, which also would be an abstract form; the vowel sequence /aa/ is also not permitted within morphemes) rather than simply having underlying /aya/ in the first place. I do see a number of disadvantages:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) By considering [y] and [w] to be vowels there would be a needless proliferation of vowel clusters in the lexicon; and, as I have tried to show above, Hindi does not tolerate three-vowel clusters, which is what Srivastava's /aia/ would be.

(b) His rules would give the correct phonetic output for words like [suər] "pig" and [swər] "sound" only by rather ad hoc means. He suggests that [suər] is actually /su:=ər/ to which his vowel laxing rule applies to give the phonetic [suər]. However, there doesn't seem to be any justification for this abstract underlying form; moreover having to mention '=ər' in the environment of the rule seems especially ad hoc.[35] More importantly, I don't know how he will handle -- in any non-arbitrary way -- words like [dua] "blessing" and [dwar] "door", etc., because [dua] will not fit the environment of his rule as stated.

I propose that within morphemes the [y] and [w] that are to appear at the phonetic level be posited at the systematic phonemic level too. Only in the cases where processes of suffixation and inflection, etc. bring about the environment for introducing a glide, should such a glide be introduced by phonological rules. For this purpose Srivastava's rule with certain modifications would be adequate. Srivastava gives his P-rule I as (135):

(16)

$$\emptyset \rightarrow \begin{bmatrix} V \\ +high \\ \alpha back \\ -tense \end{bmatrix} \Big/ \left\langle \begin{bmatrix} V \\ +high \\ \alpha back \end{bmatrix} \right\rangle + \underline{\quad} \begin{bmatrix} \left\{ \begin{bmatrix} V \\ [\alpha around] \end{bmatrix} \right\} \\ \left\{ \begin{bmatrix} V \\ [-high] \end{bmatrix} \right\} \end{bmatrix} \qquad \begin{matrix} a. \\ b. \end{matrix}$$

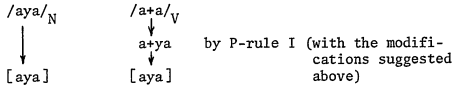
The only modification I propose is having $\emptyset \rightarrow$ y, w, i.e. having the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

zero rewrite as the glides directly and not Ø--→ i or u.

Thus I propose that the following is the derivation for the two words [aya] "maid" and [aya] "came".

(17)



In this chapter I have proposed certain modifications of the current MS theory, I have presented evidence that in certain cases "abstract" segments are apparently "psychologically real", and I have proposed a number of sequential constraints for Hindi morphemes. Let me emphasize that these SQC's are tentative. In a large number of cases their "psychological reality" still needs to be tested further.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Footnotes

1 Chomsky does not use the term "impossible" though there is the implied symmetry of "possible:impossible". Halle, however, does use the terms "possible:impossible".

2 By exceptions I mean a few items that don't undergo the rule; if the exceptions were quite large, then they might justify the writing of "minor rules" to account for them (Lightner 1968).

3 Yasui (n.d.) also notices that skl- and many other consonant sequences exist as clusters in some speakers' vocabulary but not in others'.

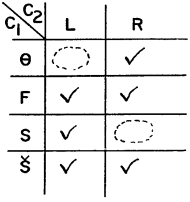
4 This is because alveolars are perceived as retroflexes and not as dentals (see Ohala, M. 1971).

5 I am grateful to John Ohala for pointing out this article to me.

6 This brings up an interesting point. It could be argued that the introduction of the šm- cluster into American English was facilitated by the prior existence of the common cluster sm-, the prior absence of šm- being considered a kind of "accidental gap". Accidental gaps in the lexicon such as /blik/ have been explicitly recognized but accidental gaps in the structural pattern of sequences have not received much attention (although cf. Yasui n.d.), let alone formal recognition. This deserves further study and testing. Neither θl- nor ðl- exist as clusters in American English. But since fl-, sl-, and šl- do exist (flick, slick, Schlitz) but vl-, zl- and žl- do not, one might expect the absence of θl- (but not ðl-) to be but an "accidental gap".

Generativists, in deciding whether a particular cluster should be ruled out or not, use "cost" as a criterion (cf. McCawley 1968:49). It would seem to me that it should be less costly to permit the clusters θl- and sr- rather than rule them out, since they are merely holes in the pattern:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Nevertheless Schane rules θl- out (1970:103). Instead of counting the "cost" of ruling θl-, etc., out, it seems to me more important to find out how native speakers of English react to θl-: do they treat it more like šm- type clusters, or more like ft- type clusters? It is not obvious that the "cost" factor will play a part in their judgment.

7 McCawley (1970) differentiates between constraints that apply at the lexical level, and constraints that apply at the phonetic level. At least for constraints at the phonetic level, he seems to equate possible morpheme with pronounceable morpheme. However this is not a technical paper, being intended for laymen, and it is difficult to understand McCawley's position from this article.

8 Kornfeld (1971) claims to have evidence that there is a difference between the acoustic signal of w that the child substitutes for r and his regular w. Kornfeld labels this conclusion highly tentative. And moreover the fact still remains that the child cannot produce the American English r and substitutes some kind of a w for it (whether it be his regular w or a different w), and won't accept an adult's non-r as a substitution for r.

9 Of course if the words were presented to them auditorily, they might not have pronounced them. However there are other problems with presenting the words auditorily, which I have mentioned on pg. 34.

10 This has also been noted by Fromkin 1971c.

11 It seems probable to me that many speakers could pronounce a word like zbornik but know it is not an English-type word (see my discussion in 2.1.1.). However McCawley's point is well taken even if his example is not.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

12 This is an interesting example for demonstrating some kind of "psychological reality" of morpheme boundaries because when this word, pot+dar, lost its occupational meaning and just remained as a last name (similar to English "Smith"), the word became [pod:ar], i.e. assimilation took place. Thus for the last name poddar a division into two morphemes would not be justifiable, in a synchronic grammar.

13 However as far as I can see, Kisseberth does not explicitly state that this is because "possible pronounciation" requires it; i.e. because of a constraint operative at the phonetic level. Thus I don't agree with McCawley when he says (presumably referring to the article cited above) that Kisseberth argues that:

> ..the constraints on possible pronunciation affect the phonological rules fully as much as the converse and that it is quite common for the details of a large number of rules to reflect the same constraint on possible pronunciation... (1970:818)

14 There are apparent exceptions to these, pointed out to me by J. Ohala. Radler one would expect to be *[ratlər] but is instead [radlər]; also, on the basis of the pair hand [hant] "hand" and handes [handəs] "hands" we would posit the underlying morpheme for both as /hand/ and expect handumdrehen to be rendered *[handumdreən], however it is usually rendered as [hantumdreən].

15 Shibatani does not give any details on how he would treat Kisseberth's example, but presumably he would retain the P-rules as Kisseberth has them and also posit an ^{m}SPC (which holds both inter- and intra-morphemically), which would state that triliteral clusters are not permitted in Yawelmani. Supposedly the ^{m}SPC would explain why the P-rules exist.

16 Shibatani "ultimately" wishes to provide an explanation for SPCs (4) and (5), in "phonetic or physiological considerations" (10). This can only be true either if SPC's are universal for all languages (which is doubtful), or if they reflect pronunciation constraints for English speakers, but this is obviously false since exceptions exist. It seems obvious that much of what he calls SPC's, are not universal constraints, nor pronunciation constraints for the given language, but are patterns that are language specific and will not have a phonetic explanation.

17 Although I show in chapter 4 the ə-deletion rule has to be more complicated, for immediate purposes this simplified version will do. In chapter 4 I also explain why I posit a ə-deletion rule and not a ə-insertion rule.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

18 As shown by the test to be discussed shortly, this is true only for some speakers.

19 See chapter 4 for the historical reasons as to why -iya is a suffix which blocks the application of the ə-deletion rule.

20 At least the dictionary (Varma 1958) doesn't list them, and they don't exist in my dialect. It is possible that for some people they are existing forms.

21 For some discussion of words of this type see 2.3.4.

22 The consistency cannot be explained by saying it was due to usage, since these responses represent previously non-existent words.

23 This is not an idiolect since there seems to be more than one speaker of it; however, I do not know if it coincides with any geographical or social boundary; this would be an interesting topic to investigate.

24 Even though historically [inkar] is a loan word, I think in a synchronic grammar of Hindi it should perhaps be considered as native. It is a commonly used word, and speakers often try to regularize it to [iŋkar].

25 [mirč] is an interesting example because it is listed (in the dictionary) as coming from Sanskrit mərič revealing that Sanskrit apparently did not have a cluster in the word.

26 It is interesting that words like [trifla] are rendered with an [f]. The word is from Sanskrit 'tri+p^{h}əla' -- Sanskrit did not have [f]. When loans from English and Perso-Arabic introduced an [f] into Hindi, Sanskrit words with [p^{h}] were frequently pronounced with [f] due to hypercorrection. Since clusters with [p^{h}] are rather rare, it seems that the pronunciation with [f] in words like [trifla] remained in the language, such that today the rendering is [trifla] and not *[trip^{h}la] in Standard Hindi.

27 I am grateful to B. Pray for the latter two examples.

28 It is pronounced [šɛhɛr] by a general rule which I will not discuss in this thesis. This problem of Vh in Hindi phonology is very interesting but beyond the scope of this study.

29 Before liquids, etc., we do get non-homorganic nasals, e.g., [imli:] "tamarind", [əmru:d "guava", etc.

30 Narang and Becker point out that the generalization holds only before non-uvular stops; however I am not mentioning "uvular"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

since the dialect of Hindi being described does not have uvular stops.

31 The fact that my examples either are of short vowel followed by a homorganic nasal+stop, or long vowel followed by a homorganic nasal+voiced stop, is not fortuitous. See chapter 3 for details.

32 Most of the following examples are from Mehrotra.

33 I am grateful to John Ohala for pointing out these articles to me.

34 We don't find these initially in Western Hindi, which has monophthongized most of the original diphthongs, e.g., Eastern Hindi has [əu:rət] "woman" but Western Hindi has [ɔrət].

35 Srivastava gives this rule as (133):

Vowel laxing rule:

$$\begin{bmatrix} V \\ +\text{tense} \\ +\text{high} \end{bmatrix} \rightarrow [-\text{tense}] \;/\; ____ \left\{ \begin{array}{ll} + (C) & V \; [+\text{tense}] \\ = \text{ər} & \end{array} \right\}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 3

3.0. I mentioned in chapter 1 that it is possible, in a large number of cases, to derive the nasal vowels from underlying oral vowels. I will discuss this in detail in this chapter.

3.1. If one examines Hindi words one finds that the nasal vowels at the phonetic level are the result of three different processes: phonetic, syntactic, and phonological.

1. Nasalization of vowels due to phonetic processes.

A vowel is nasalized when it is preceded or followed by a nasal. Examples of this type of nasalization are given in (1).

(1) [nãm] "name"
[kãn] "ear"
[mãs] "month"
[nãk] "nose"

Writers such as Hussain (n.d.), Tiwari (1966), and Kelkar (1968), have suggested that there is a difference in the degree of nasalization in words in (1) and those in (2):

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(2) [mə̃] "mother"

[mə̃s] "meat"

[mɛ̃] "I"

The degree being greater in the examples of (2). Such a difference in the degree of nasalization in words has been noted in other languages as well (Westermann and Ward 1933). Tiwari suggests that if one looked more closely one could differentiate further degrees of nasalization. Perhaps this is what Kelkar refers to when he says "A nasal preceding the vocoid ...is less influential in affecting the nasalization of the vowel than a nasal following" (38). Thus he claims that the nasalization is stronger in [kə̃n] "ear" than in [nə̃k] "nose". That greater nasalization occurs before a nasal than following a nasal has also been found in English by instrumental means (J. Ohala 1971a). This would have to be incorporated in the rules that assign scalar values for nasalization at the phonetic level. (It should be mentioned that weak and strong nasalization contrast only after nasals, never before.) There are, however, very few minimal pairs which are potentially differentiated only by the degree of nasalization on the vowel. The pair [mə̃s] "month", [mə̃s] "meat" was mentioned above; in addition there is [mɛ̃] "wine" and [mɛ̃] "I" (the latter has stronger nasalization). These are the only pairs I have been able to find. Moreover it is not clear that this difference is manifested except when speaking the two forms in consciously contrasting manner, and many speakers openly admit they do not have the contrast. Thus the status of this difference in the degree of nasalization on vowels in Hindi seems to be very marginal.

Items such as those in (1) at least, will not have their vowels

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

marked [+nasal] in the dictionary. A phonetic rule will assign nasalization to oral vowels in the environment of a nasal:

(3) V --> [+nasal] / [+nasal]

This rule says that a vowel is nasalized when preceded or followed by a nasal.

2. Nasalization of a vowel due to syntactic processes.

The words in (4) are minimal pairs differentiated only by nasalization of this origin:

(4)	[hɛ]	"is"	[hɛ̃]	"are"
	[čəli:]	"she went"	[čəlĩ:]	"they (fem.) went"
	[čəle]	"they (masc.) went"	[čəlẽ]	"shall we go" (subj.)
	[lərka]	"boy"	[lərkõ]	"boys" (pl. obl.)

For the words in the second column we again would not need to mark the vowel as [+nasal] in the lexicon. This type of nasalization will be assigned to the vowels by the rules that spell out the phonological forms of grammatical formatives.[1]

3. Nasalization due to a phonological process:

This is nasalization in lexical items not covered by the previous two categories. There is no doubt that there is a contrast between nasal and oral vowels at the phonetic level, as can be seen from such pairs of words as those in (5):

(5) [sas] "mother-in-law" [sãs] "breath"

[bas] "bad smell" [bãs] "bamboo"

We are interested in finding out if nasalization for the vowels in these items has to be marked in their lexical representation or not. That is, is the nasalization of these vowels predictable by general

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

phonological rules or not? There have been varying treatments of this problem with writers either claiming that Hindi has a set of nasalized vowels contrasting with oral ones (e.g., Hussain n.d., Qadri 1930) or that all nasalized vowels can be derived from oral ones (Narang and Becker 1971). Moreover there even seems to be disagreement (as I will shortly show) on the data, i.e. whether certain morphemes have $\tilde{V}C$ (nasalized vowel +stop) or VNC (vowel+homorganic nasal+stop), in the latter case of course the vowel would be nasalized by rule (3).

Before discussing this issue it might be worth while to first look at the history of nasalized vowels in Hindi.

3.2. Historically the long nasalized vowels came about due to a phonological change in the development of old Hindi[2] from Middle-Indo-Aryan (MIA).[3] Word-medial consonant clusters were simplified with compensatory lengthening of the preceding vowel. If the cluster consisted of a nasal followed by a consonant, the nasal was deleted, and the preceding vowel was lengthened and nasalized.[4] Formally this can be represented as follows:[5]

(6) $V \begin{bmatrix} +\text{cons} \\ +\text{nasal} \end{bmatrix} C \Longrightarrow \begin{bmatrix} +\text{long} \\ +\text{nasal} \end{bmatrix} \quad \emptyset \quad 3$

1 2 3 1 2

This rule is usually illustrated by examples of the following kind (Grierson 1922, Misra 1967):

(7)

Skt.	Pkt.	Old Hindi	
čəndrə	čəndə	čãd	"moon"

Thus Misra says "In Old Hindi all nasalized vowels are long, and there are no short nasalized ones" (203).[6]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3. As mentioned, some writers treat particular morphemes as if they were made up of $\tilde{V}C$ (distinctively nasalized vowel+stop), and others as VNC (oral vowel (rather, non-distinctively nasalized vowel) plus homorganic nasal+stop). This confusion seems to have been present even in the earliest grammars. It will be useful to review the traditional literature briefly, on this matter.

The earliest grammar that I have been able to look at which has some remarks on nasalization is The stranger's infallible East-Indian Guide[7] by J. Gilchrist (1820). He lists various vowels of Hindi (or Hindoostani as he calls it) giving their approximate English pronunciation, and gives their nasalized counterparts as well. Interestingly, most of the examples of nasalized vowels that he gives represent inflectional endings, and thus are at the end of a word. In listing the consonants he treats n as a dental nasal, and also as a representative of various kinds of homorganic nasals.

He then says:

> ...n when attenuated, to what is called the french nasal, lapses into its concomitant vowels...in the whole of which the oscillation between n and n̲ [his symbol for nasalization] is arbitrary in the extreme, and the elision or assumption of the n̲ equally so, in the different dialects of hindoostan by particular classes of people, numbers of whom may conceive, that the nasal twang stamps their language as the sterling speech of certain cities, provinces or tribes, who affect this absurd embellishment, either as a mark of their haut ton or profound erudition...(34-35)

(It is interesting to note that the nasalized vowels had a prestige value even in those days--more will be made of this point below.) So, apparently by Gilchrist's time it was no longer true that only long vowels were nasalized (which is what Misra claimed to be true for Old Hindi). In his practice (e.g., in his vocabulary lists) he uses the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

symbol for nasalization only before h, s, w, and '#'. Thus he lists:

moonh	"mouth, face"
anten	"bowels"
hunsna	"to laugh"
panw	"foot"

Elsewhere he just has n regardless of whether short or long vowel precede it:

dant	"tooth"
zunjeer	"chain"
rungna	"to color"
kampna	"to tremble"

The question thus arises as to whether his transcription really reflects the pronunciation of his day. Since he mentions that the "oscillation between n and n is arbitrary" I think it is probable that he simplified matters by having n uniformly in most of the above words.

The next work mentioning nasalization is Kellog's Grammar of the Hindi language (1965), first published in 1875. This seems to be the first book written in English which lists the Devanagari symbol for nasalization (ँ) (called anunasika) along with that of homorganic nasal (ं) (called anusvar). He mentions that in many Hindi books, especially those published by foreigners, the sign for anusvar is used for both nasalization and homorganic nasal. In his book he also uses only the anusvar.

He notices that in words taken over from Sanskrit with little or no change the anusvar represents a homorganic nasal before a stop. But in "the case of words much corrupted from the Sanskrit, after a long

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

vowel, Anusvár, even before a mute consonant, denotes, not a consonant, but a nasalization" (14). For an example of this he gives the word for "moon" as "chánd" (i.e. čãd) (though later on when writing it in Devanagari he uses the symbol for anusvar only). He does not tell us what anusvar represents in the case of tadbhav words (those words originating from Sanskrit through historical changes) following a short vowel. He seems to claim that after long vowels the anusvar represented nasalization, i.e. the vowels were nasalized; however Gilchrist had not noticed any correlation between vowel length and nasalization, or homorganic nasal (and furthermore we don't know if short vowels were nasalized or not in Kellog's time). So, either Hindi went through further change between Gilchrist's and Kellog's time leading to no cases of long vowels followed by homorganic nasals, or Kellog's grammar reflects an earlier stage of the language (i.e. Old Hindi) and not the spoken language of his time. I tend to think the latter, because of passages such as the following:

> ...Anusvár stands in the place of an original nasal consonant, preceded by a short vowel;...The question has been raised, whether the Anusvár in such cases denotes a consonant or merely nasalization. But all the Pandits that I have been able to consult insist on the inorganic [sic] character of the nasal. (14)

If there was a prestige value attached to the nasalized vowels as Gilchrist seems to suggest, then since Pundits usually belong to the erudite class, they would claim the above. Since Pundits usually seem to consider the older stage of the language "purer", their grammars tend to reflect an older stage of the language, and not the spoken language of their time.

The next grammar we come to is that of H.C. Scholberg (first pub-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lished 1940, third edition 1955). Scholberg notes that anusvar "may represent simple nasalization" (7), or a homorganic nasal. His examples contain both short vowels followed by homorganic nasals (e.g., pəŋkh "wing") and long vowels followed by homorganic nasals (e.g., si:ŋg 'horn'). He notes that before sibilants anusvar takes "the sound of the Nasal of the Dental class" (8) but "where a long vowel before 's' is nasalized it seems to be simple nasalization", (8) and notes that the same holds for long vowels before h. Further on pg. 9 he says "...speakers...learn by use whether a simple nasalization is required or a nasal letter." Since Scholberg says in the preface that he has "put down what is current use among Hindi-speaking people" (xvi), we can take it that the Hindi he is describing had both long and short nasalized vowels, and both long and short vowels followed by a homorganic nasal--a situation which reflects what was probably also true at Gilchrist's time. This seems to suggest that Kellog's grammar reflects the Hindi of many centuries earlier. Since in the rest of his book Scholberg gives examples only in Devanagari, it is unfortunately not possible to determine whether a homorganic nasal or nasalization was used in many of the words he lists.

Of the grammars written in Hindi, Kishori Das Vajpeyi's Hindi shabdanushasan (1957) claims that the tendency of Sanskrit was one of anusvar but that of Hindi is anunasika, i.e. that of nasalization. He gives as examples (his examples are in Devanagari; I give here my transliteration) Skt. əŋguṣṭʰ, Hindi ə̃guːṭʰa for "thumb"; Skt. dənt, Hindi dā̃t for "tooth". Thus his examples indicate that he thinks Hindi has a tendency for nasalization involving both short and long vowels.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

He does not deny the existence of anusvar in Hindi, though, since he lists the word for "grape" with anusvar: əngu:r. Thus according to him both short and long vowels occur nasalized, as well as with homorganic nasals. He does not mention anything about the length of the vowel as conditioning whether nasalization or homorganic nasal would occur. He says that in writing about Hindi, people should not mix the signs for anusvar and anunasika, and I can but whole-heartedly endorse this.

The problem is that for convenience in writing, authors many times use the symbol for homorganic nasal (˙) even though according to the pronunciation they should have used the symbol for nasalization (̃). Since the readers are usually native speakers this practice doesn't lead to any confusion. However it does lead to difficulties for the linguist. For example the word for "moon" is written by some with the symbol for nasalization, and by others with that of homorganic nasal (and by some both ways). Does this reflect a difference in pronunciation or not? Since purists seem to prefer the nasalization people might insist that they have a nasalized vowel, when in reality their pronunciation never has only a nasalized vowel in the word in question but rather a homorganic nasal following the vowel. Similarly many Hindi speakers claim they really do distinguish the retroflex ʂ and palatal ɕ, being influenced by the conservative orthography (which preserves these separate fossilized forms from Sanskrit).

Kamta Prasad Guru in his Hindi Vyakaran (1962) gives a more detailed account of nasalization. Noticing the confusion in writing between anusvar and anunasika, he gives some rules for when the written

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

anusvar should be pronounced as nasalization. The rules he gives are (my translation):

(a) In native words in word-final position, the written anusvar is pronounced as nasalization.

(b) The anusvar due to grammatical phenomena is pronounced as nasalization.

(c) Anusvar after a long vowel is pronounced as nasalization.

No rules are given for short vowels. However in the examples he gives, short vowels seem to occur both nasalized as well as followed by a homorganic nasal, e.g., ə̃d^{h}era "darkness", ənd^{h}er "darkness, injustice". From this it would seem that the fact that Old Hindi had no occurrences of long vowels followed by homorganic nasal is being echoed here.

However since it is not at all clear that OH lacked short nasalized vowels (contrary to Misra's claim, cf. footnote 4) maybe that is why purists have not made any such claims for short vowels, and seem to treat short nasalized vowels as if they were quite natural to Hindi.

So, nothing very clear about vowel nasalization emerges after reading most of the works on Hindi phonology which I've reviewed so far. The exact phonetic character of anusvar and anunasika is disputed. And either these writers did not provide rules for correlating the nasalization on vowels and vowel length, perhaps because the data of their time did not lend itself to this, or, if they did write rules, they usually represented a fossilized Hindi, not the spoken Hindi of the times. It is understandable that they would write rules for an older form of Hindi since in India there is "prestige" attached to any-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

thing "older" and moreover the grammarians were no doubt guided in their task by the pundits who, following Panini, believed in preserving the "beauties" of the older language, trying to combat the "corrupting" tendencies of the present.

3.4. Among recent analyses of nasalization, Kelkar (1968) has given the most phonetically detailed one. There is a wealth of detail in his analysis but his style and exposition make for very difficult reading.

He gives two sets of contrasting words to show the difference between nasalized vowels and vowels followed by a homorganic nasal.[8]

(8) /bɪ^{n}d^{h}a/ [bɪ̃^{n}d^{h}a] "made a slit"

vs.

/hIndi/ [hI̥ndi] "Hindi"

and

/da^{n}t/ [dãt] "tooth"

vs.

/šant/ [š̥ant] "quiet, peace"[9]

He goes on to say that the contrast between nasalized vowels and homorganic nasals is an unstable one because the homorganic nasals are "chiefly illustrated by Sanskrit, Persian, and English loanwords" (33) and he gives examples of this. He also notes (and here I paraphrase):

(a) Nasalized vowels are more frequent than vowels followed by homorganic nasals.

(b) Among nasalized vowels, long nasalized vowels are more common than short nasalized vowels.

(c) Among vowels followed by homorganic nasals, short vowels followed by homorganic nasals are more frequent than long vowels fol-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lowed by homorganic nasals.

He then gives examples for these in two groups, with a note that items in the second group tend to have familiar variants with homorganic nasals. I will list the examples:

(9) Group I: (vowels followed by homorganic nasals)

gand^{h}i	(surname)
manta	"accepts"
sUnta	"listens"
čəndel	(name of Rajput clan)
k^{h}ɔnča	"platter used by peddlers"
b^{h}anǰa	"sister's son"
səmǰona	"to put together"
g^{h}əmə̤nd̤	"conceit"
ə̤nt̤sə̤nt̤	"this or that, unimportant"
bed^{h}əŋga	"clumsy"
guŋga	"dumb"

(10) Group II: (V̰C)

sə^{n}b^{h}alna	"to take care of"
sa^{n}p	"snake"
ta^{n}ba	"copper"
ǰə^{n}čta	"becomes, suits"
sI^{n}čai	"irrigation"
pu^{n}ji	"capital"
pu^{n}č^{h}	"tail"
ča^{n}di	"silver"
bə^{n}t̤a	"was divided; cause to divide" (imperative)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ʧə^{n}ʈ	"cunning"
g^{h}u^{n}ʈ	"mouthful (of liquid)"
ba^{n}ʈ	"divide, distribute"
U^{n}gli	"finger"
rə^{n}gana	"to color" (tr.)
sI^{n}g^{h}ara	"water chestnut; Trapa bispinosa, Roxb."
ɖ^{h}o^{n}gi	"hypocrotical"
kəru^{n}+ga kəru^{n}ga	"I shall do"

3.5. For most Hindi speakers however, many of the above examples are not transcribed accurately. For example, while the words for "silver" or for "finger" are written in Devanagari with nasalized vowels, they are pronounced [ʧ̃andi:] and [ũŋgli:] (the vowel being nasalized in the vicinity of a nasal by rule (3)). (This is supported by instrumental evidence given below.) Kelkar's examples need to be rewritten as in table 22.[10]

Table 22

V:NC	Ṽ:	ṼNC	V̰
[tamba]	[sãp]	[səmbhalna]	[jə̃ʧta]
[pu:ñ̃ji:]	[pũ:ʧ^{h}]	[ʧə̃ɳʈ]	[sɪ̃ʧai:]
[ʧ̃andi:]	[g^{h}ũ:ʈ]	[uŋgli:]	[bə̃ʈa]
[ɖ^{h}oŋgi:]	[bãʈ]	[rəŋgana]	
[gand^{h}i:]		[siŋg^{h}ara]	
[b^{h}añ̃ja]		[ʧ̃əndel]	
[gu:ŋga]		[səñ̃jona]	
[kəru:ŋga]		[g^{h}əmə̃ɳɖ]	
		[ənʈsə̃ɳʈ]	
		[bəg^{h}əŋga]	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(I haven't marked the nasalization on the vowels in words in columns 1 and 3, but these would of course be nasalized by rule (3). I follow this same practice in the transcriptions to follow.) Items in column I all have long vowels followed by homorganic nasals, those in column II have long nasalized vowels, column III exemplifies short vowels followed by homorganic nasals, and column IV short nasalized vowels.

