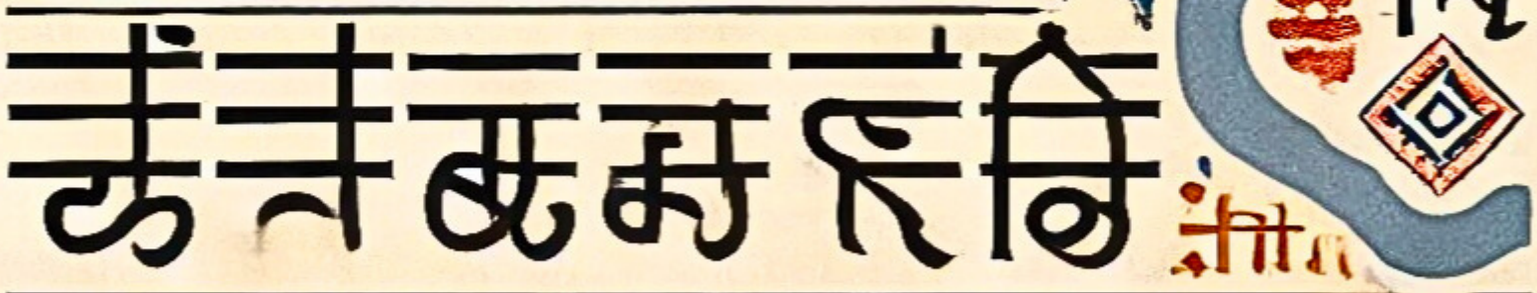
ITL-1

PROJECT REPORT

THE ESSENCE OF HINDI:

SOUND, STRUCTURE AND SCRIPT



Naman Singhal (2024114013)
CLD - IIIT Hyderabad

Acknowledgment

I express my gratitude to Dr. Rajakrishnan Rajkumar without any reservation for being my true and exceptional mentor and mentor throughout the entire course of linguistics. His encyclopedic mind and insightful recommendations greatly enhanced my learning and were of great help in the preparation of this report. This project has enabled me to appreciate the remarkable complexities of linguistics under his guidance.

Learning about dependency grammar was most enjoyable to me, as it is an alternative to the constituent grammar employed in Western language studies and used in other studies of Western languages. I am also thankful for the opportunity of studying my own mother tongue, which allowed me to look beyond the western oriented viewpoint that is commonly found in linguistic studies.

This project has been an intellectually challenging and stimulating experience and has been beneficial to me in making my transition into an academic endeavor. I thank the positive problems and the provocation that I certainly liked, which helped me expand the knowledge of linguistic plurality and different methods of considering it.

Contents

1 Abstract **5**

2 Introduction **6**

3 Phonetic Transcription **8**
3.1 Text 1- 8
3.2 Text 2- 12

4 Syllabification **17**
4.1 Introduction to Hindi Syllable Structure 17
4.2 Syllabification of selected lines from dataset 18

5 Phonology **20**
5.1 Consonants- 20
5.2 Vowels- 20
5.3 Vowel Elision 21
5.4 Consonant Doubling (Gemination) 21
5.5 Aspiration Contrast 22
5.6 Nasalization 22
5.7 Syllable Structure 22
5.8 Assimilation 23

6 Morphological Analysis **24**
6.1 Introduction to Morphology - 24
6.2 Inflectional Morphology of Verbs - 25
6.3 Inflectional Morphology of Adjectives & Determiners - 26
6.4 Derivational morphology 27
6.5 Hindi number system 29
6.6 Glossing of selected lines from the dataset- 31

7 Syntactic Analysis **33**
7.1 Introduction to Hindi sentence structure- 33
7.2 Word Trees of selected lines from dataset- 33

8 Language Family and Historical Context **39**
8.1 Historical Context- 39
8.2 Language Family: Indo-Aryan- 41
8.3 Linguistic Influences on Hindi- 41
8.4 Dialectal Variation in Hindi- 43

8.5 The Sanskritized vs Persianized Forms of Hindi- 44

9 Writing Systems and Orthography **45**
9.1 Hindi's Writing system 45
9.2 Script and Letters 45
9.3 Vowels and Matras (Diacritics) 45
9.4 Conjunct Consonants (Ligatures) 46
9.5 Additional Notable Symbols 46
9.6 Summary 46
9.7 Orthographical analysis 46

10 References **49**

Chapter 1

Abstract

This project is meant for developing linguistics through deep understanding, application and implication of a key principle of linguistics in real-world practice within the context of the Hindi language. Close analysis was carried out on the two texts written in Hindi applying a multi-layered approach toward phonetics, phonology, morphology and syntax. This project will try to obtain insight into the sound system, word-formation processes and sentence structure that are distinct to Hindi by dissecting this linguistic feature.

Following that, the historical aspect is to follow the rise and growth of the Hindi language throughout various historical epochs, which would emphasize the changes made in phonology, morphology and syntax-almost crucial segments in building this contemporary version of the language. Doing so, the historical perspective would as clearly portray to show the cultural and sociolinguistic forces that impact the development of Hindi as so enjoyed today.

Broadly, the project will directly contrast the application of theoretical concepts and analytical tools learnt during the semester to interpret and explain the interesting linguistic phenomena within a few chosen texts. The project will engage and leverage both diachronic and synchronic perspectives in any attempt to lessen the gap between theory in academic scholarship and practical linguistics analysis.

The project will deal with the peculiar features of Hindi's language particularly; conceptually lying in a broader arena of the general phenomenon of speaking and writing in Hindi. Though documentation doesn't speak out much about the history of Hindi's structure, enough of this general linguistic analysis has been done relative to the language.

Chapter 2

Introduction

Hindi is one of the officially recognized languages of India and is spoken as a mother tongue by a sizable section of its people, especially in the north including, Uttar Pradesh, Madhya Pradesh and Haryana. These areas use Hindi as the main medium of interaction and the language is the first one in the government and educational system thus making it more prominent as one of the languages connecting different groups.

The standard Hindi adopted in India is mostly based on the Khari Boli dialect, which uses the Devanagari script, which is historically used in other ancient languages such as Sanskrit. This script has expansive evolutionary aspects due to the languages it aids. Supertheoretical script also posses within the ideology and practice of the cultures it supports. As per the census conducted in 2011, about 528 million individuals residing in India consider Hindi their mother tongue, although this figure seems to be an underestimation of the actual population of Hindi speakers since it has not accounted for the vast amount of Hindi speakers residing in several nations across the globe.

The popularity of Hindi is also enhanced by Bollywood and other mainstream media which have made the language known to the people in other lands. Because of this cultural sharing, there is a bond created between Hindi speakers even with those who do not speak it, thereby emphasizing that Hindi is an inseparable part of the Indian languages and communities around the world.

Language	**Persons**	**Male**	**Female**
Magadhi/Magahi	1,27,06,825	65,90,948	61,15,877
Chhattisgarhi	1,62,45,190	81,19,432	81,25,758
Others	1,67,11,170	87,99,795	79,11,375
Rajasthani	2,58,06,344	1,33,14,254	1,24,92,090
Bhojpuri	5,05,79,447	2,60,69,236	2,45,10,211
Hindi	32,22,30,097	17,05,06,524	15,17,23,573
HINDI LANGUAGE	**52,83,47,193**	**27,66,10,187**	**25,17,37,006**

Table 2.1: Language Data by Population and Gender

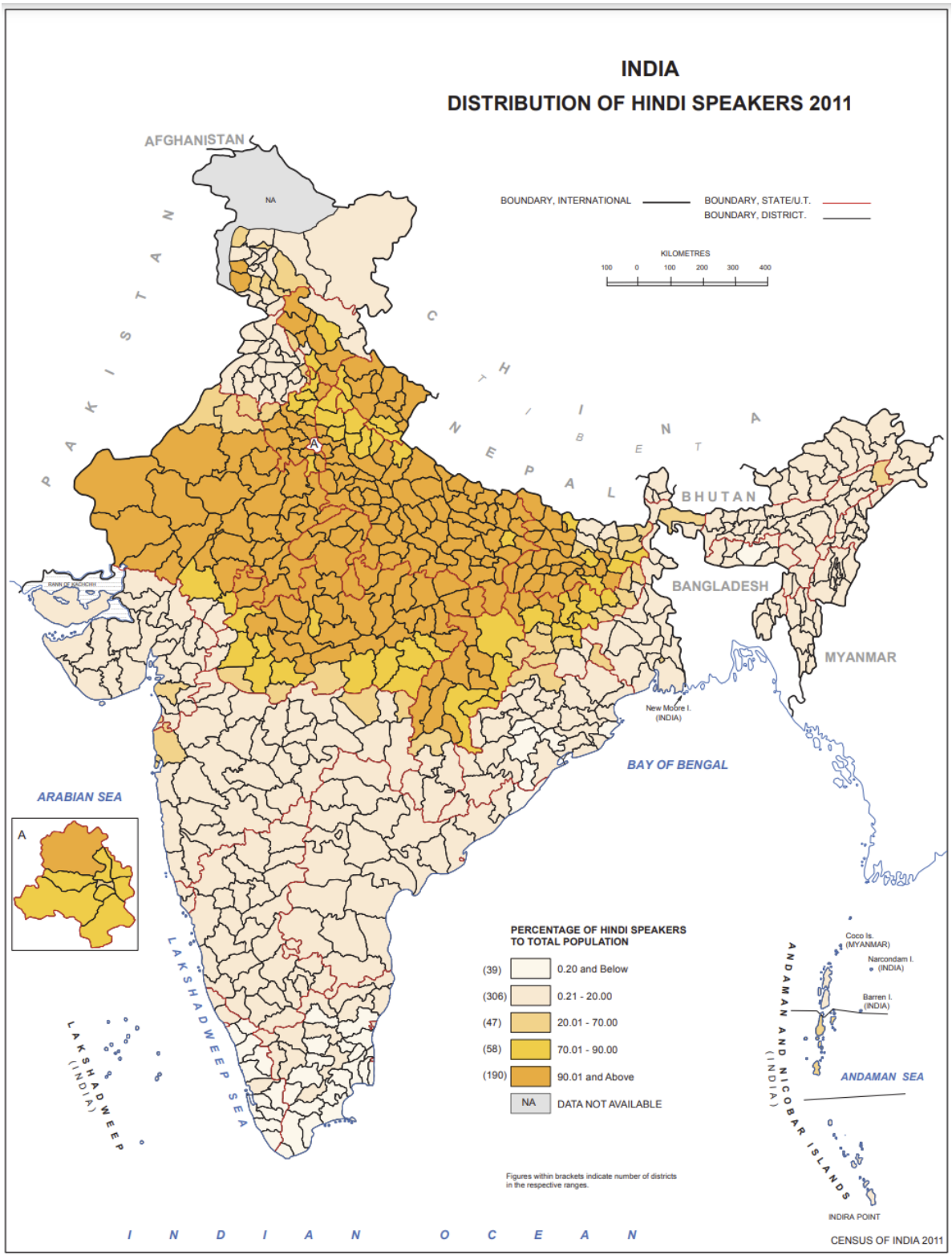


Figure 2.1: Distribution of Hindi Speakers in India

Chapter 3

Phonetic Transcription

3.1 Text 1-

Devnagri script and Roman Transcription-

मिल न पाए.. बात अलग है
Mil na paaye.. baat alag hai

पर सपनों का साथ अलग है
par sapnon ka saath alag hai

तपन तो.. मन में भर सकते हो
tapan to.. mann mein bhar sakte ho

नहीं हो सूरज.. बात अलग है।
nahin ho sooraj.. baat alag hai.

