

THE HEART OF CHANGE ISSUES ON VARIATION IN HINDI

Tatiana Oranskaia · Anvita Abbi (Eds.)

The Heart of Change

हिंदी तेरे रूप अनेक

The Heart of Change Issues on Variation in Hindi

हिंदी तेरे रूप अनेक
बदलाव के बीच में

edited by


Tatiana Oranskaia

Anvita Abbi



H A S P
HEIDELBERG ASIAN
STUDIES PUBLISHING

ORCID®

Tatiana Oranskaia  <https://orcid.org/0000-0001-7207-9190>

Anvita Abbi  <https://orcid.org/0000-0003-3163-3751>

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.dnb.de>.



This book is published under the Creative Commons Attribution License CC BY-SA 4.0. The cover is subject to the Creative Commons License CC BY-ND 4.0.



Published at Heidelberg Asian Studies Publishing (HASP), 2022

Heidelberg University/Heidelberg University Library
Heidelberg Asian Studies Publishing (HASP),
Grabengasse 1, 69117 Heidelberg, Germany
<https://hasp.ub.uni-heidelberg.de>

The electronic open access version of this work is permanently available on the website of Heidelberg Asian Studies Publishing: <https://hasp.ub.uni-heidelberg.de>
urn: urn:nbn:de:bsz:16-hasp-919-2
doi: <https://doi.org/10.11588/hasp.919>

Text © 2022 by the authors.

ISBN 978-3-948791-26-1 (Softcover)

ISBN 978-3-948791-25-4 (PDF)

Contents

Note on Transliteration, Transcription and Glossing ix

Tatiana Oranskaia and Anvita Abbi
Variation in Hindi: An Introduction xi
तत्याना ओरांस्कया और अन्विता अब्बी
प्रस्तावना

PART I GEOGRAPHIC VARIATION IN GRAMMAR AND LEXIS

Anvita Abbi
Restructuring of Grammar in the Emerging Varieties of Hindi Across India 3
अन्विता अब्बी
भारत में उभरती हिंदी भाषा की विविधताओं के व्याकरण का पुनर्गठन

Annie Montaut
On the Non-Lexical Categories of *avyay* 'Invariables' and 39
Their Grammaticalization in Pahari Languages, with a Comparison
to Standard Hindi
आनी मोंतो
पहाड़ी भाषाओं में सहायक कोटियों के रूप में अव्ययों का व्याकरणिकरण और
मानक हिंदी के साथ उनकी तुलना

Vashini Sharma
The Impact of Other Indian Languages on Dakhini 79
वशिनी शर्मा
दखनी पर अन्य भारतीय भाषाओं का प्रभाव

Heinz Werner Wessler
Cosmopolitan Hindustani Under Aurangzeb: Terminological Matters 97
in François Marie de Tours' *Thesaurus Linguae Indiae*
हाइंस वरनर वेस्लर
औरंगज़ेब के ज़माने की भूमंडलीकृत हिंदुस्तानी: फ्रोंसुआ मारी दे तूर के शब्दकोष
में पारिभाषिक शब्दावली पर एक टिप्पणी

**PART II
VARIATION IN THE GRAMMAR AND DISCOURSE
OF STANDARD HINDI**

Liudmila Khokhlova

Conative: Compleitive Contrast in Hindi-Urdu Aorist Forms 117

ल्यूडमीला खोखलोवा

हिंदी-उर्दू में सामान्य भूतकाल में प्रयत्न पक्ष और पूर्ण पक्ष का व्यतिरेक

Saartje Verbeke and Aaricia Ponnet

Animacy, Specificity, and Verb Semantics.
What Drives Differential Object Marking in Hindi? 133

सार्त्य वबेके और आरीसीया पोन्नेत्

सजीवता, विशिष्टता और क्रिया शब्दार्थ । डिफरेंशियल ऑब्जेक्ट मार्किंग
“को” हिंदी में किन कारकों से प्रभावित होता है?

Ekaterina Kostina

To Link or Not to Link? Variation in Using Discourse
Markers in Hindi 155

एकतेरीना कोस्तिना

लिंक किया जाए या नहीं? प्रोक्ति चिह्नों के प्रयोग में परिवर्तनशीलता

Tatiana Oranskaia

Hindi Clause Strings with Adverbial Clauses:
A Tentative Formalisation 187

तत्याना ओरांस्कया

हिंदी की उपवाक्य शृंखलाएँ – अंतरिम सूत्रबद्ध प्रणाली

PART III
VARIATION ISSUES IN HINDI TEACHING

<i>Rama Kant Agnihotri</i> Multilinguality and Language Teaching	213
रमा कान्त अग्निहोत्री बहुभाषिता एवं भाषा शिक्षण	
<i>Anil Kumar Pandey</i> How to Determine Grammatical Gender in Hindi	227
अनिल कुमार पाण्डेय हिंदी में लिंग निर्धारण	
<i>Sandhya Singh</i> Attitude Towards Learning Hindi in Singapore Universities: A Study	247
संध्या सिंह सिंगापुर के विश्वविद्यालयों में हिंदी सीखने का रुझान: एक अध्ययन	
On the Contributors	261

Note on Transliteration, Transcription and Glossing

The transliteration is based on the system ISO 15919 with the following deviations from it:

- if an anusvara is used in the same way as an anunasika, that is to mark a nasalised vowel, it is represented with a tilde, e.g. बूँद or बूँद *bū̃d*, आँख or आँख *ā̃kh*;
- if an anusvara marks a nasal preceding a stop, it is represented with *m̃* (not *m̄*);
- ऋ / ॠ is represented with *ṛ* (not *ṙ*).

The transcription uses the International Phonetic Alphabet (2015).

Glossing

In the articles that use morphological glosses the level of glossing varies depending on the objective of the article. Basically we follow the Leipzig Glossing Rules (<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>) occasionally complementing them with the abbreviations from the Wikipedia List of glossing abbreviations <https://en.wikipedia.org/wiki/List_of_glossing_abbreviations>, which is compiled from various academic sources.

A note is due here on the marking of postpositions in Hindi and its dialects. As their status is contestable in all positions except after pronouns, they are marked as clitics (with the preceding sign =) only after pronouns.

Tatiana Oranskaia 
and Anvita Abbi 

Variation in Hindi: An Introduction

तत्याना ओरांस्कया और अन्विता अब्बी

प्रस्तावना

Variation in language is the basis of its change. The title of the volume reflects this idea, which took form in the early sociolinguistic studies of the mid-20th century and has become nearly a truism in modern linguistic theory. The metaphor of heart in the title refers to the expression used by Stephen Levinson.¹

The topic of variation draws our attention to the fact that the more varied a language the more it is honoured and accepted in society. Like homogenous biosphere is neither sustainable nor life-giving similarly logosphere can be sustained only if language and languages are varied. English today has assumed the status of world class language because it has spread across several seas in various forms equally acceptable to non-native speakers. In fact, there are more non-native speakers of English today than the native speakers.

Linguists believe that there is a progressive relationship between the antiquity of a language and its varieties. Variety gives a language the power of adaptability and change and thus sustainability over a period. Variety also indicates its acceptability across several socio-cultural, geographical, political, and economic strata. The more various forms—regional or social—a language has, the longer the history of its existence is. A case in point is the Indo-European language family, many of whose varieties have developed into independent languages over a period, Hindi among them.

Uniformity kills a language, and this is the reason diversity of the language indicates its antiquity and sustainability. Overstandardisation of any language leads to its fossilisation. For example, classical Sanskrit is frozen today because it fell out of spoken use, unlike Prakrits, which developed into several modern Indo-Aryan

1 Interview with Stephen C. Levinson on 1 December 2017 (Max Planck Institute for Psycholinguistics 2017).

languages. A convincing evidence is the grammar of Sanskrit enshrined into Panini's *Aṣṭādhyāyī* written around 500 BCE. The principles and rules of this grammar are still used in the classroom situation to teach Sanskrit, as the grammatical structure of the classical language remains the same.

Not many notions are as broad as variation. It stretches over practically all areas of human knowledge, among which linguistics is a prominent field of scholarship, one in which reality is filtered to reduce the perception of the phenomena to a system. The Saussurean dichotomy of *langue* and *parole* is seen as the theoretical concept which originally introduced the perspective on variation within the Western science of language—one that continues to hold sway up to this day. This conceptual framework, closely preceded by the dialectological research and ideas of Alexander von Humboldt, later found notable application in the principles of the Prague school and American structuralism. The rise of sociolinguistics with the new line of inquiry introduced to research by Labov (Chambers 2002: 6) established a strong connection between social factors and variation, thereby making the latter a major topic of interest of this relatively young and blooming branch of linguistics. The present volume lies almost totally outside this trend; the studies here deal for the most part with the language per se.

Currently, variation is primarily conceived as socially or geographically determined non-standard speech usage, and more strongly related to *parole* than to *langue*. This, naturally, does not exclude the variation inherent in the language system, which belongs to the intrinsic “orderly heterogeneity” language possesses (Weinreich et al. 1968: 100). From the very moment the social constituent manifestly entered language studies it was clear that patterns of social variation do not necessarily relate to linguistic taxonomic units. The relationship between the intra- and extralinguistic factors responsible for the disproportion between a meaning and multiple means of its expression may be more precisely defined as “linguistic and social co-variation” (Wolfram 1993: 199).

The perspective on intrinsic language variation is present in the Indian linguistic tradition, which, informed by the aspiration to preserve the Vedas unaltered, focused on preserving their divine language Sanskrit unchanged. There is evidence of scholarly thought on this topic starting with the formation of the various Vedic renditions. Famously Panini's rules for correct Sanskrit also take into account variants (or ‘options’ *vibhāṣā*, ‘alternatively [applicable]’ *anyatarasyām*) in the language. Subsequently, the Pali and Prakrit grammarians follow categorisation patterns like those of Sanskrit, making use of much more variable Middle Indic data.

Most of the contributions in this volume loyally follow the intralingual variation slant of the classical Indian tradition when analysing Hindi data.

Hindi, in the broad meaning of the term, that is including numerous regional forms, surpasses the Middle Indo-Aryan languages of the preceding historical stage (as far as we know them) in the richness of its variation. Compared with Hindi in

this sense, Modern Standard Hindi provides, naturally, fewer, but still abundant sets of interchangeable linguistic items, often of diverse ancestry: Indo-Aryan, Perso-Arabic and English. In this book, the study by Heinz Werner Wessler presents a classic case of semantic variation in the target language, namely in translation of Christian theological terminology: each single Latin entry in the dictionary is rendered with several definitions, on the one hand using Persian and Arabic terms on a par with each other for Islamic notions and on the other Indo-Aryan terms for Hindu notions.

Quite a few prior studies have examined variation on different levels of the Hindi language system. Many of them are concerned with the classical sphere of variation research—phonetics, phonology, and prosody (e.g., Ohala 1976; Pandey 1989; 2020). Other works deal with variation in Hindi syntax (e.g., Vasisht et al. 2012; Kothari 2010; Davison 2015) or the interplay between syntactic and phonological structures (e.g., Patil et al. 2008). A smaller number of variation analyses explore the other subsystems.

Extralingual impetuses are responsible for numerous geographical and sociocultural code-mixed varieties of Hindi which expand its variability scope. Language contact is a multidimensional, multidisciplinary field in which interrelationships between interlocutors hold the key to the understanding of how and why people use language/s the way they do on the one hand and how structures of the contacting languages change on the other. The mixing of Hindi and English has been a favourite subject of exploration for decades, seemingly exceeding the research on the interplay between Hindi and other Indian languages. The latter, so to say, indigenous contact type, is responsible for the emergence of a plethora of grammatically simplified non-standard Hindis as languages of contact described here in the article by Anvita Abbi. Each of them functions as a regional *lingua franca* and some are even used as the language of regional administration and teaching.

The present volume does not provide a survey of the whole range of variation in Hindi. Nor does it concentrate on theoretical considerations. It sees its modest task as giving glimpses of, for the most part, grammatical but also lexical and discourse variation as well as of varying attitudes to and usages of Hindi in educational pragmatics.

The book makes use of the breadth of the notion of variation. It includes studies on diverse aspects of the phenomenon within the major official language of India considering them from different perspectives. These perspectives may be delineated as follows:

- Comprehending the paths of grammaticalisation— with its effect of increasing varieties of Hindi across several states and the possibility of bringing social cohesion, especially in contact situations (Abbi; Sharma);
- Identifying the variation as a long historical process in some dialects of Hindi (Montaut);

- Discerning constraints on variation in Modern Standard Hindi grammar and discourse (Khokhlova; Kostina; Oranskaia; Verbeke & Ponnet);
- Tracing variation in lexical diversity of translated terminology in the source and target language (Wessler) and derivational suffixation in the standard lexis (Pandey);
- Exploring sociolinguistic and ethnocultural issues in language teaching within India (Agnihotri) and particularly teaching of Hindi across borders (Singh).

The volume quashes the myth of monolithic standard language as the norm and the only variety that draws respect and acceptability (Abbi; Sharma). Language contact has been visualized here with a new perspective as contacts between languages are contingent upon ethnic, demographic, cultural, educational, and socio-economic status of communities.

The bulk of the articles deal with Hindi in India. Some of them (Abbi; Sharma) attempt to look more precisely into geographical forms of Hindi and their interaction in different contexts and types of communication. Annie Montaut's and Heinz Werner Wessler's studies add a chronological dimension to this perspective.

A similar vantage point is taken here when investigating the possibilities of and constraints upon formal variation with regard to the grammatical structures of Standard Hindi (Khokhlova; Kostina; Oranskaia; Verbeke & Ponnet). A pragmatic approach comes to the fore when discussing problems and methods of Hindi teaching in ethnoculturally mixed contexts (Agnihotri; Pandey; Singh), one point of interest being the motives among learners from different backgrounds for learning Hindi, that is, how they plan to use their new knowledge (Singh). The article by Agnihotri presents a philosophical stance on language education from the multi-lingual perspective.

The analyses use data written (including manuscripts), spoken and electronic, which allows variability patterns in Hindi to be explored both in the broad and narrow meanings of the term. In this way, the inclusivity of various forms of data of existing and ancient Hindi makes this volume rare.

Results have been obtained by consulting a wide database of Hindi (Verbeke & Ponnet), compiling and tagging a syntactic database (Oranskaia) and testing data on different subjects (Khokhlova; Kostina). The article by Heinz Werner Wessler deals with the study of the manuscripts from the early 18th century which are fruits of missionary linguistic work in South Asia. It informs readers of the antiquity of the literature available in the forms of grammars and dictionaries well archived now in European libraries.

The volume also exposes the experimental results of rather challenging issues. These are compound verbs, which continue to puzzle researchers, and discourse markers in Hindi. Ludmila Khokhlova's study of Hindi compound verbs is a meticulous analysis of their variation from the perspectives of semantics and pragmatics.

The investigation conducted by Ekaterina Kostina underpins the understanding of discourse markers as tools ensuring text integrity and coherence and emphasises that variation is their fundamental feature.

Two distinctive characteristics of the present book deserve special mentioning: first, it is concerned with variation only and exclusively in Hindi, and second, it is bilingual, with contributions in English and in Hindi.² The main text of each article is preceded by an abstract and a list of key words in both languages.

We hope that the language-specific studies may help to improve and increase our understanding of linguistic variability in general.


Heartfelt thanks are due to Ex-Vice Chancellor of the Mahatma Gandhi Antar-rashtriya Hindi Vishwavidyalaya in Wardha, Prof. Girishwar Misra, who initially came up with the idea of preparing this volume and has been supporting it with his interest and advice.


We thank most warmly the contributors to this volume for their valuable works and patient cooperation.

We also thank Heidelberg Asian Studies Publishing for taking this volume into its publication programme, and especially warmly Ms Nicole Merkel-Hilf, Ms Elizaveta Ilves, Ms Iris Matzner and Ms Daniela Jakob for their help and support.

It is with the heaviest of hearts and deepest sorrow that we think of our dear colleague and fellow contributor Prof. Vashini Sharma who will not see this book published. She left us all on 2 January 2021. Her article here remains a draft.

ORCID®

Tatiana Oranskaia  <https://orcid.org/0000-0001-7207-9190>

Anvita Abbi  <https://orcid.org/0000-0003-3163-3751>

Bibliography

Chambers, Jack K. 2002. “Studying Language Variation: An Informal Epistemology”, in: Chambers, J. K., Peter Trudgill & Natalie Schilling-Estes (eds): *The Handbook of Language Variation and Change*. (Blackwell Handbooks in Linguistics.) Malden, Mass., 3–14.

2 The variationist approach continues in the English language of the contributions. Thus, some authors use American spelling and punctuation, while others follow British English usage. In some cases, also the Hindi spelling varies.

- Davison, Alice 2015. “Hindi-Urdu: Central Issues in Syntax”, in: Kiss, Tibor & Artemis Alexiadou (eds): *Syntax - Theory and Analysis*. (Handbücher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science (HSK), 42, 3.) Berlin: De Gruyter Mouton, 1478–1517.
- Kothari, Anubha 2010. *Processing Constraints and Word Order Variation in Hindi Relative Clauses*. Stanford University, ProQuest Dissertations Publishing, 28168220. (https://stacks.stanford.edu/file/druid:vh4o2hk4761/anubha_thesis-augmented.pdf), accessed: October 17, 2020).
- Max Planck Institute for Psycholinguistics Institute for Psycholinguistics 2017. Celebrating 25 Years of Research into Language and Cognition, 14 December 2017. (<<https://www.mpi.nl/news/celebrating-25-years-research-language-and-cognition>>, accessed August 17, 2020).
- Ohala, Manjari 1976. “Nasals and Nasalization in Hindi”, in: Ferguson, Charles A., Lary M. Hyman & John J. Ohala (eds): *Nasálfest. Papers from a Symposium on Nasals and Nasalization, November 1975*. Stanford: Stanford University, 317–332.
- Pandey, Pramod 1989. “Word Accentuation in Hindi”, *Lingua* 77, 1: 37–73 (<[https://doi.org/10.1016/0024-3841\(89\)90038-7](https://doi.org/10.1016/0024-3841(89)90038-7)>, accessed: November 19, 2020).
- Pandey, Pramod 2020. “An Optimality Theoretic Account of Word Stress in Hindi”, *Lingua* 250. (<doi.org/10.1016/j.lingua.2020.102994>, accessed: January 14, 2021).
- Patil, Umesh et al. 2008. “Focus, Word Order and Intonation in Hindi”, *Journal of South Asian Linguistics* 1, 1: 55–72 (<<http://jsal-journal.org>>, accessed: January 12, 2021).
- Vasishth, Shravan, Rukshin Shaher & Narayan Srinivasan 2012. “The Role of Clefting, Word Order and Given-New Ordering in Sentence Comprehension: Evidence from Hindi”, *Journal of South Asian Linguistics* 5 (<<http://jsal-journal.org>>, accessed: January 14, 2021).
- Weinreich Uriel, William Labov & Marvin Herzog 1968. “Empirical Foundations for a Theory of Language Change”, in: Lehmann, Winfred P. & Yakov Malkiel (eds): *Directions for Historical Linguistics: A Symposium*. Austin: University of Texas Press, 95–188.
- Wolfram, Walt 1993. “Identifying and Interpreting Variables”, in: Preston, Dennis R. (ed.): *American Dialect Research*. (Centennial Series of the American Dialect Society.) Amsterdam; Philadelphia: John Benjamins Publishing Company, 193–222.

PART I

**GEOGRAPHIC VARIATION
IN GRAMMAR AND LEXIS**

Anvita Abbi 

Restructuring of Grammar in the Emerging Varieties of Hindi Across India

अन्विता अब्बी

भारत में उभरती हिंदी भाषा की विविधताओं के व्याकरण का पुनर्गठन

Abstract Intensive language contact is the genesis of diffusion of forms phonetically, morphologically, syntactically, and semantically. Both internal innovation and external pressures help in restructuring existing grammar in such a way that the language of contact is widely acceptable and prestigious. The multilingual mosaic of the country feeds the structure of the contact language in multiple ways. We observe that languages grow in layers as the model language is constantly changed by extension, innovation, and grammaticalization processes as exemplified by the contact Hindi used in Hindi-belt and non-Hindi zones.

The author discusses two different kinds of contact Hindi, the one which is used in the so-called ‘Hindi belt’, such as Bihar and Jharkhand—including the Union Territory of Andaman and Nicobar, where Hindi is the official language,—and the other in non-Hindi zones of Arunachal Pradesh marked by the multiplicity of mutually non-intelligible indigenous languages and Meghalaya, home to Mon-Khmer group of Austroasiatic languages—the two significant Northeast states of India. The author has discussed the salient features of contact Hindi and compared the similarities and differences among them. The study helps us to identify those structures of Standard Hindi which are prone to change, innovate and lose in the contact situation. It has been shown that Hindi in its non-standard variety can be seen as the biggest factor for social cohesion and mobility across the nation. The emergence of contact-induced linguistic structures makes this lingua franca a powerful contact language in the linguistically heterogeneous milieu.

Keywords standardization, language innovation, language adaptation, language acceptance, social cohesion, varieties of Hindi.

सारांश गहन भाषा-संपर्क की वजह से भाषाओं में ध्वन्यात्मक, रूपात्मक, वाक्यात्मक और अर्थ संबंधी परिवर्तनों का जन्म और प्रसार होता है। आंतरिक नवाचार और बाहरी दबाव की मदद से मौजूदा व्याकरण में ऐसे बदलाव होते हैं जिनसे संपर्क भाषा को व्यापक स्वीकार्यता और प्रतिष्ठा मिलती है। देश की बहुरंगी बहुभाषिकता संपर्क भाषा की संरचना को कई तरीकों से प्रभावित करती है। हमने पाया है कि भाषाएँ परतों में विकसित होती हैं क्योंकि मॉडल भाषा विस्तारण, नवाचार और व्याकरणीकरण जैसी प्रक्रियाओं के कारण लगातार बदलती रहती है। इसका एक उदाहरण हम हिंदी क्षेत्र और गैर-हिंदीभाषी इलाकों में संपर्क भाषा के रूप में हिंदी के इस्तेमाल में देख सकते हैं।

लेखिका ने प्रस्तुत लेख में संपर्क भाषा के रूप में हिंदी की उन दो किस्मों पर चर्चा की है जिनमें से एक तथाकथित 'हिंदी पट्टी' जैसे बिहार और झारखंड में बोली जाती है। इसमें केंद्र शासित प्रदेश अंडमान और निकोबार भी शामिल है, जहाँ हिंदी को राजभाषा का दर्जा मिला हुआ है। दूसरी ओर पूर्वोत्तर भारत के गैर-हिंदीभाषी इलाके हैं जैसे अरुणाचल प्रदेश जहाँ पारस्परिक रूप से अबोधगम्य विभिन्न जनजातीय भाषाएँ बोली जाती हैं, और मेघालय राज्य जहाँ ऑस्ट्रिक परिवार के अंतर्गत आनेवाली मोन-ख्मेर भाषाएँ बोली जाती हैं। लेखिका ने संपर्क हिंदी की मुख्य विशेषताओं पर चर्चा की है और उनके बीच की समानताओं और असमानताओं पर प्रकाश डाला है। यह अध्ययन हमें मानक हिंदी की उन संरचनाओं की पहचान करने में मदद देता है जिनमें संपर्क की स्थिति में परिवर्तन, नवाचार और लुप्त होने की संभावना है। लेख इस बात को इंगित करता है कि कैसे हिंदी की इन अमानक किस्मों को पूरे देश में सामाजिक सामंजस्य और गतिशीलता के सबसे बड़े कारक के रूप में देखा जा सकता है। संपर्क-प्रेरित भाषाई संरचनाओं का उद्भव हिंदी जैसी संपर्क भाषा को एक विषम परिवेश में शक्तिशाली संपर्क भाषा के रूप में स्थापित करता है।

मुख्य शब्द – मानकीकरण, भाषा नवाचार, भाषा अनुकूलन, भाषा स्वीकार्यता, सामाजिक सामंजस्य, हिंदी की किस्में ।

1 Introduction

Urban India in the post-independent era is connected by Hindi which serves as the only lingua franca among the uneducated and along with English for both educated and uneducated people. According to the Census 2011, Hindi is a cover term that includes 57 varieties termed as 'mother tongues'—Hindi one of them—listed under the 'Hindi language'. The total population speaking Hindi and its varieties are enumerated as 528,347,193.

These varieties have been recognized not as independent languages but as "dialects" socially, educationally, and politically in post-independent India (Abbi 2009). The reductionist attitude of the government gives us the impression that 43.63% of the Indian population speaks Hindi (Census 2011). However, the truth is that Hindi in its non-standard form has become the major lingua franca of the country used by more than 60% of the population (refer to bilingualism figure in the Appendix). Contrary to the still-dominant opinion, this is the language that is widely accepted by the communities all over India, including southern India (Abbi et al. 1998–2000). It is significant for us to understand how

various forms of contact Hindi have negotiated Hindi grammar to make it accessible and acceptable all over India.