Looking at these words and many others of the same type (which are given in tables 23, 25, and 26 below) we find that a distinct pattern emerges, leading us to make the following generalization for Hindi:

1. After short vowels, the tendency of Hindi today is to have homorganic nasals (rather than simply a short nasalized vowel).

2. In [+native] words:

(a) After long nasalized vowels we find only voiceless stops not voiced ones.

(b) A voiced stop may be preceded by a long vowel+homorganic nasal, but not just by a nasalized vowel.

Thus if it is true that at one point Hindi had long nasalized vowels only and not long vowels followed by homorganic nasal, i.e. $\tilde{V}$:C and not V:NC, then it has gone through a historical change. Therefore Kelkar's generalization that homorganic nasals are chiefly found in Sanskrit, Persian, and English loanwords no longer holds. The words in his first group of examples (given under (9)) have the structure VNC (vowel followed by homorganic nasal) and most of them are not loan words. To accurately describe the incidence of homorganic nasals, it is neccessary to be more specific about the nature of the preceding vowel (short or long) and the following stop (voiced or voiceless).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A number of the items in column I and III of Table 22 are listed in Kelkar's group II (10), and for these he claims that they have "familiar variants" with homorganic nasals (34). However these are certainly not "familiar variants" at all; they are the standard spoken forms used by native speakers. These are only "familiar variants" from the point of view of purists who would have Hindi spoken today as it was centuries ago (actually I doubt if even the purists really use these old forms in their everyday speech).

Items in column II are all pronounced as Kelkar has them, i.e. with nasalized vowels. These follow my generalization 2 (a). And to these we can add his example [dãt] "tooth" (8). Items in column I exemplify my generalization 2 (b). Items in column IV are exceptions to my generalization 1 and I will discuss them shortly.

In table 23 I will first give more examples of generalization 1, and discuss exceptions to it. These are all words that would fit into column 3 of table 22 (as mentioned above, the vowel next to the nasal in these words would be nasalized by rule (3)).

Table 23

[ɖəŋk]	"sting of a wasp or scorpion"
[ɖəŋtʰəl]	"stalk, stem"
[təmbaku:]	"tobacco"
[gənda]	"dirty"
[sindu:r]	"vermilion"
[səntra]	"orange"
[səndu:k]	"a box"
[ləmba]	"tall"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[ləŋg̣ra]	"lame"
[əm̃̆ji:r]	"fig"
[əŋgu:r]	"grape"
[kən̆̃čən]	"gold"
[d̤ən̤da]	"a stick"
[əmbiya]	"raw mango"
[səmb^{h}əv]	"possible"
[mənzil]	"stage, story"
[mən̆̃č]	"platform"
[luñ̆j]	"disabled"
[kəŋgən]	"bracelet"
[əŋgi:ṭ^{h}i:]	"an iron pot to keep fire in"

It is interesting that Kelkar notices that "all voiced stops are prenasalized after V^{n}....that is, they have a short nasal consonant as an on-glide" (24). He thus observed the facts but did not see the implications.

As I mentioned earlier, the words in column IV of table 22 are exceptions to generalization 1 since they have short nasalized vowels. However they are exceptions of a particular type---all these examples may be derived from a corresponding form with a long nasalized vowel. Thus the words in column 1 of table 24 could be said to have as their underlying forms, morphemes similar to those in column II:[11]

Table 24

[j̃ə̃̆ta]	"becomes, suits"	[j̃ăčna]	"to investigate"
[sĩ̆̃ta]	"gets irrigated"	[sĩ:̆na]	"to irrigate"
[bə̣̃ta]	"was divided"	[bə̣̃̆na]	"to distribute"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[ʈə̃kna]	"to be stitched"	[ʈə̃kna]	"to stitch"
[ʄʰə̃sna]	"to trick, or fraud"	[ʄʰə̃sna]	"to deceive"
[hə̃kvana]	"to cause to drive"	[hə̃kna]	"to drive"
[kə̃pvana]	"to cause to tremble"	[kə̃pna]	"to tremble"
[sə̃pera]	"snake charmer"	[sə̃p]	"snake"
[ũčai]	"height"	[ũ:ča]	"high"

Kelkar's example [bĩⁿdʰa] also belongs in Table 24. This is not a commonly used word, and the form from which it is derived, [bi:ndʰa], is even less common; moreover I pronounce the derived form [bĩndʰa].

Interestingly since the short nasalized vowel forms go against the general tendency of the language, many times the nasalization is lost, i.e. native speakers seem to treat these as exceptions. This is especially true in forms where the related member with a long nasalized vowel is not commonly used any more, or does not exist in modern Hindi at all. Examples are given below:

(11) [ɖə̃sna]~[ɖəsna] "to bite, to sting" ([ɖə̃s] not in use any more)

[ɖʰə̃kna]~[ɖʰəkna] "to cover" ([ɖʰə̃kna] not commonly used)

[ʄʰə̃pəkna]~[ʄʰəpəkna] "to fan"

[ə̃ʈna]~[əʈna] "to be contained in" ([ə̃ʈna] not in common use)

[cʰə̃ʈe]~[čʰəʈe] "separated" (from [čʰə̃ʈna]'to separate')

[sĩčai:] [sičai:] "irrigation" (from [sĩ:čna]'to irrigate)

Kelkar also seems to have noticed this. He mentions that short nasalized vowels followed by a voiceless stop tend to have a 'familiar variant' in short vowel followed by a voiceless stop (35). As examples he cited [ʄə̃čta]~[ʄəčta] [sĩčai:]~[sičai:] and [bə̃ʈa]~[bəʈa].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

His observation that this alternation is found only when the following stop is voiceless is explained by my analysis--if the form with the long vowel was followed by a voiced stop we would have got a long vowel followed by a homorganic nasal, thus the derived form with the short vowel would keep the homorganic nasal; it is only the forms with a following voiceless consonant that can have a short nasalized vowel in their derivatives. Kelkar also mentions that so^{n}b^{h}alna "to take care of" has an allegro variant sə̃m^{h}alna, that is interesting since this is another way in which short nasalized vowel $\tilde{\text{V}}$ is eliminated (as it is contrary to the tendency of the language). I pronounce this word as [səmb^{h}alna] and thus it is perfectly regular.[12]

There is another group of words which are not derived from forms with long vowels, which are also exceptions to generalization 1; however there are not very many words of this type. I give below all such words I could find:

(12) [pəhə̃̆cna] "to reach"
[d^{h}ə̃sna] "to get stuck in"
[hə̃sna] "to laugh"
[hə̃sli:] "collar bone"
[hə̃siya] "sickle"
[gə̃ṛasa] "an axe for cutting fodder"

Of these [hə̃sna] is an interesting example for it forms a contrast with [həns] "swan". [həns] is a [-native] word, a Sanskrit borrowing, thus it might appear that we could make a generalization to the effect that [+native] words have only short nasalized vowels before s (and not vowel+homorganic nasal), but there are enough native words like

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[p^hunsi:] "pimple" which prevent us from doing so. Moreover this would be opposite to the general tendency of the language which prefers short vowel followed by homorganic nasal. We find further support for treating these as exceptions from the fact that some speakers (mostly uneducated speakers) say [hʊ̃nsi:] for [hʊ̃si:], thus trying to make it regular. (Narang and Becker (1971) also report the regularized form [hʊsi:].)

I turn now to my generalization 2. As Kelkar rightly observes there are more examples of long nasalized vowels than of long vowels followed by homorganic nasals; however, the cases of long vowels followed by homorganic nasals are much more in number than those of short nasalized vowels. The reason is that the short nasalized vowels are the exception in the language, but long vowels followed by homorganic nasals are not. That there are fewer cases of long vowels followed by homorganic nasals, than there are of long nasalized vowels, is a matter of chance. As I have mentioned, the cases of long vowel followed by homorganic nasal occur when the following consonant is voiced, and thus they are not exceptions but are perfectly regular forms.

Kelkar's example kəru^n+ga "I shall do" (pg. 34) provides an additional piece of support for my statement that the voicing of the following consonant determines whether there is a nasalized vowel or a homorganic nasal. Most analyses list the verb 'to do' in its subjunctive form and add the suffix -ga to get the future forms thus:[13]

kər	+	ũ:	+	ga
verb stem		subjunc. marker		future marker

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But on the phonetic level we get [kərũ:ŋga] with a [ŋ]. This is accounted for by my generalization. Examples involving the future tense are the only ones I have been able to find of the generalization holding across morpheme boundaries. (I should also mention that an alternate analysis would be to treat [kəru:ŋga] etc. not as /kər+ũ:+ga/, but rather as /kər+ũ:ŋga/.)

The following are some more examples of generalization 2 (a) (these would fit in column II of table 22).

Table 25

[bẽt]	"cane, thin stick"
[sĨ:k]	"fine twig of a plant"
[ṭãka]	"a stitch"
[kẽkṛa]	"crab"
[hĨ:k]	"bad smell"
[sãp]	"snake"
[ãčəl]	"corner of a saree"
[sĨ:čna]	"to irrigate"
[ãk^{h}]	"eye"

Since in many loanwords we do find a homorganic nasal before a voiceless stop after a long vowel (examples are given below in (13)), I have restricted generalization 2 to [+native] words.

(13)	[šant]	"peaceful, quiet" (Sanskrit)
	[bæŋk]	"bank" (English)
	[k^{h}əzañči:]	"treasurer" (Persian)
	[dant]	"suppressed, one who has conquered his senses" (Sanskrit)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[prant] "province" (Sanskrit)

But I have so far not found any cases of $\tilde{V}$:C
[+voiced]

The following are examples of generalization 2 (b) (these would fit in column 1 of table 22):

Table 26

[gend]	"ball"
[čand]	"moon"
[si:ŋg]	"horn"
[taŋga]	"a horse-drawn carriage"
[kaŋgra]	"a particular valley in India"
[hi:ŋg]	"Asafoetida"
[tamba]	"copper"
[samb^{h}ər]	"a kind of Indian antelope, a kind of dish"
[čandi:]	"silver"
[and^{h}i:]	"a storm"
[j$^{\check{h}}$i:ŋgər]	"a cricket"

3.6. It remains to provide instrumental evidence for my assertion that words such as that for "moon" have a homorganic nasal, and not simply a long nasalized vowel, i.e. [čãnd] not *[čãd]. This is done in the spectograms in Figures 7-11. They show words pronounced by a male native speaker of Hindi who grew up in Lucknow and Delhi. These words were spoken in the frame /vo_____aya/, in citation style. The recording was done under ideal recording conditions in the Phonology Laboratory, University of California, Berkeley.

The spectograms in the first two figures show the clear contrast between [dant] "supressed", and [dãt] "tooth". Figures 10 and 11 show

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 7. (Calibration applies to all succeeding spectrograms, too.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

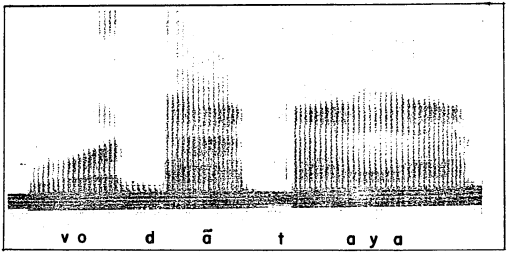


Figure 8.

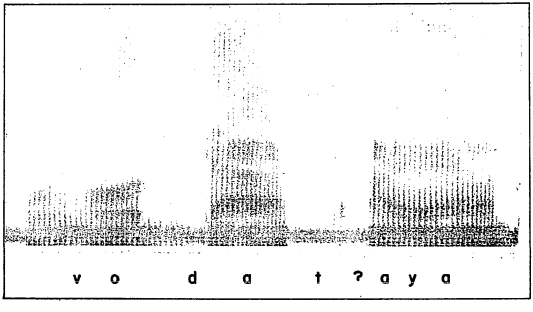


Figure 9.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

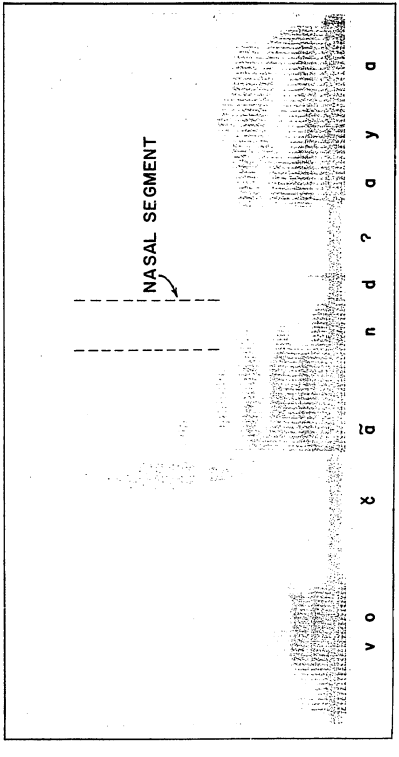


Figure 10.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

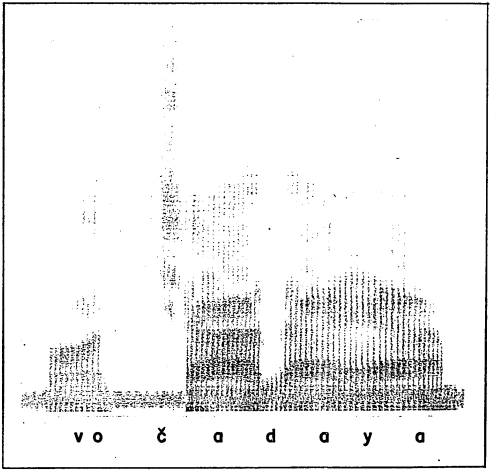


Figure 11.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

very clearly that [čand] "moon" has a homorganic nasal.

I also conducted an informal test where I pronounced the word for "moon", "cricket", and "courtyard" both ways, i.e. with just a long nasalized vowel, and with a long vowel followed by a homorganic nasal, i.e. [čād] and [čānd], [ãŋgən] and [ãgən], etc. I asked 5 subjects[14] which pronunciation they preferred. They all preferred the pronunciation with the homorganic nasal, i.e.:

(14) [čānd] "moon"
[jʰĩ:ŋgər] "cricket"
[ãŋgən] "courtyard"

3.7. I will now discuss how I propose to incorporate the observations made so far in the grammar.

Generalization 1 stated that the preferred state in Hindi is that short nasalized vowels do not occur unless they are in the vicinity of a nasal. This is incorporated in the following SQC (this was given earlier in chapter 2):

SQC 52 if: ([-nasal]) $\begin{bmatrix} +\text{syll} \\ -\text{long} \end{bmatrix}$ ([-nasal])

$\Downarrow$

[-nasal]

The above SQC states that short vowels are [-nasal] if preceded or followed by [-nasal] segments. In the vicinity of nasals of course, rule (3) will nasalize the vowel. The condition will apply at both the lexical and the phonetic level, i.e. will be an "everywhere" constraint. Morphemes such as [hə̃si:] "laughter", etc. will be marked as exceptions to the above SQC. We cannot restrict the SQC to just the lexical level for the following reasons: As mentioned earlier exceptions such as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[sĭč] "to be irrigated" are presumably derived from [sĭ:č] "to irrigate" and thus would not have a short nasalized vowel at the lexical level. If SQC 52 simply applied to the lexical level then there would be no reason to treat [sĭč] as exceptional in any respect (because the word would not exist at the lexical level with a short vowel). However, as mentioned earlier, native speakers do give evidence of treating such morphemes as exceptions, since they often regularize them. Therefore we have to state that SQC 52 applies at both levels and if the P-rule[15] that would shorten the vowel in deriving [sĭč] from [sĭ:č] produces a form that violates the SQC, that form is also treated as irregular. There are certain other examples, all involving $\underset{\sim}{\check{V}w}$, which would be treated as exceptions by the above SQC, however I have no evidence of their actually being treated as exceptions by native speakers such that they would attempt to regularize them. These examples are:

(15) [kŭ̃wər] "prince"
[čə̆wər] "a flapper made of the tail of a wild ox"
[b^{h}ə̆wər] "whirlpool"
[gə̆wa] "lose" (imp.)
[kə̆wəl] "lotus"
[sə̆wər] "decorate" (imp.)

Perhaps a few more such words exist but not very many more. I will treat them as exceptions for descriptive purposes, however I cannot claim psychological reality for this. The short vowels in the vicinity of a nasal would be nasalized by rule (3).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The following SQC will capture generalization 2:

SQC 53 if: $\begin{bmatrix}+\text{syll}\\+\text{long}\\+\text{nasal}\end{bmatrix}$ $\begin{bmatrix}+\text{cons.}\\-\text{sonor.}\end{bmatrix}$

$\Downarrow$

then: $\begin{bmatrix}\left\{\begin{bmatrix}+\text{cont.}\end{bmatrix}\right.\\\left\{\begin{bmatrix}-\text{cont.}\\-\text{voice}\end{bmatrix}\right.\end{bmatrix}$

The above condition states that after long nasalized vowels if the following segment is a non-sonorant, then it is either a fricative, or a voiceless stop. The only voiced fricative is [z]; I do not have any examples of words involving long nasalized vowels and [z], however, I intuitively feel that a nonsense word like bãz, would not be considered odd, and so I am not stating that the fricative has to be voiceless too.

Morphemes such as [čãnd] I propose entering as /čand/. The derivation would be as follows:

/čand/

↓

[čãnd] by rule (3)

In the case of morphemes such as [dãt] "tooth", there are two possibilities:

(a) we posit the underlying form to be the same as the phonetic, i.e. /dãt/.

(b) We posit an abstract underlying form /dant/, from which we derive [dãt] by deletion of the nasal, after it has nasalized the vowel.

The first solution of course claims that there are nasalized vowels at the lexical level. I am adopting solution (a)--therefore SQC 53 mentions nasalized vowels. Regardless of which of the above two solutions one adopts one has to admit nasalized vowels at the lexical

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

level in a grammar of competence, (and I do not mean in the form of exceptions) because of the following types of words:

Table 27

[d^{h}ʊ̃ã]	"smoke"
[k^{h}ɔ̤rãʊ̃:]	"wooden slippers"
[hã]	"yes"
[b^{h}ɔ̃]	"eyebrow"
[b^{h}ɔ̃ra]	"bumblebee"
[yəhã]	"here"
[kəhã]	"where"
[wəhã]	"there"
[yõ]	"in this manner"
[kyõ]	"why"
[sãĪ:]	"lord, master"
[kwãra]	"bachelor"
[bəyã]	"left"
[dəyã]	"right"
[pãw]	"foot"
[sãwla]	"swarthy"
[gãõ]	"village"
[gehʊ̃:]	"wheat"
[bãh]	"arm"
[satwã]	"seventh"
[pãčwã]	"fifth"

Positing sequences of underlying oral vowel+nasal, which nasal is deleted by a rule, would be completely ad hoc, and I do not think is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

justified even in a descriptive grammar. I will discuss this point in more detail below.

Thus I have chosen to represent [dãt], etc., with underlying nasal vowels. This is not to say that we cannot derive [dãt] from an underlying /dant/; for example, we could posit a rule which says:

(16) $\begin{bmatrix}+\text{syll}\\+\text{long}\end{bmatrix}_1 \quad \begin{bmatrix}-\text{syll}\\+\text{nasal}\end{bmatrix}_2 \quad \begin{bmatrix}-\text{son}\\-\text{cont}\\-\text{voice}\end{bmatrix}_3 \Longrightarrow 1 \quad \emptyset \quad 3$

2

This would give us the correct result if we restricted the rule to [+native] vocabulary, and if we ordered it such that it applied after the vowel was nasalized by rule (3). Thus we could have the following derivation:

/dant/
↓ native
/dãnt/ by rule 3
↓
[dãt] by rule 16

However, positing the underlying form /dant/ for [dãt] amounts to positing an abstract underlying form; in order to posit such abstract forms we would need evidence of the psychological reality of such an underlying nasal. The only evidence of this type that I have found is the following:

(a) The word for "meat" [mãs] has a few alternating forms where a nasal shows up, e.g., [mãnsahari:] "non-vegetarian" presumably from /mans+ahari:/ "meat+eater"; and [mãnsəl] "muscular" presumably from /mans+al "meat+suffix".

(b) Dixit (1963) notes the following sort of assimilation and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

elision in rapid speech:

normal style	rapid style
[pā̃č] [sɔ] "five+hundred"	[pā̃nsɔ]

One possible way of accounting for this latter case is to say that when the stop was deleted the underlying nasal showed up. It is evidence of this sort that should be looked for; Dixit gave only this one example. The evidence for the "abstract" underlying nasal in these words is rather sparse and so I have not adopted this analysis.

3.8. Before ending this chapter, let me briefly review two recent analyses of this problem in the generative framework, namely that of Srivastava (1969) and Narang and Becker (1971).

Srivastava notes the existence of a few related words such as:

dɔnt/dā̃t	"tooth"
b^{h}ɔŋg/b^{h}ā̃g	"hemp"
kɔmp/kā̃p	"trembling"
čɔndr/čā̃d	"moon"

and on the basis of these advocates setting up the same underlying forms for them, e.g., /dVnt/, deriving the phonetic representation for the second member of the pairs by his rule 26[16] which is almost identical to the one I have given as rule (6). This is a historical rule and there is no evidence of its psychological reality for a native speaker of modern Hindi. The word pairs that Srivastava mentions (given above) are not alternating forms in the usual sense of the word. They are doublets: without any question they have a common origin---a millenium or more ago---but entering the language by two entirely different routes. The first member of each of these pairs was borrowed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

from Sanskrit in modern times. The other member of these pairs ([dãt], etc.) are tadbhava forms, i.e. they have gone through historical changes in the development of Hindi from Sanskrit. Moreover the first member of these pairs only exist in the vocabulary of certain speakers and enter their vocabulary only after they receive formal education. The true alternating forms of these native Hindi words are phonetically quite different in that there is no homorganic nasal segment in them:

[kãpna]	"to tremble"	[kəpvana]	"to cause to tremble"
[dãt]	"tooth"	[dəṭu:n]	"toothbrush of fibrous twigs" (interestingly, nasalization is lost)
		and	
		[dəṭɛl]	"one with big teeth"

(Deriving [dænt] and [dãt] from the same underlying form is similar to, though perhaps less drastic than, positing a common underlying form for English tooth and dental, or queen and gynecology. These words do have a common origin—in Proto-Indo-European—but they have entered the language via different paths; the first members of the pairs having been in the language since Anglo-Saxon times, and the second being relatively recent borrowings or formations from Latin or Greek.)

Moreover, word pairs of the dænt/dãt type are few. There are a rather large number of words in Hindi today with long nasalized vowels for which the older Sanskrit forms with short vowel followed by homorganic nasal have not been borrowed. For these forms it is unclear whether or not Srivastava would posit underlying forms with short vowel+

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

homorganic nasal, deriving the long nasalized vowel by his rule 26. If he does, then it is not obvious how he will get the phonetic forms of words which have a short vowel followed by a homorganic nasal (i.e. forms listed in table 23).

Among the points Narang and Becker make about nasals and nasalization in Hindi are:

(a) "the nasalized vowels are best derived from underlying sequences of vowel+nasal consonant." (665).

Plus the following two generalizations which are to apply at "the level of systematic phonetics" (653).

(b) "While long nasalized vowels abound in Hindi-Urdu, short nasalized vowels do not occur except in a small number of words..." (654). They note four types of exceptions.

(c) "Within a morpheme, long vowels never occur before consonant clusters of which the first member is a nasal consonant. Exception: [s̆ãnt] 'quiet'." (654).

Point (a) is implemented by positing underlying V:+nasal and deriving long nasalized vowels with the following P-rule (654):

(17) $\begin{bmatrix}+\text{syllabic}\\+\text{tense}\end{bmatrix}$ [+nasal] C $\Longrightarrow$ [+nasal] Ø 3

1 2 3 1 2

and they also posit a redundancy rule which states that there are no nasalized vowels in the lexicon (665):

(MS1) [-consonantal] $\longrightarrow$ [-nasal]

Thus they have the word for "moon", which they give as [č̆ãd] as derived from underlying /čaNd/ by (17). (I have some questions regarding the accuracy (or generality) of the phonetic transcription [č̆ãd], but we

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

can ignore this for the present.) This analysis amounts to setting up abstract underlying forms, whose "psychological reality" must be supported by good evidence. The evidence they do provide for deriving nasalized vowels from underlying oral vowel plus nasal consonant, is not very convincing.

The first reason they cite is that "...a substantial savings can be achieved in the lexicon by adopting an analysis which derives the nasalized vowels from underlying sequences of vowel plus nasal consonant" (666). This analysis would make the lexicon less "costly" to describe.

Firstly it is necessary to remember that our aim is to characterize a speaker's competence; if, in order to do this more segments need to be specified in the lexicon, then that is what will have to be done. We can always write fancy rules to get rid of supposedly undesirable or "costly" characteristics of the lexicon, but do these rules have any psychological reality? It has been claimed that an appropriately developed evaluation metric would correctly assign the least "cost" to a grammar that best characterizes the native speaker's competence. We are far from that stage now. At present we are still at the point of refining the evaluation metric so that it does choose a grammer that also reflects the speakers competence. To write a grammar today according to the evaluation metric, is to me, trying to tailor the data to fit the theory, rather than the other way around.

As justification for getting rid of nasalized vowels in the lexicon they say (footnote 8): "this practice would be entirely consistent with all previous works in generative phonology." It is possible that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

this type of analysis of nasalized vowels has worked successfully for some languages before, but I am not aware that this in itself guarantees its success in an analysis of Hindi.

Another piece of supporting evidence they cite for their analysis of setting up vowels followed by a homorganic nasal in the underlying structures and deleting the nasals by a P-rule is that their ə-deletion rule treats these words as if they had clusters. For example they say that when the plural ending õ is added to the word for "courtyard" which they transcribe as [ãgən], the [ə] is not deleted. They claim that the ə-deletion rule does not apply if a cluster precedes the [ə]. As I will show in chapter 4, this is not true: in words such as [jəŋgəl] "forest", [sundər] "pretty", etc., the [ə] is deleted when the appropriate ending is added, thus giving [jəŋli:]~[jəŋgli:], [sundri:] etc. Similarly the [ə] in [ãŋgən] "courtyard" is deleted in the speech of the majority of speakers in the derived form [ãŋgnõ] (see chapter 4 for further discussion on this). Moreover it seems that for most speakers the form is [ãŋgən] and not [ãgən] (cf. the informal test reported on pg. 195). As mentioned earlier, before voiced stops we do not get long nasalized vowels without a homorganic nasal following the vowel.

Finally, they say that the stress placement rule requires this analysis. However they do not reveal this stress rule to the reader.