खुद में धूप समाना सीखो
khud mein dhoop samaana seekho

ऊंचे लक्ष्य बनाना सीखो।
oonche lakshya banana seekho.

लक्ष्य, नदी का.. घाट घाट को
lakshya, nadi ka.. ghaat ghaat ko

छूकर जाना थोड़ी है
chhukar jaana thodi hai

छोटी मोती लहरें लेकर
chhoti moti lehrain lekar

नाव डुबाना थोड़ी है..
naav dubaana thodi hai..

उसको तो सागर बनना है
usko to saagar banna hai

उसको तो बादल बनना है
usko to baadal banna hai

वही ढंग अपनाना सीखो
vahi dhang apnaana seekho

ऊंचे लक्ष्य बनाना सीखो ।
oonche lakshya banana seekho.

पांव तुम्हारे बिना थके ही
paanv tumhare bina thake hi

लक्ष्य तुम्हारा पा जाते हों
lakshya tumhara paa jaate ho

आँखें उसे देख सकती हों
aankhen use dekh sakti ho

बाण बेधकर आ जाते हों
baan bhedhkar aa jaate ho

उसको लक्ष्य नहीं कहते हैं
usko lakshya nahin kehte hain

ये इतने सरल नहीं रहते हैं।
ye itne saral nahin rehte hain.

मन को और उड़ाना सीखो
man ko aur udaaana seekho

ऊंचे लक्ष्य बनाना सीखो ।
oonche lakshya banana seekho.

मोती चुनने वाला कोई
moti chunne waala koi

तट पर कभी न मिलता है।
tat par kabhi na milta hai.

मन लहरों से ऊपर कर
mann lehraon se upar kar

लहरों के भीतर चलता है।
lehraon ke bheetar chalta hai.

जिससे टकराना मुश्किल हो
jisse takraana mushkil ho

जिस पर जाना मुश्किल हो
jis par jaana mushkil ho

उस पर धाक जमाना सीखो
us par dhaak jamaana seekho

ऊंचे लक्ष्य बनाना सीखो
oonche lakshya banana seekho.

मिल न पाए.. बात अलग है
mil na paaye.. baat alag hai

पर सपनों का साथ अलग है
par sapnon ka saath alag hai

तपन तो.. मन में भर सकते हो
tapan to.. mann mein bhar sakte ho

नहीं हो सूरज.. बात अलग है।
nahin ho sooraj.. baat alag hai.

खुद में धूप समाना सीखो
khud mein dhoop samaana seekho

ऊंचे लक्ष्य बनाना सीखो।
oonche lakshya banana seekho.

IPA Transcription-

mɪl nə pɑːeː bɑːt ələg ɦiɛ

pər səpnoː kɑː sɑːtʰ ələg ɦiɛ

təpən toː mən meː bʰər səktɛ ɦioː

nəɦi ɦioː suːrədʒ̑ bɑːt ələg ɦiɛ

kʰʊdə meː dʰuːp səmɑːnɑː siːkʰoː

uːⁿʧɛ ləkʃjə bənɑːnɑː siːkʰoː

ləkʃjə nə̤diː kɑː gʰɑːʈ gʰɑːʈ koː

ʧʰuːkər dʒ̑aːnaː ʈʰoːɖiː hɛ

ʧʰoːʈiː moːʈiː lɛɦireː lɛːkər

naːʊ ɖʊbaːnaː ʈʰoːɖiː hɛ

ʊskoː toː saːgər bənnaː hɛ

ʊskoː toː baːdəl bənnaː hɛ

ʊəɦiː ɖɦəŋ ʌpnaːnaː siːkʰoː

uːⁿtʃeː ləkʃjə bənaːnaː siːkʰoː

pɑːʊ tʊmʰɑːɾeː bɪnɑː θəkeː hiː

moːtiː tʃʊnneː ʊaːlɑː koːiː

ʈʌʈ pər kəbɦiː nə mɪltɑː hɛ

mən ləɦiəroːn seː uːpər kər

ləɦiəroːn keː bɪːʈ̪ər t͡ʃəltɑː hɛ

d͡ʒɪsseː ʈəkɾɑːnɑː mʊʃkɪl hoː

d͡ʒɪs pər d͡ʒɑːnɑː mʊʃkɪl hoː

ʊs pər ɖɦaːk d͡ʒəmɑːnɑː siːkʰoː

uːⁿtʃeː ləkʃjə bənaːnɑː siːkʰoː

mɪl nə pɑːeː bɑːt ələg hɛ

pər səpnoːn kɑː sɑːtʰ ələg hɛ

ʈ̪əpən toː mən meː bɦər səkʰteː hoː

nəɦi hoː suːɾəd͡ʒ bɑːt ələg hɛ

kʰʊd meː ɖɦuːp səmɑːnɑː siːkʰoː

uːⁿtʃeː ləkʃjə bənaːnɑː siːkʰoː

3.2 Text 2-

Devnagri script and Roman Transcription-

माना कि है अंधेर बहुत
maana ki hai andher bahut

और चारों ओर नाकामी है...
aur chaaro or naakaami hai...

माना कि थक के टूट रहे
maana ki thak ke toot rahe

और सफर अभी दूरगामी है...
aur safar abhi doorgaami hai...

जीवन की आपाधापी में
jeevan ki aapaadhaapi mein

जीने का ठिकाना छूट गया...
jeene ka thikaana chhoot gaya...

माना कि थक गए सपनों का
maana ki thak gaye sapno ka

नींदों में आना छूट गया..
neendon mein aana chhoot gaya..

माना कि हिम्मत टूट गई
maana ki himmat toot gayi

आंखों में निराशा छाई है...
aankho mein niraasha chhaayi hai...

माना कि हिम्मत टूट गई
maana ki himmat toot gayi

आंखों में निराशा छाई है...
aankho mein niraasha chhaayi hai...

माना कि चांद पे ग्रहण है
maana ki chaand pe grahan hai

रात अभी गहराई है...
raat abhi gehraai hai...

पर कृष्ण ने साफ कहा है कि..
par Krishna ne saaf kaha hai ki..

श्री कृष्ण ने साफ कहा है कि,
Shri Krishna ne saaf kaha hai ki,

बस कर्म तुम्हारा कल होगा...
bas karma tumhaara kal hoga...

श्री कृष्ण ने साफ कहा है कि,
Shri Krishna ne saaf kaha hai ki,

बस कर्म तुम्हारा कल होगा...
bas karma tumhaara kal hoga...

और कर्म अगर सच्चाई है तो,
aur karma agar sachchai hai to,

कर्म कहा निष्फल होगा...
karma kaha nishphal hoga...

हर एक संकट का हल होगा,
har ek sankat ka hal hoga,

वह आज नहीं तो कल होगा...
vah aaj nahin to kal hoga...

लोहा जितना तपता है,
loha jitna tapta hai,

उतनी ही ताकत भरता है..
utni hi taakat bharta hai..

लोहा जितना तपता है,
loha jitna tapta hai,

उतनी ही ताकत भरता है..
utni hi taakat bharta hai..

सोने को जितनी आग लगे,
sone ko jitni aag lage,

वह उतना प्रखर निखरता है...
vah utna prakhar nikharata hai...

हीरे पर जितनी धार लगे,
hirre par jitni dhaar lage,

वह उतना खूब चमकता है...
vah utna khoob chamakta hai...

मिट्टी का बर्तन पकता है,
mitti ka bartan pakta hai,

तब धुन पर खूब खानकता है...
tab dhun par khoob khanakta hai...

सूरज जैसा बनना है तो,
sooraj jaisa banna hai to,

सूरज जितना जलना होगा...
sooraj jitna jalna hoga...

सूरज जैसा बनना है तो,
sooraj jaisa banna hai to,

सूरज जितना जलना होगा...
sooraj jitna jalna hoga...

नदियों से आदर पाना है तो,
nadiyon se aadar paana hai to,

पर्वत छोड़ निकलना होगा...
parvat chhod nikalna hoga...

और हम आदम के बेटे हैं...
aur hum aadam ke bete hain...

और हम आदम के बेटे हैं,
aur hum aadam ke bete hain,

क्यों सोचे राह सरल होगा...
kyon soche raah saral hoga...

कुछ ज्यादा वक्त लगेगा पर,
kuch zyada waqt lagega par,

संघर्ष जरुर सफल होगा...
sangharsh zaroor safal hoga...

हर एक संकट का हल होगा,
har ek sankat ka hal hoga,

वह आज नहीं तो कल होगा..!!
vah aaj nahin to kal hoga..!!