The author discusses two different kinds of situations that exist in the country. The first situation arises where Hindi in the so-called ‘Hindi belt’ (includes nine states, the Union Territory of Delhi, and the geographically distanced from the main Hindi belt Union territory of Andaman and Nicobar) has restructured several grammar standardization norms, and thus emerged as a *language in contact*, imbibing some features of the closely related varieties. The second situation is represented by Northeast India marked very heavily by linguistic diversity and multiplicity of languages. Hindi in these states has emerged as a non-native variety, unfolding the processes of language acculturation, language adaptation, linguistic innovations, and above all and the most important of all, language acceptance (Sharma 2013; Abbi & Sharma 2014). In the Northeast, Hindi has a distinct structure from the standard variety, as this is the *language of contact*.

2 The status of Hindi across the country

As said earlier, according to the official census of 2011 Hindi is spoken by 528,347,193 speakers (cf. Appendix for details). If we add to this the official figure of those who reported Urdu as their mother tongue (the two languages not being very different at the spoken level), which is 50,772,631, the speakers of Hindi-Urdu will be considered as representing the highest percentage of the total population, i. e. $43.63\% + 4.19\% = 47.82\%$. There has been an increase of 4.34% in Hindi mother-tongue users since the 1991 Census enumeration, from 39.29% to 43.63% of the total population. The numerical strength of users of Hindi as a second language, or for inter-community communication across various states and union territories of India, can only be guessed, as no statistical study has been made on this.

The use of Hindi across states, and social and religious identities, has given rise to non-standard varieties of Hindi both in areas where Hindi is not the primary official language and in the states of the Hindi belt, where Standard Hindi (SH) is the official language: Himachal Pradesh, Uttarakhand, Haryana, Rajasthan, Uttar Pradesh, Bihar, Jharkhand, Chhattisgarh, Madhya Pradesh, the Union Territory of Delhi and the Union Territory of Andaman and Nicobar. Coupled with different registers, we have several varieties of non-Standard Hindi which are used in day-to-day life. The complex mosaic of the Hindi speech community is further enlarged by the emergence of cosmopolitan Hindi, which shows traits of convergence of several languages and dialects spoken in the cosmopolitan cities. This gives rise to many varieties of Hindi.

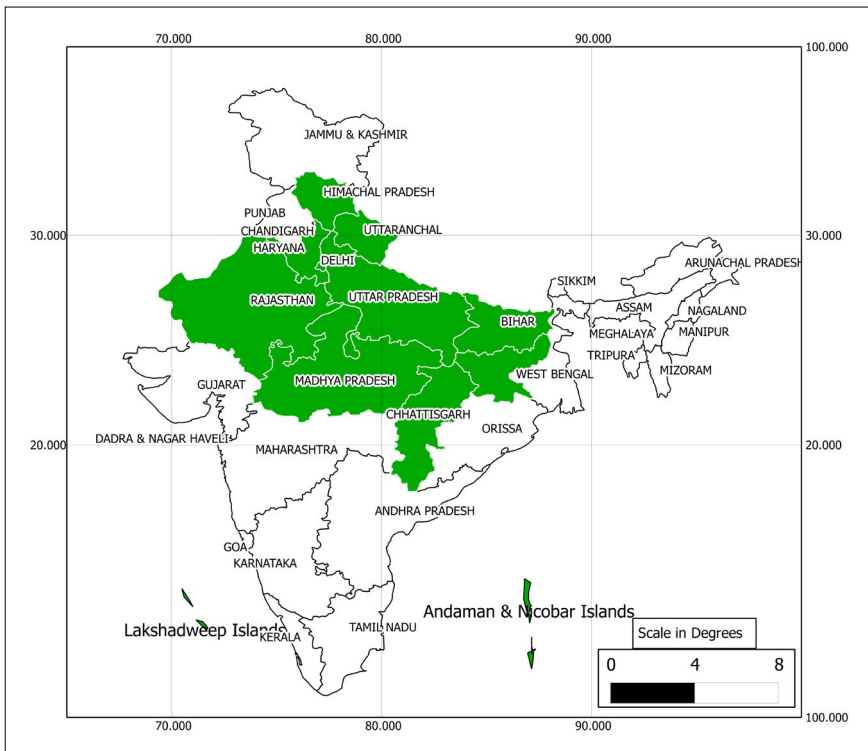


Fig. 1 Map of Hindi States (Abbi & Sharma 2014)

3 Varieties of Hindi

Although we have multiple Hindis that are operational across the country, we can roughly classify them into six varieties:

- Hindi spoken in the cosmopolitan cities
- Hindi spoken in the Hindi belt
- Hindi spoken in the non-Hindi zones
- Hindi spoken in the Northeast (is part of the non-Hindi zone)
- Hindi spoken in the Andamans which comes under the Hindi-belt but is constituted of a diverse population of India speaking various scheduled languages as mother tongues.
- Standard Hindi with several registers used by people in different professions such as teachers, writers, radio announcers, public speechmakers, politicians, and hawkers among others.

Each of these varieties is marked by a distinct grammatical structure and has emerged in different historical, political, and geographical circumstances. Extensive bi or multilingualism prevailing in each state contributes to shaping these varieties. In the Hindi belt states what are classed as varieties of Hindi (see the Appendix) are independent languages, several with a rich literary or folklore tradition, as, for instance, Magahi and Bhojpuri in Bihar, Braj and Awadhi in Uttar Pradesh, Marwari, and Mewati in Rajasthan, Kumaoni and Garhwali in Uttarakhand, Chhattisgarhi in Chhattisgarh, among others. Hindi in these regions has grown in contact with these numerous languages which are the home languages of the users of Hindi. It is worth mentioning that Hindi being the official language is taught as a compulsory subject in schools and thus, readers and listeners are exposed to the SH in schools as well as in government offices.

This situation of Hindi, being the official language or one of the official languages in the states of the Hindi belt, has given rise to Contact Hindi (CH) which has its genesis in these regional languages (57 of them) and the SH. Interestingly, when asked, speakers of tribal languages of the Hindi belt usually claim themselves to be Hindi speakers and very rarely divulge the names of their indigenous mother tongues. For instance, although tribals in Jharkhand speak non-Indo-Aryan languages such as Kharia and Ho (Austroasiatic) or Kurux and Malto (Dravidian), they have been observed to claim Hindi as their mother tongue (Abbi et al. 1998–2000; Abbi 2009). Between the competing real and claimed mother tongues arises the CH, which has traces of both. The grammatical structure of this Hindi is closer to the SH, as the speakers are exposed to this variety in schools and other official domains. Print and visual media also add to the emergence of a ‘near standard’ variety of the CH. It was observed in Jharkhand (Abbi 1997; Abbi 2009: 306) that members of the young generation are happy to forget their mother tongues and that their elders prefer to teach their children Hindi rather than their indigenous languages (Santali, a Munda language, is an exception).

The situation in the non-Hindi zone is different. The need to communicate to other non-Hindi speakers as well as to Hindi speakers outside and within the states is the motivating factor for the local population to learn and use Hindi; there is neither any serious attempt to speak a near-standard variety nor do the speakers have much exposure to the SH. The result is the emergence of a variety that is not close to the standard but has acceptability and possibly also prestige among the users. Hindi as the contact language has emerged in these states as a compulsive alternative to communicate to a larger population base. Moreover, Hindi is being seen as the instrument of social cohesion and the best reason for mitigating differences among locals and outsiders.

We have taken the area of the Northeast, especially Arunachal Pradesh and Meghalaya, for this discussion. Once known as the North-East Frontier Agency (NEFA), Arunachal Pradesh became a full-fledged state in 1987. It is the largest

among what are grouped as the seven North-eastern states or ‘the seven sisters’. Meghalaya became an autonomous region of Assam in 1970, and a separate state in 1972. The major languages spoken are Khasi, Garo, Jaintia and English.

Although government organisations such as the Central Hindi Directorate (CHD), National Council of Educational Research and Training (NCERT) and Kendriya Hindi Sansthan (KHS), and autonomous institutes such as the Vivekananda Kendra and Ramakrishna Mission have been very active in imparting knowledge of Hindi, it is the interaction with outsiders, traders, entrepreneurs, army personnel, medical professionals, and officials who have come to reside in the Northeast that has helped in forming the CH. Thus, social change and demographic restructuring in Northeast India in the mid-seventies led to the development of the CH. This reminds us of Thomason and Kaufman, who opine that “it is the sociolinguistic history of the speakers, and not the structure of their language, that is the primary determinant of the linguistic outcome of language contact” (Thomason & Kaufman 1988: 35). Aikhenvald & Dixon (2007: 47) argue against this and hold the view that:

“...[t]ypologically different linguistic structures tend to change in different ways. For each category, in order to answer the question of how diffusible it is we need to know its function, expression, and status within the language.”

The current research shows that contact-induced changes owe their existence to the interplay of the model language and the recipient language. There are some well-defined features of the model language which are more vulnerable to change and thus, these features change in any sociolinguistic situation. Then, there are typologically identifiable grammatical features of the recipient language which influence the contact language in a way that the resultant contact language assumes a distinct identity.

Linguistic structures of the Northeast Hindi have emerged in contact with typologically and genealogically distinct languages of the regions. We shall be discussing the structure of this Hindi after we discuss the structure of Hindi in the Hindi belt. As said earlier, this is the Hindi of contact. Before we begin our discussion, readers may be interested in knowing how we define the SH which is going to serve as the take-off point for the discussion.

We would like to join Yamuna Kachru in defining the concept of SH, the official language based on Western Hindi, as follows:

Hindi, the official language, is a standardized form of language that is also referred to as Modern Standard Hindi, to distinguish it from the colloquial and regional varieties of the language (Kachru 2006: 1).

It is the codified variety, primarily based on a vernacular known as Khari Boli (*kharī bolī*), which is used for official purposes and has several registers. We have taken its spoken form for the sake of comparison and point of departure for analysing Arunachal and Meghalaya Hindi.

Following is the discussion on restructured Hindi varieties in the Hindi belt, such as being used in the states of Jharkhand and Bihar as well as in the Union Territory of the Andaman and Nicobar. This is followed by the discussion on Hindi of non-Hindi zones, such as being used in the states of Meghalaya and Arunachal Pradesh in the Northeast.

4 Language in contact: Hindi in Hindi-belt, including the Andaman Islands

4.1 Contact Hindi of Bihar (CHB)

The constant process of industrialization and modernization of Bihar since the early 20th century accompanied by the introduction of educational institutions of European type and by modern communication apparatus, such as press and railways, encouraged immigration from Uttar Pradesh and other regions of Bihar into Patna, which in 1936 became the capital of the newly formed province. An amalgam of Awadhi, Chattisgarhi, Hindustani, Bhojpuri, Magahi and Maithili gave rise to a variety in the big cities of Patna and Jamshedpur. This version of the language went through the process of acculturation due to contact with Urdu first and then with SH in domains of education, governance and media. The outcome was a version of Hindi which is closer to SH but has identifiable grammatical features borrowed from contact languages.

The undivided Bihar (before the formation of Jharkhand in 2000) saw a whole century of Hindi developing as the language of education, judiciary and administration. Hindi speaking intelligentsia and upper-caste Hindus of Bihar, independent of whether Hindi was their mother tongue or not, supported the Hindi movement since 1880, coupled with the British government support. After independence, Hindi became the State Official Language of Bihar and language of instruction in all state-owned and funded schools except minority schools like Convents, *Madrasa*, *Maktab* or schools owned by any linguistic minority group. This entails that speakers of Bihar and Jharkhand had enough exposure to SH in schools, unlike the other communities of the Northeast as we shall see later. Contact Hindi, more often than not, operates in a diglossic relation with SH. Speedy developmental programmes in urban areas motivate speakers to learn SH to use in formal settings. However, Hindi in an informal setting by educated speakers is loaded with the variety of the CHB and contact Hindi of Jharkhand (CHJ). We shall consider some of the defining features of the CHB and the CHJ in the following sections.

4.1.1 Salient features of Contact Hindi of Bihar

Some of the features which keep the CHB apart from the SH are presented below using examples from Hashami (2015). It is to be noted that most of the contact Hindi omit the difference between first singular and first plural and use *həm* ‘we’ which is used for both singular and plural but refers to singular first-person addressor. This is generally considered a polite form. It is also used to give honorificity to the addressee. In the following article, I have thus, interlinearized *həm* as both singular and plural forms.

1. Ergativity is missing:

- (1) *həm* *us* *ko* *mare*
 1SG.HON/PL 3SG.OBL ACC beat.AOR.SG.HON
 ‘I beat him/her.’

The use of the first-person plural form for politeness as well as showing honour to the addressee is a very old use in many standard varieties of Hindi-Urdu. It is commonly observed in the speech of Hindi speakers of Allahabad and Lucknow.

2. Grammatical gender is not marked. However, grammatical generalization of semantic gender is visible in some varieties. Consider (2) and (3) respectively.

- (2) *na fugar* *he* *kəmzori* *ho-t-a* *he*
 No sugar COP.PRS.SG weakness.F.SG be-IPFV-M.SG AUX.PRS.3SG
 ‘No, I have diabetes, I feel weakness.’

- (3) *caj mē didi* *kja* *ḡal-t-i* *he* *bātao na*
 tea LOC elder.sister what put-IPFV-F be.3HON tell TAG
 ‘What does the elder sister put in tea? Please tell.’

Thus, biologically determined feminine noun governs gender agreement but grammatically marked feminine nouns have default agreement of 3rd masculine singular.

3. Nouns with postpositional case markers do not get inflected for oblique case (as *ləṛka* unlike SH *ləṛke*), nor is the copula ever coded for plurality. Noun ‘sisters’ *bəhən* also does not take plural marking.

- (4) *ləṛka ka* *car bʰai* *or* *do* *bəhən he*
 boy GEN four brother and two sister COP.PRS.3SG
 ‘The boy has four brothers and two sisters.’

4. The formation of the plural is by compounding. Thus, *log* ‘people’ [+human] is attached to plural animate nouns. With inanimate nouns and [-human] plurality is not marked. However, *səb* ‘all’, *sara* ‘all’, *kul* ‘all’ are used to form plurals. The suffix /-ən/ is attached with both animate and inanimate nouns to form plurals in some varieties. The finite verb again fails to agree in gender and number.

(5) *səb ləɾka log kʰela*
 all boy people.PL play.AOR.M.SG
 ‘All the boys played.’

(6) *səb ləɾk-ən kʰela*
 all boys-PL play.AOR.M.SG
 ‘All boys played.’

5. Specifier suffixes /-wa/ and /-ja/ are attached to nouns. These have a functional role of definiteness although Grierson (1903) considered them redundant.

(7) *saiɾl-ja cen-wa se bandʰ dijje*
 cycle-SPEC chain-SPEC INST fasten give.IMP.HON
 ‘Fasten up the cycle with the chain.’

6. Numeral classifiers *go* and *ʔʰo* are placed between the numerical adjective and the noun and are interchangeable, i. e. both can be used either with animate or with inanimate nouns.

(8) *həmɾ-i beɟi ke car go beɟa aur du ʔʰo*
 1SG/PL.POSS-F daughter GEN.PL four CLF son and two CLF
beɟi he ɔɾ du ʔʰo gaɾi o he
 daughter COP.PRS.3SG and two CLF cars also COP.PRS.3SG
 ‘My daughter has four sons and two daughters as well as two cars.’

7. The suffix /-e/ is used as a locative marker in the same meaning as the SH postposition *mē* ‘in’. Thus, *gʰəɾ mē* ‘in the house’ of the SH is rendered as *gʰəɾe* in the CHB.

8. *-ke* is used as a genitive postposition along with *-ka* and *-ki*. In case it is *ke* it signifies the neutralization of gender agreement. It is noticed that in many regions of Bihar genitive is dropped altogether.

9. The first-person plural form *həm* of the SH is used for both singular and plural as we saw in the sentence (1) above. Since the verb takes plural ending, as in

(1) above, it shows the agreement with the plural subject but semantically this pronoun stands for the singular subject.

10. We noticed a very productive process of verbal derivation in the CHB. From several nouns and adjectives, verbs are derived with the help of suffixes *-a* and *-ja* or *-ija*. For instance, *bat* ‘talk’ is derived as *bətijana* ‘to talk’ or *gəɾəm* ‘hot’ > *gəɾmana* ‘to heat’ etc. It has been noticed that both these verbs are also incorporated now in many standard Hindi dictionaries.
11. The suffix *-je* is used as an exclusive emphasis marker (the equivalent of the SH *hi*), and *-o* is used as an inclusive emphasis marker (the equivalent of the SH *b^{hi}*). Consider:

(9) *həm to itna səsta mē dət-je deŋge*
 1SG/PL EMPH this.much cheap in give-EMPH give.FUT.PL
 ‘I will give it (sell it) this cheap.’

12. The verb in the past tense in the speech of upper-class Muslims in Patna district is inflected with *-is* for 3rd singular non-honorific and *-in* for 3rd singular honorific. Consider:

(10) *u kəh-is tha ki car bəje ajega*
 3SG say-PST.3NHON be.PST.M.SG CONJ four o'clock come.FUT.3SG.M
 ‘He said that he would come at four o’clock.’

(11) *u-(log) k^hana k^ha lih-in*
 3-(PL) food eat take-PST.3HON
 ‘They have taken their food.’

Although a characteristic feature of the Muslims speech, one notices Hindus of elite class also using the honorific form in Patna district. Grierson also acknowledges this trait of Muslims and ascribes it to Awadhi influence (cf. Hashami 2015); he interprets *-in* as a marker of politeness. The youth of Patna no longer use these structures, as these are considered to be the speech of less-educated people.

13. Consonant clusters are strictly not allowed word-initially or finally in the contact variety.
14. The replacer phoneme, i. e. the sound that replaces the initial sound of the base word (Abbi 1991) in echo formations, is *u* unlike *v* in the SH. Thus, the SH *ro-vo* ‘crying etc’. corresponds to *ro-u* in the CHB.

4.2 Contact Hindi in Jharkhand (CHJ)

The state of Jharkhand was carved out of Southern Bihar in 2000. According to the census 2011, the population of Scheduled Tribes (ST) 8,645,042 constitutes 26.2 per cent of the total population of 32,988,134 of the state. The census lists 32 tribes in the state. The bilingual competence of tribes of Jharkhand is at a higher level at the scale of bilingual proficiency than its dominant majority communities (Abbi 1995; Abbi 1997), Hindi being the most prominent second or third language. The tribal population considers Hindi as the language of prestige and honour (Abbi et al. 1998–2000) and tries to reach the level of SH in spoken and written forms. The high bilingual proficiency has changed the grammar of the indigenous tribal languages as they have converged towards Hindi as well as other Indo-Aryan languages (Abbi 1995; Abbi 1997). When we talk of contact Hindi of Jharkhand (CHJ) we take into account the speech of both tribes and non-tribes.

Many factors are responsible for the emergence of Hindi as a contact language of Jharkhand. Among them are: (1) immigration of disparate linguistic groups from the neighbouring states in search of employment in the fast-developing Jharkhand as the industrialized state, especially in Jamshedpur, Dhanbad and Sindri; (2) immigration for higher education in Ranchi and Jamshedpur, as these two cities house schools and colleges considered among the best in the state by the locals; (3) spread of Christianity among varied tribes speaking different and distinct languages motivating religious congregation to take place in one common and easily accepted language such as Hindi and publishing religious material in SH; (4) inter-community marriages between tribes and non-tribes as well as between different non-tribal communities; (5) Hindi is the state's official language used in education, for judicial and administrative purposes. The most important factor which has been responsible for the acceptance of Hindi as the dominant contact language is the notion of dignity and honour attached to the language, especially among the tribes (Abbi et al. 1998–2000).

4.2.1 Salient features of Contact Hindi of Jharkhand (CHJ)

The CHJ shares most of the features with the CHB. Features like no grammatical gender, plural formation by compounding with *log*, default agreement with third masculine singular, absence of oblique case marking on nouns are some of these features. For detail refer to Hashami (2015: 304). We discuss below some of the significant shared and non-shared features of the CHJ.

1. Semantic gender is not recognized grammatically, and thus all nouns are treated as masculine.

- (12) *mer-a* *bēhen jāhā rēh-ta* *he*
 1SG.POSS-M.SG sister here live-IPFV.M.SG AUX.PRS.3SG
 ‘My sister lives here.’

2. The plural is formed by attaching *log* ‘people’ to animate nouns. Some speakers frequently use *jan* ‘person’, ‘people’ to pluralize human nouns. This feature is typical of the CHJ.
3. Oblique case form of 1st and 2nd person possessive pronouns is commonly used with postpositions *ko*, *se*, *tāk*, *pār*, i. e. forming *mere=ko* ‘to me’, *tere=ko* ‘to you’ instead of the SH *mujhe* and *tujhe*, or *mere=se* ‘from/by me’, *mere=pār* ‘on me’, *tere=tāk* ‘up to you (spatial sense)’ instead of SH *mujh=se*, *mujh=pār*; *tujh=tāk* respectively. These constructions are very common in cosmopolitan Hindi as spoken in Delhi and Mumbai also. However, it is noticed that the Jharkhandis also use constructions not different from SH, like *hām=ko*, *tum=ko* and *us=ko* ‘to us/you/him (or her)’ respectively, and the plural forms *hām log ko*, *tum log ko*, *ap log ko*, and *vo log ko* etc.
4. Unlike the CHB or the SH, modal *sāk* ‘can’, ‘be able to’ is used as an independent verb in the CHJ. This is a very common feature in most of the contact Hindis of the country and around the world, including the Andamanese Hindi, Fiji Hindi (Siegel 1988), Kolkata Hindi, Arunachalese Hindi and Meghalaya Hindi (see below).

- (13) *jitna sākega* *dām* *lāga-ke* *dek^h le*,
 as.much be.able.FUT.3SG.M strength apply-CVB see take.IMP
tum mere *se sāk* *hi* *nāhi payega*
 you 1SG.POSS.OBL from be.able EMPH no get.FUT.3SG.M
 Literal: As much (your) ability (is) strength apply, you from me won’t
 be able to sustain.
 ‘Put as much strength as you wish, you won’t be able to outplay me.’

5. When modal verbs *sāk*- ‘can’, ‘be able to’, *mañ*-‘to demand’ or *cah*- ‘want’ are used as the second constituent of a complex verb, the main verb takes the infinitive form ending in *-e*. For instance, *jane sākta* ‘can go’; *k^hane mañta* ‘wants to eat’ and *bolne cahta* ‘wants to speak’ etc.
6. The numeral classifier from regional languages *gotta* or *go*, which means ‘one’, is used profusely with nouns to mean both ‘one’ and ‘whole’. For instance, *gotta biskuṭ* ‘one biscuit’. It can also mean ‘one whole unbroken’, as in *car t^ho gotta mirca* ‘four CL whole chillies’. The numeral classifier *t^ho* is

borrowed from the local languages such as Sadari and *gotta* is a borrowing from the Munda languages spoken in the area.

- (14) (*həm*) *gotta-gotta saman le aje*
 (1SG/PL) one-one things take come.AOR.SG/PL
 ‘(I) brought each and everything.’

7. We don’t find the presence of *-in* for honorability in the past tense of the verb in Muslim speech, unlike in the CHB, because the Awadhi influence is absent in Jharkhand.

Features 4, 5, 6, and 7 are the features of the CHJ distinguishing it from the CHB. The two varieties also differ in lexicon marginally. For instance, ‘cloth’ is *luga* in the CHJ but *kəpɾa* in the CHB (*kəpɾe* in the SH); the word for ‘female labourer’ in the SH is *məzdurni*, but in the CHJ it is *reja* and in the CHB it is *məzdur ɔɾət* ‘working woman’; the word for ‘hen’ in the SH and the CHB is *murgi* but in the CHJ it is *caŋgni* (borrowed from Kurux). Many lexical items used in the CHJ are borrowed from Bangla and Oriya—the languages that are spoken across the border.

4.3 Contact Hindi of the Andamans (CHA)

The Andaman Islands are comprised of a cluster of approximately 550 islands, rocks and rocky outcrop running from north to south and located southeast of the Indian subcontinent in the Bay of Bengal. They are separated from the Malay Peninsula by the Andaman Sea, an extension of the Bay of Bengal, and are part of the Union Territory of the Andaman and Nicobar Islands belonging to India. Geographically, the Andaman Islands are closer to Myanmar and Indonesia than to mainland India. However, contact between the Andamanese and the populations of the neighbouring countries has not been established to date. The capital city of the Andaman Islands is Port Blair, situated in the south of the Islands at a distance of 1255 km from Kolkata and 1190 km from Chennai.

Although the first settlement in the Andaman Islands took place in September 1789 by Lieutenant (later Captain) Archibald Blair, the high death rate among the settlers discouraged any activity in the Islands for the next sixty-two years. Six decades later the first war of independence forced the British rulers of India in 1857 to reconsider the Andaman Islands for the establishment of a penal settlement to deal effectively with those who revolted against their paramount power. It was on 22nd January 1858 that Union Jack was hoisted by Capt. H. Man, the Executive Engineer, was deputed to take formal possession of the islands. The Andamans

were eventually known as the ‘Black Water Islands’ because of the dreaded Black Water Fever or malaria that was rampant there. However, these sets of islands have become popular as tourist destinations both for Indians and foreigners currently, especially after the tsunami in 2004.