In regards to getting rid of nasalized vowels in the lexicon for words like [dʰuã] "smoke", the authors say that they can do this by "the addition of two rather simple, if somewhat ad hoc, rules." (pg. 656). I had thought that in the generative dictionary "simple"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and "ad hoc" were opposites. They give their simple, if ad hoc rule as (657):

(18) $\left\{\begin{array}{l}[+\text{syllabic}] \\ [\ ] + ([-\text{consonantal}])\end{array}\right\}$ $\left[\begin{array}{l}+\text{syllabic} \\ +\text{tense}\end{array}\right]$ [+nasal] #

1 2 3 4

$\Longrightarrow$ 1 [+nasal] $\emptyset$ 4

2 3

(This rule, they say, would apply to the underlying form they give as /$d^{h}\underline{\underline{u}}\tilde{a}n$/[17] to yield [$d^{h}\tilde{u}\tilde{a}$]. This rule would be considered undesirable by any standards. Notational devices are not intended to be merely abbreviations, they are meant to be used to express some significant generalization. (There is no psychological evidence, or evidence of alternating forms for setting up a vowel followed by a nasal in the underlying form of words like [$d^{h}\tilde{u}\tilde{a}$]. True, a rule that nasalizes a vowel in the environment of a nasal is more natural than, say, one like V$\rightarrow\tilde{V}$ /--#, but why do we have to write a rule at all to predict the nasalization in these words? The fact is, that it is <u>not</u> predictable, and these words should be entered in the lexicon with nasalized vowels.

Their generalization (c) is difficult to accept on factual grounds. Narang and Becker claim that they are describing "the 'common core' language generally spoken by educated Hindus and Muslims in an area... that includes the states of Delhi, Uttar Pradesh...." (footnote 1). From this quote it would seem that they are describing what I have called Standard Hindi in the introduction, even though they do not use this term. However from the phonetic data they give it becomes apparent that they are certainly not describing the dialect of the majority of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

native speakers of Hindi.

I have lived in Lucknow, Banaras and Delhi and have not come across any native speakers of Hindi for whom Narang and Becker's generalization (b) holds (there well might be such speakers but they are not the majority). As I have demonstrated by the spectograms given earlier the majority of speakers render words such as the one for "moon" with a long vowel followed by a homorganic nasal+stop, e.g., [c̃ãnd] and similarly all the words in table 26. Unfortunately Narang and Becker do not give any information on their informants, nor do they give any instrumental data to support their generalization (c).

Regarding generalization (b), while I generally agree with Narang and Becker that short nasalized vowels are not the general case in Hindi, and that words such as [hə̃sna] "to laugh" should be treated as exceptions, I do not agree with their treatment of certain other words which they cite as being apparent exceptions to this generalization.

One group of seeming exceptions they list are words like [mɛ̃hdi:], [mɛ̃hga], [lɛ̃hga]. The majority of speakers render these as [mɛ̃hɛ̃ndi:] [mɛ̃hɛ̃ŋga] [lɛhɛ̃ŋga] etc. (the vowel being nasalized in the environment of a nasal) and thus they are not exceptions at all. Moreover for the dialect that does indeed have the pronunciations Narang and Becker list, it is better to list these as exceptions rather than to treat them as underlyingly /mAhNdII/, /mAhNgAA/ etc. as Narang and Becker do (they say that a modified rule (17) will apply to yield [mɛ̃hga] etc. (pg. 662)). (They propose treating all long vowels as underlying diphthongs---an ad hoc and most questionable proposal.)

Another group of exceptions to generalization (a) which they note

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is words such as [sĭ̆čna] "to be irrigated". They correctly observe that these are examples of a different type, since they are derived from [sĭ̆:čna] "to irrigate" etc. (I discussed such examples above.) However, Narang and Becker claimed that their generalization can be made at the systematic phonetic level (pg. 653). But since morphemes such as [sĭ̆čna] do exist at the systematic phonetic level, I fail to see how such morphemes are not exceptions. That such forms do not exist at the systematic phonemic level is really irrelevant for a generalization about the systematic phonetic level. It is true that these forms are not of themselves evidence against Narang and Becker's claim that nasalized vowels can be derived from oral ones; and it is also true that [sĭ̆čna] etc. being derived from [sĭ̆:čna] etc. 'explains' how these short nasalized vowels came about at the phonetic level; however, the fact remains that if Narang and Becker's generalization (b) is stated for the phonetic level, these are exceptions to it. Moreover as I have argued earlier, they should be treated as exceptions since native speakers seem to treat them as such.

Thus I do not accept Narang and Becker's proposal that "nasalized vowels in Hindi are best derived from underlying sequences of vowel plus nasal consonant" (666).

The analysis I have given above of vowel nasalization in Hindi, accounts for more of the data, and, as an index of its generality and validity, allows us to explain otherwise puzzling aspects of Hindi phonology and phonetics, e.g., the fact of the occurrence of homorganic nasals after long nasalized vowels before voiced consonants, instead of simply long nasalized vowels (e.g., [čãnd] "moon" [kərũ:ŋga] "I will

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

do"); the loss of nasalization in the derivation of certain words, e.g., words such as [siˇč] "to be irrigated".

To summarize briefly: The generalization can be made for Hindi that short vowels are not nasalized unless in the vicinity of a nasal. I incorporate this generalization in the grammar by an SQC which states the constraint, and a P-rule which nasalizes vowels in the vicinity of a nasal. I propose entering morphemes such as [hə̃si:] "laughter" as exceptions. I also propose that in the case of long vowels it is not the case that we have only [Ṽ:C] but rather both [Ṽ:C] and [Ṽ:NC] depending on the voicing of the stop. This is incorporated in the grammar by an SQC. I propose treating morphemes with [Ṽ:C] as containing underlying long nasalized vowels, and morphemes with [Ṽ:NC] as containing underlying /V:NC/ (by 'N' here I simply mean 'nasal' and not archisegment) with rule (3) nasalizing the vowel to obtain the correct phonetic output.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Footnotes

1 It is not clear exactly where in a grammar such rules will be. Chomsky and Halle (1968:10-11) seem to imply that this will be done by readjustment rules. Schachter and Fromkin (1968) suggest a second lexical look up. If the latter is adopted then, of course, these nasalized vowels would be in the lexicon since the phonological shape of the plural would be $\begin{bmatrix} V \\ +nasal \end{bmatrix}$. However, for the purposes of this thesis we can adopt Chomsky and Halle's suggestion.

2 This change affected most of the New Indo-Aryan languages (exceptions are Punjabi, Lahanda, and Sindhi)

3 This process could be seen as early as 1000 A.D. during the Apabhramsa period, or late MIA period. According to U. N. Tiwari (1961), by the 15th century Hindi had emerged as a distinct language, and this process was well established.

4 The exact environment for this rule is not quite clear. For instance do we need to have a [-long] condition on the first vowel? Does the rule apply vacuously if the vowel is long (the rule as I have formulated it implies this) or does it get blocked? If the latter, then we would have to specify the vowel as [-long] in our rule. This problem is not discussed in the literature (Beames (1970), Kellog (1965), Misra).

Misra also gives two additional conditions to be met before this rule could apply:

(a) that the syllable in question (i.e. the (C)VNC- in question) be the initial syllable, i.e. #(C)VNC-.

(b) that it be followed by a short vowel and not a long one.

Beames and Kellog do not mention (a) and regarding (b) Kellog says "But when long and heavy syllables follow, and in a few other cases, this compensatory lengthening of the vowel is sometimes neglected" (62). The 'sometimes' in the quote seems to indicate that the condition was not as invariable as Misra seems to suggest, and therefore I have not mentioned it in the rule.

Misra also clearly indicates that when the above conditions are not met, the nasal is not lost, i.e. there is no change. Beames and Kellog (usually cited as authoritative sources on the history of Hindi) give no clear statement on this point. Beames says "When the weak letter is a nasal and precedes the strong letter, it is changed to anuswâr; in Prakrit without, and generally in the moderns with, compensatory lengthening of the vowel" (282). Beames' 'anuswâra' (term used for homorganic nasal) is more like the 'anunâsika' (term used for nasalization) as can be seen

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

from the following quote "this anuswâr is in most languages pronounced as a distinct anumâsika or nose-intonation, as in french bon, son;" (296). The 'generally' in the qhote from Beames, and the 'sometimes' in the quote from Kellog seem to indicate that there were many exceptions to the vowel lengthening. If the nasal was retained in such cases, as Misra seems to indicate, then the result would be short non-nasalized vowels. However if the nasal was lost, we would get short nasalized vowels thus giving a number of exceptions to Misra's generalization. Since by Gilchrist's time there were enough short nasalized vowels to render it impossible to correlate vowel nasalization with vowel length, we might presume that there were quite a few cases of short nasalized vowels even in old Hindi. The problem is of course compounded by the fact that we cannot tell the actual pronunciation of a word in Old Hindi when the writer had used the symbol for anuswara in Devanagari. I will elaborate on this topic shortly when I review the literature.

5 I should mention that in some words we find a nasalized vowel, where historically there was no nasal. (E.g., Skt. ɔ̃ru "tear" Hindi [ãsu:]). This has been called "spontaneous nasalization" by writers such as Grierson (1922).

6 This state of affairs provides some evidence that needs to be considered in the development of the marking theory. See M. Ohala (1971).

7 The term 'East-Indian' here, opposes the West Indies and not West India. The language he describes is the "Grand, popular, and military language of all India" (mentioned on the title page). The title sounds like it might be a Fodor's guide to India, but it is not, it is actually a professionally-done grammar.

8 I am leaving out some of the diacritics he has used because they are not essential to my argument.

9 Kelkar uses the subscript hook on the vowels to represent partial nasalization. Referring to the word [bĩⁿdʰa] he also says "Note the prenasalization of d^h...." (33). But earlier on pg. 24 we were told that superior 'ⁿ' indicates "full nasalization of the preceding syllabic vocoid..." thus when he tells us to note the prenasalization of d^h, what symbol in his phonetic transcription are we supposed to take as representing that? And although one would suppose the tilde '˜' over the vowel indicates nasalization, one wonders what the superscript n means then.

10 Of his examples, manta "accepts" and sunta "listens", do not belong here. They are actually man+ta and sun+ta, i.e. verb root plus the participle marker.

11 This of course implies that if in intransitive/transitive verb pairs, one is to be derived from the other phonologically then the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

short vowel member, the intransitive, is derived from the long vowel member, the transitive.

12 Interestingly, the word [bə̃ṭna] "to get distributed" (a derivative of [bə̃ṭna] "to distribute") never seems to have a non-nasalized variant, presumably due to the fact that it would become homophonous with another word [bəṭna] 'to get twisted'. However another derivative, [bə̃ṭai:] 'distribution' does have a non-nasalized variant-- [bəṭai:]

13 Actually I am simplifying things because ũ: also indicates 1st. pers. sing. (second pers. would be o in the word).

14 The informants were native speakers of Standard Hindi, from Allahabad, Muzaffarnagar (U.P.), Delhi, Calcutta.

15 The exact specification of this rule is beyond the scope of this study.

16 He gives the rule as (923):

$$\begin{bmatrix}+\text{voc}\\-\text{cons}\end{bmatrix}, [+\text{nas.}], \begin{bmatrix}+\text{cons}\\-\text{voc}\end{bmatrix} \longrightarrow \begin{bmatrix}+\text{tense}\\+\text{nas}\end{bmatrix}, [\emptyset], 3$$

1 2 3 1 2

17 It seems to me that if Narang and Becker really feel the urge to put a nasal in the underlying forms of words such as [dʰũã] "smoke", they should have acted on historical evidence and put the nasal where it was historically; i.e. Skt. dʰũ:mə "smoke".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 4

4.0. In this chapter I will discuss in some detail the ə-deletion rule which I have referred to in previous chapters.

4.1. When we examine word pairs such as those in table 28, we find that they are semantically and phonologically similar. In fact we could say that the words in column I are derived from forms similar to those in column II, by the deletion of a /ə/ when the appropriate environment has been provided by a suffix or an inflectional ending.[1]

Table 28

I		II	
[pička]	"squeezed"	[pičək]	"squeeze"
[pig^{h}la]	"melted"	[pig^{h}əl]	"melt"
[ḍ^{h}olki:]	"drum" (diminutive)	[ḍ^{h}oləк]	"drum"
[təbla]	"drum"	[təbəlči:]	"drummer"
[t^{h}əpki:]	"a pat"	[t^{h}əpək]	"pat"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[həldi:]	"turmeric"	[hələ(hat^{h})]	"a ritual in weddings involving turmeric"
[dewrani:]	"brother-in-law's wife"	[dewər]	"brother-in-law"
[nəmki:n]	"salty"	[nəmək]	"salt"
[pəkṛa]	"caught"	[pəkər]	"catch"
[pu:rbi:]	"easterner"	[pu:rəb]	"east"
[p^{h}isla]	"slipped"	[p^{h}isəl]	"slip"
[bəkri:]	"goat"	[bəkər(kəsaw)]	"goatmeat seller"
[bədla]	"changed"	[bədəl]	"change"
[bərsat]	"rain"	[bərəsna]	"to rain"
[bik^{h}ra]	"scattered"	[bik^{h}ərna]	"to scatter"
[bərsi:]	"anniversary of a death"	[bərəs]	"a year"
[čipka]	"stuck"	[čipək]	"stick"
[ləkṛi:]	"wood"	[ləkəṛhara]	"wood cutter"
[ləṛka]	"boy"	[ləṛəkpən]	"boyhood"
[siski:]	"a sob"	[sisək]	"sob"
[hirni:]	"doe"	[hirən]	"deer"
[hički:]	"hiccup"	[hičək]	"hesitance, stoppage"
[u:d^{h}mi:]	"naughty"	[u:d^{h}əm]	"naughtiness"
[uk^{h}ṛa]	"uprooted"	[uk^{h}ərna]	"to get uprooted"
[uʃla]	"bright"	[uʃəlvana]	"to brighten" (causative)
[upʃau:]	"productive"	[upəʃ]	"produce"
[ulʃ^{h}a]	"tangled"	[uləʃ^{h}]	"tangle"
[ɔrtõ]	"women" (obl.)	[ɔrət]	"woman"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[sərkĕ]	"roads"	[sərək]	"road"
[ləṭko]	"hang" (2nd pers. imp.)	[ləṭəkna]	"to hang"
[tərpa]	"cause to be restless"	[tərəp]	"restlessness"
[tərsa]	"cause to long"	[tərəs]	"longing, compassion"
[pəgla]	"mad"	[pagəl]	"mad"
[pič^{h}ṛa]	"left behind"	[pič^{h}əṛna]	"to remain behind"
[pəsra]	"stretched"	[pəsərna]	"to be stretched"
[kəǰli:]	"black colored"	[kaǰəl]	"collyrium"
[p^{h}ṭka]	"instrument for carding cotton"	[p^{h}əṭəkna]	"to card cotton"
[p^{h}ərka]	"palpitated"	[p^{h}əṛək]	"palpitate"
[məčla]	"persisted in"	[məčəlna]	"to persist in"
[rəgṛa]	"the work of rubbing"	[rəgəṛ]	"rub"
[wapsi:]	"on return"	[wapəs]	"return"
[surka]	"a slurp"	[surək]	"slurp"
[hərpa]	"misappropriated, swallowed"	[hərəp]	"swallow, misappropriate"
[u:pri:]	"pertaining to the top"	[u:pər]	"top"
[ugla]	"vomited"	[ugəl]	"vomit"
[ug^{h}ṭa]	"disclosed"	[ug^{h}əṭna]	"to disclose"
[učka]	"leapt"	[učəkna]	"to leap"
[uč^{h}la]	"jumped"	[uč^{h}əl]	"jump"
[ud^{h}ṛa]	"unravelled"	[ud^{h}əṛna]	"to be unravelled"
[ubṭən]	"an unguent"	[ubəṭna]	"to anoint"
[ub^{h}ṛa]	"protruded"	[ub^{h}əṛna]	"to protrude"
[umṛa]	"increased"	[umərna]	"to increase"
[ulṭa]	"reversed"	[uləṭna]	"to be reversed"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[jʸʰijʸʰka]	"hesitated"	[jʸʰijʸʰək]	"hesitation"
[pʰudki:]	"a kind of bird"	[pʰudək]	"hop"
[bʰəbʰki:]	"a threat"	[bʰəbʰəkna]	"to inflame"
[gərǰi:la]	"thunderous"	[gərəǰna]	"to thunder"
[mət̤ka]	"wobble (something)" (transitive imper.)	[mət̤əkna]	"to wobble, coquetry"
[sənatni:]	"follower of traditional Hinduism"	[sənatən]	"traditional"
[dʰədʰka]	"fan a fire" (imper.)	[dʰədʰəkna]	"to fan a fire"
[bɛt̤ʰka]	"room for meeting visitors"	[bɛt̤ʰək]	"an assembly"

etc.

On the basis of these alternations Rule (1) is given as a first approximation of the ə-deletion rule:

(1) ə --> Ø / VC___CV

All the words given in table 28 are native words. However the ə-deletion rule also applies to non-native words as shown in table 29.

Table 29

[bəgli:]	"pertaining to the side"	[bəgəl] (PA)	"side, armpit"
[nagrik]	"citizen"	[nəgər] (S)	"city"
[təmsa]	"name of a river"	[təməs] (S)	"darkness"
[dəmniy]	"worth suppressing"	[dəmən] (S)	"suppression"
[danwi:]	"demonish"	[danəw] (S)	"demon"
[dəlti:]	"pertaining to wealth"	[dələt] (PA)	"wealth"
[dʰəwla]	"something that is white" (fem.)	[dʰəwəl] (S)	"white"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[nəkdi:]	"pertaining to cash"	[nəkəd] (PA)	"cash"
[nəzri:]	"visual"	[nəzər] (PA)	"vision"
[nɔkri:]	"job"	[nɔkər] (PA)	"servant"
[prət^{h}ma]	"nominative"	[prət^{h}əm] (S)	"first"
[profɛsri:]	"professorship"	[profɛsər] (E)	"professor"
[biradri:]	"group of people of the same caste"	[biradər] (PA)	"person of same clan"
[k^{h}ətra]	"danger"	[k^{h}ətər(nak^{h})] (PA)	"dangerous"
[k^{h}əbri:]	"pertaining to news"	[k^{h}əbər] (PA)	"news"
[k^{h}usamdi:]	"one who flatters"	[k^{h}usaməd] (PA)	"flattery"
[su:rti:]	"pertaining to the face"	[su:rət] (PA)	"face"
[gəlti:]	"mistake"	[gələt] (PA)	"wrong"
[guzre]	"passed"	[guzər] (PA)	"passing"
[čəpla]	"lightning, a proper name"	[čəpəl] (S)	"quivering"
[zək^{h}mi:]	"pertaining to the wound"	[zək^{h}əm] (PA)	"wound"
[ǰigri:]	"a bosom (friend)"	[ǰigər] (PA)	"liver"
[ǰi:wni:]	"biography"	[ǰi:wən] (S)	"life"
[zewrat]	"jewels" (pl.)	[zewər] (PA)	"jewelry" (sing.)
[nərmi:]	"softness"	[nərəm] (PA)	"soft"
[wimla]	"dirtless, proper name for girl"	[wiməl] (S)	"dirtless, name for a boy"
[b^{h}arti:]	"Indian, name for girl"	[b^{h}arət] (S)	"India"
[b^{h}rəmri:]	"female bumblebee"	[b^{h}rəmər] (S)	"bumblebee"
[muglani:]	"Moguless"	[mugəl] (PA)	"Mogul"
[rešmi:]	"silky"	[rešəm] (PA)	"silk"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[rogni:]	"greasy"	[rogən] (PA)	"grease"
[wəzni:]	"weighty"	[wəzən] (PA)	"weight"
[wičlit]	"made unstable"	[wičəl] (S)	"unstable"
[šərarti:]	"naughty"	[šərarət] (PA)	"naughtiness"
[šayri:]	"kind of poetry"	[šayər] (PA)	"poet"
[ši:tla]	"cool (fem.), name of a goddess"	[ši:təl] (S)	"cool"
[səfri:]	"pertaining to the journey"	[səfər] (PA)	"journey"
[sərla]	"without guile (fem.), name for girl"	[sərəl] (S)	"guileless"
[ekadši:]	"eleventh day of month"	[ekadəš] (S)	"eleven"
[ədalti:]	"pertaining to the court"	[ədalət] (PA)	"court"
[əmla]	"dirtless, name for girl"	[əməl] (S)	"dirtless, name for boy"
[ərbi:]	"Arabian"	[ərəb]	"Arab"
[admi:]	"man"	[adəm] (PA)	"Adam"
[suločna]	"pretty eyed, name for girl"	[suločən] (S)	"good eyes"
[kagzat]	"papers"	[kagəz] (PA)	"paper"

In my experience all native speakers drop the /ə/ in the derived forms of the native words (table 28, column I) no matter what style of speaking is used. These words are very common words, and have been in use for a long time. However the treatment of the non-native words (table 29) is slightly different. The /ə/ is not deleted in the derived forms by all speakers in all styles of speaking. I will discuss this more in 4.6.4.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.2. The ə-deletion rule given above will, however, have to be somewhat modified because the /ə/ is not deleted in the derived words such as those given in table 30.

Table 30

bə+pər̥^{h}+a	[bepər̥^{h}a]	"unread"
ə+səməy	[əsəməy]	"inopportune"
a+mər+ən̥	[amərən̥]	"until death"
su+nəyən	[sunəyən]	"pretty eyes, good eyes"
bə+k^{h}əbər	[bek^{h}əbər]	"unconscious"
ə+pələk	[əpələk]	"without blinking"
ku+səməy	[kusəməy]	"bad time"
ə+šəri:r	[əšəri:r]	"without body"
ə+kələŋk	[əkələŋk]	"spotless"
ku+šəkun	[kušəkun]	"bad omen"

As these words consist of a prefix plus a stem (plus a suffix, in some cases), the ə-deletion rule could be restricted so as not to apply across prefix boundaries. However, as I will show shortly, any morpheme boundary in the left-hand environment of the rule restricts its application. I propose the following modification of the ə-deletion rule:

(2) ə --→ Ø / VC___CV

Condition: There may be no morpheme boundary, '+', in the environment to the left of the /ə/.

The need to mention the morpheme boundary in the above condition would

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mark this as being a less general situation according to Chomsky and Halle (1968:364) since they say "processes operating within formatives normally also apply across formative boundary". For further discussion see 4.10.3.1.

4.2.1. What happens to a /ə/ in a suffix? From words such as those given in table 31 we see that it is retained.

Table 31

kəla+wət+i:	[kəlawəti:]	"one who knows the arts, name for a girl"
kari:+gər+i:	[kari:gəri:]	"craftsmanship"
sɔda+gər+i:	[sɔdagəri:]	"buyingmanship"
ekaki:+pən+a	[ekaki:pəna]	"lonesomeness, solitude"
šišu+pən+a	[šišupəna]	"childishness, childhood"
šri:+mət+i:	[šri:məti:]	"Mrs., literally 'one with wealth'"

These words will be accounted for by P-rule (2). Since the VC preceding the /ə/ in these words contains a '+' (the morphemes preceding the '+' are not prefixes) and the condition on the rule says there must be no '+' in the environment to the left, the environment of the ə-deletion rule is not met and the /ə/ is retained. This lends support to the formalization of rule 2. It accounts not only for the fact that prefixes do not provide the environment for the ə-deletion rule, but also for the fact that any preceding morpheme boundary in the environment of the rule restricts its application.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Similarly in the compounds:

(3) čəmpa+kəli: [čəmpakəli:] "Michelia champacca bud, name for girl"

gali:+gəlɔč [gali:gəlɔč] "abuse and the like"

etc.

the /ə/ is not deleted, and P-rule 2 correctly accounts for this.

In a compound like [k^hudgərzi:] "selfish" /k^hud+gərəz+i:/ the second /ə/ is deleted (at least for non-Urdu speakers) since it meets the environment of the ə-deletion rule.[2]

The question now arises as to what would happen to a /ə/ in a suffix of the form VCəCV if a vowel was added after it (as another suffix or an as inflectional ending). This question cannot be answered because of the scarcity of such suffixes to which endings are added. For example, -iyət would appear to be an example of such a suffix but when it is added to a morpheme the [i] is lost in casual speech, e.g., /kəmal+iyət/ "excellence" is [kəmalyət]. Thus the CC would prevent the deletion of the /ə/, even if an ending were added to [kəmalyət]. (Actually, as I will show shortly, the /ə/ is deleted sometimes even if a CC precedes, but in this case the three-consonant cluster that would result if the /ə/ were deleted would be an unlawful cluster and so the /ə/ would still not be deleted. Cf. section 4.6.). This point deserves further study.

4.3. It must be emphasized that the important thing which determines whether or not the /ə/ gets deleted in the words cited above, is not whether or not a linguist can posit a morpheme boundary (since he knows the history of the word) but whether or not the morpheme boundary

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

has any reality for the average native speaker. For example we find words such as those in table 32.

Table 32. Words in column 1 are historically derived from the morphemes in column 2. (Upraiti 1964).

[bawǰu:d]	"in spite of"	/ba+wəǰu:d/
[utra]	"stepped down"	/u+tər+a/
[ub^{h}ṛa]	"protruded"	/u+b^{h}əṛ+a/
[dubla]	"thin"	/du(r)+bəl+a/
[əbla]	"woman"	/ə+bəl+a/
[k^{h}əpti:]	"crazy"	/k^{h}əp+ət+i:/
[ḍerawna]	"scary"	/ḍər+a+wən+a/
[ḍ^{h}olkõ]	"drums" (obl.)	/ḍ^{h}ol+ək+õ/
[d^{h}əṛkən]	"beat"	/d^{h}əṛ+ək+ən/
[b^{h}əṛki:la]	"bright, gaudy"	/b^{h}əṛ+ək+i:la/
[əŋgṛai:]	"a stretch"	/əŋg+əṛ+ai:/
[apsi:]	"with one another"	/ap+əs+i:/

In all these words a linguist would posit a morpheme boundary on historical evidence (similar to positing English "survey" to be /sur+vey/), but the fact that the /ə/ is deleted (as can be seen from column 1 of table 32) seems to indicate that the native speaker treats these as:

(4) /bawəǰu:d/
/utər+a/
/ub^{h}əṛ+a/
/dubəla/[3]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

/əbəla/

/k^həpət+i:/

/dərawən+a/ (or perhaps /ḍər+awən+a/)

/$ḍ^h$olək+õ/

/d^hərək+ən/

/b^həṛək+i:la/

/əŋgəṛai:/

/apəs+i:/

Similarly the Sanskrit loans prəčəlit "in use" and ačərəṛ̃ "behavior" which are historically /prə+čəlit/ and /a+čərəṛ̃/ are rendered by many speakers as [prəčlit] and [ačrəṛ̃].

4.3.1. Further evidence comes from how Hindi speakers pronounce unfamiliar words written in Devanagari. If the morpheme boundary is obscure or unknown they frequently give the "incorrect" pronunciation, but in a systematic way. For example the rare word pronounced [godnəši:n] "adopted" is derived from /god+nəši:n/ "lap + sitter" but is written in Devanagari as गोद्नशीन {godənəši:n}. If a speaker knows that the word is /god+nəši:n/ he will not pronounce the 'द्' {d} of god as a CV syllable (i.e., də), but will correctly render it as simply the consonant d; he will also retain the /ə/ in nəši:n. However if he doesn't know the true morpheme boundary then he applies his ə-deletion rule from right to left and pronounces it instead as [godənši:n]. (This right-to-left application will be discussed in some detail in 4.5.1.). In fact this was the consistent response given by three subjects to whom I presented this word in Devanagari (they were all unfamiliar with the word).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.4. As stated, the ə-deletion rule applies both within morphemes and across morpheme boundaries to the right. However, all of the examples I have given so far have only involved morpheme boundaries. If we posit the underlying form of [əŋgṛai:] "a stretch" to be /əŋgəṛai:/, i.e., without a morpheme boundary as I have proposed above, then of course the ə-deletion rule would also have to apply morpheme medially to give the desired phonetic output [əŋgṛai:]. I have tried to argue for such underlying morpheme internal /ə/'s (which are then deleted by the ə-deletion rule) in 2.2. There I presented evidence from a test I conducted which showed the apparent reality of such /ə/'s in certain words. I also argued there that the problem of non-homorganic nasal+stop in Hindi is better accounted for by positing such /ə/'s (see also 2.3.5.1.). I also suggested that the child's version of the word for "lizard" ([č^{h}ipəkli:] instead of [č^{h}ipkəli:]) is better explained if we posit an underlying /ə/ within the cluster in the word, i.e. /č^{h}ipəkəli:/. I will now present some additional evidence for this point.