IPA Transcription-

maːnaː ki hɛ ʌnd̪ʱeːɾ bəhʊt̪

ɔːɾ t͡ʃaːɾoː oːɾ naːkaːmiː hɛ

maːnaː ki t̪ʰək keː ʈuːʈ ɾəɦeː

ɔːɾ səfəɾ ʌbʱiː duːɾgaːmiː hɛ

d͡ʒiːʋən kiː aːpaːd̪ʱaːpiː meː

d͡ʒiːneː kaː ʈʰɪkaːnaː t͡ʃʰuːʈ gəjaː

maːnaː kɪ t̪ʰək gəjeː səpnoː kaː

niːndoː meː aːnaː t͡ʃʰuːʈ gəjaː

maːnaː kɪ ɦɪmmət ʈuːʈ gəjiː

aːkʰoː meː nɪɾaːʃaː t͡ʃʰaːiː hɛ

maːnaː kɪ ɦɪmmət ʈuːʈ gəjiː

aːkʰoː meː nɪɾaːʃaː t͡ʃʰaːiː hɛ

maːnaː kɪ t͡ʃaːnd̪ peː gɾəɦən hɛ

ɾaːt̪ ʌbʱiː gɛɦɾaːiː hɛ

pəɾ kɾɪʃɳə neː saːf kəɦaː hɛ kɪ

ʃɾiː kɾɪʃɳə neː saːf kəɦaː hɛ kɪ

bəs kəɾm t̪ʊmʰaːɾaː kəl hoːgaː

ɔːɾ kəɾm əgəɾ sət͡ʃt͡ʃaːiː hɛ toː

kəɾm kəɦaː nɪʃpʰəl hoːgaː

ɦəɾ eːk sʌŋkʌʈ kaː ɦəl hoːgaː

ʋəɦi aːd͡ʒ nəɦiː toː kəl hoːgaː

loːɦaː d͡ʒɪt̪naː t̪əpt̪aː hɛ

ʊt̪niː ɦiː t̪aːkət̪ bʱəɾtaː hɛ

soːneː koː d͡ʒɪt̪niː aːg ləgeː

ʋəɦi ʊt̪naː pɾəkʰəɾ nɪkʰəɾt̪aː hɛ

hiːɾeː pəɾ d͡ʒɪt̪niː d̪ʰaːɾ ləgeː

ʋəɦi ʊt̪naː kʰuːb t͡ʃəməkt̪aː hɛ

mɪʈʈiː kaː bəɾt̪ən pəkt̪aː hɛ

təb d̪ʱʊn pəɾ kʰuːb kʰaːnkət̪aː hɛ

suːrədʒ dʒɛːsaː bənnaː hɛ toː

suːrədʒ dʒɪt̪naː dʒəlnaː hoːgaː

suːrədʒ dʒɛːsaː bənnaː hɛ toː

suːrədʒ dʒɪt̪naː dʒəlnaː hoːgaː

nəd̪ijː seː ɑːd̪əɾ paːnaː hɛ toː

pərʋət̪ tʃʰoːɽ nɪkəlnaː hoːgaː

ɔːɾ ɦiəm ɑːd̪əm keː beːt̪eː hɛ

ɔːɾ ɦiəm ɑːd̪əm keː beːt̪eː hɛ

kjoː soːtʃeː ɾaːɦi səɾəl hoːgaː

kʊtʃ dʒjaːd̪aː ʋəkt̪ ləgeːgaː pəɾ

səŋgʱəɾʃ dʒəruːɾ səpʰəl hoːgaː

ɦiəɾ eːk səŋkʌt̪ kaː ɦiəl hoːgaː

ʋəɦi ɑːdʒ nəɦiːi toː kəl hoːgaː

Chapter 4

Syllabification

4.1 Introduction to Hindi Syllable Structure

In Hindi, a syllable is basically a nucleus sound created with a vowel, though consonants can stand either before or after the vowel sound. Commonly used syllable structures are:

- Vowel Alone: A syllable consisting solely of a vowel, e.g., आ (a).
- Consonant-Vowel (CV): The most frequent structure, where a consonant precedes a vowel, e.g., क (ka).
- Consonant-Vowel-Consonant (CVC): Syllables ending in a consonant, e.g., कम (kam).
- Complex Structures: Syllables with more complex combinations, such as कृ (kṛ) or ग्न (gn).

Syllables in Hindi can be characterized as open (सामान्य) or closed (बंद). An open syllable ends with a vowel sound, while a closed syllable ends with a consonant sound. For example, in the word बच्चा (baccha), बच (bach) is a closed syllable and चा (cha) is an open syllable.

Hindi also allows for the use of conjunct consonants, where two or more consonants combine without an intervening vowel. For instance, in स्कूल (school), the consonants स् (s) and कू (coo) combine into a single syllable.

Words in Hindi can be monosyllabic, disyllabic, as well as polysyllabic as in the case of any other languages and the syllable structure significantly affects pronunciation, stress patterns. and rhythm. Understand syllabification of Hindi words is crucial for mastering its phonetics and phonology as the syllabification of words invariably determines how they sound in speech.

4.2 Syllabification of selected lines from dataset

A selection of sentences has been randomly chosen to illustrate specific concepts. These sentences will be consistently applied across the project. To start, we will syllabify them. From this point forward, only the IPA transcription will be used when referencing these sentences.
Each line presents the sentence with words separated by spaces, followed by its syllabified form, separated by brackets-

1. mɪl nə pɑːeː bɑːt ələg ɦiɛ

—— [mɪl][nə][pɑː][eː][bɑːt][ə][ləg][ɦiɛ]

2. pər səpnoː kɑː sɑːtʰ ələg ɦiɛ

— [pər][səp][noː][kɑː][sɑː][tʰ][ə][ləgə][ɦiaːi]

3. təpən toː mən meː bʰər səktɛ ɦioː

— [tə][pən][toː][mən][meː][bʰər][sək][tɛ][ɦioː]

4. nəɦii ɦioː suːrə$\widehat{\text{dʒ}}$ bɑːt ələg ɦiɛ

— [nə][ɦii][ɦioː][suː][rədʒ][bɑː][t][ə][ləgə][ɦiɛ]

5. kʰʊdə meː dʰuːp səmɑːnɑː siːkʰoː

— [kʰʊdə][meː][dʰuːp][səmɑː][nɑː][siː][kʰoː]

6. uːⁿʈʃɛ ləkʃjə bənɑːnɑː siːkʰoː

— [uːⁿ][ʈʃɛ][lək][ʃjə][bə][nɑː][nɑː][siː][kʰoː]

7. moːtiː tʃʊnneː ʋaːlɑː koːiː

— [moː][tiː][tʃʊn][neː][ʋaː][lɑː][koːiː]

8. ʈʌʈ pər kəbʱiː nə mɪltɑː hɛ

— [ʈʌt][pər][kə][bʱiː][nə][mɪl][tɑː][hɛ]

9. mɑːnɑː ki hɛ ʌnɖʱeːr bəhʊʈ

— [mɑː][nɑː][ki][hɛ][ʌn][ɖʱeːr][bə][hʊʈ]

10. ɔːr $\widehat{\text{tʃ}}$ɑːɾoː oːr nɑːkɑːmiː hɛ

—— [ɔːr][$\widehat{\text{tʃ}}$ɑː][ɾoː][oːr][nɑː][kɑː][miː][hɛ]

11. maːnaː ki t̪ʰək keː ʈuːʈ ɾəɦeː

—— [maː.][naː][ki][t̪ʰək][keː][ʈuːʈ][ɾə][ɦeː]

12. d͡ʒiːvən kiː aːpaːd̪ʱaːpiː meː

—— [d͡ʒiː][vən][kiː][aː][paː][d̪ʱaː][piː][meː]

13. pəɾ kɾɪʃɳə neː saːf kəɦaː hɛ kɪ

—— [pəɾ][kɾɪʃɳə][neː][saːf][kə][ɦaː][hɛ][kɪ]

14. ɦəɾ eːk sʌŋkʌʈ kaː ɦəl hoːgaː

—— [ɦəɾ][eːk][sʌŋ][kʌʈ][kaː][ɦəl][hoː][gaː]

15. loːɦaː d͡ʒɪt̪naː t̪əpt̪aː hɛ

—— [loː][ɦaː][d͡ʒɪt̪][naː][t̪əp][t̪aː][hɛ]

16. kʊt͡ʃ d͡ʒjaːd̪aː vəkt̪ ləgeːgaː pəɾ

—— [kʊt͡ʃ][d͡ʒjaː][d̪aː][vəkt̪][lə][geː][gaː][pəɾ]

Chapter 5

Phonology

5.1 Consonants-

	Bilabial		Labio-dental		Denti-alveolar		Retroflex		Palatal		Velar		Glottal	
Plosive	p	b			t̪	d̪	ʈ	ɖ			k	g		
Aspirated Plosive	pʰ	bʱ			t̪ʰ	d̪ʱ	ʈʰ	ɖʱ			kʰ	gʱ		
Affricate									t͡ʃ	d͡ʒ				
Aspirated Affricate									t͡ʃʰ	d͡ʒʱ				
Fricative	ɸ		f	v	s	z	ʂ		ʃ		x	ɤ	h	ɦ
Nasal		m				n		ɳ		ɲ		ŋ		
Tap						ɾ		ɽ						
Aspirated Tap						ɾʱ		ɽʰ						
Approximant		w		ʋ		l				j				

5.2 Vowels-

Short-

Height	Front	Near Front	Central	Near Back	Back
Close	i				u
Near-close		ɪ		ʊ	
Close-mid	e				o
Mid			ə		
Open-mid	ɛ		ʌ		
Near-open	æ		ɐ		ɔ
Open			ä		

Long-

Height	Front	Near Front	Central	Near Back	Back
Close	iː				uː
Near-close					
Close-mid	eː				oː
Mid			əː		
Open-mid					
Near-open	æː				ɔː
Open			äː		

5.3 Vowel Elision

Description: Vowel elision, otherwise called as vowel deletion, denotes the absence or non-awareness of a vowel within a word and frequently the schwa sound ([ə]) is the one lost most often in Hindi informal or rapid speech contexts.
Examples:

- **Schwa Deletion**: In the word रामन (raaman), which could theoretically be pronounced as /rɑːmn/, the schwa in the last syllable is often elided, resulting in /rɑːmn/.
- **Word Combination**: In the phrase उसके लिए (uske liye), which could be articulated as /uskə liye/, the schwa is often elided in connected speech, making it sound more like /uske liye/.

5.4 Consonant Doubling (Gemination)

Description: In phonology, consonant gemination or doubling is the lengthening of a consonant sound. In Hindi, this happens between morphemes or emphatically within speech.
Examples:

- **Morphological Context**: In the word सन्नाटा (sannaṭa, meaning "silence"), the gemination of the nasal [n] is phonemic, differentiating it from सना (sana), which has a different meaning.
- **Loanwords**: In borrowed words like कंबल (kambal, meaning "blanket"), the gemination emphasizes the foreign origin and preserves the syllable structure.

5.5 Aspiration Contrast

Description: Hindi has a four way contrast for stops consisting of voiced, voiceless, aspirated and unaspirated sounds. In this regard, aspiration is the cloud of air that is released into the open space following fasciculation of certain stops, and this contrast can change meaning in Hindi.
Examples:

- **Minimal Pairs**: The difference between ताल (taal, "rhythm") and ढाल (dhaal, "shield") shows the contrast between unaspirated [t] and aspirated [t̪^{h}].
- **Voiced Aspirates**: Voiced aspirates like [g^{ɦ}] in घड़ी (ghaṛi, "watch") versus [g] in गाड़ी (gaaṛi, "vehicle") illustrate a phonemic difference due to aspiration.