The history of the CHA is associated with the establishment of the penal colony in 1868 by the British rulers to keep the freedom fighters in captive. When British generals came to the Islands they brought along Indian officials who spoke Hindi and its major varieties. After independence in 1947 the government brought people from Bihar (now Jharkhand) and other areas for undertaking development programs in the Island. After the independence of Bangladesh in 1971, the Indian government encouraged and motivated Bengali refugees to settle down in the Islands. It also motivated members of other communities residing in the south of India and those employed by the Central government to migrate to the islands. It is not surprising to find several hamlets of Malayali, Tamil, Telugu, and Kannada speakers in Port Blair. To attract Indians from the mainland the government distributed free land in the islands—both in Great Andaman and Little Andaman, with the result that today the Andaman Islands are considered ‘mini India’ with approximately 20 Indian languages spoken as mother tongues. Hindi is revered as the link language among the local population. One can find people from every state in the Andamans who speak their native language at home but use Hindi as soon as they step out of their homes. A constant stream of visitors to the island has increased the worth of the knowledge of Hindi. It will not be an exaggeration to put it in writing that there is not a single person in the Andaman Islands (barring members of the Jarawa community in the Great Andaman and Onge community in Little Andaman) who does not speak Hindi. Hindi is not only the state’s official language; it is used in almost all domains outside the home. Great Andamanese tribes and migrants from the Hindi belt use the contact Hindi as the home language too. This is also true of the tribes brought from Bihar (before the formation of Jharkhand). They are known as ‘Ranchi’ and live in the northern part of the Great Andaman Islands. The data discussed here is the outcome of my fieldwork in the Islands in 2001 and 2005–2009.

Two interesting observations, although not very unexpected, came to light. Firstly, there were several varieties of Hindi, each marked by the influence of the local mother tongue of the speaker concerned and, secondly, Hindi used by the Great Andamanese tribes, on the contrary, did not have much influence on their indigenous language/s except in sound system. The latter is a clear indication that they had lost their competence in their respective mother tongues long ago. Since there are so many varieties of Hindi in Andamans, each marked by its speakers’ mother tongue (Telugu Hindi, Malayali Hindi, Bengali Hindi, Punjabi Hindi, Tamil Hindi etc.), we shall divide them into two major groups, i. e. Hindi used by the Great Andamanese tribes and Hindi used by others, the settlers.

4.3.1 Contact Hindi of Great Andamanese (CHGA)

As of today,¹ the population base of this tribe is rather small, as there are 74 people in the Great Andamanese community, mainly children below the age of 12. However, only four of the adult population have some competency in their native language. All converse with each other in Andamanese Hindi. Since Hindi is the state's official language, those children who go to school are exposed to SH but others are deprived of this facility. Great Andamanese Hindi is marked by few lexical items drawn from the indigenous languages but the structure of grammar is based on Hindi which the tribes hear from the locals and different migrants to the Island. For instance, *refe karo* 'have your meal' where *kər-* 'to do' is from Hindi and *refe* 'food' is a word from Jeru, the Great Andamanese language, or code-mixed Great Andamanese language with Hindi syntax. The exposure to the SH is minimal or absent, as education in a formal atmosphere was unheard of till very late. When we collected the data, most of the Andamanese children were at home. Some did go to primary school. What I give in the following pages is a sample of adult speech. It is to be noted that Great Andamanese do not use ə and substitute it in all Hindi words containing ə with a. Thus, the opposition of the SH between ə and a is neutralized in the speech of most of the Great Andamanese speakers. Exceptions are the children who go to schools in Port Blair and learn Hindi in a classroom situation. Consider:

The indigenous words in the examples are in bold.

- (15) *tele nahi kija boje me*
Attend not did marriage LOC
'did not go the marriage'
- (16) *refe kija andar me*
Food did inside in
'Had food inside (the house).'
- (17) *kjun aka-meme kuc^h nahi bol-a*
Why her-mother something not speak-AOR.M.SG.
'Why didn't her mother speak [to her]?'

1 The population figure is supplied by the Andaman Adi Janjati Vikas Samiti (AAJVS), Port Blair, Andaman and Nicobar and the competence level in the Great Andamanese language is assessed by the author who had been in constant touch with the tribe for the last 18 years.

Hindi spoken by the Great Andamanese is very close to the Bengali Hindi (such as *sək-* ‘to be able to’ used as an independent verb and the use of the verb *maŋ-* with infinitives as in *jaŋe maŋta* ‘wants to go’), perhaps because of their exposure to the Bengali language of Indian officials (baboo) who came to the island after the 1960s. As said earlier, a prominent omission in the Great Andamanese Hindi is that of mid vowel ə (symbolized in Devanagari as अ), words like Hindi *bənao* as *banao* ‘make-IMP’, and of retroflex ʒ, which is rendered as alveolar r (examples 18 and 19 below). The other interesting feature which is shared across the island is the formation of verbal attributive phrase having the passive meaning with the specifier *wala* (19). Consider:

(18) *sa:f banao*
 clear make.IMP
 ‘Make it clear.’

(19) *bara sab de dija: wala kʰajega*
 Big sahib give give.PFV SPEC eat.FUT.3SG.M
 ‘I will eat the one which was given to me by the big boss.’

Another very common feature of the Andamanese Hindi is to use several types of replacer sound for echo formations. Thus, vowel alternation as in *kʰa-kʰu* ‘eat etc.’ or dropping of the initial consonant with vowel alternation as in *pani-uni* ‘water etc.’ and replacing initial consonant with /v-/ as in *kam-vam* ‘work etc.’ are used in free variation.

An interesting aspect is also that speakers use the future tense in narration as against the SH past or present imperfect. See a sample of narration given below. The speaker, Nao Jr. aged 55 years is narrating how they light a fire. The future tense may be used to portray repeatedly occurring situations as imagined in the future as well as for the present. In Russian, it can be the same. The following transcript is in the Devanagari script so that it can be compared with the one in the IPA that follows it immediately.

(Narration in Andamanese Hindi)

पहले एक बंबू को पूरा लम्बा से छील देगा,
 पूरा अच्छा से मोटा करके उसमे जरा से एक छेद बना देगा,
 फिर पतला सा और बम्बू बना देगा।
 बना के नीचे को और बेंत का जो सड़ा किस्म का सूखा वाला, अच्छा वाला उसको नीचे
 जमा कर देगा फिर, उसको पूरा जोर से रगड़ेगा।

(Narration IPA)

*pehle ek bambu ko pura lamba se chil dega,
pura accha se moṭa karke usme jara se ek ched bana dega,
phir patla sa or bambu bana dega.
bana ke nice ko bent ka jo saḍa kism ka sukha wala, accha wala usko nice
jama kar dega phir, usko pura jor se ragrega.*

English Translation

First [I] will peel a bamboo lengthwise,
then flatten it well and then make a hole in it.
Then [I] will make another thin bamboo.
Having done [I] will collect the cane, rotten type dried one, good one
and put it in a heap.
Then, [I] will rub it vehemently.

4.3.2 Hindi used by the Andamanese settlers (CHAS)

As said earlier, Andaman is considered a ‘mini India’ as it is a multi-religious, multi-ethnic, and multilinguistic region. People from mainland India had been coming to the Islands to buy land and make the region their homes. Some were sent by the Government of India to run administrative offices and they made this Island their home. A large chunk of the population is Bangladeshi refugees and people from the South of India who were encouraged to migrate to become self-sufficient. A very recent phenomenon is that now the Andamans are considered one of the prized tourist places, especially after the tsunami in December 2004. One sees a variety of families from all over India visiting the place. All this influx has affected the contact Hindi used for inter-ethnic communication. The official languages are Hindi and English. The literacy rate is very high, as all children go to school and those employed by the government are exposed to the SH. To make a distinction between Hindi used by the Great Andamanese (CHGA) and that of the local population called “settlers” we will term the latter as the CHAS (Contact Hindi of Andamanese settlers) in the following discussion. Readers will find many similarities between the CHAS and the other contact Hindis we witnessed earlier.

Lack of agreement between adjective and noun (23), absence of ergativity as in *tom roṭi kʰaja* ‘you ate a chapati’, default agreement of 1st and 3rd person masculine singular (20, 21, 23) obliterating the difference of pronominal agreement (20), absence of progressive aspect (20, 24), absence of oblique marking on nouns with postpositions (22), use of the verb ‘to say’ as a quotative or complementizer (20, 25) are the typical features of this Hindi. Interestingly, this use seems to be present in most of the settlers who came from the south of India as it is a feature of the Dakhini

Hindi (see Vashini Sharma’s article in this volume), use of modal *sək* as an independent verb as in *məcc^hi marna cahta hē pər həm nəhi səkega* ‘[I] want to catch fish but I will not be able to’ (as in Jharkhand variety) are the salient features of the CHAS. A very common phenomenon of the CHAS is the use of construction with a conjunctive participle in place of the perfective participle in the SH (21). This is a characteristic marker of those who migrated from the South of India. Most often than not, the present imperfect tense is used for the future (22). The striking feature of both the CHGA and the CHAS is the lack of representation of ‘I’ as *mē*. Instead, first-person plural *həm* invariably marks the Andamanese Hindi. Consider:

(20) *heli nəhi jata tarasa bol-ke həm idər kam kərta*
 helicopter NEG go.PRS.M.SG Teressa say-CVB 1SG here work do.PRS.M.SG
 ‘Since no helicopter is going to Teressa Island I will be working here.’

(21) *həm jəha a-ke əbi ə^hra sal ho gija*
 1SG here come-CVB now eighteen years be go.AOR.M.SG
 ‘It has been eighteen years since I came to this place.’

(22) *sab ka g^hər mē həm ata*
 officer GEN house LOC 1SG come.PRS.M.SG
 ‘I will come to the house of the officer.’

(23) *u wala kitab mē həm sara əkfər pəhcanta*
 DEM.SG SPEC book LOC 1SG all letters recognize.PRS.M.SG
 ‘[I] recognize all the letters in that book.’

(24) *əre mama tum kja sun-ta həm kja kəh-ta*
 Oh uncle 2SG what hear-PRS.M.SG 1SG what say-PRS.M.SG
 ‘Uncle, what are you hearing and what I am saying.’

(25) *wo həm ko bola kər-ke həm aja*
 3SG 1SG DAT say do-CVB 1SG come.AOR.M.SG
 ‘[Because] he asked me to come so I came.’

(26) *məqəm ap jati*
 Madam 2SG/PL.HON go.PRS.F
 ‘Madam, will you go?’

(27) *kja səbzi dena ap ko*
 What vegetable give 2HON DAT
 ‘Which vegetable shall I give you?’

The SH verb *cah*-‘to want’ is replaced in the same meaning by the word for ‘to demand’ in many contact Hindis of the non-Hindi belt, as in Kolkata Hindi or Bombay Hindi. The following example is very much like that.

- (28) *həm ko job kərna nəhi maᅅta*
 1SG ACC job do NEG demand.PRS.M.SG
 ‘I don’t want to do any job.’

Another interesting fact is that Hindi numerals for numbers over twenty are retained only by those whose home language is Hindi or its varieties. Others use numerals in an additive fashion. Consider:

- (29) *is=ka əssi-panc de do*
 3SG.OBL=GEN eighty-five give give.IMP
 ‘Give eighty-five for it.’

4.4 Similarity of the CHA with the CHB and the CHJ

The similarity between these varieties is obvious. To summarize the discussion so far, we can list the following grammatical features of the CHB, the CHJ and the CHA which appear to be shared:

1. Absence of the grammatical gender;
2. Use of the word *log* ‘people’ as a plural marker attached to nouns;
3. Absence of oblique case marking on nouns before taking any postposition;
4. The use of modal *sək* as an independent verb;
5. Partial similarity:
 - 5a) Gender agreement with human nouns is seen in the CHAS (26), as in the CHB;
 - 5b) This variety does not mark pronominal subject agreement in number, as all sentences have finite verb marked by the singular.

Let us move now to the states which do not fall in the ‘Hindi-belt’—the states where Hindi is not the first official language. We shall consider Hindi used in the Northeast of India.

5 Hindi as the Contact Language of the Northeast

In some states of the ethnically and linguistically heterogeneous Northeast, Hindi is not merely a lingua franca but fulfils all the requirements of a major language of the community. It enjoys prestige and power, as many consider Hindi, besides English, a ladder to promotion and development in society. As one of the two official languages of the whole Indian union, Hindi is a language of priorities and is conceived as being “modern”. In the following pages, I would like to discuss the structure of this highly acceptable language of wider communication in the linguistic situation of the Northeast. The results are based on the fieldwork conducted in Arunachal Pradesh and Meghalaya (Sharma 2011–2012, 2013 under my guidance), two states of Northeast India. I have also drawn the results for these states from the research conducted by Abbi, Gupta and Gargesh (1998–2000) at the pan-India level.

As said earlier in the Introduction, the emergence of contact Hindi in the Northeast regions is need-based and exists in an environment that is marked by varying different languages drawn from the Tibeto-Burman and Mon-Khmer language families. Moreover, the constant inflow of visitors from the rest of India, mainly speakers of Bangla and varieties of Hindi—as listed in the Appendix—throws open a situation of language ecology where the SH is not a model. The need to communicate with the immigrants as well as those Hindi speakers who are employed in various government services gives rise to a variety of Hindi which is far away from the standard version. Interestingly, this variety enjoys acceptability and possibly also prestige among the users.

It would be interesting to see how Hindi in contact with genetically related Indo-Aryan languages—the so-called varieties of Hindi that we discussed earlier in the paper—and the Hindi which is spoken in an alien environment restructure the grammar of Hindi. Where do the two Hindis differ or where they are similar will be interesting to observe. Let us begin with Arunachal Pradesh, the Northeast state which touches international borders with China, Myanmar and Bhutan.

5.1 Contact Hindi used in Arunachal Pradesh (CHAP)

Arunachal Pradesh became a union territory in 1972 and a state in 1987. It is the largest among the seven North-Eastern states referred to as “the seven sisters”. This state is home to twenty-six tribes, each speaking a different language. The state is marked by mutually non-intelligible heterogeneous languages. The tribes are spread across 3,649 villages, most of which are in remote valleys and hilltops along the international borders with China, Myanmar and Bhutan. All tribes speak

mutually unintelligible distinct languages. A surge of officials from all over India came to the state communicating with locals in spoken Hindi—not necessarily of the SH variety in the post-1987 era.

The geographical inaccessibility had given rise to a lack of mutual interaction and thus, distinct languages. Fortunately, with the introduction of roads, education, and telecommunication in the last twenty-five years there has been a change in the region. People have become mobile and are willing to communicate with other tribes. English was introduced as the official language of the territory, while the Hindi variety described here serves as one of the common languages of communication along with Assamese, the language of the neighbouring state. Indigenous languages represented in Arunachal Pradesh belong to the Tibeto-Burman, Tai-Khamti and the Indo-Aryan language families. The CHAP has emerged in such a plurilingual and genetically diverse language ecology.

Examples have been drawn from the areas such as Pasighat, Naharlugun (Modi 2005), Zero as well as from Itanagar, the capital city of Arunachal Pradesh (Sharma 2013). All these areas are also home to non-indigenous languages of migrants and old-time settlers, such as Assamese, Bangla and Hindi. It was observed that Hindi is freely used by students and teachers in the schools in these regions. In general, Hindi as the contact language is used all over the state and considered to be prestigious as reported by 67% of members of the community (Modi 2005). People often aired the view that they would be happy to see Hindi as one of the official languages of the state. Hindi has served as the biggest equalizer in society and helped in mitigating differences in society (Abbi & Sharma 2014). Hindi is used between different tribes of Arunachal Pradesh, between non-tribes and tribes of Arunachal Pradesh and between the non-tribal population of Arunachal Pradesh. The domains of use are increasing every day from educational institutes to government offices, from market to social and religious functions, and from media to telecommunication. The exposure to SH has been minimal so far, as the introduction to Hindi in schools is not a very old phenomenon. The other interlocutors who speak Hindi are migrants with the base of Bangla, Assamese and some variety of Hindi from the Hindi-belt. Hence, Arunachalese are exposed to varieties of Hindi to draw their resources in lexicon and structure. Thus, CHAP has emerged after restructuring the base/source languages along with some inputs from their indigenous languages.

The most significant social factor that plays a role in mitigating differences among disparate tribes is the religious congregation that normally consists of three or four different linguistic groups, but uses the common link language Hindi, i. e. the CHAP (Modi 2005: 35). It seems that the use and wider acceptance of the CHAP have increased inter-tribe marriages in the state. As a lingua franca, it carries the maximum functional load among disparate and distinct tribal communities. Burling (2007: 223) also reports the spread of Hindi as a lingua franca in

Arunachal Pradesh. For the uneducated, this is the only language available for inter-group communication since the rapid demise of Nefamese. For the educated, this is the only language besides English that ensures upward movement in society. The structure of the CHAP is far removed from the SH, yet it is respected and accepted by all. Some of the salient features of this Hindi that are illustrated below are those which are shared across many contact Hindis including that of Meghalaya Hindi.

5.2 Contact Hindi Used in Meghalaya (CHM)

Meghalaya became an autonomous region of Assam in 1970, and a separate state in 1972. Consisting of the Khasi, Jaintia and Garo Hills, it occupies the geographical area of the Meghalaya Plateau or Shillong Plateau (Bhatt & Bhargava 2006: 15). The major languages spoken are Khasi (a language of the Austroasiatic language family), Garo (a Tibeto-Burman language), Jaintia (an Austroasiatic language) and English—Khasi and Garo being the principal languages and English the official language. The capital of Meghalaya, which has seven districts, is Shillong; the major part of the fieldwork for this research was undertaken there.

Meghalaya has people from various parts of the country and neighbouring countries, and specifically, Shillong is like a cosmopolitan city in character. The non-indigenous population is present in the form of administrative officers, teachers, traders, vendors and migrants from various regions. As they occupy public spaces where interaction with the people of Meghalaya is unavoidable, the need for a link language has developed, leading to the evolvement of the local CHM. Unlike the CHAP, the CHM is not used among indigenous peoples, but with people who have come from outside the state; the heterogeneity of languages that we witnessed in the case of Arunachal Pradesh is missing here. Khasi is the dominant language, and thus learning Khasi and, for educated people, also English is imperative.

Meghalaya Hindi draws its inputs, especially in its lexicon, from Bengali, Assamese, Nepali, Bhojpuri, Maithili, Marwari, English and, of course, Hindi. There is a formal and informal register. The former is used in classrooms and government offices, whereas the latter is being used by people in general in public places, with friends, and in audio and visual media such as radio and television. There are traders, daily wage labourers, cab drivers, hawkers, shopkeepers, etc. from various parts of India, such as the states of the Northeast, West Bengal, Assam, Bihar, Rajasthan, Uttar Pradesh, etc., and from neighbouring countries, viz. Bangladesh and Nepal, who have only one link language among themselves—and that is a non-standard version of Hindi. Another domain of the informal register is inter-regional marriages, in which spouses may be from

anywhere in the country. This is a recent phenomenon as it has increased since the turn of the century. These migrants take no time in learning Meghalaya Hindi. However, the possibility of better competence in Hindi in the migrants' community is not ruled out. Since inter-regional marriages are a recent phenomenon, we have not been able to research the language of the children born in such marriages.

The combination of the informal and the formal registers gives rise to the structure of Meghalaya Hindi, which is locally referred to as 'Bazaar Hindi'. This term has become very popular in Shillong and across Meghalaya. Since Hindi was formally introduced in schools and colleges only after the formation of the state in 1972, the population has been exposed to SH only relatively recently. For details refer to Sharma (2013) and Abbi & Sharma (2014).

Rather than discussing the individual structures of the CHAP and the CHM, I will first discuss the shared features of these two varieties and then move on to highlighting the features which are not shared between the two.

5.3 Common salient features of Arunachalese Hindi and Meghalaya Hindi

Although the base languages in Arunachalese Hindi and Meghalaya Hindi are different, viz. Tibeto-Burman in the former and Austroasiatic in the latter, there are some common features in the contact Hindis attracting our attention towards the shared mechanism of contact-induced changes. Some features adopted by both varieties prove that certain linguistic features are more vulnerable to absorption than others. Similarly, there are features of Hindi that are found to be redundant for the information structure, e.g. agreement pattern or oblique marking on nouns. The two factors, namely (a) vulnerability of absorption or adoption, and (b) redundancy of information structure, both in combination give rise to shared and common features of the Northeast contact Hindi. The contact Hindi used in Arunachal Pradesh and Meghalaya shares the following features which are adopted and then reanalysed in similar ways, although there is no large-scale contact between the people of the two states. Readers will be surprised to find that these features are also shared partly by the Hindis discussed earlier.

For details, readers may refer to Sharma (2014). I will give only some of the salient features here for discussion.

1. Formation of plural

To avoid the complexities of portmanteau plural inflection of the SH, which reflects gender, number and person of the arguments, the CHAP, like several

varieties of the so-called ‘Bazaar Hindi’, adopts the simple strategy of suffixing the noun *lok* (corresponding to the Hindi lexeme *log* ‘people’ but phonologically following Bengali and Assamese *lok*) to all types of nouns, whether animate or inanimate, and pronouns, e. g. *lar'ka lok* ‘boys’, *ma lok* ‘mothers’, *kitab lok* ‘books’, *juta lok* ‘shoes’, *ap lok* ‘you all’, *in lok* ‘them’, etc. Contrary to the SH, the plural argument does not govern plural marking on the verb. This brings us to the phenomenon of agreement. Consider:

(30) *gac-lok mār gəja*
tree-people die GO.AOR.M.SG
‘All the trees died.’

(31) *əpna ma-lok ko bulao*
REFL.M mother-people ACC call
‘Call your mothers.’

2. Lack of Agreement

The finite verb in Standard Hindi agrees with the subject or object noun of the sentence in gender and number, which is missing from the CHAP and the CHM.

(32) *mədəm klas me aja*
Madam class LOC came.3M.SG
‘Madam entered the class.’ (CHAP)

(33) *lārki-lok ləmba ho rəha hē*
girl-people tall.M.SG be PROG.M AUX.PRS.3SG
‘Girls are growing tall.’ (CHAP)

As can be seen, the verb neither agrees with the singular noun ‘madam’ in (32) nor with the feminine plural ‘girls’ in (33).

3. Use of attributive particle *-wala* for modification

We saw the use of this particle earlier in the CHB as a specifier. The CH in the two states makes use of this strategy productively and extends its use to replace the perfective participle of the SH *hu-a/-e/-i* ‘been’ by *wala* constructions to modify nouns. Thus the SH *māra hua kotta* ‘the dead dog/the dog which has died’ is rendered as *māra wala kotta*, consider:

(34) *sutta sutta kat-a-wala bad me frai karega*
small small cut-PFV-ATTR later in fry do.FUT.3SG.M
‘[You] will fry [the ones] cut into small pieces later.’ (CHM)

The genitive phrases of the SH are also replaced by constructions with *-wala* giving the passive meaning to the verb, e.g. *caj pine-wala dokan*, lit. ‘tea drinking shop’, i.e. ‘the shop [where] tea is being drunk’ in the CHM corresponding to *caj ki dokan* ‘tea shop’ in the SH. The verb here is used in the oblique form of the infinitive.

- (35) *sukan-e-wala kapra*
 Dry.INF_(Vtr)-OBL-ATTR cloth
 ‘dried clothes.’ (CHAP)

There are cases where this particle is used for deriving abstract nouns in the CHAP, as in *uska hase-wala* ‘his laughing’ instead of the SH *uski hāsi* ‘his/her laughter’ and *uska hat lik^hne-wala* ‘[by] his/her hand written’ instead of the SH *uske hath ki lik^hi hui*.

By extending its semantic domain across various grammatical forms the attributive particle is also used in place of perfective participial *hua* in SH modifying the following noun, with the difference that, unlike in the SH, the CHM *-wala* can be placed after a finite verb form in past, present and future tense to build a participial form modifying a nominal category. Hence, it has an attributive function here. Consider examples from the Meghalaya Hindi:

- (36) *kāl a-ja-wala kotta*
 yesterday come-PFV.M.SG-ATTR-M dog.M
 ‘The dog that came yesterday’
- (37) *k^hata-wala bācca*
 eat.IPFV.M.SG-ATTR.M child.M
 ‘The child who is eating’ Or ‘The eating child’
- (38) *kāl ajegi-wali citthi*
 Tomorrow come.FUT.3SG.F-ATTR.F.SG letter.F
 ‘The letter that will arrive tomorrow.’

The attributor particle *-wala* can be attached to any grammatical category and function as a nominalizer or a relativizer. Consider a few examples from CHM:

- (39) *dui-wala mār gāja te sākta*
 two-NOM die went.SG.M then be.able.PRS.M.SG
 ‘If the second one dies then they can do it [get married].’

- (40) *isa-wala kam hām ni kija nā*
 this-ATTR work 1SG/PL NEG DO.AOR.M.SG TAG
 ‘I have not done this kind of work, OK?’ (Sharma 2013)

In (39) the ordinal of the SH is replaced by the cardinal plus *-wala* strategy and in (40) two observations can be made. One, the *-wala* is redundant here and two, the phonetic shape of the SH *esa* is changed to *isa-*.