4.4.1. We find words such as those given in table 33 where some speakers have a [ə] (normally in formal style, or careful style, as in reading) and others don't, or, the same speaker may or may not delete the [ə] depending on the style of speaking.

Table 33

[kadəmbri:]	~	[kadəmbəri:]	"a novel, name for a girl"
[kundru:]	~	[kundəru:]	"a kind of vegetable"
[ǰyotsna]	~	[ǰyotsəna]	"moonlight"
[rəǰni:]	~	[rəǰəni:]	"night, name for a girl"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[wyəwhar]	~	[wyəwəhar]	"behavior, use"
[šakəmb^{h}ri:]	~	[šakəmb^{h}əri:]	"name of a goddess, name for a girl"
[ustra]	~	[ustəra]	"razor"
[dewta]	~	[dewəta]	"deity"
[pu:tna]	~	[pu:təna]	"name of a demoness"
[pundri:k]	~	[pundəri:k]	"white lotus"
[məñǰri:]	~	[məñǰəri:]	"tiny cluster of flowers, name for girl"
[godawri:]	~	[godawəri:]	"name of a river"
[surb^{h}i:]	~	[surəb^{h}i:]	"sweet smelling air, name for girl"

The above words do not involve a suffix and yet the [ə] is occasionally deleted. (We also notice in words such as [pundri:k] that the [ə] has been deleted in spite of a preceding CC. I will discuss this point in 4.6.). If we allow the ə-deletion rule to apply morpheme-medially as well as across morpheme boundaries, then we can account for the above alternations by saying that the ə-deletion rule optionally applies in certain words, the 'option' being determined by such sociolinguistic factors as formal vs. casual speech, and by the tempo of speaking.

4.4.2. More crucial evidence for the ə-deletion rule applying morpheme medially comes from the word [šəmšan] "cremation ground". This word is restructured from the Sanskrit loan šməšan; -šm-, as mentioned in chapter 2, occurs only in a few uncommon Sanskrit loans (and pronounced as such only by high Hindi speakers). Since this is the only word for "cremation ground" and is commonly used, it was restructured by speakers into /šəməšan/, by the insertion of a [ə]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

between the cluster (which is the most common method of simplifying clusters in Hindi). This word now met the ə-deletion rule's environment, and thus the form [šəmšan] resulted.[5] In casual speech in standard Hindi [šəmšan is used, [šməšan] being reserved for 'high vocabulary' or Sanskritized Hindi.

4.5. Even in the case of loanwords which contain clusters we could posit a /ə/ in the underlying form and delete it by the ə-deletion rule, e.g., we could posit /ə/'s breaking up the medial clusters in the following loans:

Table 34

[dušmən]	/dušəmən/	"enemy" (PA)
[dərzən]	/dərəzən/	"dozen" (E)
[əǰgər]	/əǰəgər/	"python" (S)
[dəftər]	/dəfətər/	"office" (PA)
[təpəsya]	/təpəsəya/	"penance" (S)
[tatpəry]	/tatəpəry/	"meaning" (S)
[šəkti:]	/šəkəti:/	"strength" (S)
[kismət]	/kisəmət/	"fate" (PA)
[dəstu:r]	/dəsətu:r/	"custom" (PA)
[təšri:f]	/təšəri:f/	"honor" (PA)

However it is not at all clear that this analysis should be adopted. There is no evidence for the 'psychological reality' of such /ə/'s.

4.5.1. Let me now discuss the right-to-left application of the ə-deletion rule, referred to earlier. If we were to posit a /ə/ in the words of table 34, then in the case of /təpəsəya/ the rule would

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

have to apply from right to left to give the correct phonetic output [təpəsya]. However there are other loans where we would get the wrong result if we posited underlying /ə/'s and applied the ə-deletion rule from right to left, e.g.:

(5)	[istəmal]~[istmal]	/isətəmal/	*[isətmal]	"use" (PA)
	[yəkšəma]~[yəkšma]	/yəkəšəma/	*[yəkəšma]	"T.B." (S)
	[watsyayən]	/watəsəyayən/	*[watəsyayən]	"a proper name" (S)
	[əstəbəl]	/əsətəbəl/	*[əsətbəl]	"stable" (PA)
	[əspətal	/əsəpətal/	*[əsptal]	"hospital" (E)
	[imtəhan]	/imətəhan/	*[imthan]	"examination" (PA)
	[ištəhar]	/išətəhar/	*[išthar]	"poster" (PA)
	etc.			

If it turns out that the ə-deletion rule does indeed apply from right to left, then these words would have to be marked as exceptions, and this would be additional motivation for not positing underlying /ə/'s in such loans -- they might just as well be exceptions to Hindi's bias against clusters at the lexical level than be exceptions to the ə-deletion rule. I propose that there are no underlying /ə/'s breaking up the clusters in the above words, therefore they are not relevant to the right-to-left application of the ə-deletion rule in my analysis.

The proposal that the ə-deletion rule applies from right to left was first made by Pray (1970). However he made this proposal because he considers the verb [nikəlva] "to have taken out" to be underlyingly /nikələva/, to which the ə-deletion rule would have to apply from right

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to left in order to obtain the correct phonetic output; otherwise we would incorrectly get *[nikləva]. However if one does not accept Pray's argument for positing underlying /nikələva/ (and I don't),[4] then the example he gives does not provide evidence for the right-to-left application of the ə-deletion rule.

The strongest evidence for the right-to-left application has been mentioned under 4.3.1., namely the way native speakers render unfamiliar words written in Devanagari. Of course this assumes that the ə-deletion rule used by Hindi speakers for reading the Devanagari orthography is the same as the ə-deletion rule they have internalized for spoken Hindi. I think this assumption is justified, since the environments in which the /ə/ is deleted are the same.

As I mentioned before, the word for "lizard" in Hindi is [ʧʰipkəli:], however children are reported to occasionally say [ʧʰipkli:]. If we assume that the underlying form is /ʧʰipəkəli:/ and that the ə-deletion rule applies from right to left, then the child's version [ʧʰipəkli:] is what we would expect; however, the word is an exception to the right-to-left application, and thus the adult version is [ʧʰipkəli:] . Such examples are, however, quite rare.

I am tentatively adopting the convention that the ə-deletion rule applies from right to left, however this should be further investigated.

4.6. Let me now discuss a point referred to under 4.4.1. in reference to some examples in table 33, namely, that the [ə] was sometimes deleted even when preceded by a consonant cluster. The example was [puṇḍri:k]~[puṇḍəri:k] which did not involve a morpheme boundary. Table 35 presents some more examples, this time words that involve

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 35

				A	B
[ɕʰuɕʰu:ndər]	"a kind of rat"	[ɕʰuɕʰu:ndri:]	"female of the species"	ⓐ	ⓐ
[ǰəŋgəl]	"forest"	[ǰəŋgli:]	"pertaining to the forest, wild"	ø	ø
[ǰələndʰər]	"name of a city"	[ǰələndʰri:]	"pertaining to Jullundhar"	ⓐ	ø
[kəŋgən]	"bracelet"	[kəŋgna]	"a ceremony at weddings"		
[kəmbəl]	"blanket"	[kəmblõ]	"blankets" (obl. pl.)		
[ḍokṭər]	"Doctor"	[ḍokṭri:]	"doctorly, doctorship"	ø	ø
[nəmbər]	"number"	[nəmbri:]	"pertaining to numbers"	ⓐ	ø
[pəñčəm]	"fifth"	[pəñčmi:]	"fifth day of month"	ⓐ	ⓐ
[bɛŋgən]	"eggplant"	[bɛŋgni:]	"purple"	ø	ø
[səmundər]	"ocean"	[səmundri:]	"pertaining to the ocean"	ø	
[sundər]	"pretty"	[sundri:]	"a pretty woman, name for girl"	ø	ø
[kuṇḍəl]	"a circular ornament"	[kuṇḍli:]	"zodiac"	ⓐ	ə
[jʰə̃ñjʰər]	"anklet"	[jʰə̃ñjʰri:]	"anklet" (dimin.)	ⓐ	ø
[ḍõŋgər]	"hill person"	[ḍõŋgri:]	"pertaining to hill people"	ø	ø
[məŋgəl]	"auspicious"	[məŋgla]	"name for a girl"	ø	ø

				A	B
[lə̃ŋgəɽ](di:n)	"lame"	[lə̃ŋgɽa]	"lame"	ø	ø
[šə̃ŋkər]	"Shiva"	[šə̃ŋkra]	"name of a raga"	ø	ø
[sə̃mb^{h}əl]	"be careful"	[sə̃mb^{h}la]	"became careful"	ø	ø
[sikəndər]	"name of a king"	[sikəndra]	"name of a city"	ø	ø
[əntər]	"difference"	[əntra]	"second part of a song"	ø	ø
[ãŋgən]	"courtyard"	[ə̃ŋgna]	"courtyard" (dim.)	ⓐ	ø
[əñ̆jən]	"collyrium"	[əñ̆jna]	"name for a girl"	ⓐ	ø
[čəñčəl]	"quick"	[čəñčla]	"name for a girl"	ⓐ	ø
[paṇḍəw]	"legendary king"	[paṇḍwõ]	"the Pandawas" (obl.)	ø	ø
[dəršən]	"philosophy"	[daršnik]	"philosopher"		
[əšṭəm]	"eight"	[əšṭmi:]	"eighth day of month"	ø	ø
[bistər]	"bedding"	[bistra]	"bedding"	ø	ø

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

suffixes or inflectional endings.[6] The /ə/ in the derived words of table 35 is deleted in spite of a preceding cluster. (In those of the above words which are Sanskrit loans, such as [sundri:], the /ə/ is deleted when the word is used as a proper name, otherwise the /ə/ is deleted in casual style but not always in formal style. This will be discussed more in 4.6.4.).

We could modify the ə-deletion rule to permit it to delete the /ə/ even if a cluster precedes, however the data are much more complicated. There are words such as those given in table 36, which involve a cluster preceding the /ə/ in which the /ə/ is not deleted.

To complicate matters further, there are words of the VCəCV type, i.e. words similar to those given in tables 28 and 29 which do not involve a preceding cluster, but where nevertheless the /ə/ is not deleted, cf. table 37. The pronunciation given in these tables is based on my dialect, but is not limited to my dialect. I conducted a few informal tests and the responses of the subjects agreed for the most part with my dialect. The results of these tests are interesting enough to suggest the desirability of a more detailed and rigorous test using more informants. In one of the tests I asked two subjects to add certain endings to the words listed in column 1 of tables 35, 36, 37, and 38. The forms in column 2 of tables 35, 36, and 37 represent my pronunciation of the derived words. The third and fourth columns represent the responses of the two informants, A and B, Ø meaning no [ə], ə̲, [ə] retained, and ⓔ neither [ə] nor Ø but either an extremely reduced ə̲ or merely an audible release of a consonant (my ear was not able to detect which). (Some words of table 35 were not presented to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 36

					A	B
1.	[kərwəṭ]	"side"	[kərwəṭə̆]	"sides"	ⓐ	ə
2.	[kərkəṭ]	"crab"	[kərkəṭi:]	"female crab"	ⓐ	ə
3.	[kudrət]	"nature"	[kudrəti:]	"natural"	ə	ə
4.	[čitwən]	"glance"	[čitwənə̆]	"glances"	ə	ə
5.	[bəlgəm]	"phlegm"	[bəlgəmŏ]	"phlegms" (obl.)	ə	ə
6.	[mətləb]	"interest"	[mətləbi:]	"selfish"	ə	ə
7.	[ədrək]	"ginger"	[ədrəkŏ]	"gingers" (obl.)	ə	ə
8.	[kəṇṭək]	"thorn"	[kəṇṭəki:]	"thorny"	ə	ə
9.	[dušmən]	"enemy"	[dušməni:]	"animosity"	ə	ə
10.	[kəsrət]	"exercise"	[kəsrəti:]	"pertaining to exercise"	ə	ə
11.	[kismət]	"fate"	[kisməti:]	"pertaining to fate"	ə	ə
12.	[k^hidmət]	"to fag for someone"	[k^hidməti:]	"pertaining to fagging for someone"	ə	ə
13.	[tərkəš]	"quiver"	[tərkəši:]	"pertaining to a quiver"	ə	ə
14.	[šərbət]	"Sherbet"	[šərbəti:]	"Sherbety"	ə	ə
15.	[sərkəs]	"circus"	[sərkəsi:]	"pertaining to circus"	ə	ə
16.	[sərgəm]	"glissando"	[sərgəmo]	"glissandos" (obl.)	ə	ə
17.	[əɟgər]	"python"	[əɟgəri:]	"pertaining to pythons"	ə	ə

Table 37

				A	B
[adət]	"habit"	[adətõ]	"habits"	ə	ə
[pi:hər]	"parents home"	[pi:hərõ]	"many such homes"	ⓔ	ə
[məhaǰən]	"money-dealer"	[məhaǰəni:]	"occupation of money dealing"	ə	ə
[musi:bət]	"difficulty"	[musi:bətõ]	"difficulties" (obl.)	ə	ə
[həzamət]	"a shave"	[həzamətõ]	"shaves" (obl.)	ə	ə
[b^hagi:rət^h]	"a sage"	[b^hagi:rət^{h}i:]	"Ganges"	ə	ə

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the informants and thus no responses have been recorded). Informant A is from Delhi and informant B from Saharanpur. Both are native speakers of standard Hindi. (Both know English.)

There are also words (table 38) (some involving a VCC$\underline{\text{ə}}$CV context and others VC$\underline{\text{ə}}$CV) in which my initial reaction was not to delete the /ə/. However after working on this topic I found myself accepting more and more the deleted $\underline{\text{ə}}$ in a number of the words. This is interesting since it shows that familiarity with the derived forms may also determine whether or not the /ə/ is deleted. I checked these with the two informants also, and the results are given in table 38.

Table 38

		A	B
gərdən+i:	"pertaining to the neck"	ⓐ	ə
takət+i:	"pertaining to strength"	ə	Ø
dik^{h}awət̤+i:	"showy"	ə	ⓐ
nirəñjən+a	"name for a girl"	ⓐ	ⓐ
nirməl+a	"name for a girl"	ⓐ	ⓐ
bənawət̤+i	"artificial"	ⓐ	ⓐ
sarəs+õ	"cranes" (obl.)	Ø	Ø
siyasət+õ	"politics" (obl.)	ə	ə
əfsər+i:	"pertaining to an officer"	ə	ə
sorət̤^{h}+a	"a kind of song"	Ø	ⓐ
amašəy+õ	"stomachs" (obl.)	ⓐ	Ø (but informant paused before answering)
di:pək+õ	"lamps" (obl.)	Ø	ə
zəru:rət+i:	"pertaining to needs"	Ø	ə
pat^{h}ək+õ	"readers" (obl.)	Ø	Ø
ǰu:ṭ^{h}ən+õ	"leftovers on plate" (obl.)	Ø	Ø
ǰ^{h}arṇən+õ	"dusters" (obl.)	ə	Ø
hi:reǰəwahərat	"diamonds and jewels"	Ø	ə

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(The last word in table 38 was elicited by asking for a translation into Hindi of the English expression; the rightmost /ə/ is the one of interest).

4.6.1. In a number of the words in table 38 the ə could be said to be part of a suffix, e.g.:

(6)	zəru:rəti:	zəru:r+ət+i:	("necessary"+nominalizing suffix+i:)
	$pa\underset{\cdot}{t}^{h}$əkõ	$pa\underset{\cdot}{t}^{h}$+ək+õ	("lesson"+agentive suff.+õ)
	ǰu:$\underset{\cdot}{t}^{h}$ənõ	ǰu:$\underset{\cdot}{t}^{h}$+ən+õ	("lie"+nom.suff.+õ)
	dik^{h}awəṭi:	dik^{h}a+wəṭ+i:	("show"+nom.suff.+i:)[7]
	bənawəṭi:	bəna+wəṭ+i:	("make"+nom.suff.+i:)
	ǰ^{h}ar̤ənõ	ǰ^{h}ar̤+ən+õ	("sweep"+nom.suff.+õ)

It is possible that the speakers who do not delete the /ə/ "know" this, but the speakers who do delete the /ə/ do not "know" of this derivational relationship.[8] The importance of whether or not the morpheme boundary is psychologically real for the speaker was discussed in 4.3. On the basis of evidence of this type we could posit, as one of the variables involved in the deletion or non-deletion of the /ə/, the following:

(a) If the speaker posits (i.e. "knows" of) a morpheme boundary in the environment to the left of the /ə/, then the /ə/ will not be deleted.

4.6.2. A detailed analysis of the forms where the ə is deleted and where it isn't reveals further determinants for the application of the ə-deletion rule.

One might think that the retention or deletion of the [ə] might be determined by how the resulting cluster would be accomodated within

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

or across the resulting syllable boundaries, but this is apparently not the case. The two words [čipək] "stick" and [adət] "habit" demonstrate this. Whereas the [ə] in [čipək] can be deleted as in [čipka] "stuck", that in [adət] cannot; *[adtē] is not allowed. No one knows how to determine where the syllable boundaries would fall in these cases (or even if it is reasonable to talk of syllable boundaries), but presumably they would be in the same place in both [cipka] and *[adtē], i.e. either VC-CV, V-CCV, or VCC-V. The first case does not help because p and d are both acceptable final consonants, and k and t are acceptable initial clusters. The second and third cases do not help because both pk and dt are unacceptable final or initial clusters. (As mentioned in chapter 2 it is for reasons such as these that I have not attempted to lump the medial sequential constraints of Hindi under syllable initial or syllable final constraints.) I will discuss below the factors involved in this deletion or non-deletion of the [ə].

We find that the words in table 35 (where the /ə/ is deleted in the derived forms) and in table 36 (where the /ə/ is not deleted) are similar in that the deletion of the /ə/ would (and does, in the case of the words in table 35) lead to the formation of three-consonant clusters. However the resulting clusters in table 35 would be different from those in table 36. This suggests strongly that the reason the /ə/ is not deleted in the words in table 36 is the nature of the clusters involved.[9] A test similar to the one reported in appendix 2 would be desirable to see what kind of medial three-consonant clusters native Hindi speakers do accept. To rule out all the clusters of table 36 would require highly specific sequential constraints (I

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

haven't stated any SQC's for three-consonant clusters except one involving homorganic nasals), but the test conducted for initial constraints referred to above indicated that some rather specific SQC's are indeed required to reflect the native speaker's competence. Further evidence of the relevance of the resulting cluster is perhaps provided by the word [əspɐtal] "hospital" (from English), where the ə-deletion rule does not apply (this would be an example of morpheme-medial application of the rule), because the resulting cluster -stp- would be unlawful. On the basis of the retention of the /ə/ in the words in table 36 (and in [əspɐtal]) we can tentatively rule out medial three-consonant clusters of the following types:

(7) liquid+glide+stop
liquid+stop+fricative
fricative+liquid+stop
fricative+nasal+stop
fricative+nasal+nasal
fricative+stop+stop
stop+liquid+stop
stop+nasal+stop
stop+glide+nasal

This would account for the retention of the /ə/ in most of the derived forms of table 36. However we still have to rule out -rkṭ-, -lgm-, -rbt-, -rgm-, -ǰgr-, and -nṭk- to account for the /ə/ not being deleted in items 2, 5, 14, 16, 17, and 8. These will have to be ruled out by rather specific constraints, since, for example, although we might want to rule -lgm- out, another liquid+stop+nasal cluster does

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

occur in the Sanskrit loan [prart^{h}na] "prayer".[10] However the liquid is [r] not [l], and the stop is [t^{h}] not [g]. If, when speakers were tested, they accepted -lgm-, -rgm- as well as -rt^{h}n- in made-up words, then of course our only recourse to account for the non-deletion of the /ə/ in words like [bəlgəmõ], etc., would be to say that usage prevents the /ə/ from being deleted in these words, i.e. they are exceptions. If, however, the speakers also rejected -rt^{h}n-, then we would have to say that liquid+stop+nasal is in fact ruled out and [prart^{h}na] exists as an exception.

In general we see that most of the words of table 35 (where the /ə/ is deleted) involve a homorganic nasal+stop+sonorant cluster.[11] Since this is the only type of three-consonant cluster that occurs in native words one would expect that speakers would have no difficulty in deleting the /ə/ in words where this type of cluster would result. Of course, as I pointed out in chapter 2, three-consonant clusters are rather few, and in general most two- and three-consonant clusters that we find at the phonetic level in native vocabulary occur due to the ə-deletion rule. As some speakers innovate and apply the ə-deletion rule to produce clusters that are not yet in the language, more and more three-consonant clusters will come to exist in the language. This again shows the spreading environment of the ə-deletion rule.

The retention of the /ə/ in the words of table 37 can also be explained by the fact that the two-consonant cluster that would result in some of them, were the /ə/ deleted, would violate the SQC's of Hindi. Thus the /ə/ in /adət+ẽ/ "habits" is not deleted, because the resultant -dt- would violate SQC 2. As indicated in table 37, the two subjects

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

whose responses are given on the right did not delete the /ə/ in this word either. The response of 27 informants[12] who were given the following test also supports this.

The subjects were asked to add the suffix -ən "adverbializing suffix", to the following words:

Table 39

			deleted	not deleted
1.	adət+ən	"due to habit"	0	27
2.	dik^{h}awət+ən	"with show"	7	20
3.	šərarət+ən	"with naughtiness"	10	17
4.	bənawəṭ+ən	"with artificiality"	6	21
5.	zəru:rət+ən	"due to necessity"	7	20

When the suffix -ən is added to these words, the environment of the ə-deletion rule is met, and thus the /ə/ can be deleted. The subject's responses are given in the columns to the right in table 39. It is interesting that the only word where the /ə/ was never deleted was /adət+ən/ where the resulting cluster as mentioned earlier would violate a sequential constraint. [adət], [šərarət], and [zəru:rət] have existing forms with -ən (I don't know whether my subjects knew this or not). [dik^{h}awəṭ] and [bənawəṭ] don't, but adding -ən would give a perfectly possible word. Any number of variables could be involved in the low number of deletions of the /ə/ in the case of words other than [adət]. Some of these were mentioned earlier under 4.6.1. For example it is possible that some speakers treat [bənawəṭ], etc., as /bəna+wəṭ/ (i.e., the morpheme boundary that is there historically is 'psychologically real'; see discussion in 4.3.) and thus the /ə/

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

would not be in an environment favorable for deletion. The second possibility is that, since -wṭ- as a cluster occurs only due to the ə-deletion rule, it is still considered a somewhat unfamiliar cluster and thus the /ə/ is retained (or perhaps it is an unlawful cluster for some speakers). In the case of speakers who sometimes did and sometimes did not delete the /ə/, I do not have an explanation at this point -- it is possible that a number of variables are involved. The 27 subjects were not tested for consistency in this part of the test. 19 subjects were consistent in either deleting the /ə/ in all words (except [adət]) or not deleting it (possibly a reflection of a dialect difference); 8 subjects had some variation.

Other words in table 37 merit individual discussion. In the case of [musi:bət] "difficulty" one possible reason for the non-deletion of the /ə/ could be that -bt- occurs only in a few rare loans and should possibly be ruled out as a lawful cluster for most speakers. Similarly -hr- occurs only in a few PA loans, and this could account for why I preferred [pi:hərõ] "parent's homes" to [pi:hrõ]. In the case of [məhaǰəni:] "occupation of dealing in money", and [həzamətõ] "shaves", it is again possible that the morpheme boundary /məha+ǰən+i:/ and /həzam+ət+õ/ (which is there historically) has some 'psychological reality' and was preventing the /ə/ from being deleted; the other possibility is that it was just usage that prevented the /ə/ from being deleted. The clusters -ǰn- and -mt- occur only in a few Sanskrit and PA loans, and this factor might be involved too.

Additional evidence can be given for the relevance of the resulting clusters to the ə-deletion rule.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.6.3. The words in table 40 involve long consonants (i.e. what other writers consider to be geminates and thus clusters) before a <u>əC#</u>.

Table 40

[čəm:əč]	"spoon"
[pət^{h}:ər]	"stone"
[ḍ^{h}ək:ən]	"lid"
[məč^{h}:ər̤]	"mosquito"
[t^{h}əp:ər̤]	"slap"
[k^{h}əč:ər̤]	"mule"
[mək^{h}:ən]	"butter"

If we add the oblique plural ending [õ] to these words the environment of the ə-deletion rule would be met. However, if the ə-deletion rule were to apply, the result would violate SQC 5 given in chapter 2 which states that long consonants are preceded and followed by vowels. Interestingly, we find that speakers either retain the /ə/ or delete the /ə/; but if they delete the /ə/, then they also make the [+long consonant][-long]. There is considerable variation, sometimes even within the same speaker at different times, as to which of the two alternatives is chosen. In some cases derived words with the ə-less forms already exist, and are in common use, e.g., /pət^{h}:ər+i:la/ [pət^{h}ri:la] "stony", /ḍ^{h}ək:ən+i:/ [ḍ^{h}əkni:] "small lid", /čəm:əč+a/ [čəmča] "spoon", etc. In these cases the informants that I checked with gave the ə-less form.[13] However, when I asked them to add the oblique plural endings to these words I got quite a bit of variation

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

between the retention or loss of the /ə/, e.g., [čəm:əčõ] [čəmčõ] "spoons" (obl.).

Thus we see that in many cases the ə-deletion rule applies even in the environment $VC^2_1_CV$. However it doesn't apply if there are three consonants preceding the /ə/, thus /ləkšmən+õ/ is [ləkšmənõ] "a name" and not *[ləkšmnõ]. This is because four-consonant clusters, as mentioned in chapter 2, are not tolerated.

All the above facts suggest the following:

(b) If the cluster that would result from the deletion of the /ə/ would violate the constraints of Hindi, then the /ə/ is not deleted.

4.6.4. Now let us consider the word [b^hagi:rət^{h}i:] "Ganges" of table 37, in which speakers do not delete the /ə/. The cluster -rt^h-, which would result if the /ə/ were deleted, is an acceptable cluster, since it does occur in Sanskrit loans, and [b^hagi:rət^{h}i:] is a Sanskrit loan. It is, however, a rather uncommon word. It appears that a /ə/ is deleted in such Sanskrit loans only if it is familiar, otherwise the /ə/ is retained. Thus a number of the Sanskrit loans in table 29 such as [suločna], [sərla], [wimla], etc., have been borrowed into Hindi as proper names for girls, similar to [sundri:] of table 35. In these cases the /ə/ is always deleted, by all speakers that I know. It would sound affected to pronounce the girl's name [sərəla] instead of [sərla]. Similarly the name of a particular goddess, whose name means "cool", is derived from /ši:təl+a/ (cool+feminizing suffix) and is pronounced [ši:tla], that is, with the /ə/ deleted. However, when not used as a name, some speakers may retain the /ə/ when speaking formally.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the case of the PA loans of table 29, Urdu speakers retain the /ə/ -- at least in formal style; I am not sure if they retain the /ə/ in casual styles of speaking. However all non-Urdu speakers delete the /ə/.

Based on the above observations we can propose the following two principles.

(c) If the word is a Sanskrit or Urdu loan, speakers aspiring to the most prestigious pronunciation will render these words as they were rendered in the original language.