5.6 Nasalization

Description: In the speech of Hindi speakers, both vowels and consonants undergo a 'nasalization' whereby a nasalized vowel is identified by the 'chandrabindu' diacritic in Devanagari. Nodalization does not amount to neutrality in terms of phonics, and it often indicates an alteration of meaning.
Examples:

- **Nasalized Vowels**: The word आँगन (aangan, "courtyard") has a nasalized vowel, differentiating it from a non-nasalized version.
- **Contextual Nasalization**: In forms like हूँ (hoon, meaning "I am"), nasalization is crucial to indicate first-person singular present tense.

5.7 Syllable Structure

Description: As a rule, the syllable structure of Hindi words is based on a (C)V(C) pattern, although sequences of consonants greater than two are found, especially in words borrowed from other languages. Such assimilations are common in borrowed and kin phrases, but native speakers typically prefer words that have an open-ended concluding vowel.
Examples:

- **Native Structure**: In मेरा (mera, "my"), the syllable structure follows a simple CV pattern (me-ra).
- **Loanword Complexity**: In स्कूल (skul, "school"), borrowed from English, the initial cluster "sk" diverges from native phonotactics, often simplified to इस्कूल (iskul) in casual speech.

5.8 Assimilation

Description: Assimilation within the scope of Hindi phonetics and phonology refers to a situation where an isolated sounds undergo considerable change so that it comes to share much in common with any of its neighbouring sounds. This may involve place or manner of articulation or both and is often allophonic, a feature describing a context derived variance in speech but without lexical change.
Examples:

- **Place Assimilation**: In उनसे (unse, meaning "from them"), the [n] sound may assimilate to the place of articulation of the following [s].
- **Manner Assimilation**: In casual speech, हैं (hain, meaning "are") may undergo nasalization assimilation, especially when followed by a nasal or voiced sound.

Chapter 6

Morphological Analysis

6.1 Introduction to Morphology -

Morphology, the analysis of the structure that differences word formation provides important understanding of hindi language. The Hindi morphology refers to the rules or patterns of combining a smaller unit of meaning with other units into complex words. Morpheme can be Roots and Affixes as well as inflections which add together with roots to express a great number of valences like tense, gender, number, case, and Aspect in Hindi.

This is highly inflectional: verbs in Hindi inflect heavily for tense, aspect and subject agreement. The root kar ('do') becomes karna in the infinitive, kiya in the past tense and kar raha hai in the continuous ("is doing"). These changes reflect the inflectional morphology of Hindi in that the form of a word shifts as dictated by syntactix context.

As with forms such as laṛka ("boy") and laṛki ("girl") that reflect the gender in Hindi, nouns also reflect agreement of number–as well. The suffix mostly gets mutated in the plural as we saw in laṛka (pl.) → laṛke (masc. pl.) or laṛki (pl.) → laṛkiya (fem. pl.). Such morphological change facilitates listeners in knowing about gender and number of a word at the earliest, which is essential due to the highly gendered and syntactically oriented nature of Hindi language.

Derivation processes, using prefixes and suffixes or compounding, when forming new words or changing the meaning of existing ones are another characteristic of Hindi morphology. For example, the addition of pra- to vidya ("knowledge"), gives pravidha ("procedure"). Another characteristic is the abundant use of compound words picking roots to associate with different meanings. One such example is vidyalay (from vidya "knowledge" + alay "place"), meaning "school".

Thus, the morphological system of Hindi is complex and open to word-formation. Its rules can help speakers and learners of Hindi to understand how meaning, syntax and grammatical relations are intricately linked —something the language is well-known for.

6.2 Inflectional Morphology of Verbs -

The Hindi verbs have inflectional morphology, where changes to the verb root express grammatical information: tense, aspect, mood, person, gender, number. The inflections on these words do not alter the meaning of the verb, but alter its form in order to conform to the relationship with the subject, and may outline temporal or modal nuances.

Tense

Hindi verbs primarily inflect to express three main tenses: present, past, and future. There are separate forms in each tense, in perfective, imperfective and progressive aspects. For example:
Present Tense: In masculine singular kar kar raha hai meaning 'is doing' and (feminine singular) kar rahi hai is same is doing(feminine).
Past Tense: Same root kar changes to kiya (masculine singular) and kiyi (feminine singular) that mean "did."
Future Tense: In Hindi, however, suffixes are used for the future action as noted in the Wikipediacited definition of karega (masculine singular) or karegi (feminine singular) meaning will do.

Aspect

Hindi verbs have auxiliary verbs and inflected participles to mark aspect, indicating whether an action is in progress, complete, or habitual. The main aspects include:
Perfective Aspect: Has the perfective participle (-a, -i, -e), perfective marking
Imperfective Aspect: Kar raha hai is a form of forms like such as kar raha tha for an ongoing action in the past ("was doing") and kar raha hai for an ongoing action in the present.
Habitual Aspect: Habitual aspect uses, karta hai(masculine), karti hai(feminine), the meaning is 'does' or 'usually does'.

Gender and Number

In fact Hindi verbs also inflect, to show the gender and number of the subject itself. This is achieved through gender-specific suffixes:
Masculine: The singular forms take -a usually (e.g., kiya for did), plural -e (e.g., kiye for did with the masculine plural subject).
Feminine: Singular has -i (e.g. kiyi for "did"), plural has -iya(e.g. kiya "did" where subject is feminine plural).

Person and Subject Agreement

Unlike Spanish or French, Hindi lacks subject specific verb conjugations, but the verb does change its ending to agree with the subject's gender and number, necessary context to know who is doing the action.

Mood

Although less defined as mood in Hindi than in many other languages, mood can be signalled via auxiliary verb forms. For example, the subjunctive mood uses ho forms (i.e. karu for "I may do"), and conditional and imperative moods have their own (kare meaning "would do", or just "do"). as a command).
Finally, the inflectional morphology of Hindi verbs provides a more elaborate system of verb forms to express tense, aspect, mood, and subject agreement. By inflecting making the language richer; speakers are able to say exactly what they mean with a great amount of precision.

6.3 Inflectional Morphology of Adjectives & Determiners -

In Hindi, inflectional morphology of adjectives and determiners plays an important role in creating agreement with the nouns they modify. Adjectives in Hindi must match the gender, number, and case of the nouns they describe, while determiners follow similar rules of agreement, allowing for cohesive and meaningful sentence structure. This system of inflections helps Hindi speakers convey specific nuances about attributes of nouns in a grammatically aligned way.

Adjectives

Hindi adjectives fall into two main categories based on their inflectional properties: There are adjectives both variable and invariable.
Variable Adjectives: The gender and number of the noun being modified determines which ending of these adjectives are chosen.
For example:
Masculine Singular: A masculine singular noun takes the adjective achha ("good") as achha laṛka ("good boy").
Masculine Plural: For masculine plural nouns, we change the adjective in to achhe, as in "good boys" - achhe laṛke.
Feminine Singular and Plural: Singular and plural feminine nouns end in i. Achhi laṛki: means good girl, Achhi laṛkiya: means good girls.

Invariable Adjectives: Adjectives listed above also do not change form according to gender or number. Such examples include adjectives like baṛa ("big"), or chhoṭa ("small"), used in an invariant fashion in certain contexts. But even those may remain under some form of requirement to concur towards the type of nouns as they utilize attributively, typically conforming of agreement without further inflection.

Determiners

Words of Hindi, which serve as determiners, indicate definiteness, proximity or quantity (vocal nouns ye ("this"), vo ("that"), sare ("all"), kuch ("some"), etc.) they rarely agree with the nouns under modification in gender and number, although they contract more in number than adjectives.

Proximity Determiners: Ye and vo change very slightly depending on the number: ye becomes ye for ”this” or ”these” and vo becomes vo for ”that” or ”those”.

Gender and Number Agreement

The inflectional morphology of adjectives and determiners aligns closely with noun characteristics:
Masculine Singular: Usually ends in -a (e.g. bara larka for ”big boy”.)
Masculine Plural: Ag. Often ends in -e (e.g. bare larke for ”big boys”).
Feminine Singular and Plural: Some usually end in –i (e.g. bari larki for 'big girl' and bari larkiya for 'big girls').

Case Agreement

Hindi adjectives and determiners also reflect case in some specific contexts, most commonly in postpositional constructions where the head of the postpositional construction is affected by an oblique case marker. Taking a possessive phrase, for example, mere bhai ka ghar (”my brother's house”); mere ('my') is the oblique form of mera ('my') agreeing with bhai.

6.4 Derivational morphology

Hindi has derivational morphology; changes of form to mean something different or change the status of a word like prefixes, suffixes, or compounding. Like inflectional morphology, derivational morphology simply suits words to grammatical needs, but produces a new words with different meanings that can express totally different meanings and shift classes of words, and enriching the vocabulary of a language.

Prefixation

The words in Hindi do have prefixes added to start them to change their meaning. Hindi takes many of its words from Sanskrit, and prefixes them for various purposes, such as negation, emphasis, direction.

Negation: We usually use the prefix na- or a- for negation. For instance:
Dosh means fault and, prefixed with ”nir”, nirdosh means honest (innocent).
adharma derives from dharma with ”a” before it meaning the opposite.

Emphasis and Direction: There are other prefixes that emphasize and give direction pra (meaning 'forward', 'intensely'), up meaning 'near', 'secondary'.
”pra” which literally means ”forward movement” is what pragati (progress) implies.
The combination of upachar — ”up” with ”char” means ”treatment”.

Suffixation

Suffixes in Hindi are used for building a new word or root word by changing the root word and sometimes word class of a word and also to convey a particular shade or meaning. However, suffixation is widely used to create nouns, adjectives as well

as forms possessed by an agent.

Noun Formation: The suffix -pan or -i makes adjectives or verbs into nouns.
Meṭhāpan means sweetness and meṭhā means sweetness. Dost ("friend") gives us dosti ("friendship") and dosti gives us dost.
Agentive Forms: Coined from the verb or noun, the suffix -wala or -kar is often attached to the end of a verb or noun to indicate an agent or somebody who performs an action.

khilana means khilaṛṛi (player) means play. gana (to sing) is the source of gayak (singer). Adjective Formation: Suffixes such as -i or -iya can be used to form adjectives, and may mean origin, or belonging.

It is from the name of India, Bharat, that the word originated in bharatiy ("Indian"). parvatiya because parvati means "mountain".

Compounding

Indepedently combining two or more words to form one additional word, is a common derivational process known as compounding in Hindi. The noun compounds can be noun-noun, verb-noun and adjective-noun; and these compounds become generally meaningful units combining separate ideas.