4. Use of modal *sək-* ‘can’, ‘to be able’ as the main verb

This feature occurs in both the CHAP and the CHM and in many other non-Standard Hindi varieties used across the country as seen earlier. It is the use of the modal *sək-* ‘to be able’ or ‘can’ as a main verb to indicate the ability or inability to do the action. Consider the following sentence as well as the (39) given above:

- (41) *traibəl-lok bahār jana pāsən nai karta jitna*
 tribal-people outside go like not do.PRS.M.SG as.much
sakta səkega
 can.PRS.M.SG. can.FUT.3SG.M
 ‘Tribal people do not prefer to go out. They do as much as they can.’
 (CHAP)

5. Adjectives and nouns can occupy the predicate slot

An adjective or a noun without any verbal constituent can take the predicate slot. This construction, reminiscent of the zero copula in the nominal sentence of Bengali and Assamese, implies that any sentence has the capability of ending in a modifier or nominal category. Consider:

- (42) *hām sutta mā bimar jasti*
 1SG/PL small mother sick much
 ‘My/our small mother (younger, i. e. step-mother) is very sick.’
 (CHM)

- (43) *it^hu bara tukri*
 Here big basket
 ‘Here is the big basket.’ (CHM)

- (44) *age admi-lok k^hātra hensām*
 Earlier men-people dangerously handsome
 ‘Earlier people were dangerously (=extremely) handsome.’ (CHAP)

6. Absence of oblique marking

All the oblique markings of nouns or pronouns preceding any postpositions in SH are absent in the contact Hindis of the Northeast, a feature we saw also in other CH. Hence, we have structures as given below with uninflected subject and object forms in the CHAP (Abbi & Sharma 2014):

- (45) *kutta ko dek-ke rona suru kəra*
 dog ACC see-CVB cry begin do.AOR.M.SG.
 ‘Seeing the dog [she/he/they] started crying.’

SH: *kutte ko*
 dog.OBL PP

- (46) *bacca-lok vaste*
 child-people for
 ‘for the children’

SH: *baccō ke vāste*
 children.OBL-BEN

- (47) *əp-əpna def me mərna he*
 REFL-REFL.DIR country LOC die COP
 ‘One has to die in one’s own country.’

SH: *əpne def mē*
 own.OBL country in

Since the CHM does not use postpositions in the normal speech, the question of inflecting nouns in oblique forms—the necessary condition for oblique marking in SH—does not arise.

7. Lexicon: some peculiarities

The lexicon is built on the sources available to the speakers from various kinds of interlocutors. As to be expected, some Hindi words are—apart from being adjusted to the phonological pattern of the borrowing language—semantically extended and are provided with new meanings. Thus, the verb *chīl-* in the SH means ‘peel’, but in the CHAP the word *k^hol* is used which includes the meaning of ‘peel’ as well as ‘open’ and ‘scrape’. Similarly, the SH verb *bet^h-* ‘sit’ is also used for extended meanings of ‘live’, ‘be alive’ and ‘be married to’ both in the CHAP and the CHM. Consider a few examples given below. There is a phonetic change in the word as it is pronounced as *bet* ~ *bet*.

- (48) *tum kis=ke sat bet-ta*
 2SG whom=GEN with sit-PRS.M.SG
 ‘Whom are you married to?’ (CHAP)

- (49) *it^hu hām sat beta*
 3SG 1SG/PL together live.PRS.M.SG
 ‘He is staying/living with me.’ (CHM)

Contact Hindis of the two regions have been observed to fill in the gaps of the SH in their respective ways. For instance, Hindi does not have a generic term for ‘cousin’. This gap is filled in the CHAP by the word *nəkli* ‘spurious’, ‘artificial’ prefixed to the kinship term as in *nəkli b^hai/behen* ‘cousin brother/sister’. The word *nəkli* is also used in the CHM for cousins. The extension of the meaning of the existing lexicon can be seen in the use of the cardinal numeral ‘two’ by the speakers of Hindi in Meghalaya. The CHM takes recourse in cardinal number *dui* ‘two’ to specify some relations, as in *dui ma* ‘step mother’, literally, ‘the second mother’.

This is a strategy to incorporate wider semantic fields within a lexeme. For detail see Abbi & Sharma (2014).

8. The grammaticalization of the verb to ‘say’

It has been observed that both the CHAP and the CHM use the verb ‘to say’ either as a complementizer or as a causal linker. Consider:

- (50) *tānda nāi e bol-ke ja-ke a gja*
 cold NEG cop say-CVB go-CVB come go.AOR.M.SG
 ‘Because it was not cold, she returned.’ (CHAP) (Sharma 2013)

- (51) *jada din re-ta ja me kam*
 many days live-PRS.M.SG here LOC work
kārta bol-ke
 do.PRS.M.SG say-CVB
 ‘I live here for many days because I work here.’ (CHM)

As is clear, in the sentence (50) the verb ‘say’ operates as a linker of the two clauses in the sentence establishing the reasoning of the action concerned. The final occurrence of *bol-ke* in the same role in the CHM (51) is not fixed and can change its position from the sentence final to sentence middle position. A similar observation is made by Vashini Sharma (this volume) for Hyderabad Hindi-Urdu and Dakhini.

9. Morphological derivation of transitive forms

Both the contact languages have dropped the causative derivational suffix *-va* of the SH. The derivational suffix *-a* or *-æ* serves all the purpose without taking into account the vowel alternation in the base morpheme or any idiosyncratic feature like insertion of *-l-* in the word *khi-l-ana* ‘to feed’ in the SH. Thus, the CHAP and the CHM form transitive which merges with causative as given below.

$p^h\partial t-$ (vitr) ‘burst, be torn’ > $p^h\partial ta-$ (vitr) ‘tear, burst’; $m\partial r-$ ‘die’ > $m\partial ra-$ ‘kill’
 (CHAP)
 k^ha- ‘eat’ > k^hae- ‘feed’; $m\partial r$ ‘die’- > $m\partial r\partial e-$ ‘kill’ (CHM)

The reason for such similarities could be the possibility of the shared model of Hindi exposed to the two communities as discussed earlier. The migrants and the outsiders that come to the region are mainly from Bengal, Assam and in the case of Meghalaya also from Bihar who bring along their versions of communicative Hindi. In addition to these external reasons, there are certain internal reasons pertaining to the grammatical structure of Hindi and its vulnerability to being modified or deleted to render it easy to grasp and use. It is not surprising that many of the shared features discussed above are also shared by many varieties of contact Hindi across the country proving beyond doubt that Hindi grammar has certain well-identified and widely attested features that can be predicted to change in a definite direction.

Despite such similarities, the two Hindis of the Northeast regions differ in the following ways owing to their differences in the duration of contact, the nature of the contact, sociolinguistic factors, and above all the difference in the major attitude towards learning Hindi.

5.4 The differences between the CHAP and CHM

Hindi used in Arunachal Pradesh has become the home language in many families due to intermarriages between different ethnic groups and thus is commonly used between siblings, friends and also parents. In the last twenty-five years, Arunachal has seen this Hindi emerging as the major lingua franca being used for wider communication. The result is that the CHAP is better structured than the Hindi of Meghalaya. Sharma (2013: 57) cites that unlike the CHM the variety in Arunachal marks the oblique form in third person pronouns, demonstrative pronouns, relative pronouns, indefinite and relative pronouns. Increased contact with the lexifier language, i.e. Hindi, has prompted the influx of the oblique marking although incomplete as yet. Contrary to the case of the CHAP, the CHM has borrowed

demonstrative pronouns *ithu*, *uthu* from Maithili which are reanalysed as bare pronouns, demonstrative and relative pronouns—all by the same lexeme, thus resulting in a single system of pronouns.

The CHM is devoid of grammatical function words most of the time. Consider the specific features in (52): the absence of genitive marking with pronoun *tum*, absence of ergative marking *ne* on subject *kutta* and of accusative marking =*ko* on object pronoun *həm*, thus leaving all the arguments in their bare forms and giving rise to a number of possible conflicting meanings, a really humorous one (‘I bit your dog’) among them. This is not witnessed in the CHAP.

- (52) *tum kotta həm kata*
 2SG-Ø dog-Ø 1SG/PL-Ø bite.AOR.M.SG
 ‘Your dog has bitten me.’

The biggest difference between the two Northeast Hindis is that the CHAP, although spoken by the speakers of the agglutinative languages of the Tibeto-Burman family, has adopted the inflectional morphology of the SH in tense, aspect and mood markings without any agreement in gender, number and person. A sentence like the following from the Arunachal Hindi is unexpected in Meghalaya Hindi.

- (53) *həm=ko mət dek-o sərəm ləg ra e*
 1PL=ACC PROH see-IMP shy feel PROG.SG aux.PRS.3SG
 ‘Do not look at me, I feel shy.’ (Sharma 2013:82)

Another feature which is not shared is that the CHM has not developed prohibitive negation *mət*. Thus, it is not surprising to get the following kinds of sentences which invite ridicule from the Hindi speakers in the town of Shillong:

- (54) *k^hao k^hao fərəm bi nei*
 Eat eat shame also NEG
 ‘Please eat, don’t be shy.’

Many other SH features such as the use of copula at the end of the sentence, conditional negative *nəhī to* ‘otherwise’, interrogative constructions of both yes/no and wh-questions, use of honorific pronoun *ap* ‘you’ etc. exist in the CHAP but not in the CHM. For the detailed study of the two Hindis, viz. CHAP and CHM, readers may refer to Sharma (2013).

Contact Hindis have spread widely in the Northeast region of India. As far as acceptability is concerned, the CHAP is acceptable as a link language at all the formal and informal domains, whereas the CHM is primarily restricted to formal spheres of life such as offices, and informal spheres as local markets, etc. While

the former is used widely among disparate ethnic and linguistic communities, the CHM serves the limited purpose of communication between the communities engaged in various occupations and the market. As far as the structure of the two varieties of Hindi is concerned, the CHAP seems to be structured whereas Meghalaya Hindi is still evolving and thus offers many conflicting and incomplete sentence structures.

Examples of conflicting structures are seen in the variation of word order because Khasi and Jaintia are SVO languages while Hindi is an SOV language. Sometimes the speaker uses the Hindi verb in the middle of the sentence and sometimes at the end of the sentence. This tendency spreads over other adjuncts too. For instance, the languages of the Austroasiatic family have prepositions while Hindi offers postpositions. In the uneducated speech, one finds structures like the following in the conversation that I (AA) had with the speaker who was a tea seller (TS) in Shillong. Not only that the speaker drops the postposition in the first clause he uses them as prepositions in subsequent clauses marked by angular brackets for clarity. Notice that the speaker does not have ordinals in his verbal repertoire so he uses only cardinal numbers:

(55) AA: *ap=ke bəcc-e kja kər-te haĩ*
 2HON=GEN child-M.PL what do-PRS.M.PL aux.PL
 ‘What do your children do?’

(56) TS *pərta ø-skul, ek <me klas> pac, do*
 study.PRS.M.SG ø-school, one <LOC class> five two
 < me klas> tin or tin me klas ek
 <LOC class> three and three <LOC class> one
 ‘[They] study in the school, the first one in class five, the second one in class three, and the third one in class one.’

Note: For the convenience of interpretation, I have underlined the ordinal indicating the class the child studies in.

To summarize the discussion so far, we can very well observe that when the SH meets other languages it changes its grammar because the users of the contact Hindi restructure it in various ways:

- (i) Features that are prone to diffusion are dropped viz. the agreement strategy and the oblique case marking or the expression of modality with the verb *sək-* ‘to be able to’ used as an independent verb are completely restructured to simplify the task.
- (ii) Semantic extension of Hindi lexicon such as the word *kʰol* ‘open’ in the CHAP and grammaticalization of Hindi lexical items such as *bol-ke* ‘say- CVB’ as

- complementizer or a causal linker are some of the innovations according to the convenience of the users and the structure of the base language.
- (iii) In non-Hindi regions, the model language is rarely the SH and thus, the model language of the immigrants influences the structure of the contact Hindi, e. g. *it^hu-ut^hu* ‘here’ from Maithili in the CHM.
 - (iv) The sociolinguistic situation of each area governs the spread and use of the contact Hindi. The larger the use, the closer the structure to the SH. The CHAP in the non-Hindi zone and the CHB in the Hindi belt are cases in point.
 - (v) Acceptance and prestige of the language are related to the attitude of the speakers of a community. Tribes in general (Abbi et al. 1998–2000) including those of Arunachal Pradesh and non-tribes of Andaman Islands consider pride in speaking Hindi. This single reason motivates the community to learn and move towards the SH.
 - (vi) Idiosyncratic features of the SH are dropped first and forms are analogized based on the forms which are statistically larger in number. In other words, exceptions and rare features are dropped in contact varieties. Thus, the intransitive-transitive-causative paradigm of verbs takes the standard derivation and ignores the exceptions.
 - (vii) Major speculation can be made that Hindi in Arunachal Pradesh will expand its scope in domains and may establish stable formal and informal registers. It may also become the mother tongue of the second generation. Contact Hindi in Meghalaya, on the other hand, has a long way to gain a stable structured grammar as the influx of English in the region is attracting the community to learn English and not Hindi. However, contact Hindi will remain to be used in the market place as the only lingua franca available for the semi-educated people of Meghalaya.

6 Conclusion

Considering the similarities of changed features of varieties of the CH one can identify a subset of those which are prone to diffusion or loss. One can safely say that these are guiding principles to give us the power of prediction as to which features of the SH will be diffused first in case of any new contact situation. The case in point is the loss of agreement phenomenon of the SH or the use of the verb ‘to be able to/can’ as an independent verb, or loss of oblique case marking on nouns.


Semantic extension of the available lexicon of the SH and reanalysis and restructuring of the grammar of the SH into new contact Hindi is inevitable. Thus,

the SH verb *bēṭh* 'sit' becomes multifunctional and the verb 'to say' operates as linker and complementizer.

To conclude, intensive language contact is the genesis of diffusion of forms phonetically, morphologically, syntactically, and semantically. Both internal innovation and external pressures help in restructuring existing grammar in such a way that the language of contact is widely accepted and is prestigious. Multilingual mosaic of the country feeds the structure of the contact language in multiple ways. Languages grow in layers as the model language is constantly changed by extension, innovation and grammaticalization processes.

The use of Hindi across states and social and religious identities has given rise to non-standard varieties of Hindi not only in non-Hindi states but also within the states of the Hindi belt. However, contact Hindi, more often than not, operates in a diglossic relation with SH in Hindi-belt. Coupled with different registers, we have several varieties of non-Standard Hindi which are used in day-to-day life. The complex mosaic of the Hindi speech community is further enlarged by the emergence of cosmopolitan Hindi, i. e. Hindi used in multilingual, multicultural and multi-ethnic communities, which shows traits of convergence of several languages and dialects spoken in the cosmopolitan cities. The shared structural similarities between various CH indicate the vulnerable linguistic features of SH that are affected in contact situations. What lies in the underlying system of these categories which make them easy to be dropped, negotiated, and changed warrants future research.

ORCID®

Anvita Abbi  <https://orcid.org/0000-0003-3163-3751>

Appendix

Data from the Census of India 2011

Total strength of speakers of Hindi 528,347,193

1. Awadhi 3,850,906
2. Baghati/Baghati Pahari 15,835
3. Bagheli/Baghel Khandi 2,679,129
4. Bagri Rajasthani 234,227
5. Banjari 1,581,271
6. Bhadrawahi 98,806
7. Bhagoria 20,924
8. Bharmauri/Gaddi 181,069
9. Bhojpuri 50,579,447
10. Bishnoi 12,079
11. Brajbhasha 1,556,314
12. Bundeli/Bundelkhandi 5,626,356
13. Chambeali/Chamrali 125,746
14. Chhattisgarhi 16,245,190
15. Churahi 75,552
16. Dhundhari 1,476,446
17. Garhwali 2,482,089
18. Gawari 19,062
19. Gojri/Gujjari/Gujar 1,227,901
20. Handuri 47,803
21. Hara/Harauti 2,944,356
22. Haryanvi 9,806,519
23. Hindi 322,230,097
24. Jaunpuri/Jaunsari 136,779
25. Kangri 1,117,342
26. Khari Boli 50,195
27. Khortha/Khotta 8,038,735
28. Kulvi 196,295
29. Kumauni 2,081,057
30. Kurmali Thar 311,175
31. Lamani/Lambadi/Labani 3,276,548
32. Laria 89,876
33. Lodhi 139,180
34. Magadhi/Magahi 12,706,825
35. Malvi 5,212,617
36. Mandeali 622,590
37. Marwari 7,831,749
38. Mewari 4,212,262
39. Mewati 85,643
40. Nagpuria 763,014
41. Nimadi 2,309,265
42. Padari 17,279
43. Pahari 3,253,889
44. Palmuha 23,579
45. Panch Pargania 244,914
46. Pando/Pandwani 15,595
47. Pangwali 18,668
48. Pawari/Powari 325,772
49. Puran/Puran Bhasha 12,375
50. Rajasthani 25,806,344
51. Sadani/Sadri 4,345,677
52. Sirmauri 107,401
53. Sondwari 229,788
54. Sugali 170,987
55. Surgujia 1,738,256
56. Surjapuri 2,256,228
57. Others 16,711,170

Bilingualism in Hindi.

Total strength of Hindi and Urdu as non-mother tongue

Number of Speakers who claimed Hindi as the second language:
138,909,608, viz. 11.47% of the entire population.

Number of Speakers who claimed Hindi as the third language:
24,307,234. Viz. 2.00% of the entire population.

Number of Speakers who claimed Urdu as the second language:
11,348,978, viz. 0.93% of the entire population.

Number of Speakers who claimed Urdu as the third language:
117,836, viz. 0.09% of the entire population.

Source: Census of India Website: Office of the Registrar General & Census Commissioner, India. Ministry of Home Affairs, Government of India. (<<https://censusindia.gov.in/2011-common/censusdata2011.html>>, accessed: August 10, 2019).

Bibliography

- Abbi, Anvita 1995. "Morphological Change in Tribal Languages of Central India", *PILC Journal of Dravidic Studies* 5,1: 1–9.
- Abbi, Anvita 2009. "Vanishing Diversities and Submerging Identities", in: Sarangi, Asha (ed.): *Language and Politics in India*. New Delhi: Oxford University Press, 299–311.
- Abbi, Anvita 1997. "Languages in Contact in Jharkhand. A Case of Language Conflation, Language Change and Language Convergence", in: Abbi, Anvita (ed.): *Languages of Tribal and Indigenous Peoples of India. The Ethnic Space*. (MLBD Series in Linguistics, vol. 10.) Delhi: Motilal Banarsidass, 131–148.
- Abbi, Anvita, R.S. Gupta & Ravinder Gargesh 1998–2000. *A Sociolinguistic Enquiry into the Acceptance Level of Hindi as a Pan-Indian Language*. (ICSSR Report.) New Delhi.
- Abbi, Anvita & Maansi Sharma 2014. "Hindi as a Contact Language of India", in: Agnieszka Kuczkiewicz-Fraś (ed.): *Defining the Indefinable: Delimiting Hindi*. Frankfurt am Main: Peter Lang Academic Research, 107–124.

- Aikhenvald, Alexandra & Robert M.W. Dixon 2007. *Grammars in Contact. A Cross-linguistic Typology*. Oxford: Oxford University Press.
- Bhatt, S.C & Gopal K. Bhargava 2006. *Land and People of Indian States & Union Territories*. Vol. 18. *Meghalaya*. New Delhi: Kalpaz Publications.
- Burling, Robbins 2007. “The Lingua Franca Cycle: Implications for Language Shift, Language Change and Language Classification,” *Anthropological Linguistics* 49, 3/4: 207–234.
- Grierson, George A. 1903. *Specimens of the Bihari and Oriya Languages*. (Linguistic Survey of India, vol. V, part II.) Calcutta: Office of the Superintendent, Government Printing, India.
- Hashami, Sabiha 2015. *Contact Hindi in Bihar and Jharkhand: Structure and Use*. PhD Dissertation (unpublished). Jawaharlal Nehru University, New Delhi.
- Kachru, Yamuna 2006. *Hindi*. (London Oriental and African Language Library 12.) Amsterdam; Philadelphia: John Benjamins Publishing Co.
- Modi, Yanky 2005. *Emergence and Impact of Hindi as a Lingua Franca in Arunachal Pradesh*. MPhil Dissertation (unpublished). Center for Linguistics, Jawaharlal Nehru University, New Delhi.
- Sharma, Maansi 2011–2012. *Field Data on Hindi as a Link Language in Arunachal Pradesh and Meghalaya* (unpublished). Jawaharlal Nehru University, New Delhi.
- Sharma, Maansi 2013. *Hindi as a Link Language in Arunachal Pradesh and Meghalaya*. PhD Dissertation (unpublished). Jawaharlal Nehru University, New Delhi.
- Siegel, Jeff 1988. “Development of Fiji Hindustani”, in: Jeff Siegel and Richard Barz (eds): *Language Transplanted: The Development of Overseas Hindi*. Wiesbaden: Otto Harrassowitz, 121–149.
- Thomason, Sarah & Terrence Kaufman 1988. *Language Contact, Creolization and Genetic Linguistics*. Oxford: University of California Press

Annie Montaut

On the Non-Lexical Categories of *avyay* 'Invariables' and Their Grammaticalization in Pahari Languages, with a Comparison to Standard Hindi

आनी मोंतो

पहाड़ी भाषाओं में सहायक कोटियों के रूप में अव्ययों का व्याकरणीकरण और मानक हिंदी के साथ उनकी तुलना

Abstract Non-lexical categories of uninflected elements are usually only listed in Pahari grammars without much explanation. In such lists, conjunctions, coordinators and discourse particles (*to*, *hī*) are packed together with lexical elements such as adverbs (*ab*, *kal*) and interjections (*are*, *hāy*) under the Sanskrit umbrella term *avyay*. Similarly, adpositions are also barely listed after the well-developed sections on nouns. Grammatical elements are however most of the time connected to other functional or lexical words in the language, and the story of their grammaticalization itself is of great significance for understanding the relations between the given language and cognate or neighbouring languages. As a matter of fact, the grammaticalization process is all the more interesting in so-called dialects, as standardization has been less variation-suppressive than in the so-called major languages: dialects, particularly those close to the regional major language they 'belong' to, are then the best standpoint for looking at the grammaticalization in major languages, since they are a kind of language sanctuary, where variation is preserved.

This article bears on this understudied aspect, with special reference to the so-called Central Pahari (Garhwali and Kumaoni) in relation to Standard Hindi, showing that the very diversity of forms can help to understand the nature and evolution of the equivalent material in Standard Hindi. Our theoretical framework combines historical grammar, theories of grammaticalization, functionalist typology, and occasionally contact linguistics.

Keywords case markers, diachrony of postpositions, Garhwali and Hindi dialects, quotative.

सारांश आम तौर पर पहाड़ी भाषाओं के व्याकरणों में अव्यय या *uninflected elements* बिना किसी व्याख्या के केवल एक सूची के रूप में प्रस्तुत किए जाते हैं। इन सूचियों में समुच्चय बोधक (तो, ही), क्रिया-विशेषण (अब, कल), विस्मयबोधक (अरे, हाय) आदि को संस्कृत के पारिभाषिक शब्द “अव्यय” के अंतर्गत ही प्रस्तुत किया जाता है। इसी प्रकार से, परसर्ग भी संज्ञाओं की विस्तृत व्याख्या के बाद एक सूची भर के रूप में ही दिए जाते हैं। व्याकरणिक तत्व अक्सर भाषा के अन्य क्रियात्मक या कोश-विषयक शब्दों से संबंधित होते हैं और किसी भाषा के सजाति या भौगोलिक रूप से नज़दीकी दूसरी भाषाओं से संबंध को समझने के लिए उनके व्याकरणीकरण का इतिहास बहुत महत्व रखता है। असल में बोलियों में व्याकरणीकरण की प्रक्रिया और भी रुचिकर होती है क्योंकि बोलियों का मानकीकरण मुख्य कही जाने वाली भाषाओं की तुलना में कम विविधता निषेधात्मक होता है। बोलियाँ, विशेष रूप से जो क्षेत्रीय संबद्ध भाषा के समीप होती हैं मुख्य भाषाओं में व्याकरणीकरण को समझने का सर्वश्रेष्ठ दृष्टिकोण प्रदान करती क्योंकि वे एक ऐसा भाषा अभयारण्य हैं जिनमें विविधता संरक्षित होती है।

यह लेख इस अल्पअध्यायित (अध्ययनाधीन) पहलू के बारे में है। मध्य पहाड़ी भाषाओं (गढ़वाली और कुमाऊँनी) और मानक हिंदी के विशेष संदर्भ में, यह दर्शाता है कि रूपों की विविधता हमें मानक हिंदी में समकक्ष सामग्री की प्रकृति और उद्भव को समझने में मदद दे सकती है। इस लेख में प्रस्तुत सैद्धांतिक संरचना ऐतिहासिक व्याकरण, व्याकरणिककरण के सिद्धांत, क्रियात्मक वर्गीकरण और संपर्क भाषाविज्ञान को संघटित करता है।

मुख्य शब्द – कारक सूचक चिह्न, परसर्गों का ऐतिहासिक उद्भव, व्याकरणिककरण, गढ़वाली की हिंदी से तुलना, क्लेटेटिव।

1 Introduction: scope and goal of the paper

1.1 Garhwali language in its linguistic and cultural environment

Garhwali is an Indian living language with twofold gender distinction, spoken by about 3 million people in the Western part of Uttarakhand (including important speaking communities in Himachal Pradesh, Haryana and Punjab).¹ Although it is noncontroversially considered an Indo-Aryan language, naturally incorporating non-Indo-Aryan words, as do all other IA languages including Sanskrit, another opinion exists outside of the milieu of linguists. In some non-professional essays in the social media platforms of Uttarakhand, Garhwali is deemed not related to Sanskrit or Vedic, and having no more relation to them than a reciprocal one: Old Garhwali or Kumaoni, the Khas languages, had, as stated in the essay, been the

1 It is, however, difficult to correctly estimate the number of speakers (whether of L1 or L2), as mentioned in Ethnologue: linguistic activists tend to overestimate the numbers while the UNESCO report on endangered languages gives hardly more than 500 000 speakers, see *Atlas of the World's Languages in Danger* by Moseley (2010).

instrument of lexical creation in Sanskrit.² That may, of course, appear an exaggerated vision for any IA language, whose syntax and lexicon share more with OIA than with any substrate. But the reasons given by the author are quite interesting since his wish to dissociate non-standardized languages, such as Garhwali, from a codified language fixed by literary or religious canons (hence unchangeable and as such unable to give birth to any language) points to a reality: standardized modern languages represent in his terms a “cauterization” of the living proliferation of spontaneous continuous creativity which is typical of “dialects” of all times.