(d) /ə/'s tend to be deleted in casual style of speaking and retained in formal style, other things being equal.

Some of the other variables that are involved in the deletion or retention of the /ə/ are the following:

(e) /ə/'s tend to be deleted in a faster tempo of speaking. For example in the alternations illustrated in table 33 this could be one of the variables involved.

(f) If the environment in which the ə-deletion rule is to apply is supplied by one of the rule-blocking suffixes such as -iya or -izm, then the /ə/ is not deleted (see chapter 2, section 2).

We thus see that the deletion or non-deletion of the /ə/ is a complicated matter, and the formalism of the rule cannot involve just the immediate phonetic environment.

Since some or a number of the variables given above play a part in whether a speaker deletes the /ə/ or not at any given time, we expect to find variation amongst speakers, and within the same speaker in the case of a number of words.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This is not to say, however, that there are not certain areas where the deletion (or non-deletion) of the /ə/ is not fully determinate. In the case of forms derived from verbs, e.g., [nikəl] "come out", [nikla] "came out", etc., the /ə/ seems to be lost without question for all speakers I know of. These forms have been in the language for a long time. Also almost all[14] the verbs of Hindi are native words, and therefore the question of 'status' involved in retaining the /ə/ does not arise.

4.6.5. As a final point we might note that if a consonant cluster (of any type) follows the /ə/, the /ə/ is never deleted (here the non-deletion is fully determinate), as can be seen from the following words:

(8)	/pak^{h}ə̤n̤d̤+i:/	is	[pak^{h}ə̤n̤di:]	*[pak^{h}n̤d̤i:]	"hypocrite"
	/pələŋg+õ/	"	[pələŋgõ]	*[pəlŋgõ]	"beds" (obl.)
	/šətrəñǰ+õ/	"	[šətrəñǰõ]	*[šətrñǰõ]	"chess sets" (obl.)
	/kələŋk+i:/	"	[kələŋki:]	*[kəlŋki:]	"one with a spot"
	/prəsəŋg+õ/	"	[prəsəŋgõ]	*[prəsŋgõ]	"contexts"
	etc.				

This fact has already been incorporated in the ə-deletion rule proposed by the formulation of the environment following the /ə/ as CV.

4.6.6. To summarize: I propose that the ə-deletion rule applies within a morpheme and across morpheme boundaries to the right. It applies to both native and loan vocabulary. It applies even if certain types of clusters of two consonants precede but does not apply if a cluster follows. I tentatively accept the proposal that the rule

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

applies from right to left. The following factors determine the deletion or non-deletion of the /ə/:

(a) Whether the resulting cluster violates the constraints of the language.

(b) Whether the speaker has posited a morpheme boundary in the environment preceding the /ə/.

(c) Whether the word is a loan from Sanskrit or PA and is used in 'high vocabulary', and whether the speaker is aspiring to prestigious speech.

(d) Whether the style of speaking is formal or casual.

(e) Whether the tempo of speaking is fast or slow.

(f) Whether or not the suffix providing the environment is a blocking suffix, i.e. a suffix marked [-ə-deletion rule].

To incorporate all of the above, I propose the following formalism of the ə-deletion rule:

rule (9) $$ə \longrightarrow \emptyset \; /VC_1^2 ____ CV/ \; \left\{ \begin{bmatrix} +\text{loan} \\ +\text{casual speech} \end{bmatrix} \right\}$$
$$\begin{bmatrix} +\text{normal tempo} \end{bmatrix}$$

Condition 1: There may be no '+' in the environment to the left.

Condition 2: The output of the rule will not violate the sequential constraints of Hindi.

Convention: The rule applies from right to left.

In regard to condition 2 one might note that if there are certain constraints that hold both at the phonetic and the systematic phonemic level, it seems reasonable that they will not be violated by rules in the language. This point should be investigated in other languages too. The 'rule conspiracy' that Kisseberth (1970) refers to (see chapter 2 for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

details), seems to be due to such a constraint that holds at both the phonetic and the systematic phonemic level.

The above rule takes care of variables (a), (c), (d), and (e). Variable (b) would of course be accounted for, because the morpheme in question either will or won't be entered with a morpheme boundary '+' in the lexicon. And similarly (f) is accounted for because the suffix that blocks the ə-deletion rule will be entered as such.

There might, of course, be certain words where none of the above variables are involved and yet the [ə] is not deleted. This might be because usage has dictated the form of that word. That is, certain words might be exceptions, and will have to be so marked in the lexicon.

4.7. The question arises as to how one verifies a statement that a /ə/ has been deleted, phonetically. Is it necessary to have vocal cord vibration with the accompanying formant structure of a vowel, or would it suffice to have simply a release of an obstruent, with no obvious formant structure in the resulting turbulent noise? I propose that there has to be the formant structure of a vowel for it to be considered a /ə/. The release of an obstruent even if it contains voicing will not count as a /ə/ unless there are concomitant vowel formants.

I can illustrate what I mean with instrumental data. The data is from a male Standard Hindi speaker, a native of Agra (who also knows English). Figure 12 shows an oscillographic representation of the word [bɛṭ^{h}ək] "a meeting". The parameters recorded are from the top: the (raw) audio microphone signal, the rectified integrated audio

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

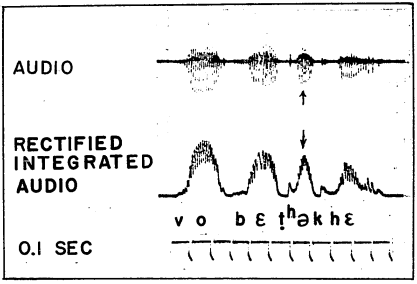


Figure 12. In this and the following figures the top line shows the raw microphone signal and the second line the rectified, integrated microphone signal. Here there is a clear short vowel between the [$t̪^h$] and the [k] (see arrow).

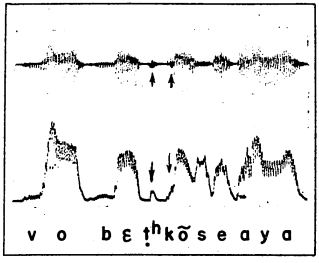


Figure 13. In the word [$bɛt̪^hkõ$] there is clearly a release of the [$t̪^h$] (see first arrow), but there is no voicing between that and the release of the [k] (see second arrow).

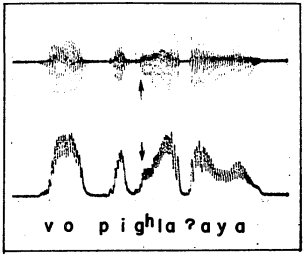


Figure 14. In the word [$pigʰla$] the [$gʰ$] has a voiced release (see arrow) but the [l] begins immediately afterwards. Thus there is no intervening [ə].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

signal, and a 0.1 sec. time standard. Here one can clearly see a vocalic element after the [ṭʰ]. Figure 13 shows [bəṭʰkõ] "meetings". Here one can identify the release of the [ṭʰ] but there is nothing that can be identified as a vowel, between the [ṭʰ] and the [k]. Figure 14 shows the word [pigʰla] "melted", where one can see that the release of the [gʰ] has voicing, however this cannot be identified as a vowel (in fact it is released directly into the [l]). The spectrographic record in Figure 16 shows the word [ɟəŋgəl] "forest". Here we can identify the [ŋ] and something that looks like a [g], and then the [ə] before the [l]. Figure 15 shows the word [ɟəŋgli:].[15] Here we do not see any evidence of a ə-like vowel.

4.8. I have mentioned that the ə-deletion rule is still undergoing change. This is clear from a brief examination of its history.

Historically the ə-deletion rule is quite an old rule, although originally it was a short-vowel-deletion rule and did not solely involve a /ə/. It originally applied to Old Hindi medial short vowels, i.e. i, u, and ə (the data on which the above statement is based is from Misra 1967), as follows:

Rule (10) $\begin{bmatrix} V \\ -long \end{bmatrix} \longrightarrow \emptyset \ / \ (C)VC___CV:$

(Where long vowels were a i: u: e ɛ o ɔ).

Misra gives examples of the following sort:

Old Hindi	Modern Standard Hindi	
bəgula	[bəgla]	"a bird"
kəṭʰinai:	[kəṭʰnai:]	"difficulty"
ukəsana	[uksana]	"to incite"

Misra further says that whereas the ə was always lost in the above

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

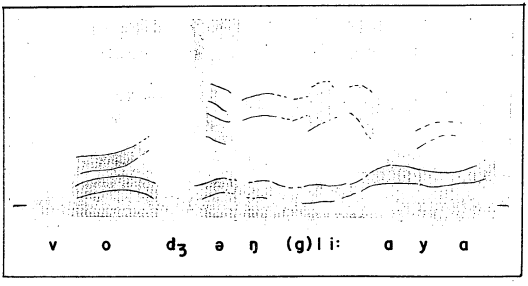


Figure 15. (Slightly retouched: formants outlined.)

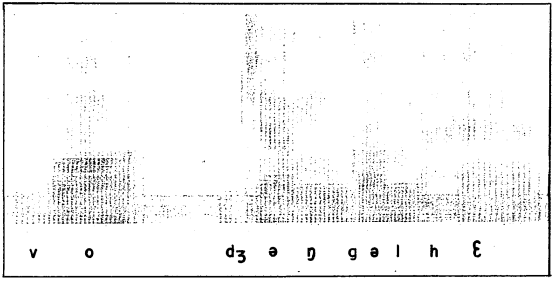


Figure 16.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

environment, in the case of i and u there were some exceptions, thus Old Hindi (OH) pəhũčə "reached" remains unchanged in Modern Standard Hindi (MSH).

The rule in MSH of course applies only to a /ə/. Another change is that in pre-MSH the following vowel had to be long, but in MSH, as can be seen from examples such as /d^{h}ə̤rək+ən/ [d^{h}ə̤rkən] "a beat", the rule has become more general and applies regardless of whether the following vowel is short or long. However, this generalization of the environment is not complete, and certain suffixes with short vowels block the ə-deletion rule, e.g., -iya.[16] Thus it is possible that originally the reason why these suffixes blocked the ə-deletion rule was due to their having short vowels, but now it is not the short vowel but the suffix as a whole that does the blocking.

However, it seems that this vowel deletion rule is still in the process of change, and thus for some speakers (as mentioned in chapter 2) suffixes like -iya block the rule and for others they do not. Thus the dialect for which it is not a blocking suffix probably represents the innovating dialect, where the environment of the rule has been generalized so far as to eliminate exceptions. Another change that the rule's environment is undergoing is that now it applies even when a consonant cluster precedes. Thus it seems that the environment is becoming even more general.

The ə-deletion rule is interesting for another reason: it creates consonant clusters (at the phonetic level). By the time of late Middle-Indo-Aryan (MIA), most Sanskrit medial clusters had been

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

simplified so that there were just geminates (if these are considered clusters) and clusters of the type: homorganic nasal+stop. The general sound change in late MIA which -- with few exceptions -- simplified even these clusters (with a compensatory lengthening of the preceding vowel), left the language (now Old Hindi) with almost no consonant clusters medially except in loans. But then, as more and more loans came into Hindi from other Indo-Aryan languages (such as Punjabi) and from Perso-Arabic, Sanskrit, English, etc., more and more clusters entered the language. And, of course, the short vowel deletion rule of pre-MSH created clusters even in native words. The ə-deletion rule of modern Hindi is creating more clusters. In fact, certain clusters such as -wṭ- exist only due to its application. In Hindi, at the present time, three-consonant clusters are still rather rare, but since for a number of speakers the ə-deletion rule applies in the environment VCC__CV, three-consonant clusters are bound to increase (at the phonetic level). At the lexical level, however, the tolerance of clusters still seems to be imperfect.

4.9. The question can be raised, why a ə-deletion rule instead of a ə-insertion rule?[17] That is, given a typical word pair [čipək] "stick" [čipka] "stuck", why could we not just as well derive the first word from the second by inserting a /ə/ rather than by deleting the /ə/, i.e. derive the second from the first?

First, all writers to treat the subject so far, Kellog (1965), Guru (1962), Srivastava (1969), Pray (1970), and Narang and Becker (1971), have treated it as a ə-deletion rule. This is not a weighty argument since linguists can be wrong, but it must be granted that all

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of these writers could not have overlooked something this obvious.

Second, as stated in 4.8., the origin of this alternation, historically, was a short vowel (including ə) deletion rule. Of course, strictly speaking, historical facts are not relevant as such to a synchronic grammar, since a speaker does not carry around the history of his language in his head. However, the fact that it was a ə-deletion rule once suggests that it may be so today. Moreover, it is entirely likely, as mentioned in the previous section, that this same rule (somewhat modified as to the vowels it deletes and to the environment in which it applies) is still productive today.

But the most important argument against a ə-insertion rule is that it would overly complicate the situation.

(a) Obviously in typical pairs such as [čipək] [čipka] the one with the [ə] is the basic form and the one without the [ə] the derived form. Thus we cannot derive [čipək] from [čipka]. If we were to posit a stem /čipk/ (from which both forms were to be derived), and a ə-insertion rule in order to obtain [čipək] we would be claiming that the underived, basic form is more complex than the derived form (since no change would be necessary in the stem in order to get [čipka]). This is intuitively undesirable.

(b) With a ə-insertion rule we would presumably give the underlying forms of [bistər] [bistra] as /bistr/ and then have a /ə/ inserted after the first two consonants in order to derive the first word. But we would have to specially mark as [-ə-insertion rule] certain other words where a /ə/ is not to be inserted, e.g. [influɛnza] "influenza" where we don't want *[infəluɛnza] or *[inəfluɛnza], or

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[dəst] "diarrhea" where we don't want *[dəsət]. This certainly is undesirable, especially in the case of [dəst], because this makes it appear as if [dəst] is in some way an exception, where the fact is that [dəst] contains a permissible cluster and thus is regular. One never encounters speakers attempting to render this as [dəsət] which one would expect if it really were exceptional. Moreover we cannot simply have the ə-insertion rule restricted such that it not insert a /ə/ between st clusters, because of alternations like [ɔsət] "average" [ɔstən] "average"+adverbializing suffix. (Speakers vary between deleting the /ə/ or not deleting it in this word, but at least for speakers who do delete the /ə/ my argument applies). Here if we were to posit /ɔst/ and derive [ɔsət] from this by the ə-insertion rule, we have to allow the ə-insertion rule to insert a /ə/ in some st clusters but not others.

(c) There would be a large number of words with underlying forms with clusters that violate the sequential constraints that I have stated. In other words some of the constraints could no longer be stated for the lexical level, e.g., the constraint that nasals are homorganic before stops, which is a highly plausible rule. If we posit a ə-insertion rule a form such as [čəmək] "brightness" would have the underlying form /čəmk/ with a final -mk.

Morpheme-medially, the generalization that in the majority of cases nasals are homorganic before stops could not be reflected in the grammar at all. As mentioned in chapter 2 this tendency does not hold morpheme-medially at the phonetic level. However if we posit a ə-insertion rule, we cannot posit morphemes such as [kənk^{h}i:]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

"sideways glance" to be at the lexical level /kəəkʰi:/ since we would not have a ə-deletion rule to give us the correct phonetic output. Therefore our grammar could not reflect the fact that a large number of the words have homorganic nasals before stops. This is undesirable, especially since we can no longer account for the different treatment speakers give to [inkar] "denial" as opposed to [kəŋkʰi:] discussed in chapter 2 (namely, that they sometimes render the former with a homorganic nasal but never the latter).

(d) As mentioned before, the ə-deletion rule is used as an orthographic-to-phonetic conversion rule. The Devanagari orthography frequently inserts (or implies the presence of) /ə/'s not actually pronounced. Native speakers use a ə-deletion rule to derive the correct pronunciation of words when they learn to read Devanagari. When readers come across unfamiliar words which are written in Devanagari without a cluster, they still read them with a cluster if the environment of the ə-deletion rule is met. Of course sometimes the result is an "incorrect" pronunciation if they didn't know that the word contained a morpheme boundary. Since they learn to speak before they learn to read, we can probably say that they also use the internalized ə-deletion rule for reading Devanagari. However, if we posit a ə-insertion rule, we are claiming that upon beginning reading they have to learn to use a completely different rule: the ə-deletion rule.

4.10. Before ending the discussion on the ə-deletion rule I should perhaps indicate why I haven't adopted those versions of the ə-deletion rule proposed by other linguists.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Grammarians writing in Hindi (e.g., Guru 1962), and writers such as Kellog 1965, have treated the deletion of the /ə/ under so-called pronunciation rules, i.e. rules on how to read Devanagari. I will not go into any details of these works here; for a discussion of Kellog's rules see Narang and Becker.

4.10.1. Among recent proposals of the ə-deletion rule in the generative framework, Srivastava's (1969:919) formulation was perhaps the first. He gives it as:

$$(11)\ \begin{bmatrix} V \\ -\text{tense} \\ +\text{low} \end{bmatrix} \rightarrow \emptyset \ /\ (C)VC___\left(C\ \begin{bmatrix} V \\ +\text{tense} \end{bmatrix}\right)\ \#$$

His rule thus applies in two environments:

(a) /VC___C V [+tense] #

(b) /VC___#

Srivastava considers all Hindi words to have a word-final vowel. In the case of words which phonetically end in a consonant, he posits a final /ə/ at the lexical level. Thus the word [g^hər] "house" is claimed to be bisyllabic at the lexical level: /g^hərə/. Therefore in the above rule he has to mention the word boundary '#', since he also uses the above rule to delete the word-final /ə/ to obtain the correct phonetic output. He gives no reason or evidence for positing this underlying word-final /ə/ which does not occur at the phonetic level. Therefore I have not adopted this part of his analysis.

Narang and Becker (1971) correctly point out that, as stated, Srivastava's rule (part (a)) will not work with words such as /nikəl+ega/ "come out"+future tense. The rule can of course easily be modified to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(12) ə --> Ø / (C)VC___ $\left\{ \begin{matrix} V \\ C\,[+\text{tense}] \\ \# \end{matrix} \right\}$

and this would meet their criticism. Thus part (a) of Srivastava's rule could best be re-written as:

(a') /VC__C $\begin{matrix} V \\ [+\text{tense}] \end{matrix}$

This formulation now is very similar to the historical rule (Rule 10) given under 4.8. for pre-MSH. However, as I have pointed out, the condition that the following vowel be 'tense' does not hold for modern Hindi, as can be seen from examples such as:[18]

(13)	/d^{h}ərək+ən/	[d^{h}ərkən]	"a beat"
	/čipək+ən/	[čipkən]	"stickiness"
	/dəršən+ik/	[daršnik]	"philosopher"
	/wičəl+it/	[wičlit]	"moved"
	/ərəb+istan/	[ərbistan]	"Arabia"

4.10.2. The next formulation of this rule in the generative format is that of Pray (1970:93), who gives it as:

(14) a --> Ø / VC___CV

(Pray uses a for [ə]). Pray states that the rule applies first to the right-most /ə/. I have discussed this right-to-left application of the ə-deletion rule under 4.5.1. and tentatively accepted this convention.

Pray also suggests that the rule should first apply within a stem and then to the stem+ending. The example he gives is /kərəwəṭ+ẽ/, where first the rule has to apply to kərəwəṭ to give [kərwəṭ] "side" and then to /kərwəṭ+ẽ/ to give [kərwəṭẽ] "sides". The /ə/ that Pray posits in the underlying form of [kərwəṭ], between the /r/ and the /w/,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is an 'abstract' /ə/ since it does not occur at the phonetic level, and there are no alternations where it shows up. Pray gives no evidence (or reasons) for positing this underlying /ə/. I have tried to provide evidence in chapter 2 for the abstract /ə/ in certain types of words; however, the test I designed would not work for words of the kərwəṭ type because the suffix added would only affect the /ə/ on the right, i.e. the third /ə/ in /kərəwəṭ/. If some evidence can be found for positing a /ə/ in [kərwəṭ] and similar words, then, of course, Pray's suggestion of the ə-deletion rule first applying to the stem would have to be adopted. For the present I am not adopting it since there is no evidence for the abstract /ə/ in words of this type. Therefore it seems simpler to posit the underlying form of [kərwəṭ] to be identical to the surface form; upon addition of the inflectional [ĕ] the final /ə/ of [kərwəṭ] is not deleted because the resulting three-consonant sequence would be an unlawful cluster.

Neither Pray nor Srivastava mention that the rule has to be restricted so that it does not apply across morpheme boundaries to the left.

4.10.3. The most detailed discussion of the ə-deletion rule is that of Narang and Becker (1971:648). They give the rule as:

(15) $\begin{bmatrix}+\text{syllabic}\\+\text{compact}\\-\text{tense}\end{bmatrix} \longrightarrow \emptyset \;/\; VC__C+V$

4.10.3.1. Narang and Becker do note that prefixes do not provide the environment for the ə-deletion rule, however they handle this differently from the way I proposed in 4.2. They posit words such as [amərəy̆ṛ] "until death" to be /a#mər+əy̆ṛ/ "On the same 'general

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

syntactic' grounds which prompt Chomsky and Halle .... to analyze English resolve 'solve anew' as /rE#sɔlv/" (651).

There are two possible criticisms of this solution. Firstly, Chomsky and Halle retain the word boundary '#' around lexical categories or replace it with a formative boundary '+' by the use of ad hoc readjustment rules. The motivation for this is to restrict application of certain phonological rules only to the "right" words, the right words being marked by which boundary they contain (see Chomsky and Halle pg. 12-13). Thus they posit /rE#sɔlv/ "solve anew" and /rE=sɔlv/ "determine" with different boundaries because this is required to make certain rules work (e.g., the 's'-voicing rule, the stress rule, etc.). This is all quite ad hoc as Chomsky and Halle admit, and amounts to the diacritic use of boundary features. Narang and Becker apparently advocate the same ad hoc procedure, but with much less evidence, e.g., they do not mention any other rules in Hindi phonology which would require this.

If prefixes do not in fact provide the environment for the ə-deletion rule, it seems to me more explanatory to posit this as a condition on the rule as I have proposed, rather than to use boundaries as diacritic features to accomplish this. The use of diacritic features does not make a less general situation more general.

Secondly, a more damaging argument against Narang and Becker's proposal is that it won't work -- or more exactly, in its present formulation it is less general than it could be and so fails to account for certain phonological patterns that it should account for. Narang and Becker claim it is a boundary '#' between the prefix and the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stem in /a#mər+əṛ̃/ that blocks the ə-deletion rule. But as I have pointed out, it is not only a prefix boundary that prevents the rule from applying but any morpheme boundary to the left. Thus in /kəla+wət+i:/ "one who has fine arts, name for a girl" [kəlawəti:] we find that the /ə/ of the suffix wət is not deleted. Since Narang and Becker set up a morpheme boundary (and not a '#') between mər and əṛ̃ of [amərəṛ̃], to be consistent they would have to set up a morpheme boundary '+' even between kəla and wət of [kəlawəti:], but now their ə-deletion rule would apply and give the wrong result *[kəlawti:]. The ə-deletion rule I have proposed to handle forms such as [amərəṛ̃] will also handle forms such as [kəlawəti:] and no extension will be needed, nor will there be any need to change the underlying forms of the morphemes so as to accomodate the rule.

4.10.3.2. Narang and Becker's requirement that the CV following the /ə/ contain a morpheme boundary '+' (i.e. that the /ə/ is deleted in the environment VC__C+V) also requires comment. Because of this condition their rule cannot apply morpheme-medially (see 4.4. for discussion on the desirability of having the rule apply morpheme-medially). For example, we would be missing a generalization if we failed to account for the alternation [rəɲni:] [rəɲəni:] "night" depending on formal/informal speech without using the ə-deletion rule.

The reason they give (in their footnote 3) for positing a C+V is that otherwise the /ə/ in [waraṛ̃əsi:] "Benares" would be incorrectly deleted. However, the non-deletion of the /ə/ in this word can be accounted for in a number of other ways:[19]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) The cluster that would result if the /ə/ were deleted, ṉs, is not a permissible cluster, and this accounts for the non-deletion of the /ə/. Of course, a large number of speakers probably pronounce the word with a [n] (i.e. a non-retroflex nasal) and thus ns as a cluster would be permissible. I do not know if these speakers delete the /ə/ or not.

(b) The elite speakers who know Sanskrit (the word is a Sanskrit loan, and in spite of the Government pushing it, most people still use the native word [bənarəs] for the city) would not delete the /ə/ in the word since Sanskrit had the /ə/ (the derivation of the word in Sanskrit is something like /wara+ṉəs/ "water"+suffix, however I am not sure if speakers are aware of this boundary), i.e. they would simply treat the word as an exception to the ə-deletion rule.

The principle that the applicability of the ə-deletion rule depends on the permissibility of the consonant cluster that would result can also be applied to reveal an error in one of the examples Narang and Becker give. They list [adət] "habit" and [adtẽ] "habits" (647), but as I have pointed out [adtẽ] is not a lawful word for Hindi because of the unpermissible cluster -dt- (we either hear [adətẽ] or [at:ẽ]). This is supported by the responses of the subjects in the informal test reported in 4.6.2. The fact that the resulting clusters determine in part whether or not a /ə/ is deleted was noticed by Hoenigswald in 1948 who said "when a consonant precedes, the treatment [of the ə] varies according to the consonant cluster which results or would result" (142, footnote 7).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.10.3.3. None of those who previously treated the ə-deletion rule noticed that, for a number of words such as those given in table 35, the /ə/ is lost in spite of a preceding cluster (instrumental verification for which was given in 4.7.). Narang and Becker give [sundəri:] "beautiful woman" as an example of a word where the [ə] does not get deleted (648). However, they qualify this in their footnote 21 and note that there is an "optional rule which allows most speakers of Hindi-Urdu to delete a medial [ə] before a sequence of liquid plus vowel. As a result of this optional deletion rule, the word for 'beautiful woman' which we have transcribed as [sundərĩ] is frequently heard as [sundrĩ]." They also claim that this 'optional deletion rule' is not the same as their P-1 (the ə-deletion rule). First, as is obvious from table 35, it is not simply the fact that only a liquid+vowel allows the /ə/ to be deleted; there are many other environments as well (e.g., /paṇḍəw+õ/ is [paṇḍwõ] "Pandawas" (obl.)).

Second, they miss an obvious generalization by claiming that this optional ə-deletion rule is different from the main ə-deletion rule. I have shown above that they are the same rule.

Third, the rule is not really optional in this environment, since more can be said about the conditions under which it applies. I have indicated above the kind of sociolinguistic factors relevant to the application of this rule in this and other phonological environments.

4.10.3.4. In general it is difficult to verify Narang and Becker's statements as to whether the /ə/ is or is not deleted, because they do not tell us how we would verify such a statement, i.e. when they do or

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

do not consider a /ə/ to be deleted phonetically.

They also fail to give important information on their informants. E.g., how many informants? Their background? Did they speak any other languages besides Standard Hindi? Were they literature specialists? All these factors are important. For example, Narang and Becker further state in their footnote 21:

> we are aware of the fact that some speakers of Hindi-Urdu, unlike the majority of our informants, delete the shwa even in such forms as [ãgənõ] 'courtyards' (obl.pl.).... We suggest that such speakers, while having the same underlying representations as those who retain the shwa, apply rules P1 and P2 in the order P2-P1. (666)

They claimed earlier in footnote 1 that they are describing the Hindi-Urdu spoken in cities, in Uttar Pradesh, Delhi, etc. I have lived in Banares, Delhi and Lucknow, and in my experience all speakers delete the /ə/ in the words given in casual speech -- the only informants I found who do not delete the /ə/ were literature specialists (for example the fourth informant mentioned in footnote 13), and the speech was not casual since I was trying to elicit data (and even this fourth informant deleted the /ə/ in /paṇḍəw+õ/).[20] It is, of course, possible that some speakers do not delete the /ə/ in the word for "courtyards" even in casual speech; but then this is a minority dialect, certainly not the "common core" that the authors claim they are describing.