Noun Compounds:
These are two nouns combined to say a certain thing.

Verb Compounds:
Often verb compounds are used for concrete nuances on the action extending to expression of completion, repetition, and intensity.

Adjective Compounds:
Specific descriptive words can be produced by adjectives compounds with nouns,

Reduplication

Hindi uses reduplication, a derivational process in which a word is repeated, over and over again in order to make an emphasis or bring out a nuance of meaning. It is used to describe continuity, variety, or distribution.

Emphasis or Plurality:
Reduplication has the sense of 'intense', as in:
dheere dheere ("slowly slowly") meaning 'very slowly,' or 'gradually.'
to convey the diminutive quality, by adding chhote-chhote referring to tiny.

Distributive or Repetitive Sense: This often means that something is occurring repeatedly. kabhi-kabhi (sometimes sometimes)

6.5 Hindi number system

Introduction

This document presents an analysis of the Hindi number system in comparison to the English system. Both systems are decimal-based, but Hindi uses unique linguistic patterns, especially for the teens, multiples of ten, and compound numbers.

1–9: Basic Numbers

Like in English, each number from 1 to 9 in Hindi has a distinct word. These form the basis for larger numbers.

Numeral	Hindi (Devanagari)	Transliteration	English
1	एक	Ek	One
2	दो	Do	Two
3	तीन	Tin	Three
4	चार	Char	Four
5	पाँच	Panch	Five
6	छह	Chhah	Six
7	सात	Sat	Seven
8	आठ	Ath	Eight
9	नौ	Nau	Nine

10--19: Unique Teen Structure

In English, numbers from 11 to 19 have a "teen" suffix from thirteen onward. In Hindi, however, each number in this range has a distinct term.

Numeral	Hindi (Devanagari)	Transliteration	English
10	दस	Das	Ten
11	ग्यारह	Gyarah	Eleven
12	बारह	Barah	Twelve
13	तेरह	Terah	Thirteen
14	चौदह	Chaudah	Fourteen
15	पंद्रह	Pandrah	Fifteen
16	सोलह	Solah	Sixteen
17	सत्रह	Satrah	Seventeen
18	अठारह	Aṭharah	Eighteen
19	उन्नीस	Unnis	Nineteen

Multiples of 10 (20--90): Distinct Terms with Some Pattern

In English, multiples of ten share a "-ty" suffix (e.g., twenty, thirty). Hindi uses unique terms for each multiple of ten, without consistent suffixes but with some regular sound structure.

Numeral	Hindi (Devanagari)	Transliteration	English
20	बीस	Bis	Twenty
30	तीस	Tis	Thirty
40	चालीस	Chalis	Forty
50	पचास	Pachas	Fifty
60	साठ	Saath	Sixty
70	सत्तर	Sattar	Seventy
80	अस्सी	Assi	Eighty
90	नब्बे	Nabbhe	Ninety

21--99: Compound Number Formation

In English, numbers like "twenty-one" are formed by hyphenating the multiple of ten with the unit. Hindi, however, uses distinct forms for compound numbers, particularly from 21 to 29, 31 to 39, etc., where each compound has a specific term.

For example:

- 21 in Hindi is इक्कीस (ikkis), not a direct combination of बीस (bis, "twenty") + एक (ek, "one").
- 22--29 and similar ranges each have unique names without hyphenated combinations.

Additionally, numbers like 71 (इकहत्तर, ikhattar) have unique names that are not directly derived from combinations.

100 and Beyond: Regular Patterns with Unique Names

For 100 and above, Hindi follows a systematic approach similar to English:

- 100 is सौ (sau).
- Multiples like 200, 300, etc., follow the pattern number of hundreds + सौ, such as दो सौ (do sau, "two hundred").

Abbreviation	Full Form	Description
AUX	Auxiliary	A helping verb that supports the main verb.
ERG	Ergative	Indicates the subject of a transitive verb in certain languages.
INF	Infinitive	A non-finite verb form used to express the base meaning of a verb.
NOM	Nominative	Marks the subject of a verb.
PERF	Perfective Aspect	Indicates a completed action or state.
PL	Plural	Indicates more than one.
PRS	Present Tense	Refers to an action occurring in the present.
REFL	Reflexive	Refers back to the subject of the sentence.
SG	Singular	Indicates one entity.
3SG	Third Person Singular	Refers to a single third-person entity (he/she/it).

6.6 Glossing of selected lines from the dataset-

1. mɪl nə pɑːeː bɑːt ələg ɦiɛ

meet-INF NEG possible-NEG matter-NOM different-NOM AUX.3SG.PRS
”It’s a different matter we can’t meet.”

2. pər səpnoː kɑː sɑːtʰ ələg ɦiɛ

but company-NOM of dreams-NOM different-NOM AUX.3SG.PRS
”But the company of dreams is different.”

3. nəɦi ɦioː suːrəd͡ʒ bɑːt ələg ɦiɛ

no is sun matter-NOM different-NOM AUX.3SG.PRS
”There’s no sun, that’s a different matter.”

4. kʰʊdə meː dʰuːp səmɑːnɑː siːkʰoː

yourself into sunlight engulf-PURP learn-IMP
”Learn to engulf sunlight into yourself.”

5. moːtiː tʃʊnneː ʋaːlɑː koːiː

pearl-NOM touch-INF one-POSS someone
”Someone who touches pearls.”

6. mɑːnɑː ki hɛ ʌn̪ᵚᵚᵚ

8. d͡ʒiːʊən kiː ɑːpɑːd̪ʱɑːpiː meː

life-NOM of hustle-NOM in
”In the hustle of life.”

9. ɦiər eːk sʌŋkʌʈ kɑː ɦiəl hoːgɑː

every one problem-NOM of solution-NOM will_be
”There will be a solution to every problem.”

10. pər krɪʃɳə neː sɑːf kəɦɑː hɛ kɪ

but Krishna-NOM ERG clear say-PERF AUX.3SG.PRS that
”But Krishna has clearly said that.”

Chapter 7

Syntactic Analysis

7.1 Introduction to Hindi sentence structure-

Syntactic Analysis Hindi sentence structure is an examination of how words combine to form grammatically correct phrases and sentences. Hindi syntax is marked primarily by a word order of subject-object-verb, although there are numerous instances where an emphasis or rhetorical motivation might place the verb before an object. In general, the analysis also comprises research about noun conflation with adjectives, postposition placement, and even the roles verbal auxiliaries play with tense, aspect, and mood. This process helps to understand the hierarchical relationships between the different constituents within a sentence.

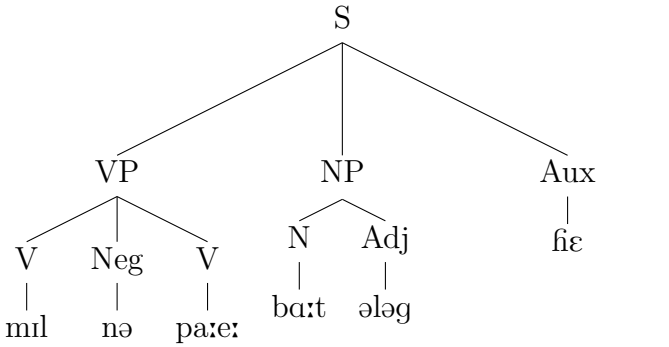
7.2 Word Trees of selected lines from dataset-

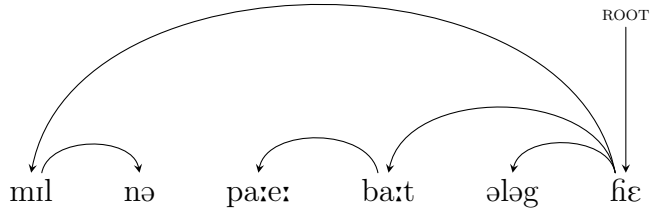
Abbreviation	**Meaning**
S	Sentence
NP	Noun Phrase
VP	Verb Phrase
V	Verb
N	Noun
Aux	Auxiliary verb
Neg	Negation
Adj	Adjective

Table 7.1: Grammar Abbreviations Used in Word Trees

1. mɪl nə paːeː baːt ələg ɦiɛ

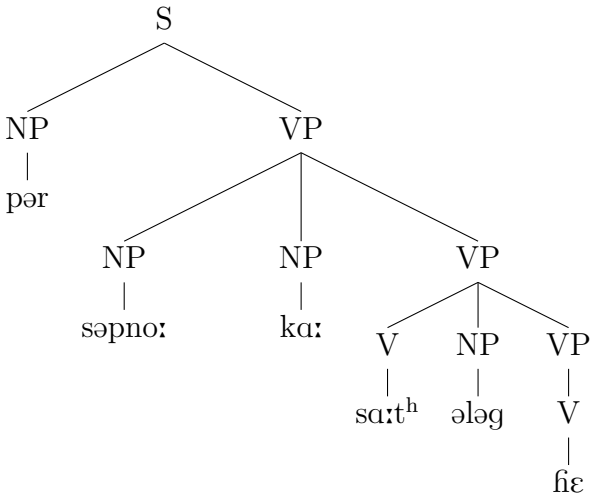
— [S [VP [V mɪl] [Neg nə] [V paːeː]] [NP [N baːt] [Adj ələg]] [Aux ɦiɛ]]

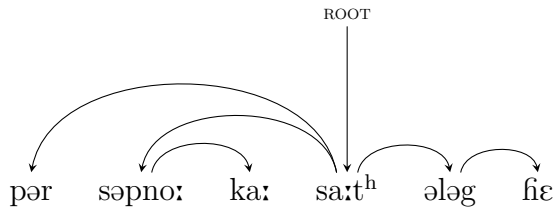




2. pər səpnoː kɑː sɑːt^{h} ələg ɦiɛ

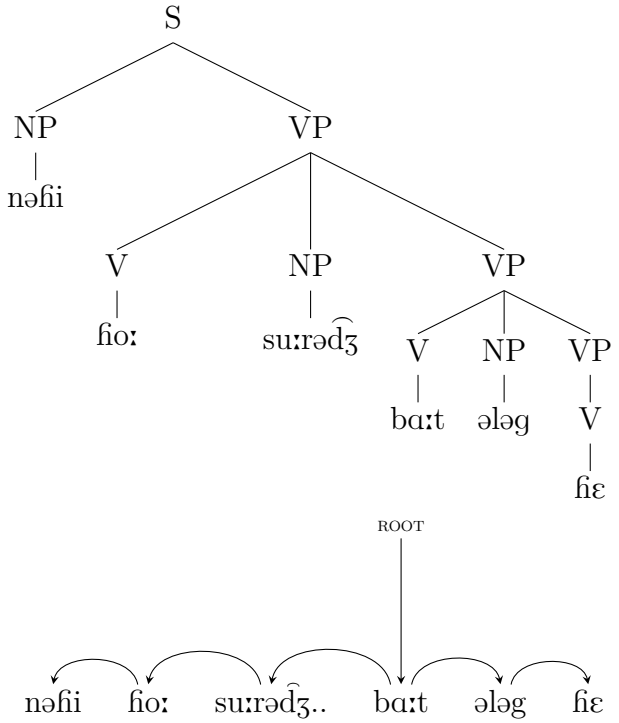
— [S [NP pər] [VP [NP səpnoː] [NP kɑː] [VP [V sɑːt^{h}] [NP ələg] [VP [V ɦiɛ]]]]]