Garhwali is deemed vigorous and developing by Ethnologue, with a positive attitude among its speakers, although many might object that the turn to Hindi in the younger generation suggests a not so positive attitude. The UNESCO *Atlas of World Languages in Danger* (Moseley 2010) mentions it, on the contrary, as rapidly shrinking and classifies it in the “unsafe” category, and the UNESCO Silk Road Programme estimates the number of speakers at 279 500.³ My own field observations lead me to the same conclusion (apart from language activists who obviously exhibit strong linguistic pride, most speakers grant the topmost status to English, followed by Hindi).

A non-written language until recently, Garhwali is now enjoying a number of publications, including creative writing (fiction and poetry) and cultural publications such as the magazine *Dhād*, founded in 1971 by linguistic activists.

The many similarities of Garhwali and Rajasthani are consistent with Garhwal’s political history since the influence of migrated Rajputs has deeply impacted the local (Khasa) culture and language; “a somewhat corrupted form of Rajasthani”, says Grierson (1916: 279).

As any minor language considered a dialect, Garhwali offers interesting interactions with the local material culture, which can also be considered endangered, and with it the precious knowledge of the local environment and biodiversity.⁴

2 “Vedic Sanskrit or Sanskrit Are not Mother of Garhwali and Kumaoni Languages”, reversely Old Garhwali and Khas languages can be regarded as mothers of Vedic and Sanskrit (Kukreti s.d.).

3 UNESCO Silk Road Program.

4 The same is of course true for lexical elements, particularly those related to traditional labor, tools and natural environment, and many scholars have dealt with this aspect of “eco-linguistic” conservation (starting with David Harmond’s “Losing Species, Losing Languages: Connections between Biological and Linguistic Diversity” (1996) and Peter Mühlhäusler’s “The Interdependence of Linguistic and Biological Diversity” (1995). It is necessary to mention the great works of Anupam Miśra (1993; 1998) on the cultures and languages related to water conservation: *Rājasthān kī rajat būdē* (The Radiant Raindrops of Rajasthan) and *Āj bhī khare hāi tālāb* (The Lakes are Still Alive). Nettle and Romaine (2000: 166) write, “Delicate tropical environments in particular must be managed with care and skill. It is indigenous peoples who have the relevant practical knowledge, since they have been successfully making a living in them for

The corpus used for this study includes notes from four sessions of fieldwork carried out between 1997 and 2016 in Srinagar, Pauri, Dehradun, and Western Kumaon (the examples from the field research are not referenced), as well as the available collections of Garhwali epics (*lokgāthāē*) and folksongs (*lokgīt*), such as Bābulkar 1996, Nauṭiyāl 1996 and Cātak 2000. Section 2 analyses the main case markers in core functions, other case markers are considered in section 3, and the less discussed marker *bal* in section 4.

1.2 Grammatical words within the wider category *avyay* ('invariables') or uninflected items

Why should we focus on grammatical elements rather than adverbs or interjections, and, among grammatical items, why on adpositions and the quotative marker rather than on other conjunctions, coordinating or subordinating?

The main reason why interjections have been excluded from this study is that they convey also, maybe mainly, onomatopoeic features, and thus would require a different methodology, including tools for dealing with sound symbolism and expressivity (Dingemanse 2012). Furthermore, they differ not so much from Hindi, whose properties are contrasted with the Garhwali facts in this study. Also most particles are the same as those used in Hindi, the reason why they will be only briefly alluded to below: Hindi *to* for instance is conveyed by its Garhwali cognate *ta* in most of its uses, both as a conjunction ('then', 'so'), a resumptive in the correlative system and as a discourse particle with a contrastive and topicalizing meaning... < Skt. *tavat*, commented in Montaut (2016b). An exception has to be made for *bal*, unknown in Hindi, which is sometimes classified as a discourse particle (*vācak avyay* Cātak 1966: 139–141) and sometimes as a case marker (Jošī 2011: 13), but is closer to the meaning and behaviour of evidential markers in languages that have them. It will be studied in the last section.

Adverbs are a more interesting category for they sometimes radically differ from Hindi; as a class, however, they offer fewer opportunities for studying the various paths of grammaticalization, which is the main goal of this paper. An example of such striking difference is the creation of time adverbials: the well-known use of the same word in Hindi for referring to 'tomorrow' and 'yesterday'

hundreds of generations. Much of this detailed knowledge about local ecosystems is encoded in indigenous languages and rapidly being lost". Mühlhäusler (2003: 60) describes how "[t]he rapid decline in the world's linguistic diversity thus must be regarded with apprehension by those who perceive the interconnection between linguistic and biological diversity".

is not attested in Garhwali; it is also unknown in many Indian languages including Indo-Aryan languages,⁵ such as Dakkhini Hindi/Urdu or Bengali. The fact that Garhwali displays the two distinct words *byāle/byālī* 'yesterday' and *bho! / bhow* 'tomorrow'⁶ points to the originality of Hindi and closely related languages (Urdu, Punjabi), which use the same word *kal* (< Skt. *kalyam* 'at daybreak').

Conjunctions introducing subordinate clauses, such as *ki* 'that' and complex conjunctions including it (*kyōki* 'because', *tāki* 'in order to') are common in modern Garhwali, as in modern Hindi, but quite rare in traditional Garhwali where the correlative system prevails as in Old Indo-Aryan (OIA). Even the hypothetic system (*jū ... ta ...*, equivalent to the Skt. basic correlative structure *ya-... ta-*)⁷ is less frequent in Modern Garhwali than the all-purpose *cā/jā* (< *cāhe*) 'if, whether'. Correlation, according to Minard (1936), was the only device in Old Indo-Aryan equivalent to the modern subordinating system, with the basis *ya-* (> MIA and NIA *j-*) in the first member of the diptych and the resumptive pronominal basis *ta-* in the second one, with various endings depending on the meaning of the relation. This system is still echoed in Modern Hindi's correlative devices with a renewal of resumptive *ta-* as *va-* (*jab ... tab* 'when', *jahā ... tahā/vahā* 'where', *jo ... vah* 'who', *jaise ... vaise* 'such as', *jitnā -...- utnā* 'as big/numerous as', while properly strictly subordinating devices are borrowed from Persian; they all include the all-purpose complementizer *ki* 'that' (*cūki* 'since', *tāki* 'so that', *hālāki* 'even if, although').

Coordinating conjunctions are the same in Garhwali and Hindi (*ar/aur* 'and' < Skt. *aparam, avaram*), except *dī* 'and', possibly derived from *ādi* 'etc., and others',⁸ which is occasionally substituted with *ar* in Garhwali.

- 5 The fact that Dakkhini uses *sabā* for 'tomorrow' as distinct from *kal* for 'yesterday' has often been interpreted as a calque from Dravidian neighboring languages, which all have two different words.
- 6 *Byālī* is generally derived from MI *viāla* < Skt. *vikāla* 'evening' (Turner 1966), or alternately derived from Skt. **vibhāne* 'shining.loc' > *vihāne* (Cātak 1966: 137); *bhol* is derived by Turner (1966) from OI **bhōlā/bhōrā* 'daybreak', and by Cātak (1966: 137) from the attributive phrase containing the future passive participle of *bhū* 'to become, be' and the noun 'time' *bhavya velā*. See the Hindi word *bayalu* 'leftover from yesterday' referring to the breakfast which comprises of food leftover from the previous night, a possible borrowing from Garhwali, according to Anvita Abbi – whom I thank for the remark.
- 7 See for instance *jū tū nī kardo ta māi karlū* (if 2SG NEG do.PRS.2SG then 1SG do.FUT.ISG) 'If you do not do [it], I will do [it]'. In Old Hindi also and up to the 20th century, 'if' was expressed by *jo* (< Apabhramsha *jau*) and *jo* still occurs in this function in Modern Hindi.
- 8 *tū dī māi jaulā* [2SG and 1SG go.FUT.PL] 'You and me will go' (Cātak 1966: 139).

1.3 State of the art

Descriptions of the Garhwali language are not very many and are all in Hindi but some of them provide very useful comments, particularly the first fully-fledged one, *Garhṽālī vyākaraṇ kī rūp rekḥā* by Abodh Baṁdhu Bahugūṇa published in 1960, followed by Govind Cātak's pioneering work *Madhya pahārī kā bhāṣāśāstrīy adhyayan* (Linguistic Study of Central Pahari) published in 1966, with many reprints up to now. Goviṁd Juyāl published in 1967 a comparative grammar of Central Pahari (Garhwali and Kumaoni) and Hindi, *Madhya pahārī bhāṣā (garhṽālī kumaonī) kā anuśīlan aur uskā hindī se saṁbaṁdh*, also quite enlightening. The latest Garhwali grammar I have been able to consult is Rajnī Kukretī's *Garhṽālī vyākaraṇ*, published in 2010, rich in examples and contemporary uses, yet unfortunately lacking in scholarly references and ignoring the work done by previous scholars. The closely related language Kumaoni has been magisterially described by D. D. Sharma in his two volumes *The Formation of Kumauni Language* (1987), a most useful companion for the study of Garhwali too.

Besides these specialized works, more general grammars, particularly those with a diachronic perspective, most of them provided by the best philologists of the earlier generation, also mention Garhwali forms. Prior to them, the pioneering Hindi grammar published in 1876 by Reverend Samuel Henry Kellogg covered the entire range of regional languages along with “High Hindi” and lavishly drew from early Hindi: this *Grammar of the Hindi Language*, significantly subtitled, “In which are Treated the Standard Hindī, Braj, and the Eastern Hindī of the Rámáyan of Tulsī Dás, Also the Colloquial Dialects of Marwar, Kumaon, Avadh, Baghelkhand, Bhojpur, Etc.; with Copious Philological Notes”. Among the best philological works, initiated by the magistral *Origin and Development of Bengali Language* by Suniti Kumar Chatterji (1926, many times reprinted), are Udaynārāyan Tivārī's historical description of Hindi grammar (*Hindī bhāṣā kā udgam aur vikās*, first published in 1955), which gives reliable information about case markers from all the languages of the Hindi belt; Baburam Saksena's description of Awadhi (1937); Hardev Bāhrī's work on the so-called “dialects” of Hindi *Grāmīṇ Hindī boliyā* (1966) is also a useful tool.

Many presentations of the Garhwali culture and literature also frequently provide a grammatical sketch of the language and an introduction to its history, such as ‘Sailesh’ Bhaṭṭ's *Garhṽālī bhāṣā aur uskā sāhitya* (Garhwali Language and Literature), published in 1976, or Janārdan Prasād Kālā's *Garhṽālī bhāṣā aur uskā lok sāhitya* (Garhwali Language and Folklore, 1959).

2 Main case markers

What will be meant by “main” in this section are the most frequent forms for marking the functional cases, that is, those conveying core grammatical functions,⁹ such as ergative, dative/object/experiencer, ablative/instrumental and locative. Equivalents of Hindi *ke lie* ‘for’ marking beneficiaries are for instance not considered as “main” case markers for dative, nor equivalents of Hindi *par* ‘on’, *ke ūpar* ‘above’, *ke nīce* ‘under’, *ke bīc mē* ‘inside’, *ke pās* ‘close, near’ for locative, *ke sāth* ‘with’ for comitative. These markers will be discussed in the next section. Phonetic variations will also be left aside in the simplified presentation below and will be studied in the subsection dealing with etymology. In this way, Table 1 (from Cātak 1966: 100), with the addition of some forms present in Grierson’s samples (1916: 300ff), presenting a huge proliferation of forms, can be simplified as presented in Table 2, by dispensing with the merely phonetic variants as well as the less frequent forms or forms used for non-grammatical cases. The discussion in the next subsection will be mainly based on this simplified table.

Table 1 The main case markers in Garhwali.

Function	Form
Ergative/Agent (<i>kartā</i>)	<i>na/la</i>
Accusative/Patient (<i>karm</i>)	<i>ka, ku, kū, kaī, saṅī, haṅī, gaṅī, kaṅī, khuṅī, chanaī, taī, thaī</i>
Instrument(al), cause (<i>karaṅ</i>)	<i>na, -n, se, sī, tī</i>
Dative (<i>sampradān</i>)	<i>kaī, taī, tāī, taī, thaī, laī, laii (लैइ), lāī, ka, saṅī, haṅī, gaṅī, kaṅī, khuṅī, kū, caī</i>
Ablative (<i>apādān</i>)	<i>na, tī, te, tāī, biṭe, baṭī, se, sī, paran</i> ¹⁰
Genitive (<i>saṃbamdh</i>)	<i>ko, kā, kī, rū, rā, rī, no</i>
Locative (<i>adhikaraṅ</i>)	<i>par, mā, mu, mang, māje, tanaī, mathe, undū</i>

9 The issue of how abstract a grammatical case is, how much ‘concrete’ meaning remains in dative or ablative will be deliberately ignored in this section and left open until the discussion of grammaticalization paths. Strictly speaking only the ergative could be properly regarded as a grammatical or functional (abstract) case, and it could even be argued that transitivity in Indo-Aryan languages is largely semantic. As for the accusative, it is well known that not all direct objects are marked, and that the marking is largely sensitive to discourse factors.

10 Equivalent of Hindi *par se* (periphrastic ablative).

Table 2 The most frequent forms of the main case markers in Garhwali.

Agent (<i>kartā</i>)	<i>na</i>
Patient (<i>karm</i>)	<i>ku, saṅī, khuṅī, taī</i>
Dative (<i>saṃprādān</i>)	<i>ku, saṅī, khuṅī, taī</i>
Instrument/cause (<i>karaṅ</i>)	<i>na, sī, tī</i>
Ablative (<i>apādān</i>)	<i>na, sī, tī, baṭī/biṭe</i>
Genitive (<i>sambamdh</i>)	<i>ko, rū, kar</i>
Locative (<i>adhikaraṅ</i>)	<i>mā</i>

2.1 Use of case markers

2.1.1 Ergative/instrumental

A striking difference with Hindi is the distribution of the agent (ergative) marker (1a–b), which also occurs in Garhwali as an instrumental, conveying the meaning ‘cause’ (2a) or ‘means’ (2b), and even as an ablative (3), whereas Hindi would display the instrumental *se* in the latter examples (Hindi parallels are given between square brackets for the sake of comparison):

- (1a) *mi=na ū duī (ū duyū thāī) khujyāi*
 1SG=ERG 3PL two (3PL two ACC) search.AOR¹¹
 [H. *māī=ne un donō ko khojā*]
 ‘I looked for these two.’

- (1b) *ve=na satū saṅī ve talau mā ḍāl dinyā*
 3SG=ERG sattu ACC that lake in throw give.AOR
 [H. *us=ne sattu ko us jhīl mē ḍāl diyā*]
 ‘He threw the sattu (ground barley or rice) in the lake.’ (Grierson 1916: 287)

11 The ending *-i/ī* (sometimes *-e*) is an invariable tense marker for simple past (definite past or preterit or aorist) which will be glossed as AOR and not by the more usual PFV since it is not really a marker for perfectivity (Montaut 2016c). Personal endings also optionally occur in this tense (*-au* for 1st person, *ān* for plural 3rd person, etc.). Abbreviations other than the standard ones used: HONOUR for Honorific, GRD for Gerund, PRSUMP for the Presumptive mood; H. for Hindi, U. for Urdu. Postpositions are transliterated according to the way as they are written in the Devanagari texts and glossed as separate words after nouns and as clitics after pronouns (and many nouns in Garhwali). Examples indicated without source are from personal field notes.

- (2a) *bhūk=an mare* [H. *bhūkh se mare*]
 hunger=INS die.AOR.3PL
 'They died (because) of hunger.' (Cātak 1966: 101)
- (2b) *dātu=n khāda* [H. *dātō se khātā hai*]
 teeth=INS eat.PRS.3SG
 'He eats with (his) teeth.' (ibid)
- (3) *vakh=an āe* [H *vahā se āyā*]
 here=ABL come.AOR
 'He came from there.' (ibid)

Both functions, instrumental and agent, can occur in the same clause:

- (4) *mi=na nauno (naunā sañī) bēṭ na māri*
 1SG=ERG child (child ACC) cane INS strike.AOR
 [H. *maĩ=ne bacce ko bēṭ se māra*]
 'I hit the boy with a cane.'¹²

As in Hindi, ergative agents have the referential properties of subjects: they clearly control reflexivation and converb constructions; thorough research has yet to be done on the topic.

- (5) *jai=na o apñā khetu mā bhejyo*
 REL=ERG 3SG¹³ REFL field in send.AOR
 [H *jis=ne use apne khet mẽ bhejā*]
 'Who sent him to his field.' (Grierson 1916: 284)

A peculiarity of Garhwali (6a) and Kumaoni (6b) contrasting with Standard Hindi/Urdu (6c) is the use of the ergative marker for obligation, which will be dealt with in a more detailed way in section 3.

- (6a) *maĩ=na (maĩ-la) ājj barat rakhña*
 (6b) *maĩ=le āj barat rakhña*
 1SG=ERG today fast keep.INF
 'I have to fast today.'

12 The past form is either invariable (*-i/-e*), as in Kumaoni (*-e*), or variable, see (1b) and (5). Grierson for the same contents gives the sentence *maĩ=n nauno bēṭ-an māre* (1916: 300).

13 Note the absence of the accusative marker, although the pronoun has a human reference, in contrast with Hindi.

- (6c) *mujhe āj vrat rakhnā hai* H/U
 1SG.DAT today fast keep.INF is
 ‘I have to fast today.’

The same peculiarity has been noticed in Hindi/Urdu (H/U) of the Punjab, where the ergative marker *ne* alternates with the dative marker *ko*:

- (7a) *maĩ=ne jānā hai* PANJAB H/U
 1SG=ERG go.INF is
 ‘I must go.’ or ‘I want to go.’
- (7b) *mujhe jānā hai* STANDARD H/U
 1SG.DAT go.INF is
 ‘I must go.’

As regards the alternation between (7a) and (7b), Butt (2006: 86) suggests that (7a) conveys a particular emphasis on future and the deliberate will to act. This hypothesis is strongly rejected for punjabized Hindi and Punjabi by Khokhlova (2013: 95), who argues that a single speaker does not use both constructions and when using the ergative postposition does not particularly emphasize deliberate will, since in his dialect it is the standard construction for obligation.

2.1.2 The dative/accusative marker

The most frequent standard postpositions for differential object marking in the modern conversational register are *taĩ* and *sañĩ*, frequently alternating with *kũ*, the Garhwali counterpart of Hindi *ko*:

- (8a) *tũ anĩ rāvat tain/sañĩ/kũ jāndi cha?*
 2SG Anil Rawat ACC know PRS
 ‘(Do) you know Anil Rawat?’
- (8b) *mi=na yĩ film sañĩ/taĩ dekhi/dekhe*
 1SG=ERG this film ACC saw
 (cf. *ĩ film dekhi* ‘this film saw’, where the object is unmarked)
 ‘I saw this film.’
- (8c) *ĩ bāt taĩ yād rākhi*
 DEM thing ACC memory place.IMP
 ‘Remember this thing.’

As is well known, differential object marking was unattested in Sanskrit and Prakrits and was still not a fully-fledged system in early Hindi, where even proper nouns could remain unmarked as objects, although they started being marked either by means of the adposition *ku/kau* or by the mere inflectional ending *-i/i/hi* (Montaut 2018). In the 14th century poet Kabir, object marking seems to obey only discourse strategies or metric reasons, such as in (11) for the common nouns denoting the unique entities ‘sun’ and ‘sea’ (usually marked in Standard Hindi)¹⁵:

- (11) *ulaḥī gaṅga saṃmudra-hi soṣai, saṣihara sūra grāsai*
 reversed Ganga ocean-ACC dry.PRS.3SG moon sun swallow.PRS.3SG
 ‘The reversed Ganga dries up the ocean, the moon swallows the sun.’
 (Kabir, pad 185.1, in Callewaert 2000: 293)

One may consider that Garhwali is in a transitional phase, less advanced than Hindi, which requires the marking of common animate nouns and displays a rather grammaticalized behaviour in this area (although less grammaticalized for inanimates), yet more advanced than Kumaoni where differential object marking is blocked by the ergative marking of the agent.

The postpositions marking object also occur as dative markers — for standard beneficiaries (*tvai-ku nī āūn* ‘I won’t come for you’), as in (12–13), and experiencers (14–16), as in all Indo-Aryan languages.¹⁶ The marker *saṇī* (12) alternates with *tai/taī* (14) for the beneficiary dative, as well as with *thaī* (15) or *kuṇi*, a variant of *kaṇi*, for the experiencer dative (16):

- (12) *bābā=jī birsat mā-n jo mero hisā cha,*
 Father=HON property in-from REL my part be.PRS
 [H. *pitā jī birāsat* (‘heritage’) *mē jo merā hissā hai*
so māi=saṇī de devā
 DEM3SG 1SG=DAT give give.IMP
vo mujhe de do]
 ‘Father, give me that part of the property which is mine.’ (Grierson 1916: 300)
- (13) *tvai=taī¹⁷ dendū māi māletho ko gaū*
 2SG.OBL=DAT give.PRS.1SG 1SG Maletha of village
 [H. *tumhē detā hū māi mālethā kā gāv*]
 ‘I give you the village [of] Maletha.’ (Naṭṭiyāl 2000: 148)

15 One finds similar alternations for proper names in the *Rāmcaritmānas* of Tulsidas (Montaut 2018).

16 Except Bengali, well known for using the genitive for most of the experiencer subjects.

17 The form *taī* in the original may stand for a variant of *taī*, may be a typing mistake, or be due to the non-standardization of the language (in the same couplet we find *tvai saṇī*

- (14) *nauni taĩ pata bi nĩ chau*
 girl DAT knowledge.M.SG even NEG be.PRS.3SG
 [H. *larĩ ko patā bhĩ nahĩ hai*]
 'The girl does not even know.'
- (15) *yu sabbi mĩ=thaĩ sadani yāḍ rālu*
 these all 1SG=DAT always memory stay.FUT
 [H. *ye sab mujhe hameśā yāḍ rahegā*]
 'I will always remember this all.'
- (16a) *tvĩ=kuṇi sūraj kanĩ unĩḍā paṛĩ ca*
 2SG=DAT Suraj what.kind sleep fall PRF
 [H. *tujhe sūraj kaisĩ nĩḍ paṛĩ hai*]
 'How asleep you have gone, Suraj?' (Naṭṭiyāl 1997: 111)
- (16b) *maĩ=kuṇi jiyā bvai āj supĩno hvege*
 1SG=DAT mother mum today dream be.go.AOR
 [H. *mujhe mā āj svapn ho gae*]
 'I have dreamed of mother today.' (Naṭṭiyāl 1997: 111)

The unaspirated form *kuṇi* alternates with the aspirated *khuṇi* in a quite usual way in Garhwali, in the same way as *taĩ* with *thaĩ*, and one finds frequent alternate forms for occlusive consonants (*tābakhū/tambāku* 'tobacco', *par/phar* 'on', etc.).

2.1.3 Instrumental/ablative

The Hindi marker *se* is used in Garhwali as well as its parallels *sū*, *sĩ*. Besides, one finds the marker *na/an*, same as in the agentive use (2a, 2b), and the postposition *te/tĩ*. As a marker for non-canonical agent (incapacitive passive for instance) one finds both the Hindi *se* or its Garhwali counterparts and the instrumental/ergative marker *na* (see below 3.2). But, unlike Hindi, Garhwali has also a specifically ablative marker *baṭi* and its alternative forms *baṭe*, *baṭĩ*, *biṭe*.