Moreover Narang and Becker ascribe this 'difference of dialects', in deleting or not deleting the /ə/, to the difference in the order of their P1 and P2 in informants, i.e. whether the order is P1,P2 or P2,P1. (P1 is their ə-deletion rule, and P2 is V:NC--> $\tilde{V}$:C). However, they give no evidence for making this claim.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As I have mentioned before, the deletion or non-deletion of the /ə/ depends in part on various sociolinguistic factors.

Narang and Becker do say in a footnote that there is some variation in the ə-deletion rule, in that in PA loans involving a /ə/ Urdu speakers do not delete the /ə/. However, they do not mention that 'high Hindi' speakers similarly do not delete the /ə/ in Sanskrit loans (this was discussed in 4.6.4.).

For the reasons discussed above I found all the versions of the ə-deletion rules proposed by the above-mentioned writers inadequate, and have proposed the formulation given in Rule (9). It is quite possible that, as more investigation is done, even Rule (9) would be found to be inadequate; however, so far I think it accounts for all the facts in the most explanatory manner.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Footnotes

1 Some of the words also involve other types of vowel changes which need not concern us here.

2 There are some words such as [milni:] "a ceremony at weddings" which seem to be exceptions, and indeed would be if it were derived from /mil+ən+i:/, rather than the correct /mil+ni:/, i.e. [milni:] is not derived from /milən/ (which itself is /mil+ən/) but from /mil/. The suffix -ni: occurs in words like [or^{h}ni:] "scarf" from /or^{h}+ni:/, [məŋgni:] "engagement" from /maŋg+ni:/, etc. (see Upraiti 1964).

3 Perhaps [dubla] "thin", [əbla] "woman" are considered to be /dubəl+a/ and /əbəl+a/. In the case of the former we do get changes like [dubla] [dubli:] depending on the gender of the noun that "thin" is modifying. This gets into the question of whether we should posit the stem to be /dubəl/ to which /a/ or /i:/ is added, or posit /dubəla/ with the final /a/--> /i:/ in the appropriate environment (i.e. in the case of the following noun being feminine). Discussion of this question is beyond the scope of this study.

4 Pray posited the underlying form /nikələva/ for the following reason: He noted that "Verb stems consist of alternate vowel and consonant segments" (91). (Certain apparent exceptions such as [ləṭka] "hang", he notes are actually /ləṭək+a/ and thus not really exceptions.) He further says

> In general, all occurrences of CC within verb stems ... are to be interpreted as CaC [Pray uses a for ə and aa for a] from which the a is later deleted. Thus the causative stems in vaa, e.g. ... /nikalvaa/ 'to have taken out', [by the '7' Pray does not mean underlying forms] are assumed to have underlying representations ... nikalavaa .... (93)

I do not see any need for adopting Pray's analysis, because I fail to see why [nikəlva] is to be considered a stem. [nikəlva] consists of the stem nikəl plus the suffix va, it is not an exception to the statement on the verb stem structure, and thus there is no need to posit the abstract underlying /ə/ which has to be deleted by the ə-deletion rule.

5 This is an interesting example of what looks superficially like a change due to metathesis but is really something else.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6 The 'A' and 'B' columns on the right of the table refer to responses of informants 'A' and 'B'. For the present we can ignore these columns. I will discuss their responses in detail shortly.

7 The morpheme break /dik^{h}a+wəṭ+i:/ in [dik^{h}awəṭ] and [bənawəṭ] are from Upraiti. In the case of [bunawəṭ] "the knittage", Upraiti gives /bun+awəṭ/. He does not give any reasons for the different treatment of these words, i.e. why we couldn't derive [dik^{h}awəṭ], etc., from the intransitive verb /dik^{h}/ plus a suffix "to be shown" /dik^{h}+awəṭ/, instead of deriving it from the transitive verb /dik^{h}a/ "to show" plus /wəṭ/. If it were the case that some speakers posited /dik^{h}+awəṭ+i:/ and others /dik^{h}a+wəṭ+i:/, this would be another explanation for the variation between [dik^{h}awəṭi:]~[dik^{h}awṭi:], because in the case of /dik^{h}+awəṭ/ there is no '+' in the environment to the left, whereas in the case of /dik^{h}a+wəṭ/ there is.

8 In the case of [di:pək] "lamp" we have two words for "lamp", [di:pək] and [di:p]; both are loans from Sanskrit (the native word is [diya]). One can speculate that possibly some speakers treat them as two separate words whereas others derive [di:pək] from /di:p+ək/; in the latter case the /ə/ would not be deleted.

9 It is also the case that the three-consonant clusters that would result medially, if the /ə/ were to be deleted in the words in table 36, do not occur in other words of the language. However this in itself could not be an inhibiting factor, since a few of the three-consonant clusters of table 35 (e.g., -štm-) also do not occur in words except those to which the ə-deletion rule has applied.

10 I should mention that perhaps a morpheme boundary '+' could be posited in [prart^{h}na], since we have related words such as [prart^{h}it] "prayed for, desired", which might suggest that we have a morpheme /prart^{h}/ and thus the -rt^{h}- is not really an example of a morpheme-medial cluster, but a morpheme-final cluster, and could be left out of the discussion on clusters relevant to the ə-deletion rule. This is certainly a possible analysis which has the advantage of eliminating the troublesome -rt^{h}- (which occurs only in this word) from this discussion, i.e. we could then say that the /ə/ is not deleted in [b^{h}agi:rət^{h}i:] because it is an impermissible cluster. However I think only speakers who have the form [prart^{h}it] in their vocabulary, and who know Sanskrit, would 'know' of this morpheme boundary.

11 Except for a few Sanskrit and English loans such as [pəŋkti:] "line" (S), [inspɛktər] "inspector" (E), the third consonant in three-consonant clusters always seems to be a sonorant.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Also -ǰg- and -gm-, which would occur as part of the three-consonant cluster in items 5, 17, and 16, occur only in very few loans -- the former only in Sanskrit loans (I have only one example, the word [əǰgər] "python") and the latter only in a few PA loans.

12 These informants are the same as those used in the tests reported in chapter 2.

13 I checked this with four informants. Two of these are informants 'A' and 'B' mentioned earlier. Of the other two, one is from Agra (and knows English besides Standard Hindi), and the other is from Banaras (he knows Bhojpuri, German and English, besides Standard Hindi).

14 I say almost all, because there are a few (very few) verbs such as [filma] "to film" (derived from the noun 'film') which are non-native.

15 Interestingly there does not seem to be an obvious [g], i.e. in certain types of clusters the [g] (and possibly other stops also after homorganic nasals) is not present phonetically. Of course we would still have to posit it at the underlying level, because when a vowel follows as in [ǰəŋgəl], the /g/ is phonetically manifested; we also need to posit the /g/ in order to state the nasal assimilation to following stops.

16 This type of blocking only occurs with suffixes with initial short vowels, since we would not find a morpheme- or word-final short vowel in Hindi; as mentioned in chapter 2, short vowels do not occur in final position.

17 I am not referring to the ə-insertion rule that is used by speakers to break up clusters that they cannot pronounce, e.g., the rendering [səṭeśən] for English [steśən] "station" by some speakers. Inserting a [ə] is the favorite method of breaking up such clusters, however not the only one: the other is to delete one of the consonants. Thus some speakers say [ṭeśən]. And as far as initial clusters are concerned, a third way is to make the cluster medial; thus some speakers say [isṭeśən]. The above simplification, for example by inserting a [ə], reflects a pronunciation constraint.

18 Narang and Becker also note in their footnote 3 that the vowel following the /ə/ need not be 'tense', but the example they give to demonstrate this, is the pair [mu:rət] "picture" [mu:rti] "idol". This is a bad example since the final i is tense for all speakers that I know -- except purists trying to pronounce the final i as short because it was short in Sanskrit.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

19 I should mention that [waraɾ̃əsi:] is the only example of this kind that I have been able to find.

20 Also as I have already indicated in chapter 3, the word for "courtyard", for the majority of speakers, has a homorganic nasal [ãŋgən] and not [ãgən] as Narang and Becker have it.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

Arun, V. B. 1961. A comparative phonology of Hindi and Panjabi. Ludhiana: Panjabi Sahitya Akademi.

Bach, E. 1968. Two proposals concerning the simplicity metric in phonology. Glossa 2.128-49.

Beames, J. 1970. A comparative grammar of the Modern Aryan Languages of India. New Delhi: Munshi Ram Manohar Lal. (Reprint of the 3 volumes originally published in London, 1872.)

Bellugi, U. 1964. The emergence of inflections and negation systems in the speech of two children. Paper presented at the New England Psychological Association.

Berko, J. 1958. The child's learning of English morphology. Word 14.150-77.

Bhatia, K. C. 1964. Consonant-sequences in Standard Hindi. Indian Linguistics 25.206-12.

Brame, M. 1969. On the abstractness of phonology. Unpublished manuscript.

Bright, W. O. 1965. Languages of South Asia. To appear in: The Indian sub-continent: Peoples and cultures of South Asia, ed. by D. P. Sinha and Purnima Sinha for the East-West Center, University of Hawaii.

Brown, G. 1969. Syllables and redundancy rules in generative phonology. Journal of Linguistics 6.1-17.

Chomsky, N. 1964. Current issues in linguistic theory. The Hague: Mouton.

_____, M. Halle, and F. Lukoff. 1956. On accent and juncture in English. In: For Roman Jakobson. The Hague: Mouton. 65-80.

_____, and M. Halle. 1968. The sound pattern of English. New York: Harper & Row.

Crothers, J. 1971. On the abstractness controversy. Project on Linguistic Analysis Reports, Second Series, no. 12. CR1-CR29. Berkeley: Phonology Laboratory, University of California.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Dixit, R. P. 1963. The segmental phonemes of contemporary Hindi. Austin: University of Texas master's thesis.

Esper, E. A. 1925. A technique for the experimental investigation of associative interference in artificial linguistic material. Language Monograph 1.

Fromkin, V. A. 1971a. The non-anomalous nature of anomalous utterances. Language 47.27-52.

_____. 1971b. In defense of systematic phonemics. Journal of Linguistics 7.75-83.

_____. 1971c. Simplicity is a complicated question. Working Papers in Phonetics 17.76-84. Los Angeles: University of California.

Gilchrist, J. B. 1820. The stranger's infallible East-Indian guide. 3rd ed. London: Black, Kingsbury, Parbury, and Allen.

Gleitman, L. R. and Gleitman, H. 1970. Phrase and paraphrase, some innovative uses of language. New York: W.W. Norton and Company, Inc.

Greenberg, J. H. and Jenkins, J. J. 1964. Studies in the psychological correlates of the sound system of American English. Word 20.157-77. Reprinted in: Readings in the psychology of language, ed. by L. A. Jakobovits and M. S. Miron. Englewood Cliffs: Prentice-Hall, Inc. 186-200.

Grierson, G. A. 1922. Spontaneous nasalization in the Indo-Aryan languages. Journal of the Royal Asiatic Society. 381-88.

Gumperz, J. J. and Naim, C. M. 1960. Formal and informal standards in the Hindi regional language area. Linguistic diversity in South Asia, ed. by C. A. Ferguson and J. J. Gumperz. Bloomington: Indiana University. 92-118.

Gupta, M. L. 1966. The system of consonantal clusters in modern Hindi. The study of sounds, 12. Phonetic Society of Japan. 525-40.

Guru, K. M. 1962. Hindi Vyakaran. Varanasi: Nagari-pračarini Sabha.

Halle, M. 1959. The sound pattern of Russian. The Hague: Mouton.

_____. 1964. Phonology in a generative grammar. In: The structure of language, ed. by J. A. Fodor and J. J. Katz. Englewood Cliffs, N.J.: Prentice-Hall. 334-52.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

_____, and Stevens, K. N. 1971. A note on laryngeal features. Quarterly Progress Report no. 101. M.I.T. Research Laboratory of Electronics. 198-213.

Harte, B. F. 1937. Mliss. In: The luck of roaring camp, and other sketches. The Spencer Press. 105-38.

Hoenigswald, H. M. 1948. Declension and nasalization in Hindustani. Journal of the American Oriental Society 68.139-44.

Husain, M. n.d. A phonetic & phonological study of the word in Urdu. (Department of Urdu studies, I.) Aligarh, India: Muslim University.

Hyman, L. 1970. How concrete is phonology? Language 46.58-76.

Kelkar, A. R. 1968. Studies in Hindi-Urdu, I: introduction and word phonology. (Building Centenary and Silver Jubilee series, 35.) Poona: Deccan College.

Kellog, S. H. 1965. A grammar of the Hindi language. London: Routledge & Kegan Paul Ltd. (First published in 1875).

Kiparsky, P. 1968. How abstract is phonology? Unpublished manuscript.

_____. 1971. Historical linguistics. In: Survey of linguistic science, ed. by W. O. Dingwall. University of Maryland. 576-649.

_____. n.d. Metrics and morphophonemics in the Rigveda. Unpublished manuscript.

Kisseberth, C. W. 1969. On the abstractness of phonology: the evidence from Yawelmani. Papers in Linguistics 2.248-82.

_____. 1970. On the functional unity of phonological rules. Linguistic Inquiry 1. 291-306.

Kornfeld, J. R. 1971. What initial clusters tell us about a child's speech code. Quarterly Progress Report no. 101. M.I.T. Research Laboratory of Electronics. 218-21.

Ladefoged, P. 1971a. The limits of phonology. In: Form and substance, phonetic and linguistic papers presented to Eli Fischer-Jørgensen, ed. by L.L. Hammerich, R. Jakobson, E. Zwirner, Akademisk Forlag. 47-56.

_____. 1971b. Phonological features and their phonetic correlates. Working Papers in Phonetics 21.3-12. Los Angeles: University of California.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Langacker, R. W. 1969. Mirror image rules II: lexicon and phonology. Language 45.844-62.

Lehiste, I. 1970. Suprasegmentals. Cambridge: M.I.T. Press.

Lightner, T. M. 1968. On the use of minor rules in Russian phonology. Journal of Linguistics 4.69-72.

Maddah, M. M. Khan (ed.) 1959. Urdu-Hindi Shabdakosh. Uttar Pradesh: Sučna Vibhag, Prakashan Shakha.

McCawley, J. D. 1968. The phonological component of a grammar of Japanese. Monographs on linguistic analysis, 2. The Hague: Mouton.

_____. 1970. Meaningful noises. The Times Literary Supplement. July 23, 817-8.

McGlone, R. E. and Shipp, T. 1971. Comparison of subglottal air pressures associated with /p/ and /b/. Journal of the Acoustical Society of America 50.139.

McNeill, D. 1970. The acquisition of language: the study of developmental psycholinguistics. New York: Harper and Row.

Mehrotra, R. C. 1959. Hindi syllabic structure. Indian Linguistics 20.231-37.

_____. 1964. Hindi phonemes. Indian Linguistics 25.234-46.

Misra, B. G. 1967. Historical phonology of Modern Standard Hindi: Proto-Indo-European to the present. Cornell University dissertation, unpublished.

Narang, G. C. and Becker, D. A. 1971. Aspiration and nasalization in the generative phonology of Hindi-Urdu. Language 47.646-67.

Netsell, R. 1969. Subglottal and intraoral air pressures during the intervocalic contrast of /t/ and /d/. Phonetica 20.68-73.

Ohala, J. 1970. Aspects of the control and production of speech. Working Papers in Phonetics 15. Los Angeles: University of California.

_____. 1971a. Monitoring soft palate movements in speech. Paper delivered at 81st meeting of the Acoustical Society of America, April 1971. Abstract in: Journal of the Acoustical Society of America 50.140.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

_____. 1971b. The role of physiological and acoustic models in explaining the direction of sound change. Project on Linguistic Analysis Reports, Second Series, no. 15.25-40. Berkeley: Phonology Laboratory, University of California.

_____. 1971c. How to represent natural sound patterns. Paper delivered at the winter meeting of the Linguistic Society of America, 1971. To appear in Project on Linguistic Analysis Reports, Second Series, no. 16. Berkeley: Phonology Laboratory, University of California.

Ohala, M. 1971. Conflicting expectations in the direction of sound change. Paper presented at the first All California Linguistics Conference, May 1971. To appear in Project on Linguistic Analysis Reports, Second Series, no. 16. Berkeley: Phonology Laboratory, University of California.

Öhman, S. 1966. Co-articulation in VCV utterances: spectrographic measurements. Journal of the Acoustical Society of America 39. 151-68.

_____. 1967. Numerical model of coarticulation. Journal of the Acoustical Society of America 41.310-20.

Pray, B. R. 1970. Topics in Hindi-Urdu grammar. (Research monograph 1.) Berkeley: Center for South and Southeast Asia Studies, University of California.

Qadri, S. G. M. 1930. Hindustani phonetics. Villeneuve: L'Union Typographique.

Saciuk, B. 1969. The strata division of the lexicon. Papers in Linguistics 1.464-531.

Schachter, P. and Fromkin, V. 1968. A phonology of Akan: Akuapem, Asante & Fante. Working Papers in Phonetics, no. 9. Los Angeles: University of California.

Schane, S. 1970. Phonology. Unpublished manuscript.

Scholberg, H. C. 1955. Concise Grammar of the Hindi Language. London: Oxford University Press. (First published in 1940.)

Shibatani, M. 1971. The role of surface phonetic constraints in generative phonology. Project on Linguistic Analysis Reports, Second Series, no. 13.MS1-MS54.

Srivastava, R. N. 1969. Review of Studies in Hindi-Urdu, I, by A. R. Kelkar. Language 45.913-27.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

_____. 1970. The problem of the Hindi semivowels. Indian Linguistics 31. 129-37.

_____. n.d. Some 'constructs' of prosodic theory and generative phonology. Unpublished manuscript.

Stanley, R. 1967. Redundancy rules in phonology. Language 43.393-436.

Tiwari, B. N. 1966. Hindi Bhasha. Allahabad: Kitab Mahal.

Tiwari, U. N. 1961. Hindi Bhasha ka Udgam aur Vikas. Allahabad: Bharati Bhandar.

Upraiti, M. L. 1964. Hindi me pratyaya vičar. Agra: Vinod Pustak Mandir.

Vajpeyi, K. D. 1957. Hindi Shabdanushasan. Banares: Nagari-pračarini Sabha.

Varma, R. C. (ed.) 1958. Sankshipt Hindi Shabdasagar. Kashi: Nagari-pračarini Sabha.

Weinreich, U., Labov, W., and Herzog, M. I. 1968. Empirical foundations for a theory of language change. In: Directions for historical linguistics, ed. by W. P. Lehmann and Y. Malkiel. Austin: University of Texas Press. 97-195.

Westermann, D. and Ward, I. C. 1933. Practical phonetics for students of African languages. London: Oxford University Press.

Yasui, M. n.d. Consonant patterning in English. Tokyo: Kenkyusha.

Zimmer, K. E. 1969. Psychological correlates of some Turkish morpheme structure conditions. Language 45.309-21.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix 1

THE PROBLEM OF ASPIRATION IN HINDI PHONETICS*

Hindi and some other Indo-Aryan languages (e.g., Marathi) are unique among the world's languages in possessing contrasts between aspirated and non-aspirated stops in both the voiceless and the voiced series, i.e., /p/ vs. /p^h/ and /b/ vs. /b^h/.[1] The exact physical correlates of the aspirates vis-a-vis the non-aspirates have been a subject of considerable interest to linguists for over 2000 years. In this paper we propose to test a recent description of the Hindi aspirates by Chomsky and Halle (1968).

Statement of the problem

Specifically citing Hindi, Chomsky and Halle state that heightened subglottal air pressure (HSAP) is a necessary characteristic of all aspirated consonants, i.e., /b^h/ and /p^h/, a possible but non-essential characteristic of voiceless unaspirated consonants, i.e., /p/, and must necessarily be absent from voiced unaspirated consonants, i.e., /b/. Chomsky and Halle assume HSAP is an independent variable and is not a function of any laryngeal or supralaryngeal adjustment.

To test this hypothesis the following test was performed.

Experimental procedure

While a native speaker of Hindi (the first author of the present article) pronounced selected Hindi phrases, the speaker's subglottal air pressure was sampled using a large needle (approximately 0.8 mm

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

inner diameter) which was inserted through the cricothyroid membrane into the trachea such that the open end rested about 1 cm below the vocal cords (see Fig. 17). A pressure transducer, an FM-type condenser microphone, was connected to the outside end of the needle with a very small coupling space, and the electrical signal from this was fed to an FM tape recorder along with the signal from a microphone recording the subject's speech wave. These recorded signals were later written on paper using an ink-writing oscillograph after appropriate low pass filtering to remove the high frequency variations in the signal due to voicing; for this an integrator with a time constant of 22 ms was used.[2]

The following are the nonsense phrases spoken during the experiment:

Table 41

1.	vo ama laya	(1)
2.	lala lab^{h} aya	(3)
3.	lala mal aya	(2)
4.	lala pal aya	(2)
5.	lala lah aya	(3)
6.	lala b^{h}al aya	(3)
7.	vo aba laya	(3)
8.	vo ap^{h}a laya	(3)
9.	vo ab^{h}a laya	(2)
10.	vo apa laya	(1)
11.	vo ala laya	(1)
12.	lala lap^{h} aya	(3)
13.	vo aha laya	(2)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

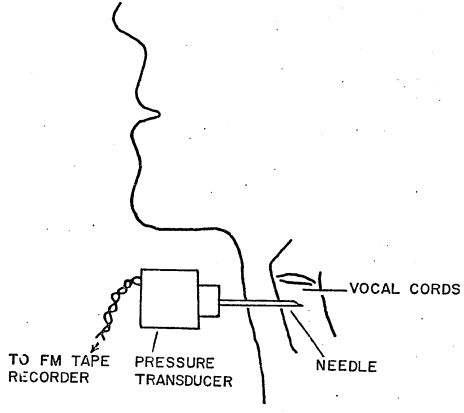


Fig. 17. The experimental setup used in the present study: subglottal air pressure was recorded via a large needle inserted through the subject's crico-thyroid membrane into the trachea; the needle was attached to a pressure transducer whose output signal was recorded on magnetic tape by an FM-type tape recorder. A microphone placed in front of the subject's mouth (not shown here) picked up the speech wave, and this audio signal was simultaneously recorded on the same magnetic tape.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

14.	lala hal aya	(3)
15.	lala lam aya	(4)
16.	lala lab aya	(3)
17.	lala lap aya	(3)
18.	lala bal aya	(2)
19.	lala p^hal aya	(5)

These were constructed so that the sounds b, b^h, p, p^h, m, l, h were in initial, medial, and final position in a word and such that the consonant of interest was preceded and followed by the vowel [ə]. For the experiment these phrases were spoken 5 times each in a randomized order (the order given in Table 41). It was discovered after the experiment that the signal recorded during some of the samples contained artifacts due to intermittent blocking of the bore of the needle by an accumulation of mucus. The number of usable tokens obtained for each phrase is indicated by the number in parentheses after each phrase as listed in Table 41. Calibration of the pressure signal at the end of the experiment indicated a slow DC leak in the coupling between the needle and the pressure transducer. There is a possibility that this leak was present during the experiment, giving rise to quantitative errors that can be corrected by a roughly constant multiplicative factor for the observed pressure values. Since the data collection could not be repeated, it was unfortunately not possible to determine this factor; but even if this were the case, it could affect only the scale for pressure. Therefore, the subglottal pressure records obtained must be of sufficient quality to support the qualitative conclusions we derive from them.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Results

Some of the results are illustrated by the graphic traces in Fig. 18. In general, HSAP was occasionally noted during the closed portion of the aspirated consonants (see point labelled 3 in Fig. 18B) but it was not found consistently (see point 1 in Fig. 18A). Moreover HSAP could sometimes be found on the voiced unaspirated /b/, where it was not supposed to be found (see point 5 in Fig. 18C). One consistent finding for /b^h/ and /p^h/ was the sudden decrease of air pressure upon the release of the consonant closure (see point 2 and 4 in Fig. 18A and 18B). There was also a consistent decrease of air pressure during /h/. In other respects the pressure curves are similar to those found for other languages (Ohala 1970).[3]

Discussion

Chomsky and Halle's claim that HSAP is a necessary correlate of aspirates was not confirmed. If we consider the pressure values during the actual moment of aspiration, i.e., when the stop is released, the subglottal air pressure is momentarily lower. The same is true for the /h/ which, like the "aspirates," is characterized by heavy air flow. Furthermore, HSAP appears inconsistently during the closure portion of both the aspirates and the voiced non-aspirate /b/, the latter of which the authors claimed may not possess HSAP.

In fact, it is curious that they should have made these claims since there is no indication in the classical or modern phonetic literature that HSAP accompanies aspiration. On the contrary, two recent studies dealing with subglottal pressure during the English voiced and voiceless aspirated sounds /p/ vs /b/ and /t/ vs /d/ show

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

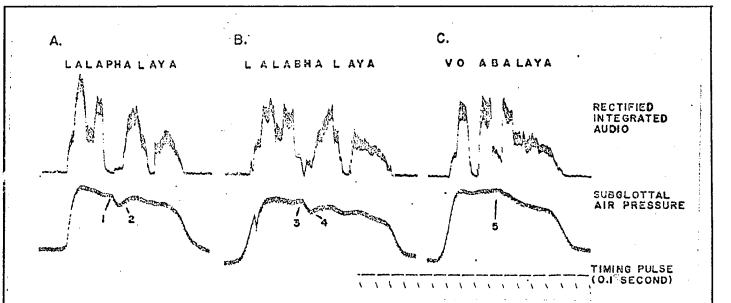


Figure 16. Examples of signals obtained in this study. At top, the rectified and integrated audio signal from the microphone; at the bottom, the subglottal air pressure. The phrases spoken are: A. [lala p^hal aya], B. [lala b^hal aya], and C. [vo aba laya].

that there is no significant difference in the subglottal pressure during the two stops (Netsell 1969, Shipp and McGlone 1971). The increased air flow characteristic of the aspirates is not due to any increased activity of the respiratory muscles, rather it is simply due to the fact that during aspirates, unlike other speech sounds, there occurs a moment when the air under pressure in the lungs encounters very little resistance. That is, during the obstruent closures the air is blocked by an oral constriction, and during normal voicing, as during most sonorants, including vowels, the air is partially blocked by the vocal cords. But during /h/ and upon the release of the aspirated stops there occurs a moment when there is no oral constriction and when the glottal resistance is markedly lower than that for normal voicing. Given such lowered resistance to the lung air, the air naturally rushes out in great volume. And consequently the air pressure just below the glottis is momentarily lowered.