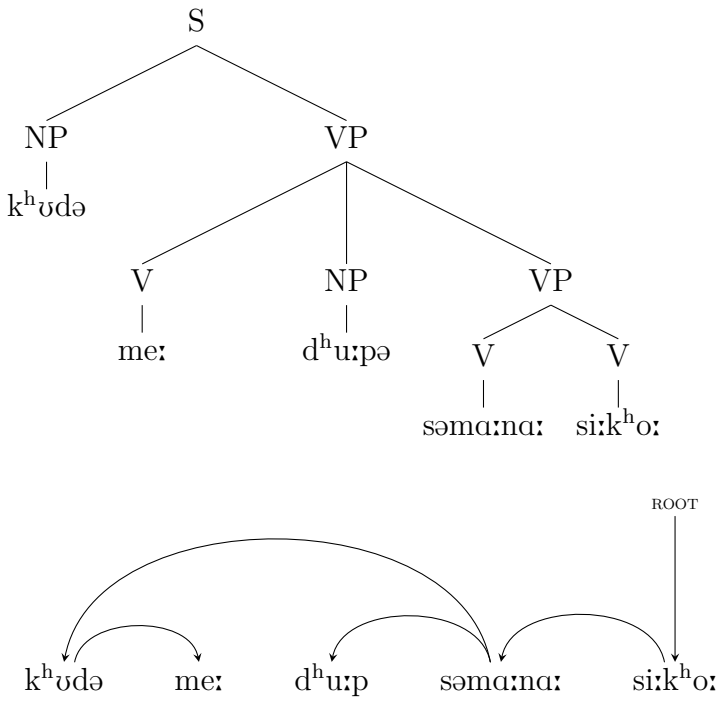
3. nəɦii ɦoː suːrəd͡ʒ baːt ələg ɦɛ

— [S [NP nəɦii] [VP [V ɦoː] [NP suːrəd͡ʒ] [VP [V baːt] [NP ələg] [VP [V ɦɛ]]]]



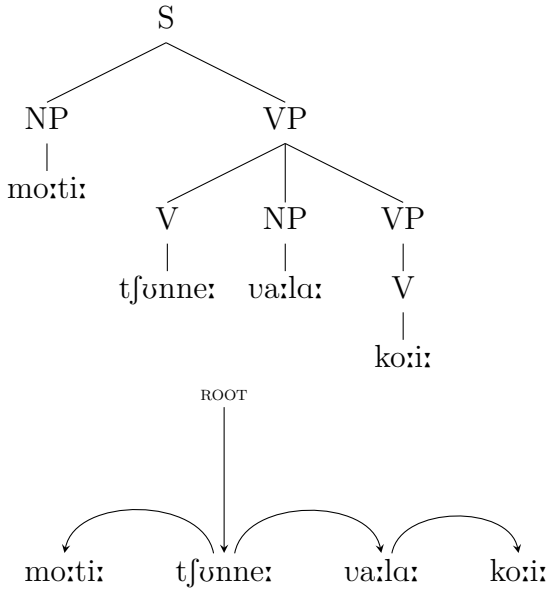
4. kʰʊdə meː dʰuːp səmaːnaː siːkʰoː

— [S [NP kʰʊdə] [VP [V meː] [NP dʰuːpə] [VP [V səmaːnaː] [V siːkʰoː]]]]



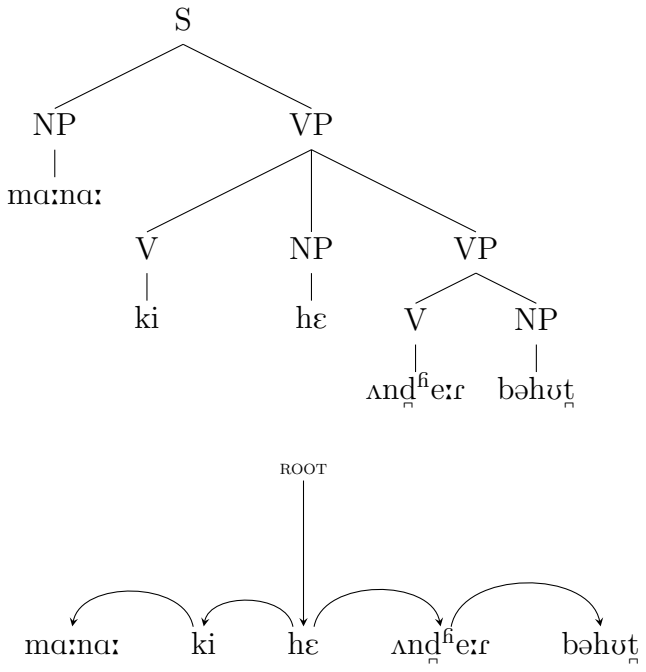
5. moːtiː tʃʊnneː vaːlɑː koːiː

— [S [NP moːtiː] [VP [V tʃʊnneː] [NP vaːlɑː] [VP [V koːiː]]]]



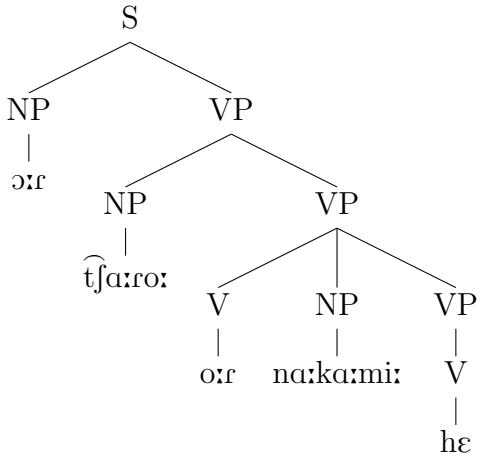
6. maːnaː ki hɛ ʌnd̪ʱeːr bəhʊt̪

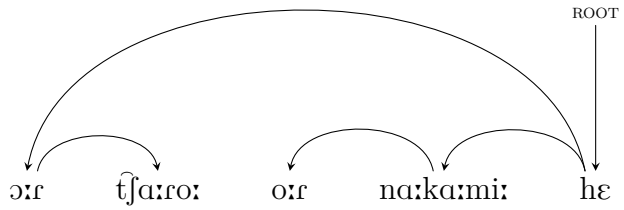
— [S [NP maːnaː] [VP [V ki] [NP hɛ] [VP [V ʌnd̪ʱeːr] [NP bəhʊt̪]]]]



7. ɔːɾ t͡ʃaːɾoː oːɾ naːkaːmiː hɛ

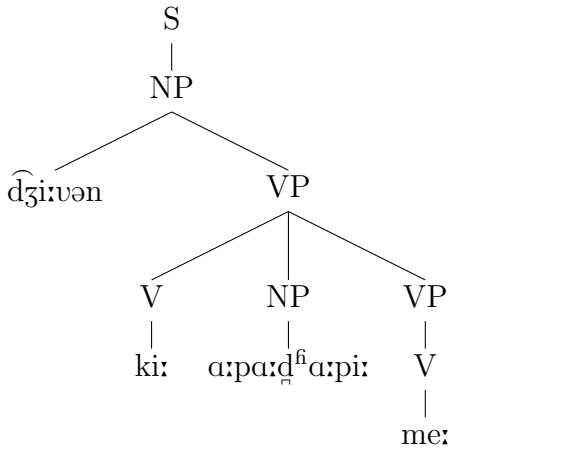
— [S [NP ɔːɾ] [VP [NP t͡ʃaːɾoː] [VP [V oːɾ] [NP naːkaːmiː] [VP [V hɛ]]]]]

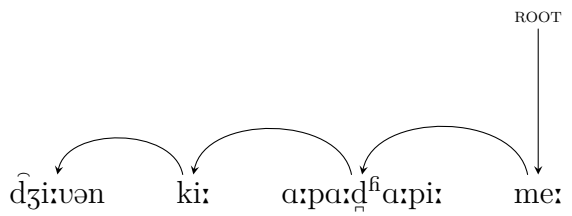




8. d͡ʒiːvən kiː aːpaːɖᵑaːpiː meː

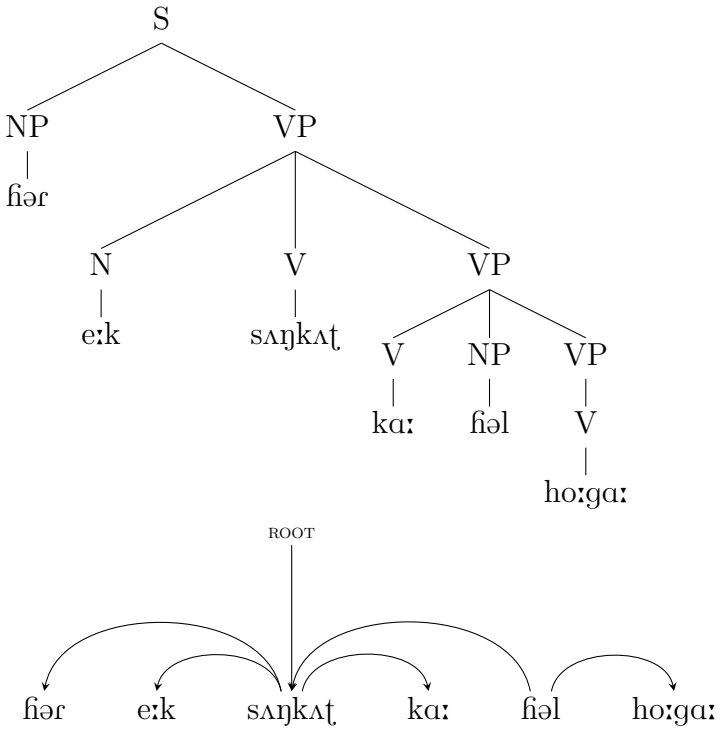
— [S [NP d͡ʒiːvən] [VP [V kiː] [NP aːpaːɖᵑaːpiː] [VP [V meː]]]]