- (17a) *tū sab=tĩ syālĩ chai*
 2SG all=ABL beautiful be.PRS.2SG
 'You are the most beautiful of (from) all.'

in two words for the same meaning 'to you') or the metric constraints. Similarly, *naunĩ* 'girl' is alternately written with long or short *i*.

- (17b) *kakh baḥī hvā tyār byoh?*
 where ABL be.AOR your marriage?
 ‘Where from did you marry (was your marriage)?’
- (17c) *mi byāḥī dehradū baḥī aū*
 1SG yesterday Dehradun ABL come.AOR
 [H. *māḥ kal Dehradun se āyā*]
 ‘I came from Dehradun yesterday.’

2.1.4 Genitive and locative

Both cases only deserve a brief mention because they present less variation and are formed on the same base as their Standard Hindi equivalents and can only be considered as core cases in possessive clauses.

The genitive, or relator (*sambandh*), has only two forms, either the *k-* postposition like Hindi, which may be considered (Kukreti 2010) as a Hindi borrowing, or its enlarged form *ker/kar-*, or the *r-* postposition as in Rajasthani, mentioned by Cātak (1966: 103) as dominant in the region of Rawalti.¹⁸ Both inflect for gender and number as in Hindi (hence not a real *avyay*, yet a grammatical item), with different endings, same for *k-*, *r-*, and *kar-*: *kau/ko/kū* for masculine singular direct case, *kā* for masculine plural and oblique, *kī* for feminine:

- (18a) *cācā karaū/kū dero*
 uncle GEN.M.SG house.M.SG
 ‘The house of uncle/uncle’s house.’ (Cātak 1966: 103)
- (18b) *anīl rau nauno*
 Anil GEN.M.SG boy.M.SG
 ‘Anil’s boy’

The most frequent locative marker is *mā* (cf. examples (1b) and (5) above).

18 Cf. *Garhvālī loggīt* (Cātak 2000: 112–27). The *kera* form is found in many other languages, such as Awadhi, Eastern Hindi; see also in Kabir, for instance, *pānī kerā budbudā* ‘the sound of water’.

2.2 Origin and scope of the core case markers

2.2.1 Case marking: a recent evolution

During the earliest stage of Hindi as well as of other New Indo-Aryan languages, the inflexional endings of Sanskrit are in the process of being replaced by adpositions (nominal category) and auxiliaries (verbal category). Yet this process is far from being completed in 14th c. Hindi, and the ordinary situation in the discourse is the absence of clear relators, the few oblique cases maintained in the language being used for various syntactic purposes: the *-i* locative for the agent in past transitive processes, and a fused oblique *-hi* (*ehi, eĩ, aĩ*) derived from the fusion of the old dative/instrumental (already achieved in Middle Indo-Aryan) for all kinds of obliques including agents of transitive verbs:

(19a) *guri* *dīyā* *palītā* (SANT BHASHA)
 guru.OBL/LOC give.AOR.M.SG stick.M.SG
 ‘The guru gave the stick.’ (Kabir, pad 8, in Callewaert 2000: 126)

(19b) *virrsih joysīyā* *bhāṣ* *pāi* (OLD KUMAONI)
 Virsingh Joshi.OBL proclamation/bond.F.SG get.F.SG
 ‘Virsingh Joshi (the king) received the bond.’ (from Pant 2009 in Stroński 2014: 281)

Most of the time nouns are unmarked, the syntactic and semantic sequences being simple enough for the meaning to be clear. The *-hi* ending was the most frequent marker for differentially marked objects in Kabir, while the postpositional marking (*ku/kau*) just starts appearing (Strnad 2013: 325); the markers are used similarly for experiencers:

(20a) *kāmĩ* *amĩ* *na* *bhāvai*
 lascivious.OBL nectar NEG please.PRS.3SG
 ‘Lascivious men do not like nectar.’ (Kabir, pad 20, in Vaudeville 1957: 38)

(20b) *so* *bhakta* *merai* *mani* *bhāvai*
 DEM devotee POSS.1SG.OBL heart.OBL please.PRS.3SG
 ‘That bhakta (devotee) is dear to my heart.’ (Kabir, pad 65, in: Strnad 2012: 50)

(20c) *premiĩ* *kaũ* *premiĩ* *milai*,
 lover DAT lover meet/find.PRS.3SG
tab *sab* *biṣ* *amrit* *hoi*
 then all poison nectar be.PRS.3SG
 ‘[When] the lover finds the lover, all poison becomes nectar.’
 (Kabir, pad 43, in: Vaudeville 1957: 64)

In Garhwali, unlike Modern Hindi and like in Kabir, inflectional case markers are still in use in certain dialects, and Cātak (1966: 102) mentions an *-a* ending dative (21a), an *-ā* ending locative (21b), an *-u* ablative (21c), and an *-ai* instrumental/ablative (21d):

- (21a) *naunā-a miṭhāī lāyū*
 boy-DAT sweets bring.AOR.1SG
 [H. *larke ke lie miṭhāī lāyā hū*]
 ‘I brought sweets to the boy.’
- (21b) *ghar-ā pānī nā, cull-ā āg nī*
 house-LOC water NEG fireplace-LOC fire NEG
 [H. *ghar mẽ pānī nahī, cūlhe mẽ āg nahī*]
 ‘No water in the house, no fire in the fireplace.’
- (21c) *vaṇa-u (vaṇau-u) āye vo ghar* [H. *vah van se ghar āyā*]
 forest-ABL come.AOR he home
 ‘He came home from the forest.’
- (21d) *ḍamḍ-aī māre* [H *ḍamḍe se mārā*]¹⁹
 stick-INS strike.AOR
 ‘He struck with a stick.’

Pronouns in the beneficiary (22a–b) and experiencer (22c) functions are often in the oblique case, such as *maī* or *mu* for the first person singular. In folk songs these pronominal forms may appear without case markers.

- (22a) *mu dayā putaro ko var*
 1SG give.IMP son GEN boon
 ‘Give me the (divine) gift of a son.’ (Cātak 2000: 116)
- (22b) *sabi bainiyō gainī dinī maī dine nāk nathūlī*
 all sister.F.PL jewels give.AOR 1SG give.AOR nose ring
 ‘To all [my] sisters you gave jewels, to me you gave a nose ring.’
 (Cātak 2000: 121)
- (22c) *mu lage bāpū kī būr mu lage amī kī būr*
 1SG touch dad of memory 1SG touch mum of memory
 ‘I remember (miss) dad, I remember mum.’ (Cātak 2000: 124)

Yet the agentive (ergative) marker in modern Garhwali is never a mere inflection but the adposition *na*, nor is it an inflection in the traditional songs, sometimes

19 Other examples mentioned by Cātak contain forms still in use in some dialectal varieties: *ām ḍālai.i* (ढलैइ) *bhvīṅṅ paṅe* (mangoes tree.ABL ground fall.AOR) ‘The mangoes fell on the ground from the tree’ [H. *ām peṅ se zamīn par gire*] (*ḍālo* also means ‘tree’ in Garhwali).

deemed more conservative, while we frequently find *-i* oblique marking for agents in the Sant Bhasha and *-ã* in Old Pahari (ex. 19).

2.2.2 Etymologies

The ergative marker *na*

Indeed, today the most wide-spread form of the ergative case markers is *ne* (*ni*, *nai*, *nē*, *ne*), found in Hindi/Urdu, Punjabi, Marathi, Gujarati, and it is derived from the reconstructed locative form of the noun *karna* 'ear', with the pronominal ending: **karnasmin* (classical Sanskrit locative *karnē*) (Tessitori 1914: 65ff; for more details and examples see Montaut 2016a; 2017). Then this form underwent reduction along the following lines: *karnasmin* > *kaṇṇahī* > *haī* or *kaṇṇai* > *naī* > *nai* > *nē* > *ne*. Tessitori was the first scholar who identified the correct origin of *ne*, but already Trumpp (1872: 401) had traced *kane* ('near, at the edge', then 'to') to the Sanskrit noun *karna* 'ear'. The origin of the ergative marker has later on been accepted by all traditional grammarians (Tivārī 1961; 1966; Saksena 1937; Chatterji 1926; Cātak 1966). Tessitori (1914: 68–70) gives examples of Old Rajasthani such as (23a),²⁰ with a clearly locative meaning, (23b) with an allative meaning, (23c) with a dative meaning and (23d) with an agentive meaning:

- (23a) *cārāi nai nirmala nīra*
road LOC pure water
'A limpid lake close by the road.'
- (23b) *āvyā rā kaṇṇhai*
come.M.PL king LOC/ALL
'[They] came to the Raja.'
- (23c) *te savihū nai karaū paranām*
3PL all.OBL LOC/ALL do.PRS.1SG salutation
'I bow to all of them (in front of/ for).'
- (23d) *adiśvara nai dikṣā lidhi*
Adishvara LOC/ERG consecration.F.SG take.F.SG
'The Adishvara took the consecration.'

20 With the longer form closer to the etymology: *mithyādr̥ṣṭī loka kaṇhai sravai vasiraunahī* (false.look people LOC hermit.M.SG dwell.PRS.3SG NEG) 'A shravaka (hermit) should [does] not live near heretics' (Tessitori 1914, loc.cit.).

The dative/accusative postposition *tāĩ*, *ta.ĩ* or *taĩ* (ताई, तई, तै) is usually derived from Skt. *tāvati* ‘so far’ > *tāmahĩ* > *tāvahĩ* > **tāāĩ* (Cātak 1966: 102²¹, quoting Tessitori 1914: 251), rather than from Skt. *tarite*, which conveys the verbal notion of going through, or *prati* ‘against, toward’ also suggested by Cātak. One still finds in certain contemporary dialects of Garhwali the originally directional meaning of ‘till’ (‘the limit up to which’): *āj tāĩ* ‘till today’.²² Saksena (1937: 224) derives the Awadhi ablative *taĩ/tē* from the instrumental **tatena* of the deictic *tan*. As for *thaĩ* (थै) or *tha.ĩ* (थई), more frequent in Kumaoni, it has also been derived from different sources, mainly from Skt. *tasmai* or *tasmin*, locative forms of the third person or deictic pronoun, and from *tahĩ*, a “periphrastic locative”, as Sharma (1987: 50) calls it, meaning probably a lexical item used as a locative marker. Furthermore, it is also associated by some scholars with the Sanskrit verbal root *sthā* ‘stand’, ‘stay’, or *sthāne* (place.LOC) in the sense ‘near by’ when used as an adverb (Sharma 1987: 136). Interestingly, the same form behaves as an instrumental/ablative in Old Rajasthani (*nakṣatra thaĩ* ‘from the constellation’) and its “extended form” *thānnu* exists in Konkani; Khatre (1966: 124) derives it from *sthā*. The diversity of opinions regarding etymological sources is itself interesting, since it shows the weak semantic link between source and target, and at the same time the important phonetic erosion makes different sources plausible.

The widely used *sañĩ*, which also occurs for both accusative and dative in Kumaoni as well as in Garhwali, is usually derived from the Sanskrit noun *saṃga* ‘company’ in the oblique case (Apabhramsha *saṃge*). Tivārĩ mentions this etymology (1961: 343), whereas Sharma (1987: 50) suggests a possible derivation (which he seems to favour) from Sanskrit *samāna* ‘equal’, with an “analogical extension after the style of *hunĩ*”, another dative-accusative marker derived from the present participle of verb *bhũ* ‘be’, *hunte* ‘being’ (Apabhramsha *hunto/hunte*, cf. 3.1.2). Both grammaticalization paths (from *saṃg* or *saman*), equally plausible, suggest that a same base can result in quite opposite grammatical meanings, since they both are also considered in connection to the derivation of *se*, an instrumental/ablative marker.

Interestingly, the more typical Garhwali marker *khunĩ* (DAT), mentioned by Sharma (1987: 50) as the “Garhwali correlate” of Kumaoni *sañĩ*, is derived by Cātak (1966: 102) from the Sanskrit noun *karṇe* ‘ear’ (**karṇasmin* > Apabhramsha *kañnahĩ*).²³ The same grammaticalization path is also attested for the agentive marker *ne*, and again, points to opposite grammaticalized meanings (source *vs*

21 तावति > तामहिं > *तावहिं > *ताअइ > ताई.

22 Similar meaning in Rajasthani.

23 Cātak mentions that the form *kana.ĩ* is used as a directional marker (*diśābodhak śabd*): *bhaijĩ, kanaĩ chā jāñā* [H. *bhāĩ jĩ, kidhar ko jā rahe ho*] ‘brother HON where to be go] ‘Brother, where are you going?’.

target). Cognates such as *kane/kaṇe* are also found in Gujarati as dative/accusative markers, *kanaī* in Mewati and *kaṇi* in Kumaoni (Cātak 1966: 102) or Eastern Garhwali.

The marker *ku/kū/kū*, a reflex of Standard Hindi *ko*, is consensually (Chatterji 1986; Tivārī 1951; Cātak 1966; Sharma 1987) derived from the Sanskrit word *kakṣa-* 'armpit', a word also considered as the origin of the spatial interrogative *kakh* 'where' in Garhwali, supported by a fusion with the interrogative/indefinite base *k-* common to all Indo-Aryan languages.²⁴

The instrumental/ablative postposition *se*, common to Hindi and Garhwali, or its variant *sī* (INS/ABL) is sometimes derived from *śakyate* 'is able', as specially used to convey the force causing action: *karaṇ kārak mē viśeṣ prayog, śakti ke sūcak* (Cātak 1966: 100). The fact that languages from Rajasthan and Gujarat display a palatal reflex *śū*, along with other alternate forms such as *sū*, *siū*, provides evidence for accepting the derivation from *sākam* suggested by Pischel (1900: 206). Although some have traced its origin to *santo/sunto* 'existing' from the active participle of Skt. *as* 'to be', this opinion is discarded in Tivārī (1961: 342), as well as a derivation from *saṃga* (Kellogg 1876: 132). The most convincing origin according to Cātak (1966), Tivārī (1951), Kellogg (1972/1876) and Chatterji (1986/1926) is from the instrumental form of *sama* 'equal, on par' (> *samena* > *samē* > *saē*).²⁵

The instrumental or ablative *te, tī* (also found in Marathi and Konkani as attested by Khatre 1966, para 218: 123) is sometimes derived from the present participle of the verb *bhū* 'be', a semantically even more empty basis (Apabhramsha *hunti/hunte* < *hontai*). But most of scholars follow Hoernle (1880: 226), who derives it from the locative of the Skt. past participle *tarite* (> *ta:ie* > *te, taī*) 'passed to', hence meaning 'up to', 'upon' and 'from upon' (accordingly, 'from').

The ablative case marker *baṭi* is specific of Pahari languages and also occurs in Kumaoni, with its cognates *biṭe, bāṭ, baṭe, bai*, and Garhwali *baṭikh, biṭa, bai*, always with the same meaning of origin or point of departure. Standard Hindi has the noun *bāṭ* 'way', 'road', mainly used in expressions such as *bāṭ johnā* 'wait', 'look for' but it is never grammaticalized into a directional marker. The noun has remained widely used in Garhwali and it occurs in the meaning of 'way, road':

(25a) *bīc bāṭ mā kani hoye* (Nauṭiyāl 1997: 208)
 middle way LOC how happen.AOR
 'How did it happen in the middle of the way?'

24 And more generally IE indefinite/interrogative markers (**k^vo-*, **k^ve-*, *k^vā-*).

25 The Old Bengali (15th c.) suggests a derivation from *sama*, a word distinct from the one used for 'company' (*saṃga*) in the same verse: *toe sama kariba māi saṃga* (2.OBL with *do.ba* (M.SG) 1SG.OBL company.M.SG) 'I shall have union (do company) with you' (Chatterji 1926: 966).

The meaning of the source noun evokes allative rather than ablative (and so the Hindi expression ‘wait’ < ‘look towards/at the road’), whereas Garhwali as well as Kumaoni use it in a specifically ablative meaning. However, its cognate *bai* is used in Nagari (spoken in Naggar, a small town in Himachal, Kullu district) as an equivalent of Hindi *ko*, for experiencers:

- (25b) *mu=bai pār śobhle lāge*
 1SG=DAT mountains nice touch.AOR
 ‘I liked mountains.’

Both apparently opposed grammatical meanings are plausible in relation to the etymological origin of the marker, since it comes from the verbal root *vṛt* ‘turn’, ‘be’, ‘happen’: *varte* < *vartate* (Sharma 1987: 53 derives it from the participial form *vartmani*). As in the case of *te*, or *se* (see above), or in the grammaticalization path of Kumaoni *huni* suggested in Kellogg (1876: 131) from noun of action of ‘be’ verb, the etymon has a very abstract and vague meaning and grammaticalized into a more precise one. Yet the original meaning of the point of departure (*baṭi* ‘from’) may also convey extended meanings, such as the way through which one comes (*ched baṭi* ‘through the hole’).

The locative *mā* is, like the Hindi *mē*, from the Sanskrit word *madhyam* ‘middle’, which is particularly clear in its Garhwali reflexes *māje*, *māhe*. As for the two genitive markers, *kau* (*kī*) /*kerau* (*kerī*) and *rau* (*rī*), they are consensually all derived from the participle of the verb *kṛ/kar* ‘do’ (< *kia* or *kerā* < *kṛta* ‘done’), implying the genitive of agent. This explains why they agree in gender and number with the head noun.

2.3 Conclusion

Two striking facts can be emphasized: first, the locational origin of most markers, which further specialized into allative or ablative functions. Their specialization is weaker in Garhwali than in Standard Hindi, and we find many markers for the same function as well as various derivations from ‘be’ verb grammaticalized into mutually opposite functions. Secondly, the very abstract, semantically empty origin of many markers may account for the surprisingly diverging grammaticalization paths (*se* an instrumental in Hindi, *saṅi* a dative/accusative in Garhwali/Kumaoni, *thaṅ* a dative/accusative in Garhwali and Kumaoni, an ablative in Old Rajasthani). The same result obtains with extremely precise and concrete source terms such as ‘ear’, which grammaticalized into ergative/instrumental (Garhwali *na*) and dative (Garhwali *khuṅi*, Gujarati *ne*, Punjabi *nū*), a grammaticalization path unattested in Heine and Kuteva (2002). As for the variety of derivations suggested by reliable scholars, it can be explained by the phonetic erosion, particularly intense in words most frequently

used. As mentioned by Bloch (1935: 179), “[t]he effect of this deterioration is that the etymological interpretation of the postpositions is often difficult or impossible”.

3 Other adpositions and other functions of core markers

3.1 Forms and origin of other adpositions

3.1.1 Ergative *la*

Besides the ergative marker *na*, the postposition *la*, echoing the Kumaoni ergative marker *le*, is also used in the same function in some parts of Garhwal, including in Pauri, although the city is not very close to the eastern limit of the district next to Kumaon:

- (26) *mi=la ī film dekhi*
 1SG=ERG this film see.AOR
 ‘I saw this film.’

This alternative agentive marker *la* (as also its Kumaoni cognate *le*) is generally assumed to derive from the Skt. verbal root *lag*, meaning ‘touch, be stuck to, placed’. Some scholars, like Tivārī (1961), have suggested the Skt. verb *labh* ‘to get, obtain’ as an alternative derivation. The regular path runs as follows: *lagya* ‘having come in touch with’ > *lage* > *lai*, *lai* (*le*). Interestingly, this base, glossed by Juyāl (1967: 117) as ‘for the sake of’, ‘with the object of’ is widely used as a dative marker in Indo-Aryan (e.g. Marathi *lā*, Old Hindi *lāi*, *lagi*, *lāgi*), and it also occurs in that function in Garhwali, which illustrates once more the apparently contradictory grammaticalization paths for lexical units of the same origin.

3.1.2 Target markers *tak*, *bānā*

The postposition *bānā* is mentioned by Cātak (1966: 102) as a causal or instrumental marker (*kāraṇ*), but it is used as a synonym of *khātar* (Hindi/Urdu *khātir*, from Arabic) ‘for the sake of’, hence in many contexts ‘because of’.

- (27a) *terā bānā choṛe māi=na bhainā!*
 your sake give.up.AOR 1SG=ERG brother-in-law
 ‘What did I (not) give up for your sake (because of you) brother-in-law!’
 (Nautīyāl 1997: 209)

- (27b) *terī khātar chore syālī bākī bagūrī*
 your sake give.up.AOR sister.in.law Banki Baguri
 ‘I have given up Banki Baguri for you, sister-in-law.’ (Bhaṭṭ 1976: 186)

No etymology is suggested for *bānā*, but one may relate it to the Awadhi beneficiary marker *bādi* (*raurā bādi* ‘for your lordship’, ‘for You’), derived by Saksena (1937: 230) from the Sanskrit verb *vand-* ‘praise’.

As for *tak*, the postposition expressing the limit to reach or point of arrival – ‘up to, till’, it is common to many if not all Indo-Aryan languages, sometimes with the alternate form *talak*. Its grammaticalization from the verbal root *tāk* ‘see, look at’ – according to Sharma (1987: 57), who mentions Turner’s etymological dictionary of Nepali as his source,— presents an interesting grammaticalization path: ‘look at’ > ‘up to’ for allatives, but lacks solidity.²⁶

3.1.3 Dative/accusative markers

The postposition *chāī, chā.ī* (छै, छइ), which occurs both in Garhwali and Kumaoni as an accusative marker, “[i]s the palatalized version of *thāī*” (Sharma 1987: 50), and most probably *ce* and *te* display a similar connection.

The postposition *huṇi*, which also occurs in Kumaoni and alternates with *haṇī*, is derived (Sharma 1987: 50) from the present participle of the Skt. verb *bhū* ‘be’ (Aphabramsha *hunto, hunte*), one more example of a very abstract meaning of an etymon grammaticalizing into an abstract case marker, yet with a definitely allative meaning when used in Garhwali and Kumaoni in such expressions: *ghar huṇi* ‘to home’, *itha huṇi* ‘to this side’.

However, other Indo-Aryan languages present strong counter-examples to the allative grammaticalization path, since they grammaticalized the same word into an ablative marker, such as Old Marathi: *gharhuni* ‘from house’, further transformed into *gharauni*, where the marker is close to a mere case affix.²⁷

As for *kaṇī* and its variant *gaṇī*, they are derived by Sharma (1987: 49) from the Sanskrit word *karṇa* ‘ear’ (**karṇebhi* > *kaṇṇahi*), which is also the basis from which the Garhwali and Hindi ergative/instrumental marker is derived, as well as the dative/accusative *khuṇi*. This raises a problem for cognitive interpretations

26 Heine & Kuteva (2002) do not list verbs of vision as a source for grammatical words. However, one finds it as a source for the presentative in French (*voici* ‘here is’, *voilà* ‘there is’, from the verb *voir* ‘see’). It can also be observed in Hindi *cāh-* ‘want’ derived from the Skt. verbal root *caṣ-* ‘see/look’ at the origin of the modal auxiliary *cāhie* ‘should’.

27 Cf. *ethuni* ‘hence’ (lit. ‘that.from’).

of grammaticalization paths, since the most commonly attested derivations tend to group target markers including dative and often differential accusative markers together as opposed to source markers, e.g. instrumental/ergative/ablative (Malchukov & Narrog 2009). The aspirated form *haṇi*, an alternate form of *saṇi*,²⁸ is accordingly derived from either Skt. *saṃga* 'company' or Skt. *sama* 'equal'.

Grierson (1916: 312ff.) acknowledges the problem of the contradictory behaviour of a given case marker (source *vs* target) when he observes that *gaṇī* like *khuṇi* is an equivalent of Hindi *ko* and therefore behaves as a dative but, "curiously enough", also as an ablative. However, the examples he gives for the ablative meaning suggest a translation from Hindi where the multi-purpose *se* would translate *gaṇī*: *ek banyā gaṇī* 'from a shopkeeper' [H. *ek banīe se*], *kāi gaṇī* 'from whom' [H. *kis-se*] (Grierson 1916: 314), which one can interpret as 'at a shopkeeper', 'at whom', a locative meaning. Whenever the meaning of ablative (source) is non-ambiguous, Grierson gives examples with other case markers, for instance *mākoī*, which is given by Cātak (1966: 104) as an ablative: *cullā māko khānā* 'food from the oven', equivalent of Hindi *cūlhe mẽ se/ cūlhe kā khānā*, or *uṃḍe* 'from': *kuwā-uṃḍe pānī* 'water from the well'.