It is also possible to account for the occasionally found increase of subglottal air pressure during the closure portion of stops without assuming any increase in the activity of subglottal respiratory system. Let us assume that the task of the subglottal respiratory system is to provide a relatively constant pressure to the air in the lungs. Since air is constantly escaping from the lungs, at least during sonorants, this can be accomplished by producing a constant rate of volume decrease of the chest cavity. However when the air is completely blocked, as it is during stop closures, this continuing decrement in lung volume would cause the air pressure in the system to increase momentarily.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Summary

The distinctive character of the aspirates including /h/, as observed in the time course of the subglottal air pressure, is not increased subglottal air pressure as proposed by Chomsky and Halle. Rather, the occurrence of reduced glottal resistance during a period when there is no accompanying oral constriction invariably causes a momentary lowering of the pressure. Those momentary increases in subglottal air pressure that do occasionally occur during the closure portion of stops, whether aspirated or not, can be explained without reference to any different behavior on the part of the respiratory system.[4]

Acknowledgements

The research reported here was performed primarily at the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo, when the second author was a NSF post-doctoral Fellow there during the years 1969 and 1970. We are grateful to Dr. Osamu Fujimura, director of the Research Institute, for his help and interest in this research, and to the members of the staff of the Institute who assisted in the experiment, in particular Dr. Hajime Hirose who performed the needle insertion. Part of the data analysis was done at the Phonology Laboratory, Department of Linguistics, University of California at Berkeley, funded in part by the National Science Foundation. We thank Mr. Robert Krones for his help in this phase of the study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Footnotes

* This is an English translation and a slightly revised version of an article written jointly with John Ohala, which will appear in Hindi in Bhāshā. This paper is also reprinted in the Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, Univ. of Tokyo, as well as in the POLA reports no. 16, Berkeley.

1 Although we just mention the bilabials here and elsewhere our comments apply as well to the stops at other points of articulation.

2 This part of the study was performed at the Phonology Laboratory, University of California, Berkeley.

3 There was also an occasional pressure increase during the closure portion of /p/, but there was never any pressure change during /l/ or /m/.

4 After the research reported here was undertaken, Halle and Stevens (1971) apparently rejected the use of the feature HSAP for aspirates and posited instead the new feature composition

$$\begin{bmatrix} \text{+spread vocal cords} \\ \text{-constricted vocal cords} \end{bmatrix}$$

as being common to the aspirates, the [b^h] being differentiated from [p^h] by means of additional features thus:

	b^h	p^h
stiff vocal cords	-	+
slack vocal cords	+	-

We believe the results in the present paper are useful in showing that the original feature of HSAP was wrong, because although it was adopted by Chomsky and Halle without empirical evidence, so it was discarded by Halle and Stevens also without any evidence. The present data reveal that [b^h], in accord with the Halle-Stevens proposal, is correctly considered to have 'slack' vocal cords and that [p^h] has 'spread' vocal cords; however there is no evidence (nor is it necessary to assume) that [p^h] has 'stiff' vocal cords and [b^h] has spread vocal cords. However more experimental data rather than mere speculation are needed to settle this point.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix 2

The experiment reported here was conducted to see if native speakers of Hindi have an awareness of the sequential constraints of their language, in particular the constraints on initial consonant clusters.

Thirty words were presented to the subjects in the Devanagari script. Of these thirty words, six were words that exist in Hindi. The remaining twenty-four words were constructed by me; of these, 12 were words that I hypothesized to be acceptable Hindi words, i.e., accidental gaps, and 12 were words that I hypothesized to be words containing impermissible or unlawful clusters for Hindi (see 2.3. for the detailed SQCs). Of the 6 existing words, 2 were very common native words, 2 were common Sanskrit loans, and 2 were rare Sanskrit loans. This was partly to see how the subjects would treat uncommon loans, and partly so that the less familiar Sanskrit loans would prepare them for the strangeness of the clusters in the words containing impermissible or unlawful clusters. In the 'nonexisting' forms I included a variety of clusters, e.g., stop+stop, fricative+stop, stop+fricative, etc.[1] The words used are given in table 42.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 42

Existing words

Devanagari	phonetic	gloss	source
प्रेम	prem	love	Native
क्या	kya	what	Native
क्षमा	kšəma	forgiveness	common Skt. loan
ग्राम	gram	village	" " "
म्लान	mlan	weak	rare " "
स्मिता	smita	smiling	" " "

Accidental gaps

Devanagari	phonetic	source of cluster	existing words having same cluster[2]
ग्वेक	gwek	Native	[gwɛ̃ṭʰna] (old Hindi) "to twist", [gwala] "cowherd"
त्यूप	tyu:p	"	[tyũ:] "in that manner" (vernacular), [tyohar] "festival"
ड्योम	ḍyom	"	[ḍyoṛʰi:] "porch", [tyag] "sacrifice"
फ्यूल	pʰyu:l	"	[pʰyas] "dandruff"
प्यम	pyəm	"	[pyar] "love", [pyəṇḍ] "a round piece" (old Hindi), [pyosər] "newly calved cows milk"
क्यी	kyi:	"	[kya] "what", [kyõ] "why"
स्क्यूप	skʰu:p	Skt.	[skʰəlit] "fallen", [sku:l] "school" (E)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Devanagari	phonetic	source of cluster	existing words having same cluster
स्टेप	step	Skt.	[sten] "thief", [stu:p] "stupa", [stuti:] "prayer"
क्रूप	kru:p	"	[kru:r] "cruel", [krod^{h}] "anger"
भ्राक	b^{h}rak	"	[b^{h}rata] "brother", [b^{h}rəm] "confusion"
ह्रू	hru:	"	[hras] "shortage", [hridɛ] "heart"
स्टाल	stal	Eng.	[stak] "stock" (in some dialects), [stešən] "station", [sti:mər] "steamer"

Words containing impermissible sequences

Devanagari	phonetic
क्फाल	kfal
थ्रूग	t^{h}ru:g
ब्मीन	bmi:n
च्सूल	čsu:l
ज़्फ़ीक	zfi:k
म्सोक	msok
थ्गेन	t^{h}gen
प्दील	pdi:l
म्कोल	mkol
स्भाल	sb^{h}al
ल्झाम	lj^{h}am
ड्लेश	dleš

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

These words were randomized and presented to 25 informants. The informants were told that the list contained 3 types of words:

(a) words that they will recognize as Hindi words;

(b) words that they may not recognize but which are still Hindi words used in technical vocabulary;

(c) words which just cannot be Hindi words.

Their task was to tell me for each word which of the 3 categories it belonged to. For existing forms that they recognized as such, their response was to be "yes"; for forms they didn't recognize but thought might be Hindi words, the response was to be "maybe"; and for forms that they didn't recognize and didn't think could possibly be Hindi words, their response was to be "no".

Only informants who claimed their mother tongue was Standard Hindi were used. Of the 25 informants, 4 were men and the rest women. Eight of the subjects were in the 30-45 year age group and the rest in the 18-30 year group. Most of the informants were from 2 women's colleges in Delhi. Their hometowns were in various parts of the area where Hindi is spoken -- Lucknow, Gaya, Jhansi, Delhi, Hyderabad (Deccan), etc.[3]

All but one of the informants pronounced the words before giving their judgment even though this was not asked of them.[4] In recording their responses, I also noted if they broke the initial cluster in pronunciation by inserting a [ə]. If the subject broke up the initial consonant cluster by inserting a [ə] and replied with a "yes" or "maybe" his response was not counted, or rather was included in the "other response" category. If the subject broke up the cluster with a [ə]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

but replied "no", his response was counted as usual. The reason for this was that in the former case there was then no way of knowing whether they were reacting to the word given, e.g., bmi:n, or to the word bəmi:n which no longer contains the initial cluster of interest. Table 43 summarizes the responses of the subjects. Table 44 gives the total and the percentages. Figure 19 presents the results plotted.

Table 43

Responses:	yes	maybe	no	other
Category I				
1. prem	25	0	0	
2. kya	25	0	0	
3. kšəma	25	0	0	
4. gram	25	0	0	
5. smita	19	5	1	
6. mlan	13	4	6	2 "yes", inserted [ə]
Category II				
7. sk^{h}u:p	0	5	19	1 "maybe", inserted [ə]
8. stal	4	8	6	1 "no", "it's Eng."; 6 "yes", "it's Eng."
9. step	5	9	11	
10. gwek	1	6	17	1 "yes", inserted [ə]
11. kru:p	5	10	10	
12. b^{h}rak	2	13	10	
13. hru:	5	11	9	
14. p^{h}yu:l	0	7	15	1 "maybe", inserted [ə]; 2, "it's Eng."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

		yes	maybe	no	other
15.	pyəm	0	9	16	
16.	tyu:p	2	6	17	
17.	dyom	1	6	16	1 "maybe", inserted [ə]
18.	kyi:	2	4	15	1 no response; 2 "maybe", inserted [ə]; 1 "maybe it's a Punjabi word"
Category III					
19.	kfal	1	3	19	1 "it's an Urdu word"; 1 "maybe it's an Urdu word"
20.	t^hru:g	1	4	20	
21.	bmi:n	0	1	24	
22.	čsu:l	0	5	20	
23.	zfi:k	0	4	19	1 "maybe", inserted [ə]; 1 "maybe Urdu"
24.	msok	0	5	19	1 "maybe", inserted [ə]
25.	t^hgen	0	1	23	1 "maybe", inserted [ə]
26.	pdi:l	1	5	19	
27.	mkol	0	4	18	2 "maybe", inserted [ə]; 1 "yes, exists as Urdu word"
28.	lj^ham	1	3	20	1 "maybe", inserted [ə]
29.	sb^hal	3	11	7	3 "maybe", inserted [ə]; 1 "maybe, like [səmb^hal]", (take care of)
30.	dleš	3	8	14	

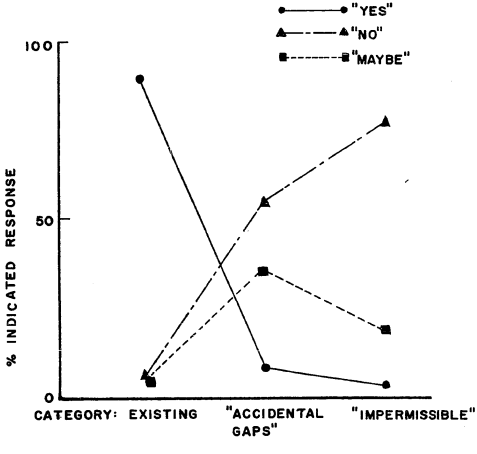
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 44

	yes	maybe	no	total
Category I	132 (89%)	9 (6%)	7 (5%)	148
Category II	27 (9%)	104 (36%)	161 (55%)	292
Category III	10 (4%)	54 (19%)	222 (77%)	286

Note: The percentages reflect the total of "yes", "no", and "maybe's"; "other responses" have been excluded.

Figure 19



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Discussion: The subjects gave the expected overwhelming acceptance to the existing words (Category I) and the expected overwhelming rejection of those forms hypothesized to contain clusters that violate the SQCs of Hindi (Category III). And although the number of "maybe" responses is, as expected, larger for the presumed "accidental gap" words (Category II) than it is for either of the other categories, the greater number of "no" responses than "maybe" responses for these words go against my predictions. There are a number of possible reasons for this. First, it is possible that to a certain extent the test did not adequately reflect speakers' true tacit awareness of the morpheme structure constraints of their language. Perhaps speakers' judgments have to be elicited a little more subtly (see Esper 1925 and Greenberg and Jenkins 1960). However, the fact that there was nevertheless an overwhelming rejection of Category III seems to argue against this possibility. Still, an improvement in the test would involve monitoring how much time was taken for each response, and not just the response itself (or perhaps asking the subjects to give scalar values as Greenberg and Jenkins did). One would predict that little response time would be required for existing words, or words that were far out on the acceptability scale. On the other hand more response time would be taken for the 'accidental gap' words, or words that were just marginally irregular.

Second, it is also possible that speakers' judgments as to what is a permissible word in their language depends to a large extent on usage, i.e., whether or not they have ever heard the word. In T. Smith's test (personal communication) some English speakers rejected

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sclerosis as a possible English word. One improvement on my test, to take into account this second possibility, is to give speakers another test consisting of words which are presumed 'accidental gaps' but which do not contain any clusters, e.g., ti:m. If they accepted all such words, then it would show that speakers were not just reacting to usage in my test. If they rejected all such words, it would show that usage is the governing factor. However, if the results were mixed, then the test would not show anything.

Third, it is possible that what I hypothesized as permissible clusters in Hindi were not really permissible. That is, they may be found in a Hindi dictionary but still not be accepted as truly native by most Hindi speakers. This is an interesting possibility; however, the fact that the clusters in native words under Category II had greater rejection (e.g., items number 10, 14, 15, 18, etc.) than some of the clusters occurring only in loans argues against this.

It is also possible that speakers' judgments are quite tenuous -- that they are random in their responses. Perhaps another improvement on my test would be to test the speakers for consistency.

Thus it is possible that any one of the above factors, or a complicated interplay of all of them, was at work to give the responses the subjects gave to words in Category II.

The test also revealed some other interesting points. Forms that violated only one sequential constraint, e.g., bmi:n (which violates SQC no. 16, see section 2.3) received slightly more "no's" than did forms such as zfi:k which violated 4 constraints (SQCs 7, 17, 21 and 23).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

One of the by-products of the experiment is that it reveals that speakers have some awareness of the fact that some clusters occur only in loans. Thus a number of speakers recognized st- as occurring only in English loans (as can be seen by looking at the 'other responses' for stal in table 43. Also a few informants reported mkol and zfi:k as being Urdu words. Fricative+fricative clusters and nonhomorganic nasal+stop clusters are quite common in Urdu words -- however only medially and finally, not initially -- but these informants who were not well versed in Urdu identified these clusters as being Urdu-type clusters anyway.

It might also be useful to discuss some of the other responses. I had hypothesized kyi: (item no. 18) to be an accidental gap in the vocabulary, since a number of Native words exist with the cluster ky- (for examples see table 42). However the responses of my informants to this word made me take a second look. I think perhaps kyi: might be an unlawful form: there is perhaps a constraint in Hindi which does not permit a high front vowel after a C_2 which is y. The fact is that no study of Hindi so far takes into account restrictions on the vowels that can follow clusters. Unfortunately this study doesn't include such constraints either.

The large number of "no" responses to other words in the 'accidental gap' category surprised me, too. I have no explanation for it, as common native words exist containing clusters exemplified by words like 10, 14, 15, 16, 17, etc., in table 43 (examples are given in table 42). I should add, however, that morphemes containing initial clusters are statistically rather few.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In Category III of table 43 the two words sb^{h}al and dleš were treated more like the words in Category II, i.e., the 'accidental gap' category. Perhaps this might be due to their resemblance in fast speech to the words [dileš] "a proper name", and [səmb^{h}al] "to take care of" -- in fact, as can be seen from table 43, "other responses", one informant did indicate that "[səmb^{h}al] is a word so perhaps sb^{h}al is also a word." It is possible that an English informant might not rule out *ftu: on the basis of its similarity to a rapidly spoken version of "for two" [ftu:]. This factor should be controlled for in future tests.

Perhaps at this point one might ask the question, what are the chances that my subjects were responding to orthography, that is, that they were reacting to the well formedness of the orthographic form? First, since the Devanagari orthography is almost phonetic,[5] I don't think it matters even if the informants were reacting to the script. Second, as I mentioned, the informants pronounced these words.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Footnotes

1 I should mention that Native words permit only certain varieties of stop plus liquid or glide clusters initially. Urdu words do not permit initial clusters, and even Sanskrit and English loans have clusters that are quite restrained. For details see 2.3.

2 As far as possible I have given existing words containing the same vowel as the "constructed" words, however this has not always been possible (partly due to initial clusters being statistically few). In most cases I do not think it made a difference in the made-up words that the vowel following the cluster differed from the vowels which can be found in the same environment in existing words. However in one case (namely that of kyi:) I think this was a problem; I discuss this on page 290.

3 I also noted if the informants said they knew any other language but Hindi -- however I do not know what their proficiency in these languages is. All the informants said they knew English. Five said they knew some Punjabi, 2 knew some Bengali, 1 some Marathi, 1 some Sanskrit, and 1 some Russian.

4 This was helpful and perhaps should be explicitly required in future tests; however, the fact that they pronounced the words even though they were not asked to is interesting in itself.

5 That is, a speaker can write phonetically in Devanagari whatever he wants to. There are, however, also certain conventional spellings, e.g., the symbol for nasalization may be used even when homorganic nasal is meant; or in the environment where the ə-deletion rule applies a consonant cluster need not be written as a cluster, but still will be pronounced as such. Some people follow these conventions; others prefer to use Devanagari phonetically. In my experiment the words were written in Devanagari phonetically, and there was no spelling rule that could have possibly applied anyhow, since these clusters were word initial.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix 3

The following examples are in broad phonetic transcription. Phonetic details such as those mentioned on pg. 88 of chapter 2 have not been included.

Examples of initial clusters:

ky-	kyõ (N) "why"
kw-	kwãra (N) "bachelor"
kr-	krəm (S) "order"; kristan (E) "christian"
kl-	klant (S) "tired"; klәb (E) "club"
kš-	kšəma (S) "mercy"
k^hy-	k^hyal (N) "thought"
k^hw-	k^hwab (P) "dream"
gy-	gyarɛ (N) "eleven"
gw-	gwala (N) "cowherd"
gr-	grahək (S) "customer"; grauṇḍ (E) "ground"
gl-	glani: (S) "shame"; glas (E) "glass"
g^hr-	g^hrit (S) "clarified butter"
čy-	čyəwənpraš (S) "a kind of medicine"
ǰy-	ǰyõ (N) "as" (adverb)
ǰw-	ǰwar (N) "high tide"
tr-	tren (E) "train"
ḍy-	ḍyor^{h}a (N) "2½ times"
ḍr-	ḍrama (E) "drama"
tr-	triṛ̃ (S) "bit of dry grass"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ty-	tyohar (N) "festival"
tw-	twəča (S) "skin"
dy-	dyu:t (S) "gambling"
dw-	dwar (S) "door"
dr-	droh (S) "malice"
d^hy-	d^hyan (S) "meditation"
d^hw-	d^hwəǰa (S) "banner"
d^hr-	d^hruv (S) "permanent"
py-	pyar (N) "love"
pr-	prit^{h}wi: (S) "earth"; prɛs (E) "press"
pl-	pli:ha (S) "spleen"; pleṭ (E) "plate"
by-	byah (N) "marriage"
br-	brəmha (S) "a god"; briged (E) "brigade"
bl-	blæk (E) "black"
b^{h}r-	b^{h}rəm (S) "confusion"
šy-	šyama (S) "a name"
šw-	šwas (S) "breath"
šr-	šrəm (S) "labor"
šl-	šlok (S) "a kind of verse"
šm-	šməšan (S) "cremation ground"
sy-	syar (N) "jackal"
sw-	swaŋg (N) "disguise"
sr-	srišṭi: (S) "creation"
sl-	sleṭ (E) "slate"
sk-	skənd^h (S) "shoulder"; sku:l (E) "school"
sk^h-	sk^{h}ələn (S) "stumbling"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

st̤-	st̤ešən (E) "station"
st-	stu:p (S) "stupa"
st^{h}-	st^{h}an (S) "place"
sp-	spərš (S) "touch"; spirit̤ (E) "spirit"
sp^{h}-	sp^{h}ət̤ik (S) "a crystal"
sn-	snan (S) "bath"
sm-	smərə̃r̤ (S) "memory"
zy-	zyada (P) "more"
nr-	nrity (S) "dance"
ny-	nyota (N) "invitation"
my-	myãũ: (N) "meow"
mr-	mrig (S) "deer"
ml-	mlan (S) "weak"
wr-	wrindawən (S) "a place"
wy-	wyəkti: (S) "person"
hr-	hridɛ (S) "heart"

Three consonant clusters:

str-	stri: (S) "woman"
spr-	sprišt̤ (S) "touched"; spriŋg (E) "spring"
spl-	spli:n (E) "spleen"
skr-	skri:n (E) "screen"
smr-	smriti: (S) "remembrance"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Examples of medial clusters:

-kṭ- dakṭər (E) "doctor"; kukṭi: (N) "a kind of cotton"

-kt- čəkti: (N) "a round piece"; mukta (S) "pearl"; nukta (P) "a dot"

-kd- təkdi:r (A) "fate"

-kb- əkbər (A) "name of a king"

-kč- ukčən (N) "a kind of flower"

-k^{h}ṭ čok^{h}ṭa (N) "a frame (for a door, etc.)"

-k^{h}t- tək^{h}ta (P) "wooden shelf"

-k^{h}b- ək^{h}bar (PA) "newspaper"

-gṭ- prəgṭana (N) "to make known" (caus. inf.)

-gt- b^{h}ugtan (N) "payment in full"

-gd- ǰəgdəmba (S) "a goddess"

-gd^{h}- mugd^{h}a (S) "a young shy girl"

-g^{h}ṭ- ug^{h}ṭa (N) "one who indulges in unpleasant talk"

-ṭk- čəṭki:la (N) "bright (color)"

-ṭ^{h}k- bəṭ^{h}ka (N) "room in the house for meeting visitors"

-tk- gətka (N) "a kind of club"; čəmətkar (S) "miracle"

-tp- tatpəry (S) "meaning"

-tb- rutba (A) "rank"

-dk- p^{h}udki: (N) "a kind of bird"; sədka (A) "gift"

-dg- udgəm (S) "source"

-dg^{h}- udg^{h}aṭən (S) "opening ceremony"

-db- tədbi:r (A) "scheme"

-db^{h}- ədb^{h}ut (S) "wonderful, strange"

-d^{h}k- dəd^{h}kana (N) "to fan (a fire)" (inf.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-pk- čʰipkəli: (N) "lizard"; upkar (S) "favor"

-pṭ- čipṭa (N) "flat"

-pt- gupta (N) "a last name"; səptah (S) "week"; kəptan (E) "captain"

-pd- updeš (S) "advice"

-pč- upčar (S) "remedy"; nɛpču:n (E) "neptune"

-pǰ- upǰau: (N) "productive"

-bg- i:səbgol (P) "a medicinal seed"

-bṭ- ubṭən (N) "a paste for rubbing on the body"

-bt- əbtər (PA) "scattered"

-bd- šətabdi: (S) "century"; təbdi:l (A) "transfer"

-bdʰ- əbdʰi: (S) "ocean"

-bǰ- kubǰa (S) "hunch backed"

-bʰkʰ- bʰəbʰki: (N) "a threat"

-čk- hički: (N) "hiccup"

-čṭ- učṭana (N) "to feel indifferent, fed up" (inf.)

-čʰt- pəčʰtana (N) "to regret" (inf.)

-ǰg- əǰgər (S) "python"

-ǰd- səǰda (P) "bowing down"

-ǰb- məǰbu:r (A) "helpless"

-ǰʰkʰ- ǰʰiǰʰki: (N) "hesitation"

Stop+liquid

-kr- ṭokri: (N) "basket"; prəkriti: (S) "nature"; təkrar (A) "dispute"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-kl- čəkla (N) "round wooden board"; təkli:f (PA) "trouble"; kakli: (S) "melodious tune"

-kṛ- kə̃kṛa (N) "crab"

-k^{h}r- ək^{h}roṭ (N) "walnut"; ik^{h}raẏ (A) "to take out"

-k^{h}l- ok^{h}li: (N) "mortar and pestle"; bɔk^{h}lana (PA) "to get angry"

-k^{h}ṛ- uk^{h}ṛa (N) "uprooted"

-gr- gəgri: (N) "vessel"; sugri:w (S) "a name"; ḍigri: (E) "degree"; ẏigri: (PA) "pertaining to the heart"

-gṛ- pəgṛi: (N) "turban"

-gl- pəgla (N) "crazy"

-g^{h}r- g^{h}ag^{h}ra (N) "long skirt"

-g^{h}l- pig^{h}la (N) "melted"

-g^{h}ṛ- sug^{h}ṛapa (N) "cleverness"

-ṭr- pəṭra (N) "platform"; miliṭri: (E) "military"

-ṭl- poṭli: (N) "small bundle"

-ṭ^{h}r- koṭ^{h}ri: (N) "room"

-ṭ^{h}l- gúṭ^{h}li: (N) "seed"

-tr- čəkotra (N) "grapefruit"; šətru: (S) "enemy"; k^{h}ətra (A) "danger"

-tl- titli: (N) "butterfly"; mətləb (A) "meaning"; ši:tla (S) "a disease"

-t^{h}r- pət^{h}ri:la (N) "stony"

-t^{h}l- ut^{h}la (N) "churned up"

-t^{h}ṛ- či:t^{h}ṛa (N) "rag"

-dr- bidri: (N) "a kind of craft"; nidra (S) "sleep"; kudrət (A) "nature"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-dl-	bədla (N) "changed"; kədli: (S) "banana"; təbadla (A) "transfer"
-dṛ-	gudṛi: (N) "patchwork quilt"
-dʰr-	sudʰra (N) "reformed"
-dʰṛ-	udʰṛa (N) "untwisted, ripped"
-pr-	u:pri: (N) "top, surface"; čəprasi: (P) "peon"; wipri:t (S) "opposite"; propraiṭər (E) "proprietor"
-pl-	popla (N) "toothless"; čaplu:s (P) "flatterer"; čəpla (S) "lightning"; ḍiploma (E) "diploma"
-pṛ-	ɲʰɔ̃pri: (N) "hut"
-pʰṛ-	pʰepʰṛa (N) "lungs"
-pʰr-	əpʰra (N) "a disease"
-br-	gubrəla (N) "cowdungish"; librəl (E) "liberal"; ibrani: (A) "name of a language"
-bṛ-	ǰəbṛa (N) "jaw"
-bl-	dubla (N) "thin"; əbla (S) "woman"; ǰubli: (E) "jubilee"; mukabla (P) "competition"
-bʰr-	wibʰrəm (S) "confusion"
-bʰṛ-	ubʰṛa (N) "swelled up, erupted"
-čr-	kəčri: (N) "a vegetable"
-čl-	kučla (N) "squashed"; əčla (S) "stable"
-čṛ-	kʰičṛi: (N) "a dish"
-čʰl-	məčʰli: (N) "fish"
-čʰṛ-	pičʰṛa (N) "left behind, lagging"
-ǰr-	guǰrat (N) "name of a provin:e"; muǰrim (A) "the accused"
-ǰl-	biǰli: (N) "light, lightning"; məǰlis (A) "meeting"
-ǰʰl-	məǰʰla (N) "middle"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stop+fricative

-ks- čəksi: (N) "watchfulness"; nuksan (A) "harm"; wiksit (S) "open, blooming"; ṭæksi: (E) "taxi"

-kš- əkšər (S) "letter"; nəkša (A) "map"; kənekšən (E) "connection"

-kf- wakfiyət (A) "acquaintance"

-gz- kagzi: (A) "pertaining to paper"

-ṭf- plæṭfɔrm (E) "platform"

-ts- utsəw (S) "festival"

-tf- nutfa (A) "offspring"

-ds- hadsa (PA) "accident"

-dš- əkadši: (S) "eleventh day"; badšah (PA) "king"

-ps- ləpsi: (N) "a kind of paste"; əpsəra (S) "nymph"

-bš- həbši: (PA) "negro"

-bz- səbzi: (PA) "vegetable"

stop+glide

-kw- čəkwa (N) "a kind of bird"; bəkwas (PA) "nonsense"; pəkwašəy (S) "stomach"

-k^{h}y- prək^{h}yat (S) "famous"

-gy- prəgya (S) "intelligence"

-gw- b^{h}əgwan (N) "god"

-ṭw- pəṭwari: (N) "village accountant"

-ty- mrityu: (S) "death"

-tw- pətwar (N) "oar"; čətwər (S) "quick"

-t^{h}y- rət^{h}ya (S) "path"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-tʰw- prithwi: (S) "earth"

-dy- widya (S) "knowledge"

-dw- widwan (S) "learned"

-dʰy- ədʰyəyən (S) "study"

-dʰw- madʰwi: (S) "a name"

-py- rupya (N) "rupee"; kripya (S) "please"

-pw- əpwad (S) "exception"

-ph- uphar (S) "gift"

-bh- šubha (PA) "suspicion"

-ɣw- əɣwain (N) "a spice"; uɣwəl (S) "bright"

-ɣh- iɣhar (PA) "to proclaim"

stop+nasal

-km- čəkma (N) "swindle"; rukmiṇi: (S) "a name"; tukma (PA) "a kind of buttonhole"

-kn- čikna (N) "greasy"; məknati:s (PA) "magnet"

-kʰm- zəkʰmi: (PA) "wounded"

-kʰn- ṭəkʰna (N) "ankle"; yəkʰni: (PA) "meat stock"

-gm- nəgma (PA) "a song"

-gn- ɣugnu: (N) "glow worm"; əgni: (S) "fire"

-ṭn- gʰuṭna (N) "knee"

-tm- atma (S) "soul"; itminan (PA) "peace, confidence"

-tn- kitna (N) "how much"; pu:tna (S) "a name"; fitna (PA) "a perfume"

-tʰm- prətʰma (S) "first"

-tʰn- tʰu:tʰni: (N) "snout"; kətʰni:y (S) "worth telling"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-dm- pədmini: (S) "a kind of lotus"; mukədma (PA) "law suit"

-dn- wedna (S) "sorrow"; sudni: (PA) "fate"

-d^{h}m- u:d^{h}mi: (N) "naughty"; əd^{h}man (S) "

-d^{h}n- sad^{h}na (S) "penance"

-pm- upma (S) "simile"

-pn- sepna (N) "dream"; gopni:y (S) "worth hiding"

-bm- səbməri:n (E) "submarine"

-čm- ačmən (S) "a ritualistic drinking of water"

-čn- kəčnar (N) "a plant"; su:čna (S) "notice"

-č^{h}m- ləč^{h}mi: (N) "a goddess"

-č^{h}n- kəč^{h}ni: (N) "a kind of wearing apparel"

-ǰm- aǰmana (PA) "to test"; yəǰman (S) "priest"

-ǰn- rəǰni: (S) "night"; əǰnəbi: (PA) "stranger"

fricative+stop

-sk- siski: (N) "sob"; puresker (S) "prize"; baiskop (E) "bi-scope"; məska (PA) "butter"

-sk^{h}- məsk^{h}əra (PA) "joke"

-st̤- poste̤ǰ (E) "postage"

-st- səsta (N) "cheap"; k^{h}əsta (PA) "crisp"; kənəstər (E) "canister"; nastik (S) "atheist"

-st^{h}- prəst^{h}an (S) "departure"

-sd- təsdi:k (A) "proof"

-sp- wənəspəti: (S) "vegetation"; əspətal (E) "hospital"; čespã (P) "stuck"

-sp^{h}- wisp^{h}ot̤ (S) "to burst due to heat"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-sb- kəsba (A) "locality"

-sǰ- məsǰid (PA) "mosque"

-šk- muškil (A) "difficult"; puškər (S) "name of a place"

-šg- məšgu:l (A) "absorbed"

-šṭ- drišṭi: (S) "sight"

-šṭ^{h}- nišṭ^{h}a (S) "decision"

-št- təštəri: (P) "platter"

-šp^{h}- nišp^{h}əl (S) "fruitless"

-šp- nišpət:i: (S) "completion"

-šč- aščəry (S) "astonishment, surprise"

-zk- təzkira (A) "talk, conversation"

-zd- nəzdi:k (A) "near"

-zb- məzbu:t (A) "strong"

-fg- əfgan (A) "afghan"

-ft- dəftər (P) "office"

fricative+liquid

-sr- du:sra (N) "other"; k^{h}əsra (PA) "measles"; prəsraw (S) "the flow"

-sl- g^{h}ösla (N) "nest"; fasla (P) "distance"

-šr- pərišrəm (S) "labor"; məšrik (A) "east"

-šl- əšli:l (S) "obscene"

-zr- həzrət (PA) "Mr."