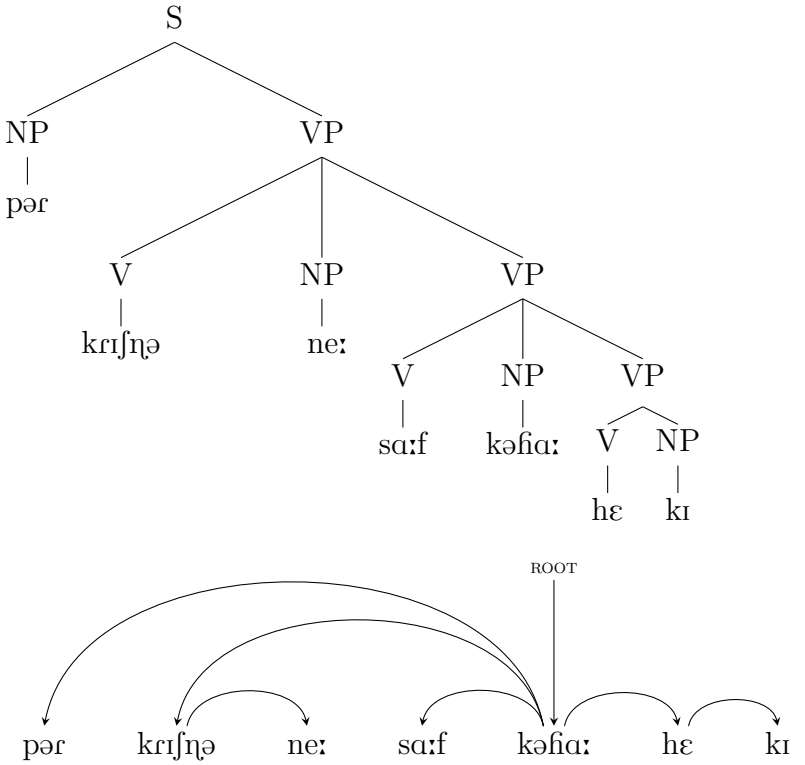
9. ɦər eːk sʌŋkʌt̪ kɑː ɦəl hoːgɑː

— [S [NP ɦər] [VP [N eːk] [V sʌŋkʌt̪] [VP [V kɑː] [NP ɦəl] [VP [V hoːgɑː]]]]]



10. pər krɪʃɳə neː saːf kəɦɑː hɛ kɪ

— [S [NP pər] [VP [V krɪʃɳə] [NP neː] [VP [V saːf] [NP kəɦɑː] [VP [V hɛ] [NP kɪ]]]]]



Chapter 8

Language Family and Historical Context

8.1 Historical Context-

Hindi is one of the officially recognized languages of India. It is one of the languages most widely spoken globally. Hindi is heir to a very rich history and culture which has gone thousands of years. It originated from assimilation among various types of different languages, cultures, and historical influences within the Indian subcontinent. For the formation of the Indo-Aryan branch of the family of the Indo-European languages.

Ancient Period: Proto-Indo-European to Vedic Sanskrit-

Hindi originated from the Proto-Indo-European (PIE) tongue of peoples estimated to have resided in the Eurasian steppe circa 4000-3500 BCE. The Indo-Iranian branch stemming from PIE and the Indo-Aryan languages containing a sub-group within this branch. Of all the Indo-Aryan languages, Vedic Sanskrit (c. 1500–500 BCE) was the oldest attested form; it was the language of the Vedas, India's oldest sacred texts.

Hindi has origins as ancient as Proto-Indo-European, the far back peoples — as far back we know from historical framework — living around 4000-3500 BCE in the Eurasian steppe. Their parent language evolved into this branch of the Indo-Iranian languages, and indeed, the so-called Indo-Aryan languages developed in fact

Middle Period: Prakrits and Apabhramsha-

Towards the end of Vedic period leading on to later Sanskrit period around 600 BCE the spoken forms of Sanskrit transitioned into Prakrit languages. These regional lingua franca of sorts were popularly spoken versions of middle Indo-Aryan languages, with a grammatically and lexicon less complex than classical sanskrit, they became the speech for like the majority of n. Indias population.

Prakrits had branched into regional languages by the first millennium CE, one such language being Apabhramsha, which is a direct ancestor of modern Hindi. Is

an Apabhramsha dialect then which were spoken over northern India from the 6th to the 13th centuries and is a precursor for Hindi, and many other modern Indo-Aryan languages such as Bengali; Marathi; Gujarati; and Punjabi.

Mughal Period: Persian Influence and the Emergence of Hindi-Urdu-

The opening up of the Delhi Sultanate (13th–16th centuries) and subsequently the Mughal Empire (16th–19th centuries) in northern India, brought with it significant linguistic and cultural shifts that influenced both Sanskrit as well as regional languages. Persian as the administrative and literary language of the Mughal court intermeshed with localized Indo-Aryan dialects and formed a type of Hindustani, incorporating Hindi elements with Persian vocabulary. The language used in this region around Delhi came to be known as 'Khari Boli' and ultimately became the basis of present day Standard Hindi and Urdu.

Focusing on this period, Hindi started taking a large volume of loanwords from Persian, Arabic and Turkic languages which were instrumental in building the modern-day oral dictionary of Hindustani. Urdu evolved as a dialect of Hindustani that absorbed the Persian script and substantial Persian and Arabic vocabulary, whilst Hindi, using the Devanagari script, retained more Sanskrit features.

Colonial Period: Standardization and the Hindi-Urdu Divide-

Only when the British colonial rule founds its roots in 18th and 19th century, standardization of Hindi was made inevitable. The standardization of the vernaculars for administrative and educational purposes was one of several processes promoted by British administrators that resulted in the formal separation of Hindi from Urdu. It was in the Devanagari script that Hindi became standardized, while attempting Sanskritization by removing Persian and Arabic loanwords – Cleaning up Hindi, especially in modern formal & literary context. At the same time, Urdu continued with its present Persian script and vocabulary, especially in fields of literature and culture.

The 19th century also saw the rise of Hindi as a symbol of national identity during the Indian independence movement. Hindi and Urdu, which were once considered two registers of the same language, increasingly came to be seen as separate languages, reflecting the growing religious and cultural divide between Hindus and Muslims.

Post-Independence: Hindi as the National Language-

Following Indian independence in 1947, Hindi became one of the official languages of the Union along with English. This was done as a part of the larger nationwide initiative to propagate one nation, many identities idea, but it also led to the rising linguistic unrest in non-Hindi-speaking areas (especially southern regions) in India. In the face of these obstacles, however, Hindi continued to flourish and today occupies a central space in the culture of large sections of northern India as the dominant

language of government, education, media and popular culture.

8.2 Language Family: Indo-Aryan-

Hindi/Hindustani is a part of the Indo-Aryan group within the Indo-European language family, which also includes nearby major languages, including Bengali, Marathi, Gujarati and Punjabi as well as Urdu. Indo-Iranian is a major language family that consists of the Indo-Aryan languages, as well as the Iranian languages such as Persian and Pashto.

Major Stages in Hindi's Linguistic Evolution:

Old Indo-Aryan (c. 1500 BCE – 600 BCE)-
The chief linguistic family that existed was Vedic Sanskrit, the language of early Indian scriptures. This inflected nature of Sanskrit provided the foundation for later grammatical systems in Indo-Aryan languages, Hindi among them.

Middle Indo-Aryan (600 BCE – 1000 CE)-
In the aftermath of Classical Sanskrit, as the Prakrits (spoken vernaculars) emerged on the scene, grammar was more or less simplified and vocabulary became diversified. Languages as Pali and Shauraseni Prakrit started to develop, used in literature, drama and religious texts during this time. Apabhramsha dialects, which were crucial in the evolution of Hindi, were being spoken during this time.

New Indo-Aryan (after 1000 CE)-
During the medieval period, local dialects, like Braj Bhasha and Khari Boli became predominant and converged afterward to produce contemporary Hindi. With the advent of Turkic and Persian into these dialects with Muslim conquests in India starting from 1192 CE around the time of the Delhi Sultanate and continuing till Mughal period, an enormous vocabulary with different syntax was added.

8.3 Linguistic Influences on Hindi-

Over the course of its rich and diverse history, Hindi absorbed linguistic influences which helped it develop into a language with a wide variety of possible vocabulary, syntax, and even pronunciation.

Sanskrit (Branch of Indo-Aryan)-

The core vocabulary and grammar of Hindi mostly comes from Sanskrit. Sanskrit (tatsama words) which today makes a large part of Hindi, in terms of formal, technical and literary idioms, in addition to the many more tatbhava words that link Sanskrit with popular use of the language through sporadic pangs of linguistic simplification. The morphological system of Hindi is heavily influenced by Sanskrit,

especially its noun and verb conjugations.

Example-

- Sanskrit: *pustakam* → Hindi: *pustak* (book)
- Sanskrit: *mitram* → Hindi: *mitr* (friend)

Persian and Arabic (Perso-Arabic Influence)-

The Persian and Arabic influences on Hindi stem primarily from the Mughal era. Words related to administration, law, art, and culture were borrowed extensively from Persian and Arabic. These loanwords are especially prevalent in Urdu, but they also form a significant part of Hindi's lexicon, particularly in everyday speech.

Example-

- Persian: *duniya* (world), *sabzi* (vegetable)
- Arabic: *qanoon* (law), *kitab* (book)

Turkic Influence-

The Turkic influence on Hindi is less pronounced than Persian or Arabic, but some Turkic words, especially those related to military and governance, entered Hindi via Persian during the Mughal period.

Example-

- *topee* (hat)
- *sipahi* (soldier)

Portuguese Influence-

The Portuguese colonization of parts of India in the 16th century introduced several loanwords into Hindi, particularly related to food, clothing, and daily life.

Example-

- *almirah* (cupboard)
- *baldi* (bucket)
- *sabun* (soap)

English Influence-

British colonization left an indelible mark on Hindi, particularly in domains like technology, governance, and education. English words are frequently used in contemporary Hindi, either in their original form or with slight phonetic modifications. Example-

- *train*, *station*, *doctor*
- *file*, *manager*, *school*

8.4 Dialectal Variation in Hindi-

Hindi is spoken across a vast region, and several dialects reflect local cultural and linguistic diversity. Some dialects are mutually intelligible with Standard Hindi, while others, like Bhojpuri or Marwari, are considered distinct languages.