- (28) *ve=mākoī vo rupyā le le*
 3SG=ABL these rupies take take.IMP
 'Take the rupees from him!' (Grierson 1916: 314)

Interestingly, Grierson also gives an example of *gaṇī* as a locative marker, which confirms the originally locative meaning (consistent with its origin) of its more frequent function of dative/allative marker:

- (29) *nānā kuṛā=gaṇī ran*
 small house=LOC stay.PRS.3SG
 'He lives in a small house.' (Grierson 1916: 313)

The use of *gaṇī* as an allative like *saṇi* (in such expressions as "to turn to", "throw at"), apart from being a DAT/ACC marker, and initially as locative echoes the grammaticalization path of Skt. *kaṅṣa* 'side', which also served initially as a locative marker and further grammaticalized into allative. It also echoes the grammaticalization path of its etymon *karṇa* 'ear' from the initial locative form, acquiring subsequently the function of allative or dative (as Gujarati *ne*), and later grammaticalized into agent (Garhwali, Hindi). It also suggests that the seemingly paradoxical grammaticalization of Skt. *karṇe* into both an agent marker (*ne*) and at the same time and in the same language as a dative/allative marker (*khuṇī*) is not an isolated

28 Although some scholars relate it to *gaṇī*.

case. Whatever the origin of *saṇi* and its variant *haṇi*, it similarly illustrates symmetrically opposed paths, since it is used in Garhwali Kumaoni as a dative marker (as well as the dialectical form *hnī* in the region of Ramoli²⁹), while the cognate forms *san*, *sañ* are used in Braj and Awadhi as instrumental markers. This confirms our first conclusions reached regarding core case markers (see 2.3).

3.1.4 The comitative and relator equivalents of Hindi *se*

The same form as Garhwali and Kumaoni comitative — *daḡar*, *daḡarī*, *daḡārā*, *daḡare*, *diḡar* — is found in Awadhi as a postposition too but with the meaning ‘through’, ‘by means of’³⁰ and as an independent noun meaning ‘way’ (Saksena 1937: 128). In the latter meaning it is registered also in Chattisgarhi (*ḡaḡr.e* ‘on the way’). No etymology is given for the Garhwali comitative *daḡare* and its cognates, derived from the “deśi” word with retroflex *ḡaḡar* (Sharma 1987: 54), but the Garhwali noun *daḡūrī* ‘society, company’ (30c) and the Kumaoni noun *daḡariyāna* ‘companions’ suggests the shift from the meaning ‘way’ to the comitative meaning. Again, one may notice the diverging paths borrowed by the synonymic words *bāṭ* and *daḡar* ‘way’, the former grammaticalized into an ablative (source), the latter into a comitative marker.

(30a) *tyār daḡar ar kū kū rahandan*
2SG.GEN with and who who stay.PRS.3PL
‘Who else are living with you?’

(30b) *mañ daḡarā kar devā guru jī rājulā saukyān*
1SG with make give.IMP guru HON Rajula Saukyan
‘Join Rajula Saukyan with me Guru ji!’ (Nautiyāl 1997: 29)

(30c) *chore rāṇiyō kī daḡūrī*
abandon.AOR queen.OBL.PL GEN society
‘I left the society of queens.’ (Bhaṭṭ 1976: 186)

The word *daḡar* is also used as a mere relator in verbal locutions involving reciprocity in the same way as the relator *se*:

(31a) *meru ū daḡar kvī ristā / āṇu-jāṇu uṭhan-baiṭhan nī ca*
1SG.GEN 3SG with any relation / come-go rise-sit NEG is
[H. *merā us=se koī ristā / āna-jānā uṭhnā-baiṭhnā nahī hai*]
‘I have no relation, no frequenting with him.’

29 See for instance *tve hnī dine* [2SG DAT gave] ‘he gave you’ (Cātak 1966: 102).

30 Standard Hindi has *dvārā* (< Sanskrit *dvāra* ‘door’) in the meaning of ‘by (means of)’.

- (31b) *merū dakṭār dagar bāt hoge*
 1SG.GEN doctor with speech be.go.AOR
 [H. *merī dākṭar se bāt ho gāī*]
 'I spoke to the doctor.'

Besides, the words *neṛe* and *nal*, also used as adverbs (32a), indicate vicinity, as does Hindi *pās* (< Skt. *pārśva* 'side'), the *-l* inflection for locative in (32b) and (36) below being possibly derived from *nal*.

- (32a) *aurū nerū aige*
 ECHO.WORD close come.go.AOR
 'He came in the vicinity.' (Nautiyāl 1997: 209)
- (32b) *tarāḍā ghām mā jītū chai-l baiṭhige*
 scorching sun LOC Jitu shadow-LOC sit.go.AOR
 'In the scorching sun Jitu sat in/at the shade.' (Nautiyāl 1997: 209)

A number of other postpositions also differ from Hindi, such as *ūḍo* or *udo* 'below' from Sanskrit *adhaḥ*, according to Cātak (1966: 138), *bhuṭ* 'down, under' from Sanskrit *bhūmi* (see example in footnote 19), *mathe* 'above, on' from Skt. *masta-* 'skull, head' (> Pkr *mattha*, the final *-e* resulting from a later development), *aic* 'above' from Sanskrit *uccais* 'high', *ubbo* 'on top of/above', from Skt. *ūrdhva* 'upright'. This adds to the sufficient evidence that the language is deeply related to Indo-Aryan, having evolved, as all New Indo-Aryan languages, from Sanskrit through the Middle Indo-Aryan (Prakrits) stage.

Quite a few postpositions are common in Garhwali, Hindi and some other NIA, such as *par* < Skt. **uppari* (and its variant *phar*) as well as the related locutional postposition *paran* 'from above' (H. *par se*).

To sum up, this brief review of non-core case markers confirms the preceding findings: some markers are derived from very concrete words, such as parts of the body or landmarks, a common path for forging directionals by an abstraction process ('ear', 'head', 'armpit', 'earth'), others on the contrary derive from very abstract, almost lexically empty words, such as the various participial forms of the verb 'be'.

3.2 Other functions of core markers

3.2.1 *na* as a marker for subjects of obligation or modal future

We saw above in passing that in Garhwali, the so-called ergative marker can mark the subject in obligative clauses, as in (6a), in contrast with Standard Hindi, which uses the dative marker for this purpose. I am therefore maintaining the gloss

ERG for the Garhwali examples with *-na*, a case marker only used for agents and instruments. Ergative-marked subjects are the standard obligative constructions in Garhwali, they occur with both transitive (33b) and intransitive verbs (33a) and they control coreference with the converb, as in (33b).

In each sentence in (33) a different ending of the same infinitive form occurs: *-ā* in (a), *-a* in (b), and *-Ø* in (c):

- (33a) *mī=na jāṇā ijā bvai tātā lūhāgarḥ*
 1SG=ERG go.INF today mother Tata Luhagarḥ
 [H. *māi jāṅgā /mujhe jāṇā hogā tātā lūhāgarḥ*]
 ‘I will/must go today, mother, to Tata in Luhagarḥ.’ (Nautiyāl 1997: 111)
- (33b) *maī=na tū biyā-ik rāḡilī vairāt lāṇa*
 1SG=ERG 2SG marry-CVB Rangili Vairat bring.INF
 ‘I will/must bring you, Rangili, to Vairat (after) marrying you.’ (Nautiyāl 1997: 28)³¹
- (33c) *maī=na jīṇ marūṇ jiyā saukyānī des jāṇ*
 1SG=ERG live.SBJV.1SG die.SBJV.1SG mother Saukyani country go.INF
 ‘Whether I live or die, mother, I have to go to the Saukyan’s country’.
 (Nautiyāl 1997: 29)

The most frequent obligative patterns involve the subjunctive (from the synthetic Sanskrit present with personal endings),³² a finite mood form which in Modern Garhwali expresses both obligation and wish with a nominative subject: *āgyārthak aur icchārthak rūp bhinn nahī hai* ‘the forms conveying obligation and wish do not differ’ (Cātak 1966: 131). Significantly in the non-nominative pattern as illustrated in (33), the verb has no personal ending, and such constructions are particularly frequent in the traditional folk songs and epics but they also occur in colloquial Garhwali, cf. (6a). The ending in *-aṇ* is analyzed by Cātak (1966: 130–132) as a shortening of the future [passive] participle (*bhaviṣya kāl kā krdant*) in *-ṇya*,³³ which agreed with the object, if any, in gender and number and used to convey, along with a future meaning, a stronger obligative meaning (*kartavya-akartavya kā vicār viśeṣ mahatva kā hotā hai*).³⁴ I have no explanation

31 Similarly, also with a pronominal human object 3rd person singular, oblique form unmarked (of 3SG.F: *vī/vī*), in the same context: *maī.na vā vivāik lauṇa, ḍolā ḍhasakaik* (Nautiyāl 1997: 28), in the meaning of a strong will: ‘I will bring her after marrying, installed on a palanquin.’

32 First two verbs of (33c), conveying also eventuality, as in Hindi.

33 Obviously derived from Skt. *-anīya*.

34 With the following example: *hamārī bāt kai=na nī jāṇ* (our speech INDF=ERG NEG know) [H. *hamārī bāt kisī=ko nahī jānī hai/koī nahī jānegā*] ‘Nobody should/will/could know our conversation.’

for the *-ā* ending in (33a) except that it is from a separate piece in short verses and might be due to metrical factors.

Almost all occurrences of this latter pattern involve a 1st-person subject and convey a strong involvement of the speaker — personal intention and will of doing what (s)he considers a personal obligation, whereas the bare subjunctive pattern (see above) conveys a simple obligation or otherwise the future, which can also be translated by the Hindi obligative construction <S *ko* (DAT)... *hai* (is)>, for instance in the mother's reply to the daughter's requirement to go (34) in the same context as (33b–c). The speaker, far from agreeing, since she does not want her daughter to go, shows of course no personal involvement, hence the future:

- (34) *tū jālī merī pothlī huniyō kā des !*
 2SG go.FUT.2SG.F my girl Hunyas of country
 [H. *tū jāegī merī beṭī huniyō ke deś*]
 'You will go my daughter to the country of the Hunyas!' (Nautiyāl 1997: 29)

These facts, far from being a bizarre peculiarity of the language, are strongly reminiscent of the history of the future and perfect in Indo-Aryan. The origin of the ergative pattern is well-known and stems from a nominal sentence with past participle and instrumental agent, a case to be replaced later on by various postpositional markers, generally locative in their origin. Lesser known is the fact that the *-b/-v-* future of Eastern New Indo-Aryan languages also comes from a nominal sentence with oblique agent and a future passive participle (or gerundive) and has long conveyed its initial obligative meaning, until it shifted to potential and future. With the shift to a purely temporal meaning, the syntactic pattern with oblique agent shifted towards a nominative pattern, but languages maintaining the modal meaning also tended to hold on to the oblique agent, such as Marathi, the only language in the West of the Aryandom, which retained the original syntactic pattern — with the verbal form continuing the Sanskrit gerundive — for potential, and Garhwali for obligation/will (more details in Montaut 2016a; Montaut 2017).³⁵

35 More generally, the Romance branch of Indo-European languages also displays the same evolution from nominal sentence with participle and non-canonical agent, but the shift towards a nominative pattern was associated to the use of the 'have' auxiliary, in the future as well as in the perfect. Latin and Sanskrit patterned similarly (oblique agent—past participle or gerundive) for both the periphrastic past and the obligative construction, except that the case used for marking the agent was the dative, the case also used in possessive constructions (hence the notion of possessive meaning in Benveniste's (1966) pioneering paper on the possessive meaning of the construction with the perfect participle of a transitive verb).

3.2.2 Other modal meanings merging volitional future and inability: *se/na/baṭi*

Other non-finite verbal forms occur with the same case marker *na*, presenting again clearly not a canonical ergative. The verb ending in (35a) is the same as in (33) above, the meaning more akin to incapacitive. However, in (35b), very close in meaning and similarly displaying a non-canonical subject, the verb has the future participle form (*-ṇya*) and the subject (incapacitated agent) is in the genitive under the postposition *karī*:

- (35a) *maĩ=na tvai=ku nĩ auṇ*
 1SG=ERG/INS 2SG=DAT NEG come
 ‘I won’t/can’t/am not to come for (the sake of) you.’ (Nautiyāl 1997: 149)
- (35b) *terā karī nĩ honya*
 your GEN NEG be.GRD
 ‘You won’t be able/willing to do/you are not to do.’ (‘It is not to be done by/ of you.’) (Cātak 1966: 101)

It is noticeable that Hindi uses in such cases infinitive with the genitive *kā* (*maĩ tere lie nahĩ āne kī* 1SG 2SG for NEG come.INF GEN.F.SG), a relatively semantically bleached case marker, whereas the instrumental is also used for more clearly incapacitive meaning with intransitives both in Hindi and Garhwali, the latter employing the marker borrowed from Hindi:³⁶

- (36) *saberĩ uṭhĩk pāñĩ ko bāṭhā mūḍ mā tharĩ ganga-!*
 morning get.up.CVB water of pot head on set.CVB Ganga-LOC
jaik laṇnũ maĩ=se nĩ hoṇu (Nautiyāl 1997: 148)
 go.CVB bring.INF 1SG=INS NEG be.INF
 [H. ... *pāñĩ kā gharā ... lānā mujh=se nahĩ hogā* (lit. *honā*)]
 ‘I won’t be able (refuse) to get up in the morning and bring the pot of water on my head from the Ganga.’

36 This instrumental marking of the non-canonical agent in modal passive constructions or with intransitive verbs is also attested in medieval Hindi, for instance, in Kabir:
sāĩ sũ sab hota hai, bande thaĩ kuch nāhĩ (pada 38, Vaudeville 1957: 58)
 Lord INS all be PRS.3.SG slave INS something NEG
 ‘The Lord can [do] everything, the slave is not able of [doing] anything.’

The ablative *baṭi*, which is used as the reference or starting point in comparative patterns (as Hindi *se*),³⁷ and also used as an instrumental for a secondary agent with causative verbs (as Hindi *se*), is also used as a non-canonical agent of medio-passive intransitive verbs (as Hindi *se*) to mark an agent devoid of deliberate will. The path seems to be the following: from a more or less spatial origin with the word 'hand' (by/from hand) in (37a) to the straightforward meaning of non-volitional, non-intentional agent in (37b):

- (37a) *merā hāth bāṭ tamro gilās photyau*
 my hand ABL/INS your glass be.broken.AOR
 'I broke your glass by mistake/your glass was broken by/through my hand.'
- (37b) *mā bāṭ bhūl gayo*
 1SG ABL/INS forget went
 [Punjabi Hindi *mujhse bhūl gayā*]
 'I forgot.' (unwillingly)

Both Kumaoni (38) and Garhwali may also use the dative marker *huṇi* in alternation with *thaṭ* before *baṭi/bai* to mark the secondary agent, as if the dative marker such as 'to be' were a simple oblique form with no special directional meaning (like Hindi *us* before the ablative marker *se*). This is consistent with the high poly-functionality of such case markers and with their non-directional origin ('being', 'existing', 'way'):

- (38) *tum vi thaṭ baṭi / huṇi bai ituk kām iai*
 2SG 3SG.OBL ACC/DAT ABL ACC/DAT INS this.much work even
ni karnai śaka
 NEG do can.AOR
 'You could not get even this much work from him.' (Sharma 1987: 54)

3.3 Conclusions

Summarizing evidence gathered in section 3, we observe in Garhwali a consistent tendency to use adpositional markers for non-canonical agents: ERG/INS for obligation+will, ABL for non-volitional/non-deliberate agents, and various dative

37 As well as in Kumaoni: *māṭ bai/baṭi thul chi* [1SG ABL tall be.PRS.3SG] 'He is taller than me' (Sharma 1987: 54). The form *tē/tī* is naturally also used for comparative and superlative: *nauniyō māje=n tū sab=tī syāñī chai* (girls among=INS/ABL 2SG all=ABL be.PRS.2SG) 'You are the most beautiful of (from among) all the girls.' (Cātak 1966: 104).

postpositions for the non-canonical subjects such as dative experiencers (see section 2). The most noticeable contrast with Hindi is the ERG/INST marking of subjects of obligation patterns, yet the modal future meaning associated with the construction echoes the original pseudo-ergative modal constructions in the earlier phase of Indo-Aryan (Montaut 1997).

Another striking fact, confirming the conclusions in 2.3, is the diverging grammaticalization paths from a given lexical item, across dialects or across languages (*lag*, for instance, grammaticalizes into a dative/beneficiary/target in Marathi and many Hindi dialects, but as an ergative and source in Kumaoni Garhwali). It may to a large extent be accounted for by the semantic content of the etymological source terms, in many cases very vague or even abstract, but concrete in case of the locative postpositions stemming from the nouns ‘ear’ and ‘armpit’, initially grammaticalized into locative markers. When Hindi and its “dialects” started emerging from of an early New-Indo-Aryan phase where the basic marking of non-nominative nouns was a syncretic oblique, the sentence was a predication either with unmarked nouns for functional cases (nominative and accusative) or with loosely related obliques, that is, with the nominal entities which were not canonical subject or canonical object. Any marker derived from any locative or abstract base, such as ‘to be’, or from ‘equal’, could do the job, until it took on more specific roles as the language evolved. One still finds remnants of this transitional state (before role specialization) in Bangaru, a non-written language spoken in Haryana, related to Punjabi more than to Hindi. In the northern variety of this language the same marker *nae* (cognate of Hindi *ne*) works as accusative, instrumental, or ergative marker, and expresses target roles as well as source roles, whereas in the southern variety *tī* (ablative in Garhwali) also expresses both source and target roles (yet not ergative agent):

- (42a) *kutte nae ḍaṁḍe nae mārya*
 dog ACC stick INS strike.INS
 ‘He struck the dog with a stick.’ (Singh 1970: 69)
- (42b) *balkā nae toriyā hongē*
 child.M.PL ERG break PRSUMP.3M.PL
 ‘The children have probably broken [it].’ (Ibid.)
- (42c) *rupay tī us=tī le lo*
 money ACC 3SG=ABL take take.IMP
 ‘Take the money from him.’ (Tivārī 1961: 177)

4 The “evidential” *bal*

The marker *bal* has in Garhwali a wide range of functions, although not much commented in most grammars and never labelled a quotative nor an evidential. However, since all the functions I could trace ultimately belong to the well-known array of meanings conveyed by most of the evidential markers in cross-linguistic studies, I suggest to identify it as an evidential marker, rather than a quotative (a category well known in Dravidian languages), since most Indian languages which have so-called quotative markers use them primarily to report speech, thought, and introduce conditionals.

The only work devoted to *bal*, to my knowledge, a three pages presentation by Deveś Jośī in 2011, emphasizes its discourse properties: *bal* “conveys conviviality, rhythm, natural fluency, a pleasant feeling” (*chapchapī, arthāt tarāvātjanya sukhānubhūti*, p.11),³⁸ rather than the classical meanings associated with quotatives. The article is significantly entitled *Garhvālī kā bal hai bal* (The strength of Garhwali is *bal*), with a play on words, since *bal* as a noun means ‘strength’. Cātak (1966: 139ff), however, with only a few lines devoted to the marker scattered in the section on *sambandhvācak* (conjunctions),³⁹ primarily identifies it as a marker of reported speech (*kisī aparokṣ vyakti kī vānī ko uddhṛt karne ke lie*), a function also mentioned in Jośī (2011), yet with no particular emphasis: *bal* is “used for quoting examples” (*udāharaṇ dene ke lie*) or as a synonym of *iti* “for giving the feeling of somebody else’s statement” (*vaktā dvārā anya puruṣ kathan kā bhāv prakāṣ karne ke lie*), as in (43):

- (43) *mantrī=jī tumrā gaūmā ayā bal*
 Minister=HON POSS.2 village.LOC come.AOR *bal*
 ‘[I heard /it’s said that] the minister came to your village.’

4.1 Range of uses and functions

Bal is, however, not so often used to simply introduce reported speech, as in Dravidian languages where a quotative is the standard device corresponding to complementizer *ki* ‘that’ of Persian origin, which is used in Indo-Aryan after speech verbs. Whenever *bal* occurs in this meaning in Garhwali, far less

38 Deveś Jośī (2011) wrote this paper in a deliberately non-academic way in a collection of papers intended for a wide audience for the promotion of the region, its environment, tourism and culture.

39 He gives a list of twelve items, including such words as *ar* ‘and’, *phir* ‘then’, *ta* ‘then, so’, *cā* ‘if/whether’, *ki* ‘that’, *kilai* ‘why’, *jū/ji* ‘if’, *par* ‘but’.

frequently than the finite verb *bolṅu* (*bole* ‘said’, *boldi* ‘says’), with or without the complementizer *ki*, it always conveys some additional meaning such as indirect knowledge of the reported information with some doubt regarding its authenticity:

- (44) *tvai=na bole bal ki maĩ=na terā rupyā deṅan*
 2SG=ERG say.AOR *bal* that 1SG=ERG your rupees give.GRD.PL
 [Somebody told me that] ‘You said that I should give (back) your money.’

This example, given by Cātak as an illustration of the indirect speech, is all the more interesting as it shows, on the contrary, that quoting is provided by the verb *bol* ‘speak/say’ followed by the complementizer *ki* introducing the content of the reported speech (‘I should give back...’) while *bal* adds some epistemic modality (doubt about the origin and validity of the fact that ‘you said it’). This is confirmed by Jošī (2011: 12) who suggests that *bal* is used for conveying doubts about the credibility of some content (*kisī kathan kī viśvāsnīyatā par sandeh vyakt karne ke lie*), with the following example:

- (45) *mantrī=jī cha bal āyā tumrā gaūmā (kyā kyā*
 Minister=HON AUX *bal* come.PFV POSS.2 village (what what
ghoṣṇa kari gayā)
 declaration do.CVB go.AOR)
 ‘The Minister came (really?) to your village (left having made what announcements).’

The speaker’s doubt about the reported speech he presents as a hearsay is part of the core meanings of evidentials in languages with this category (Aikhenwald 2004), as well as the use of the evidential marker to introduce tales or dreams (things which are not related with factual truth), another function of the Garhwali *bal*:

- (46) *ek cha bal rajjā. Rajjai chai bal dvī rāṅī*
 one be.PST.SG *bal* king king.OBL be.PST.PL *bal* two queens
 ‘(Once upon a time) there was a king. The king had two queens.’ (Jošī 2011: 12)

Bal occurs quite rarely in traditional folk songs, and still more rarely with the hearsay meaning:

- (47a) *tumhārī bataīdī bal vā boṅ śāvnī ka hāth jure*
 your speaking *bal* DEM sowing Shavani of hand be.tied.AOR
 ‘Hearing what you say, these sowings are in the hand of Shauni.’
 (Bhaṭṭ 1976: 183)

But we significantly find it four times in the few lines commenting the dialogue between the Saint Gorakhnath and the hero Malusha, who has left his kingdom and wife and travelled to the great Guru to find a way for marrying Rajula (Naṭtiyāl 1997: 31):

- (47b) *bal sun mālūsāī ḍola baṇaulā* [H. *sun Mālūsāhī, ḍolī banāḍge*]
bal listen.IMP Malushāī palkin make.FUT
 ‘Listen Malushahi, we will make a marriage palanquin,’
vakh par kvī nauṇī baiṭhauḷo ar nau dharlā rājulā
 there on some girl make.sit.FUT and name put.FUT Rajula
 ‘install any girl on it and name her Rajula.’
bal gurujī nau ta dharlā rājulā,
 [H. *gurujī, nām to rakh lēge rājulā*]
bal guru.HON name PTCL put.FUT Rajula
 ‘Guruji, we will give the name Rajulā.’
par nakh sikh kakhān laulā vī janā?
 but toe.nails head (= features) where.from bring.FUT her like
 ‘But where from will you bring features like hers?’
bal acchā tvai triyā bauḷ lagī? bal hā
bal well 2SG.DAT women craze touch.AOR *bal* yes
 [H. ‘*acchā, tujhe baulapan ho gayā na?*’ – ‘*hā*’]
 ‘*Bal* well, you have got crazy about women, right? *Bal*, yes.’
ta le merī boksārī vidyā pahle
 then take.IMP my boksari knowledge first ...
 ‘Then take first my Boksari knowledge ...’

The dialogue between the two starts before this exchange, without any occurrence of *bal*, and there is obviously no hearsay involved and no doubt about the validity of the speech content, so that we may consider this heavy density of the marker in this crucial section as a marker of emphasis, a function mentioned by Jošī (2011) “to emphasize a given expression” (*kisī śabd par zor dene ke lie*) with the following example, a function close to the mirative meanings, very often associated with evidentials:

- (48) *maī=na aṇṇī ākhyū dekhī bal*
 1SG=ERG REFL eye.OBL.PL see.AOR *bal*
 ‘I have **seen** that with my own eyes, (I swear it).’ (Jošī 2011)

Jošī also mentions the optional use (*anāvāśyak prayog*) for adding glow and life (*lālitya ke moh ke kāraṇ*) with an example (49) where the hearsay is conveyed not by *bal*, but by the periphrastic expression ‘comes to hear’, *bal* only making the narration more lively:

- (49) *ṭeṣṭ ta dīlī bal par agnāi apaṇu apaṇu bhāg*
 exam PTCL give.FUT *bal* but afterwards REFL REFL destiny
 ‘He sure will take (lit. give) the exam *bal* but afterwards each one’s one’s fate.’
thorā bhot sifāris bhī chaīch par sunan-mā auṇu ki
 little much bribe too throw.AOR but hear.INF-in come.PRES that
noṭ chan bal calaṇā
 note be.PAST *bal* go
 ‘He has also done some bribing but it is heard that the money *bal* was gone.’