-zl- nəzla (PA) "cold"

-fr- nəfrət (PA) "hate"; əfri:ka (E) "Africa"

-fl- əflatu:n (A) "Plato"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fricative+glide

-sy- təpəsya (S) "penance"

-sw- təpəswi: (S) "sage"; təswi:r (A) "picture"

-šy- wešya (S) "prostitute"

-šw- rišwət (A) "bribe"; i:šwər (S) "god"

-šh- məšhu:r (A) "famous"

-zh- məzhəb (A) "religion"

-fw- əfwah (A) "rumor"

fricative+fricative

-fs- əfsos (PA) "regret"; əfsər (E) "officer"

-sf- təsfiya (A) "agreement"

fricative+nasal

-sm- kismət (A) "fate"; əkəsmat (S) "suddenly"; bəptisma (E) "baptism"

-sn- čusni: (N) "sucker, pacifier"; məsnəd (A) "a kind of pillow"

-šm- čəšma (P) "spectacles"; rəšmi: (S) "ray"

-šn- rošni: (P) "light"; mišnəri: (E) "missionary"; prišni: (S) "a kind of cow"

-zm- hazma (A) "digestion"

-zn- wəzni: (A) "heavy"

-fn- dəfnana (PA) "to bury" (inf.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

glide+stop

-yk- gayki: (N) "style of singing"; zayka (PA) "taste"

-yt- rayta (N) "a dish"; kifayti: (PA) "a careful user or spender"

-yd- fayda (PA) "benefit"

-yb- gaybana (PA) "in the absence of"

-yč- ilayči: (N) "cardomom"

-wk- kewka (N) "a kind of food"; əwkaš (S) "leisure"

-wṭ- dik^{h}awṭi: (N) "that which is for show"

-wḍ- əwḍer (N) "confusion"

-wt- dewta (S) "deity"; dawti: (PA) "feastly"

-wd- dewdar (S) "a tree"

-wd^{h}- əwd^{h}u:t (S) "ascetic"

-hk- təhki:k (A) "research"

-ht- muhtəməl (A) "suspicious"

-hd- ohda (A) "position"

-hb- səhbət (PA) "company"

glide+liquid

-yr- šayri: (PA) "a kind of poetry"; ḍayri: (E) "diary"

-yl- koyla (N) "coal"

-wr- dewrani: (N) "sister-in-law"; godawri: (S) "a river"; muhawra (PA) "idiom"

-wl- newla (N) "mongoose"; ləwli: (S) "a tree"

-wṛ- kewṛa (N) "a sweet smelling plant"

-hr- təhri:k (A) "to change something"

-hl- prəhlad (S) "a name"; təhli:l (A) "to melt"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

glide+fricative

-ys- ǰaysi: (N) "a name"

-ws- əwsər (S) "opportunity"

-hs- əhsan (PA) "obligation"; sahsi: (S) "courageous"

-hš- wəhši: (PA) "savage"

-hf- tohfa (PA) "gift"

glide+glide

-wh- wyəwhar (S) "behavior"

-hy- guhyək (S) "a deity"

-hw- ǰihwa (S) "tongue"; kəhwa (A) "coffee"

glide+nasal

-yn- nəyna (S) "one with pretty eyes"

-wn- nəwni:t (S) "butter"

-hn- məhni: (S) "attractive"; sihna (A) "constable"

-hm- muhməla (A) "a kind of Arabic letter"

liquid+stop

-rk- kirka (N) "a small bit of dust"; tərkari: (P) "vegetable"; marketṭ (E) "market"; kərkəš (S) "harsh"

-rk^{h}- tərk^{h}an (N) "carpenter"; kark^{h}ana (P) "factory"

-rg- girgiṭ (N) "chameleon"; k^{h}ərgoš (P) "rabbit"; durga (S) "a goddess"

-rg^{h}- kərg^{h}a (N) "loom"

-rṭ- parṭi: (E) "party"; kərṭək (S) "a kind of plant"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-rṭʰ- soɾṭʰa (N) "name of a metre"

-rḍ- bordiŋg (E) "boarding"

-rt- surti: (N) "eating tobacco"; mu:rti: (S) "statue"; kurta (PA) "a kind of dress"

-rtʰ- ərtʰi: (S) "bier"

-rd- gərdən (P) "neck"; šarda (S) "a goddess"

-rdʰ- gowərdʰən (S) "name of a mountain"

-rp- kʰurpi: (N) "spade"; dərpəṇ (S) "mirror"; tarpi:n (E) "turpentine"; bərpa (P) "present"

-rb- pu:rbi: (N) "eastern"; sorba (P) "broth"; bərbər (S) "barbaric"

-rbʰ- surbʰi: (S) "fragrant wind"

-rč- mirča (N) "chilies"; pərča (P) "paper"; ku:rčika (S) "brush"

-rčʰ- pərčʰaĭ: (N) "shadow"; mu:rčʰa (S) "unconscious"

-rǰ- kʰərǰu:r (S) "date"; fərǰi: (P) "a kind of dress"

-rǰʰ- murǰʰana (N) "to wither"

-lk- ḍolki: (N) "drum"; bilkul (A) "absolutely"; əlka (S) "a name"

-lg- pʰalgun (S) "a month"; bəlgəm (PA) "phlegm"

-lṭ- gilṭi: (N) "gland"; pulṭis (E) "poultice"

-lḍ- ḍalḍa (N) "vegetable shortening"; holḍər (E) "holder"

-lt- alta (N) "a red paint for the hands"; malti: (S) "a creeper"; iltiǰa (PA) "request"

-ltʰ- paltʰi: (N) "cross-legged"

-ld- həldi: (N) "turmeric"; ǰəldi: (PA) "quickly"

-ldʰ- kəldʰu:t (S) "silver"

-lp- kəlpana (N) "to cause to lament"; kəlpit (S) "imaginary"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-lb- čilbil (N) "a tree"; bulbul (PA) "a bird"; əlbəm (E) "album"

-lč- lalči: (N) "greedy"; kulča (P) "a kind of bread"; kəlčuri: (S) "name of a princely family"

-lč^{h}- kəlč^{h}i: (N) "a ladle"

-lǰ- šəljəm (PA) "turnip"; əlǰəbra (E) "algebra"

-lǰ^{h}- ulǰ^{h}a (N) "entangled"

-ṛk- k^{h}iṛki: (N) "window"

-ṛk^{h}- kəṛk^{h}a (N) "song sung in war"

-ṛt- pərtal (N) "investigation"

-ṛp- təṛpana (N) "to cause to palpitate"(inf.)

-ṛč- əṛčən (N) "obstacle"

liquid+fricative

-rs- sərsõ (N) "mustərd"; kursi: (A) "chair"

-rš- dəršən (S) "philosophy"; maršəl (E) "marshall"; iršad (PA) "to permit"

-rz- dəṛzen (E) "dozen"; mərzi: (A) "inclination"

-rf- əšərfi: (P) "gold coin"

-ls- əlsi: (N) "linseed plant"; tulsi: (S) "oscimum sanctum"; ǰəlsa (A) "meeting, convention"

-lš- ɡulšən (P) "garden"; bolšewik (E) "bolshevik"

-lz- ilzam (A) "offense"

-lf- kulfa (P) "a plant"

-ṛs- sərsi: (N) "a kind of tongs"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

liquid+liquid

-rl- birla (N) "rare"

-lr- j^{h}alrõ (N) "fringes" (plural oblique)

liquid+glide

-ry- məryada (S) "limit"

-rw- pərwəl (N) "a vegetable"; dərwaza (P) "door"; sərwis (E) "service"; čarwak (S) "a name"

-rh- ərhəṭ (N) "buckets for taking water out of wells"; pərhez (P) "abstinence"; garhest^{h}y (S) "householder"

-ly- kəlyaṇ (S) "benefit"

-lw- təlwar (N) "sword"; həlwa (PA) "a dish"; silwər (E) "silver"; wəlwəl (S) "a name"

-lh- du:lha (N) "bridegroom"; ilhad (A) "atheism"

-ṛw- kəṛwa (N) "bitter"

liquid+nasal

-rm- pərməl (N) "a kind of grain"; gərmi: (P) "summer"; nərməda (S) "name of a river"; j̆ərmən (E) "German"

-rn- kərnal (N) "name of a place"; p^{h}irni: (P) "a sweet dish"

-lm- məlməl (N) "muslin"; muselman (P) "muslim"; walmi:ki: (S) "name of a sage"; əlmuñyəm (E) "aluminum"

-ln- č^{h}əlni: (N) "strainer"

nasal+nasal

-mn- samna (N) "frontage"; kamna (S) "desire"; čimni: (E) "chimney"

-nm- unmət (S) "intoxicated"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

nasal+homorganic stop

-ñǰ- bʰañǰa (N) "nephew"

-ɳṭ- gʰəɳṭa (N) "hour"

-ɳṭʰ- ḍəɳṭʰəl (N) "stem"

-ɳḍ- ḍəɳḍa (N) "stick"

-ñčʰ- pəñčʰi: (N) "bird"

-nd- bəndər (N) "monkey"

-ndʰ- əndʰa (N) "blind"

-mb- ləmba (N) "tall"

-ŋkʰ- pəŋkʰa (N) "fan"

-ŋg- dəŋga (N) "riot"

-ŋgʰ- kəŋgʰa (N) "comb"

-mp- čəmpa (N) "Michelia Champacca"

-mbʰ- əčəmbʰa (N) "surprise"

-nt- əntəṛi: (N) "intestines"

-ŋk- kəŋkəṛ (N) "pebble"

-ñč- kəñčən (N) "a proper name" (from Sanskrit)

etc.

nasal+non-homorganic stop

-mk- ǰʰumka (N) "a kind of earring"; mumkin (A) "possible"

-mg- təmga (PA) "medal"

-mṭ- čimṭa (N) "tongs"

-mt- gomti: (S) "name of a river"; imtihan (PA) "examination"

-md- nəmda (PA) "a kind of bedspread"

-mdʰ- səmdʰi: (N) "in-law"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-mč- čəmča (N) "spoon"

-mǰ^{h}- səmǰ^{h}a (N) "understood"

-nk- tinka (N) "bit of dry grass"; inkar (PA) "denial"; menka (S) "a nymph"

-nk^{h}- tənk^{h}a (PA) "pay"; kənk^{h}i: (N) "corner of the eye"

-ng- čingari: (N) "spark"

-nb- kunba (N) "family"

nasal+fricative

-ms- təmsa (S) "dark" (fem.)

-mš- šəmšan (N) "cremation ground"; šəmšer (P) "sword"

-mz- rəmzan (A) "a muslim holy month"

-ns- insaf (PA) "justice"; pensil (E) "pencil"; bənsi: (N) "flute"; mənsa (S) "name of a goddess"

-nš- mənša (A) "desire"; penšən (E) "pension"; ənšən (S) "hunger strike"

-nz- mənzil (PA) "goal, story"

nasal+liquid

-mr- t^{h}umri: (N) "a kind of song"; əmrit (S) "drink of the gods"; kumri: (A) "a kind of bird"

-ml- imli: (N) "tamarind"; wimla (S) "dirt-less"; mamla (PA) "matter"

-mṛ- čəmṛa (N) "leather"

-nr- čunri: (N) "scarf"

-nl- sənlap (S) "conversation"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

nasal+glide

-mh- kumhar (N) "potter"

-nw- kinwani: (N) "water droplets"; tənwi: (S) "slender girl"

-nh- nənha (N) "tiny"; tənhai: (P) "solitude"

Three consonant clusters:

-krw- čəkrwak (S) "a bird"

-kšm- yəkšma (S) "tuberculosis"

-kšw- ikšwaku: (S) "name of a king"

-tkr- utkrant (S) "rising"

-tsn- ǰyotsna (S) "moonlight"

-tsy- watsyayən (S) "a name"

-tpr- utprekša (S) "a kind of figure of speech"

-tkr- utkriṣṭ (S) "excellent"

-stm- istmal (P) "use"

-str- ustra (P) "razor"

-rtg- purtgal (E) "Portugal"

-rtr- kərtrik (S) "done"

-rt^{h}n- prart^{h}na (S) "prayer"

-mbr- nəmbri: (N) "pertaining to numbers" (from Eng. "number")

-mb^{h}l- səmb^{h}la (N) "become careful"

-mb^{h}r- səmb^{h}rant (S) "respectable"

-mpr- empres (E) "empress"; səmprəday (S) "sect"

-ŋkt- pəŋkti: (S) "line"

-ŋk^{h}y- səŋk^{h}ya (S) "number"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-ŋgṛ- leŋgṛa (N) "lame"; kaŋgṛa (N) "name of a mountain district in India"

-ŋgr- kiŋgri: (N) "a kind of instrument"

-ŋgl- uŋgli: (N) "finger"

-ŋgy- səŋgya (S) "noun"

-ŋg^{h}r- ʈ^{h}uŋg^{h}ru: (N) "tiny bells for the ankle"

-ŋkr- səŋkranti: (S) "transition"

-ñčm- pəñčmi: (N) "fifth day" (from Skt. pəñčəmi:)

-ñj^{h}r- ɟ^{h}añj^{h}ri: (N) "anklets"

-ñj^{h}l- ɟ^{h}uñj^{h}lana (N) "to be irritated"

-ñǰṛ- piñǰṛa (N) "cage"

-ñǰr- məñǰri: (N) "tiny cluster of flowers" (from Skt. məñǰəri:)

-ñšč- puñščəli: (S) "harlot"

-ñšy- iñšyorəns (E) "insurance"

-nṭr- kəṇṭrol (E) "control"

-ndr- faundri: (E) "foundry"

-ndl- piṇḍli: (N) "a bone of the body"

-nd^{h}l- d^{h}and^{h}la (N) "swindle"

-ntr- məntri: (S) "minister"; səntra (N) "orange" (from Portuguese)

-ntw- santwəna (S) "pacification"

-ndr- kundru: (N) "a vegetable"; pəndrəh (N) "fifteen"

-nd^{h}r- ǰələnd^{h}ri: (N) "of a city called Julandhar"

-nd^{h}y- sənd^{h}ya (S) "evening"

-nfl- influ:enza (E) "influenza"

-nsk- sənskərən (S)

-nst^{h}- sənst^{h}a (S) "institution"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-nsṭ- kɔnsṭəbəl (E) "constable"

-nsp- inspɛkṭər (E) "inspector"

-nsm- sənsmərən (S) "memoir"

Examples of final clusters:

-kṭ əkṭ (E) "act"

-kt rəkt (S) "blood"; wəkt (PA) "time"

-kf vəkf (PA) "wealth given for religion"

-kš mokš (S) "nirvana"; nəkš (PA) "engraved"

-ks nuks (PA) "fault"; ṭæks (E) "tax"

-ky vaky (S) "sentence"

-kw pəkw (S) "cooked"

-kr čəkr (S) "circle"; fikr (PA) "worry"

-kl šukl (S) "white"; əkl (PA) "intelligence"

-km rukm (S) "gold"; hukm (PA) "order"

-k^{h}t sək^{h}t (PA) "hard"

-k^{h}š bək^{h}š (PA) "forgive"

-k^{h}s šək^{h}s (PA) "person"

-k^{h}y muk^{h}y (S) "main"

-k^{h}m zək^{h}m (PA) "wound"

-gd^{h} dəgd^{h} (S) "burnt"

-gz məgz (P) "brain"

-gy yəgy (S) "sacrifice"

-gr əgr (S) "ahead"

-gn nəgn (S) "naked"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-gm tigm (S) "hot"

-g^{h}r wyag^{h}r (S) "tiger"

-g^{h}n kritəg^{h}n (S) "ungrateful"

-čy udi:čy (S) "eastern"

-ǰy raǰy (S) "kingdom"

-ǰr wəǰr (S) "thunderbolt"

-ṭy naṭy (S) "act"

-ḍy ǰaḍy (S) "stupidity"

-ts wi:b^{h}əts (S) "cruel"

-ty paščaty (S) "western"

-tw tətw (S) "matter"

-tr čitr (S) "picture"; itr (A) "perfume"

-tl kətl (A) "murder"

-tm əd^{h}yatm (S) "spiritual"

-tn rətn (S) "jewel"

-tf lutf (A) "pleasure"

-t^{h}y ket^{h}y (S) "speakable"

-dy k^{h}ady (S) "eatable"

-dr səmudr (S) "ocean"

-dm pədm (S) "lotus"

-d^{h}y məd^{h}y (S) "center"

-d^{h}w məd^{h}w (S) "a sect"

-d^{h}r grid^{h}r (S) "vulture"

-pt tript (S) "satisfied"; zəpt (PA) "seized"

-py gopy (S) "hidden"

-pr sipr (S) "moon"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-pn swəpn (S) "dream"

-bǰ kañykubǰ (S) "name of a place"

-ps həps (PA) "humid, sultry"

-bd šəbd (S) "word"

-bd^{h} ləbd^{h} (S) "obtained"

-bz səbz (P) "green"

-br kəbr (A) "coffin"

-bl kəbl (A) "before"

-b^{h}r šub^{h}r (S) "white"

-b^{h}y ləb^{h}y (S) "worth getting"

-šk šušk (S) "dry"; k^{h}ušk (P) "dry"; məšk (A) "practice"

-šṭ klišṭ (S) "difficult"

-šṭ^{h} əŋgušṭ^{h} (S) "thumb"

-št gošt (PA) "meat"

-šp pušp (S) "flower"

-šy drišy (S) "invisible"

-šw wišw (S) "universe"

-šr mišr (S) "mixed"; nəšr (A) "spreading of news"

-šm čəšm (P) "eye"; gri:šm (S) "summer"

-šn prəšn (S) "question"; jəšn (P) "festival"

-sk mənəsk (S) "state of the mind"

-sṭ kəmyu:nisṭ (E) "communist"

-st ləst (N) "tired"; čust (P) "tight"; həst (S) "hand"

-st^{h} əntəst^{h} (S) "situated in the interior"

-sp dilčəsp (P) "interesting"

-sk^{h} məsk^{h} (A) "change of form"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-sf nisf (A) "half"

-sy rəhəsy (S) "sector"

-sw sərwəsw (S) "everything"

-sr əǰəsr (S) "permanent"; misr (A) "name of a country"

-sl gusl (PA) "bath"

-sm b^{h}əsm (S) "ashes"; kism (A) "kind"

-sn husn (PA) "beauty"

-zb ǰəzb (A) "attraction"

-zm kəmyu:nizm (E) "communism"; bəzm (P) "assembly"

-zn vəzn (A) "weight"

-ft muft (A) "free"

-fr kufr (A) "ungratefulness"

-fs nəfs (A) "soul"; ləfs (A) "word"

-fn dəfn (A) "laid in a grave"

-wr ti:wr (S) "sharp"

-wy kəwy (S) "poetry"

-hy grahy (S) "grasp"

-rk kərk (S) "a zodiacal sign"; klərk (E) "clerk"; fərk (A) "different"

-rk^{h} mu:rk^{h} (S) "stupid"; $surk^{h}$ (P) "a red color"

-rg marg (S) "path"; buǰurg (P) "elderly"

-rg^{h} di:rg^{h} (S) "long"

-rč k^{h}ərč (P) "expenditure"

-rǰ sərǰ (S) "create"; čarǰ (E) "charge"; dərǰ (A) "written"

-rṭ šərṭ (E) "shirt"

-rḍ garḍ (E) "guard"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-rt	muhu:rt (S) "a time"; šərt (PA) "bet"
-rd	mərd (PA) "male"; gonərd (S) "a kind of bird"
-rt^{h}	ərt^{h} (S) "meaning"
-rd^{h}	ərd^{h} (S) "half"
-rp	dərp (S) "pride"
-rb	k^{h}ərb (S) "a counting measure"
-rb^{h}	gərb^{h} (S) "womb"
-rz	fərz (A) "duty"
-rš	wərš (S) "year"; fərš (A) "a kind of ground"
-rf	bərf (PA) "ice"
-rs	hirs (A) "greed"
-ry	ary (S) "Aryan"
-rw	pu:rw (S) "east"
-rh	dəšarh (S) "name of a king"
-rm	d^{h}ərm (S) "religion"; šərm (P) "shame"
-lk	šulk (S) "fee"; silk (E) "silk"; mulk (A) "country"
-lk^{h}	təlk^{h} (PA) "bitter"
-lṭ	fɛlṭ (E) "felt"
-lḍ	fi:lḍ (E) "field"
-ld	ǰəld (PA) "fast"
-lp	šilp (S) "art"
-lp^{h}	gulp^{h} (S) "part of the body"
-lb	kəlb (A) "dog"
-lb^{h}	prəgəlb^{h} (S) "stubborn"
-lf	zulf (PA) "tendril of hair"
-lw	wilw (S) "a kind of fruit"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-ly	mu:ly (S) "price"
-lm	zulm (PA) "crime"; gulm (S) "a kind of plant"
-nš	ənš (S) "portion"
-ns	həns (N) "swan"
-nz	tənz (A) "satire"
-nh	činh (S) "sign"
-nf	sinf (PA) "race"
-nm	ǰənm (S) "birth"
-my	gəmy (S) "go-able"
-mt	simt (A) "direction"
-mr	tamr (S) "copper"; umr (A) "age"
-ml	əml (S) "sourness"
-mn	d^{h}rišṭədyumn (S) "a name"; zimn (A) "topic"
-ms	ləms (A) "touch"
-mz	ləmz (A) "to burn"
-mh	brəmh (S) "God"

homorganic nasals+stops

-ŋg^{h}	ǰaŋg^{h} (N) "thigh"
-ŋk	rəŋk (S) "pauper"
-ŋg	maŋg (N) "part in the hair"
-ŋk^{h}	pəŋk^{h} (N) "feather"
-ñč	məñč (S) "platform"
-ñǰ	mu:ñǰ (N) "a kind of grass"
-ñ$ǰ^{h}$	bañ$ǰ^{h}$ (N) "barren"
-nṭ	čənṭ (N) "cunning"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-n̤d	jʰun̤d (N) "group"
-nt	sənt (N) "saint"
-ntʰ	pəntʰ (S) "path"
-nd	gend (N) "ball"
-ndʰ	sindʰ (N) "name of a province"
-mp	bʰuːkəmp (S) "earthquake"
-mb	kədəmb (S) "a kind of tree"

etc.

Three consonant clusters:

-kšy	upekšy (S) "negligible"
-kš̤n	tiːkš̤n (S) "sharp"
-kš̤m	pəkš̤m (S) "eye lashes"
-gdʰy	vɛdəgdʰy (S) "cleverness"
-try	čaritry (S) "character"
-tsy	mətsy (S) "fish"

-š̤tr	raš̤tr (S) "nation"
-š̤tʰy	oš̤tʰy (S) "labial"
-str	əstr (S) "weapon"
-stʰy	garhəstʰy (S) "duties of a householder"
-sty	puləsty (S) "name of a sage"
-rky	tərky (S) "argumentative"
-rkʰy	mɔrkʰy (S) "foolishness"
-rgʰy	dɛrgʰy (S) "expanse"
-rǰy	wərǰy (S) "that which is not permitted"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-rty mərty (S) "mortal"

-rtm wərtm (S) "path"

-rt^{h}y prart^{h}y (S) "worth praying for"

-rdr ardr (S) "damp, wet"

-rd$^{\dot{h}}$w u:rd$^{\dot{h}}$w (S) "above"

-ršw paršw (S) "side"

-rhy gərhy (S) "graspable"

-rṇy wərṇy (S) "worth describing"

-rmy hərmy (S) "beautiful palace"

-lky yagywəlky (S) "name of a sage"

-ṇḍy dəṇḍy (S) "deserving punishment"

-ṇḍr puṇḍr (S) "white lotus"

-ŋk^{h}y saŋk^{h}y (S) "a philosophy"

-nty ənty (S) "last"

-ntr təntr (S) "thread, government"

-ndr čəndr (S) "moon"

-ndw prətidwəndw (S) "competition"

-ndy wəndy (S) "venerable"

-nd^{h}y trisənd^{h}y (S) "a kind of time"

Four consonant clusters:

-rtsy wərtsy (S) "alveolar"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.