Major Dialects of Hindi-

Khari Boli (Standard Hindi):
Spoken in the Delhi region, Khari Boli forms the basis of Modern Standard Hindi. It became the standardized form during British rule and is now the language of education, government, and media in much of northern India.

Braj Bhasha:
Historically associated with the region around Mathura and Vrindavan, Braj Bhasha was the language of Krishna Bhakti poetry. Though its prominence has diminished, it remains an important literary dialect.

Awadhi:
Awadhi, the language of classical epic poetry such as Ramcharitmanas by Tulsidas and spoken in eastern Uttar Pradesh It preserves its cultural and historical essence of the hindi heartland.

Bhojpuri:
Bhojpuri is a language spoken in the state of Bihar and eastern UP; sometimes considered a separate language, it largely exhibits Hindi features. Widely spoken throughout the Indian diaspora of the Caribbean and Mauritius.

Marwari:
Marwari — Spoke in Rajasthan and although it has a separate identity, Hindi is somewhat related. And it's widely applied in folk songs And customs of Rajasthan.

8.5 The Sanskritized vs Persianized Forms of Hindi-

The influence of Sanskrit and Persian on Hindi has given it two distinct styles that underscore its diverse heritage.

Sanskritized Hindi:

Heavily influenced by Sanskrit, this form is used in formal contexts such as government documents, school texts and ancient poems. Sanskritised Hindi focuses on "pure" Indic vocabulary, employing tatsama words (direct borrowings from Sanskrit) and identified with poise and ritual. Consider prashasan or administration and vidyalaya for school, both commonly used words in official documents. The form also enhances religious discourse and patriotic literature, adding a note of dignity and tradition.

Persianized Hindi:

Persianized Hindi influenced by Persian, Arabic and Turkic languages arose during the Mughal period which speaks in daily life conversation and also in popular cultures including Bollywood. It adds warmth and accessibility with calls like duniya (world), bhalek, and zindagi (life), that we readily understand. Perhaps more powerful and better utilized for romance or poetry, Persianized Hindi softens the lyrical edge that drips from everyday conversation.

Modern Usage:

Today, speakers alternately bounce between both forms depending on context — Sanskritised Hindi for the lofty and nationalistic and Persianised Hindi for daily speaking in an even quasi-relatable manner. Such flexibility emphasises adaptability, maintaining its roots that are steeped in culture as well as ancient texts while also absorbing and encompassing the cultural exchange, and this contrast of aspects of identity which India reflects into one language.

Chapter 9

Writing Systems and Orthography

9.1 Hindi's Writing system

Hindi uses a script called Devanagari, which is a member of the Brahmic script family tree. This script is mainly used for writing Hindi and many other Indian languages like Sanskrit, Marathi, and Nepali amongst others.

Devanagari is unique in that it is an abugida that is, in this system, each consonant grapheme is associated with an inherent vowel, which is often the short a. This vowel can also be changed or omitted completely depending on the diacritics that are used, which in this case are called matras. These shapes cling onto the consonants and tweek their sounds so that they are able to make different vowel sounds.

Therefore, the arrangement of letters in the Devanagari script allows the language to be written phonemically, which makes it flexible yet orderly. All these features help enhance the Hindi language and more so encourage an effective crisp writing for better statements.

9.2 Script and Letters

The Devanagari script is composed of 46 base symbols, namely 33 consonants and 13 vowels. Consonants in Hindi are normally pronounced with an 'a' sound attached, but this is altered with the use of vowel marks, or matras. Most of the Hindi writing is said to be phonetic since sound corresponds proportionally well with each character that is written.

9.3 Vowels and Matras (Diacritics)

Hindi has independent vowels that stand alone and vowel diacritics (matras) that modify consonants. Here is a list of Hindi vowel diacritics and examples:

English Term	Hindi Character	Matra Form	Example Word
Long “a” sound	आ	ा	काम (kaam)
Long “i” sound	ई	ी	पीना (peena)
Short “i” sound	इ	ि	किताब (kitaab)
Long “u” sound	ऊ	ू	फूल (phool)
Short “u” sound	उ	ु	कुल (kul)
E sound	ए	े	एक (ek)
Ai sound	ऐ	ै	बैल (bail)
O sound	ओ	ो	तो (to)
Au sound	औ	ौ	और (aur)

Hindi uses a virama, or halant (्), to indicate that a consonant does not carry an inherent vowel sound, often when forming conjunct consonants.

9.4 Conjunct Consonants (Ligatures)

In Devanagari, when two consonants appear together without a vowel, they form a ligature called a conjunct consonant. These are visually distinct forms that blend the consonant shapes together. Examples include:

- क्त (kta), as in क्त in संयुक्त (sanyukt, meaning “united”)
- त्र (tra), as in त्र in त्रास (traas, meaning “terror”)

9.5 Additional Notable Symbols

- Anusvara (ं): This represents a nasal sound added to vowels, as in हंस (hans, meaning "swan").
- Chandrabindu (ँ): This nasalizes a vowel, as in आँख (aankh, meaning "eye").
- Visarga (ः): Adds a breathy “h” sound, commonly seen in Sanskrit loanwords, like दुःख (duḥkh, meaning "sorrow").

9.6 Summary

The Devanagari script used for Hindi is an abugida system, where each consonant inherently carries a vowel sound, typically ”a.” Through the use of matras (diacritical marks), this vowel can be modified to produce a range of sounds. Devnagri also features conjunct consonants or ligatures, which combine multiple consonants into unique symbols for complex sounds. Overall, Hindi’s orthographic system allows for clear phonetic representation, making it both versatile and systematic.

9.7 Orthographical analysis

Devanāgarī alphabet for Hindi

Vowels (स्वर) and vowel diacritics

अ	आ	इ	ई	उ	ऊ	ऋ	ए
a	ā	i	ī	u	ū	ṛ	e
[ʌ]	[a]	[i]	[iː]	[u]	[uː]	[ɾɪ]	[e]
क	का	कि	की	कु	कू	कृ	के
ka	kā	ki	kī	ku	kū	kṛ	ke
ऍ	ऐ	ओ	ऑ	औ	अं	अः	अँ
ê	ai	o	ô	au	aṅ	aḥ	ãṃ
[e]	[æː]	[o]	[o]	[ɔː]	[aŋ]	[əh]	[ãː]
कॅ	कै	को	कॉ	कौ	कं	कः	काँ
kê	kai	ko	kô	kau	kaṅ	kaḥ	kāṃ

Consonants (व्यंजन)

क	ख	ग	घ	ङ	च	छ	ज	झ	ञ
ka	kha	ga	gha	ṅa	ca	cha	ja	jha	ña
[kə]	[kʰə]	[gə]	[gʱə]	[ŋə]	[ʧə]	[ʧʰə]	[ʤə]	[ʤʱə]	[ɲə]
ट	ठ	ड	ढ	ण	त	थ	द	ध	न
ṭa	ṭha	ḍa	ḍha	ṇa	ta	tha	da	dha	na
[ʈə]	[ʈʰə]	[ɖə]	[ɖʱə]	[ɳə]	[tə]	[tʰə]	[də]	[dʱə]	[nə]
प	फ	ब	भ	म	य	र	ल	व	
pa	pha	ba	bha	ma	ya	ra	la	va	
[pə]	[pʰə]	[bə]	[bʱə]	[mə]	[jə]	[rə]	[lə]	[ʋə]	
श	ष	स	ह						
śa	ṣa	sa	ha						
[ʃə]	[ʂə]	[sə]	[ɦə]						

Variant letters

Some letters are two forms: the Classical, Northern or Kalikata (Calcutta) form is used in the north of India; while the Modern, Southern or Mumbai (Bombay) form is used in the south India and has become the standard form.

Classical	অ	আ	ও	ঔ	ঝ	ণ	ল
Modern	अ	आ	ओ	औ	झ	ण	ल
	a	ā	o	au	jh	ṇ	l

A selection of conjunct consonants

There are about a thousand conjunct consonants, most of which combine two or three consonants. There are also some with four-consonant conjuncts and at least one well-known conjunct with five consonants.

क्क	क्ख	क्क	क्ण	क्त	क्त्य	क्र	क्र्य	क्त्व	क्न	क्न्य	क्म
kka	kkha	kca	kṇa	kta	ktya	ktra	ktrya	ktva	kna	knya	kma
क्य	क्र	क्र्य	क्ल	क्व	क्व्य	क्ष	क्ष्म	क्ष्य	क्ष्व	ख्य	ख्र
kya	kra	krya	kla	kva	kvya	kṣa	kṣma	kṣya	kṣva	khya	khra
ग्य	ग्र	ग्र्य	घ्न	घ्न्य	घ्म	घ्य	घ्र	ङ्क	ङ्क्त	ङ्क्त्य	ङ्क्य
gya	gra	grya	ghna	ghnya	ghma	ghya	ghra	ṅka	ṅkta	ṅktya	ṅkya
ङ्क्ष	ङ्क्ष्व	ङ्ख	ङ्ख्य	ङ्ग	ङ्ग्य	ङ्घ	ङ्घ्य	ङ्घ्र	ङ्ङ	ङ्न	ङ्म
ṅkṣa	ṅkṣva	ṅkha	ṅkhya	ṅga	ṅgya	ṅgha	ṅghya	ṅghra	ṅṅa	ṅna	ṅma
ञ्य	च्च	च्छ	च्छ्र	ञ्च	च्म	च्य	छ्य	छ्र	ज्ज	ज्ज्ञ	ज्ञ
ñya	cca	ccha	cchra	cña	cma	cya	chya	chra	jja	jjha	jña
ज्ञ्य	ज्म	ज्य	ज्र	ज्व	ञ्च	ञ्म	ञ्य	ञ्छ	ञ्ज	ञ्ज्य	ट्ट
jñya	jma	jya	jra	jva	ñca	ñcma	ñcya	ñcha	ñja	ñjya	ṭṭa

Numerals

०	१	२	३	४	५	६	७	८	९	१०
0	1	2	3	4	5	6	7	8	9	10

Chapter 10

References

1. Hindi Census Data of 2011

2. Language Atlas 2011

3. Poem "Dar Hai Yadi Gir Jane Ka To, Shikhar Tumhare Liye Nahin Hai"
-by Kavi Sandeep Dwivedi

4. Poem "Har ek sankat ka hal hoga vah aaj Nahi to kal hoga"
-by Shubham Shyam

5. Hindi Structure and History

6. Hindi alphabet and pronunciation

7. Hindi ipa chart

8. Omniglot

9. An Introduction to Language - Victoria Fromkin

10. History and linguistic classification of hindi

11. Information on hindi language

12. Latex Stack Exchange

13. Wikipedia

14. Leipzig Glossing Slides