As a marker of conviviality in intersubjective exchange, this use of *bal* echoes the phatic function (cf. expressions such as ‘you see’, ‘you know’, ‘figure it out’, ‘just imagine’, etc.), which, although not mentioned as a common one in the literature on evidentiality, is not totally inconsistent with its emphatic or mirative meanings, and could account also for what Joṣī finally calls a feeling of indifference or detachment (*taṭasthatā kā bhāv*):

- (50) *sarkār dām ca bal yakh baṇaulī. – hā bal.*
 government dam is *bal* here make.FUT yes *bal*
 ‘The government say they will make a dam here. – Yes *bal* (so what/might be).’
tumāru pānī tumhī=taī becṇai iskīm ca bal baṇṇī
 your water 2PL=DAT sell.INF scheme.F.SG is *bal* be.made.INF.F.SG
 ‘They say there is to be a scheme for selling your water to you.’
phyer tumāru hamāru bhaviṣya kyā holū? – Prabhū icchā bal!
 then your our future INT be.FUT Lord desire *bal*
 ‘Then what will be our future? – [It is] God’s will *bal!*’ (Joṣī 2011: 13)

Interestingly, *bal* seems to be taught in schools as a *kāarak*. The definition by the school boy, asked by the school inspector is that “*bal* is such a case *bal* that it brings a change in the situation of a thing *bal* or tries to do so *bal*”. This enigmatic definition may point to the subjective emphasis (“a change”) added by the word in the everyday speech of Garhwali speakers.⁴⁰

40 *bal* vo *kāarak* hai *bal* jo kisī vastu kī sthiti mē parivartan kartā hai *bal* yā parivartan karne kī kośīś kartā hai *bal*. A quite surprising definition, as well as the teacher’s comment: ‘I had to make clear that the first *bal* in the definition of the school boy is a Tibetan (*bhautik*) *bal*, the others Garhwali *bal*.’ Meaning probably that the *bal* after “is a case” suggests that one is quoting some authority, whereas the others are adding emphasis and strength to the affirmation.

4.2 Origin of *bal*

The contact with Dravidian languages is often called upon to explain the presence of quotatives in Indo-Aryan languages such as Marathi or Bengali (Meenakshi 1986). This explanation must however be ruled out for Garhwali for geographical but also for empirical reasons: no conditional form or meaning is attested; as opposed to Tamil *enru* and Dakhini *bolke*, *bal* does not behave as a complementizer with speech verbs, nor does it occur in final position in the complement clause before the finite main verb.⁴¹

Bal is derived by Cātak (1966: 141) from the Sanskrit *manye* (> Prakrit *bane/bale*), but one cannot totally discard its relation to the verb *bol* 'speak', especially if we consider the Kumaoni evidential markers. There are two of them, clearly explained by Manohar Śyām Jośī (1995: 7) in the glossary of his Hindi novel *Kasap* (a Kumaoni untranslatable word meaning something like 'don't know'), which makes a lavish use of both forms in his Kumaoni Hindi. One is *kahā* (51a), a marker of insistence (*āgrah kā sūcak*), the other is *bal* (51b), a marker of hearsay (somebody else said so, *aisā kisī aur ne kahā*). Both examples are given with the author's Hindi gloss, (51c) is from the novel itself:

- (51a) *baṛī sundar dikhtī hai kahā*
 great.F beautiful seem PRS *kahā*
 [H. *māi kah rahā hū̃/, vah baṛī sundar lagtī hai*
 1SG say PROG PRS.3SG great beautiful seem.F.SG PRS]
 'I say/believe me, she really does look very beautiful.'
- (51b) *baṛī sundar dikhtī hai bal* [H. *sunā, vah baṛī sundar lagtī hai*]
 great.F beautiful seem PRS *bal* (heard 3SG great beautiful seem PRS)
 'It is said / I heard that she looks very beautiful.'
- (51c) *īsq mē aisā bhī hone vālā thahrā bal*
 love in such too be IMMEDIATE.FUT stay.AOR *bal*
 (H. *aisā bhī honā thā*)
 'One says, it is bound to happen so in love.' (Such a thing too does really happen in love.)

The neighbouring Himalayan languages display more classical forms of quotatives, used to complement a verb 'say' or 'ask', and placed at the end of the complement clause and just before the following verb of the main clause, in Nepali (52a), as well as in Tibetan (52b). In both languages the form of the quotative is clearly related to the verb 'say' (Nepali *bhanera* is a gerund of verb base *bhan-* 'say', and Tibetan *ze* is derived from verb *zer* 'say, say to oneself').

41 *o sabā ātū bolke bolyā* 3SG come.ISG QUOT say.AOR 'He said he will come tomorrow.'

- (52a) *ma sahar janchu bhanera bhanyo (... sodhyo)*
 1SG city go.PRS.1SG QUOT say.AOR.3SG. ask.AOR.3SG
 ‘He said that (he asked if) he was going (I’m going) to the city.’
 (Riccardi 2003: 577)⁴²
- (52b) *khong lhasa-r ‘gro-gi-yin ze lab-song*
 3SG Lhasa-to go-FUT-1SG QUOT say-CONSTATIVE.3SG
 ‘S/he said that s/he will go to Lhasa.’ (lit. s/he said ‘[I] will go to Lhasa’)
 (personal communication from a Tibetan colleague in Inalco, Paris)

Darjeeling Nepali however displays mixed constructions, with both a complementizer *ki* after the ‘say’ verb, and final *bhanera* in the complement clause, a position also found in Garhwali, although with a different meaning:

- (52c) *us=le sodhyo ki ma sahar jānchu bhanerā*
 3SG=ERG ask.AOR that 1SG city go.PRS.1SG QUOT
 ‘He asked whether he (I) should/may go to the city.’
 (personnal communication of a native speaker)

The Garhwali marker *bal* behaves more as a discourse particle with evidential meanings than as a quotative, which makes its derivation from Skt. *manye* not unpalusible despite the phonetic problems it raises. More research on the topic is obviously needed before coming to a sound conclusion both about its origin and its real function in the language.

5 Conclusion

The study of the proliferating Garhwali case markers, only part of which are related to Hindi markers, has shown that various etymological bases can grammaticalize into practically any case function, the most spectacular being the word ‘ear’, grammaticalized into an ergative/instrumental (*na*), and a dative (*khuni*). The large number of markers derived from participial forms of the verb ‘be’ is on the other hand a striking example of an etymological source so vague and abstract that one expects widely diverging results of the grammaticalization process. These facts, curious in appearance, are however confirmed by a wider comparison with other Indo-Aryan languages, where a common set of basic markers

42 The form *bhanne* occurs as complementing a noun of speech:
pradhān mantri nepāl pharkanubhayo bhanne samācār āyo
 prime minister Nepal return.3SG say.INF/QUOT news come.3SG.AOR news hear.1SG.AOR
 ‘The news came that the Prime Minister had returned to Nepal.’

ends up conveying very different and often opposed semantic roles, a situation unknown in other Indo-European languages. Whereas the study of case markers tells us a story similar in Garhwali, in Hindi and in other Indo-Aryan languages, the study of the evidential *bal* suggests an altogether unique evolution, since no equivalent of its major functions is found in Standard Hindi nor in any Indo-Aryan language, even those with quotatives, whatever the apparent analogy of forms may be (if we accept a *bol* > *bal* etymology). The attempt made here to disintricate the complexity of *avyay* words is a first step, in the hope of further research by other scholars.

Bibliography

- Aikhenwald, Alexandra 2004. *Evidentiality*. Oxford: Oxford University Press.
- Bābulkar, Mohanlāl 1996. *Garhvālī lok-sāhitya kā vivecnātmak adhyayan* (A critical Study of the Garhwali folklore). Prayāg: Hindī Sāhitya Sammelan.
- Bāhrī, Hardev 1966. *Hindī grāmīṇ boliyā* (Rural Dialects of Hindi). Ilāhābād: Kitāb Mahal.
- Bahugūṇa, Abodh Baṃdhu 1960. *Garhvālī vyākaraṇ kī rūprekhā* (An Outline of Garhwali Grammar). New Delhi: Garhwali Prakashan.
- Benveniste, Emile 1966. “La construction passive du parfait transitif”, in: *Problèmes de linguistique générale*. T.I. Paris: Gallimard, 176–186.
- Bhaṭṭ, Haridatt ‘Śaileś’ 1976. *Garhvālī bhāṣā aur uskā sāhitya* (The Garhwali Language and its Literature). Lucknow: Uttar Pradesh śāsan.
- Bloch, Jules 1935. *Indo-Aryan from the Vedas to Modern Times*. Paris: Adrien-Maisonneuve.
- Butt, Myriam 2006. “The Dative-Ergative Connexion”, in: Bonami, Olivier & Patricia Cabredo Hofherr (eds): *Empirical Issues in Syntax and Semantics* 6, 69–92 (<http://www.cssp.cnrs.fr/eiss6/index_en.html>, accessed: March 2, 2019).
- Callewaert, Winand (in collaboration with Swapna Sharma & Dieter Taillieu) 2000. *The Millenium Kabīr Vānī. A Collection of pad-s*. Delhi: Manohar.
- Cātak, Goviṃd 1966. *Madhya pahārī kā bhāṣāśāstrīy adhyayan* (Linguistic Study of Central Pahari). Delhi: Radhakrishna Prakashan.
- Cātak, Goviṃd 2000. *Garhvālī loggīt* (Garhwali folk songs). (Ādivāsī bhāṣā sāhitya prakalpa/Indian Literature in Tribal Languages.) Delhi: Sāhitya Akādāmī.
- Chatterji, Suniti Kumar 1986 [1926]. *The Origin and Development of the Bengali Language*. Delhi: Rupa.
- Dalrymple, Mary & Irina Nikolaeva 2011. *Objects and Information Structure*. (Cambridge Studies in Linguistics 131.) Cambridge: Cambridge University Press.

- Dingemans, Mark 2012. “Advances in the Cross-Linguistic Study of Ideophones”, *Language and Linguistic Compass* 6 (10): 654–672.
- Grierson, George Abraham 1967 [1916]. *Specimens of the Pahārī Languages and Gujurī. Linguistic Survey of India*, vol. 9, part IV. Delhi: Motilal Banarsidass.
- Harmond, David 1996. “Losing Species, Losing Languages: Connections Between Biological and Linguistic Diversity”, *Southwest Journal of Linguistics* 15: 89–108.
- Heine, Bernd & Tania Kuteva 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hoernle, August F. Rudolf 1991 [1880]. *A Comparative Grammar of the Gaudian Languages with Special Reference to the Eastern Hindī*. Delhi: Asian Educational Services.
- Jošī, Deveś 2011. “Garhvālī kā bal hai bal (The strength of Garhwali is the [functional word] bal)”, in: *Ghughutī na bās, Uttarākhamḍ kī bhāṣā, saṃskṛti, paryātan evaṃ paryāvaraṇ viṣayak lekh* (Turtle Dove. Papers on the Topic of the Language, Culture, Tourism and Environment of Uttarakhand). Dehradun: Winsar Publishing Company, 11–13.
- Jošī, Manohar Śyām 1995 [1982]. *Kasap*. Nāī Dillī: Rājkamal Prakāśan.
- Juyāl, Goviṃd Guṇānand 1967. *Madhya Pahārī bhāṣā (Garhvālī Kumāñī) kā anuśīlan aur uskā hindī se sambandh* (Analysis of the Central Pahari Language (Garhwali and Kumaoni) and its Relation with Hindi). Lucknow: Navyug Gramthāgār.
- Kālā, Janārdan Prasād 1959. *Garhvālī bhāṣā aur uskā lok sāhitya* (The Garhwali Language and Folk Literature). Dehradun: Samadaya Sakshaya.
- Kellogg, Samuel Henry 1972 [1876]. *A Grammar of the Hindī Language: in Which Are Treated the High Hindī, Braj, and the Eastern Hindī of the Rāmāyan of Tulsī Dās, also The Colloquial Dialects of Rājputānā, Kumāon, Avadh, Rīwā, Bhojpūr, Magadha, Maithila, etc.; with Copious Philological Notes*. Delhi: Oriental Book Reprints.
- Khatre, Sumitra M. 1966. *The Formation of Konkani*. Poona: Deccan College Postgraduate and Research Institute.
- Khokhlova, Liudmila 2013. “Obligational Constructions in New Indo-Aryan Languages of Western India”, *Lingua Posnaniensis* LV (2), 91–107.
- Kukreti, Bhisham s.d. “Vedic Sanskrit or Sanskrit is not the mother of Garhwali and Kumaoni languages” (<<https://groups.google.com/forum/#!topic/uttranchalkalasangam/bVK3VghrdUk>>, accessed: January 10, 2021).
- Kukreti, Rajni 2010. *Garhvālī bhāṣā kā vyākaraṇ* (A Grammar of the Garhwali Language). Dehradun: Winsar Publications.
- Malchukov, Andrej & Heiko Narrog 2009. “Case polysemy”, in: Malchukov, Andrej & Andrew Spencer (eds): *The Handbook of Case*. Oxford: Oxford University Press, 518–534.

- Meenakshi, K. 1986. "The Quotative in Indo-Aryan", in: Krishnamurthi Bh. et al. (eds): *South Asian Languages: Structure, Convergence and Diglossia*. Delhi: Motilal Banarsidass, 209–218.
- Minard, Armand 1936. *La subordination dans la prose védique*. Paris: Les Belles Lettres.
- Montaut, Annie 2016a. "Why Ergative Case Marker in Modal Clauses? The Historical Evolution of Aspect, Modality, Ergative and Locative in Indo-Aryan", in: Dahl, Eystein & Krzysztof Stronski (eds): *Ergativity in Indo-Aryan and Beyond*. (Typological Studies in Language 112.) Amsterdam: Benjamins, 135–165.
- Montaut, Annie 2016b. "The Discourse Particle *to* and Word Ordering in Hindi", in: Fernandez-Vest, Jocelyne & Robert Van Valin, Jr. (eds): *Information Structuring of Spoken Language from a Cross-linguistic Perspective*. Berlin; Boston: De Gruyter Mouton, 263–282.
- Montaut, Annie 2016c. "The Verbal Form V-ā in Hindi-Urdu: An Aorist with Aoristic Meanings", in: Guentcheva, Zlatka (ed.): *Aspectuality and Temporality*. Amsterdam: Benjamins, 313–344.
- Montaut, Annie 2017. "Grammaticalization of Participles and Gerunds in Preterit, Future and Infinitives", in: Bisang, Walter & Andrej Malchukov (eds): *Unity and Diversity in Grammaticalization Scenarios*. (Studies in Diversity Linguistics 16.) Berlin: Language Science Press, 97–136 (<https://langsci-press.org/catalog/book/152>), accessed: December 15, 2019).
- Montaut, Annie 2018. "The Rise of Differential Object Marking in Hindi and Related Languages", in: Seržant, Ilja & Alena Witzlack-Makarevich (eds): *Diachrony of Differential Argument Marking Patterns of Variation*. (Studies in Diversity Linguistics 19.) Berlin: Language Science Press, 281–313. (<https://langsci-press.org/catalog/book/173>), accessed: October 27, 2019).
- Moseley, Christopher (ed.) 2010. *Atlas of the World's Languages in Danger*, 3rd edn. Paris: UNESCO Publishing.
- Mühlhäusler, Peter 1995. "The Interdependence of Linguistic and Biological Diversity", in: Myers, David (ed). *The Politics of Multiculturalism in Oceania and Polynesia*. Darwin: University of the Northern Territory Press, 154–161.
- Mühlhäusler, Peter 2003. *Language of Environment, Environment of Language: A Course in Ecolinguistics*. London: Battlebridge.
- Nauṭiyāl, Śivānaṃd 1997. *Uttarākhaṃḍ kī lokgāthāē* (Folk ballads of Uttarakhand). Almora: Book Depot.
- Nettle, Daniel & Suzanne Romaine 2000. *Vanishing Voices: The Extinction of the World's Languages*. Oxford: Oxford University Press.
- Pischel, Richard 1999 [1900]. *A Grammar of the Prakrit Languages*. Delhi: Motilal Banarsidass.

- Riccardi, Theodore 2003. “Nepali”, in: Cardona, George & Dhanesh Jain (eds): *Indo-Aryan Languages*. London; New York: Routledge, 538–580.
- Saksena, Baburam 1937. *Evolution of Awadhi: A Branch of Hindi*. Allahabad: Indian Press.
- Sen, Sukumar 1960. *A Comparative Grammar of Middle Indo-Aryan*. Poona: Linguistic Society of India.
- Sharma, D.D. 1987. *The Formation of Kumauni Language*. New Delhi: Bahri Publications.
- Singh, Jag Dev 1970. *A Descriptive Grammar of Bangaru*. Kurukshetra: Kurukshetra University Press.
- Strnad, Jaroslav 2012. “A Note on the Locative/Instrumental/Ergative Case in Old Hindī”, *Oriental Archive* 80: 41–64.
- Strnad, Jaroslav 2013. *Morphology and Syntax of Old Hindī*. Amsterdam: Brill.
- Stroński, Krzysztof 2010. “Variations of Ergativity Patterns in Indo-Aryan”, *Poznan Studies in Contemporary Linguistics* 46 (2): 237–253.
- Stroński, Krzysztof 2014. “On the Syntax and Semantics of the Past Participle and Gerundive in Early New-Aryan. Evidence from Pahari”, *Folia Linguistica Historica* 35: 275–305.
- Tagare, Ganesh Vasudev 1948. *Historical Grammar of Apabhramsha*. Poona: Deccan College.
- Tessitori, Luigi Pio 1914. “Notes on the Grammar of the Old Western Rajasthani, With Special Reference to Apabhramsha and to Gujarati and Marwari”, *The Indian Antiquary* XLIII: 55–63, 84–91 (discontinuous pagination).
- Tivārī, Udaynārāyaṇ 1961. *Hindī bhāṣā kā udgam aur uskā vikās* (The Origin and Development of the Hindi Language). Prayāg: Lok Bhārati Prakāśan.
- Tiwari, Udai Narain 1966. *The Origin and Development of Bhojपुरi*. (The Asiatic Society Monograph Series, vol. X.) Calcutta: The Asiatic Society.
- Trumpp, Ernest 1872. *Grammar of the Sindhi language*. London: Trubner and Co.; Leipzig: Brockhaus.
- Turner, Ralph 1966. *A Comparative Dictionary of the Indo-Aryan Languages*. London; New York; Toronto: Oxford University Press.
- UNESCO Silk Road Program (<<https://en.unesco.org/silkroad/silk-road-themes/languages-and-endanger-languages/garhwali>>, accessed: January 10, 2021).
- Vaudeville, Charlotte 1957. *Kabīr Granthāvalī*. (Publications de l’Institut Français 12.) Pondichéry: Institut Français d’Indologie.

Vashini Sharma

The Impact of Other Indian Languages on Dakkhini

वशिनी शर्मा

दखनी पर अन्य भारतीय भाषाओं का प्रभाव

Abstract Dakkhini, an Indo-Aryan language, has been in constant contact with Telugu and several other Dravidian languages for more than four centuries. As a result, it has acquired a number of syntactic features that are not found in Hindi-Urdu but found in any one of the Dravidian languages that it has been in contact with. This syntactic change is generally labelled as convergence which results in a Linguistic area. The study shows that Dakkhini and Telugu are closer to each other than Dakkhini and Urdu or Telugu and Urdu. The data have been verified using some YouTube programs in current Dakkhini.

Keywords convergence, linguistic area, contact language, code switching, agreement, honorifics.

सारांश लगभग चार शताब्दियों से भारतीय आर्य भाषा दखनी द्रविड़ की अनेक भाषाओं के विशेष रूप से तेलुगु के संपर्क में रही है। परिणामस्वरूप दखनी की वाक्य संरचना में ऐसी कई विशेषताएं हैं जो कि हिंदी या उर्दू में नहीं पाई जातीं पर द्रविड़ भाषा परिवार की इस खास भाषा, तेलुगु में पाई जाती हैं। इस तरह के वाक्य परिवर्तन भाषाई अभिसरण के नाम से जाने जाते हैं जो किसी भी भाषाई-क्षेत्र की उपज हो सकते हैं। यह लेख इसी प्रकार की संरचनाओं के बल पर इस निष्कर्ष पर पहुंचता है कि वाक्य संरचना की दृष्टि से दखनी भाषा तेलुगु के ज्यादा पास हैं बजाय हिंदी या उर्दू के। इस आलेख का आधार हैदराबाद के वे स्थानीय निम्न और मध्यवर्गीय मुस्लिम परिवार हैं जहाँ दखनी बोली जाती है। आजकल के दखनी संदर्भों की पुष्टि के लिए यू-ट्यूब से वीडियो सामग्री की सहायता भी ली गई है।

मुख्य शब्द – भाषाई अभिसरण, भाषाई क्षेत्र, संपर्क भाषा, कोड स्थान्तरण, अन्विति, सम्मानसूचक।

१ परिचय

हैदराबाद की भाषाई पृष्ठभूमि बहुत रोचक है। यहीं कुछ सदियों पहले दखनी हिंदी का आविर्भाव हुआ। उर्दू के प्रचलन के साथ-साथ उसका रूप बदलता रहा। पुरानी और वर्तमान दखनी में बहुत अंतर है। आधुनिक बोलचाल की भाषा मानक हिंदी-उर्दू के अधिक निकट है। निज़ाम के शासनकाल में उर्दू यहाँ की राजभाषा थी और शिक्षा के माध्यम की भी भाषा थी। आंध्र प्रदेश बनने के बाद भी पूरे तेलंगाना में दखनी बोलचाल तथा संपर्क का माध्यम बनी रही। इस कारण दखनी में एक ओर उर्दू से प्रभावित भाषिक विशेषताएँ हैं तो दूसरी ओर क्षेत्रीय भाषाओं का प्रभाव भी स्पष्ट दिखाई पड़ता है।

दखनी हैदराबाद में पली और निज़ाम राज्य के पूरे इलाके में असें तक संपर्क का माध्यम बनी रही। हैदराबाद असें तक आंध्र प्रदेश की राजधानी रहा, पर अब नव निर्मित तेलंगाना प्रदेश की भी राजधानी है जिसकी आज राजभाषा तेलुगु और उर्दू है। यहाँ की जनता महानगरीय है, जिसमें तेलुगु बोलने वालों की बहुतायत है। इस कारण दखनी में द्रविड परिवार की कई भाषिक विशेषताएँ भी हैं। तेलंगाना के पश्चिमी जिलों में तथा हैदराबाद शहर में भाषा के आधार पर प्रदेशों के पुनःसंगठन से पहले अधिक संख्या में मराठी भाषी थे। तेलुगु और मराठी पड़ोसी भाषाएँ होने के नाते उनमें कुछ सामान्य विशेषताएँ विकसित हुईं जिससे दखनी हिंदी में ये विशेषताएँ भी दिखलायी पड़ती हैं। इस तरह ये विशेषताएँ किसी खास भाषा-समुदाय की विशेषताएँ नहीं, बल्कि शहर की सामान्य बोलचाल की भाषा में न्यूनाधिक रूप में लोगों के भाषा-कोश में विद्यमान है। यहाँ के स्थानीय हिंदी भाषियों में भी ये विशेषताएँ कमो-बेश दिखाई देती हैं। शिक्षा तथा मानक हिंदी के प्रचार के कारण सभी लोगों में सभी विशेषताएँ नहीं दिखायी पड़ती। सामाजिक संदर्भों में पढ़े-लिखे लोगों में कोड-अंतरण (Code-Switching) देखा जा सकता है, जहाँ एक व्यक्ति औपचारिक क्षेत्रों, वर्गीय संपर्कों में मानक हिंदी का प्रयोग करे और अन्य स्थानीय व्यक्तियों के साथ दखनी का।

२ दखनी हिंदी के प्रकार

भाषिक पृष्ठभूमि के आधार पर संपर्क भाषा के रूप में दखनी बोलने वाले व्यक्तियों के दो वर्ग हो सकते हैं -

(अ) मातृभाषा के रूप में दखनी का प्रयोग करने वाले स्थानीय हैदराबादी निवासी जो अधिकांश मुस्लिम समाज से हैं और कुछ अन्य प्रदेशों से आकर बसे हुए परिवार हैं। सामाजिकता और आर्थिक दृष्टि से भी यह उच्च वर्गीय खालिस उर्दू बोलनेवालों से भिन्न वर्ग है।

(आ) संपर्क भाषा के रूप में हिंदी भाषा का प्रयोग करने वाले द्विभाषी समुदाय के लोग, जिनमें प्रमुख तेलुगु और मराठी भाषी हैं। इनकी भाषा को मैं दखनी का एक रूप हैदराबादी हिंदी मानती हूँ जो दखनी से कुछ अर्थों में भिन्न है -पर है दखनी की ही एक शैली (शर्मा 1984)।

दक्षिण के विभिन्न क्षेत्रों में वहाँ की स्थानीय भाषाओं के प्रभाव से दखनी के हैदराबादी, बैंगलोर, औरंगाबादी रूप भी मिलते हैं। मातृभाषा भाषी वर्ग में स्थानीय भाषा से मानक हिंदी तक का एक क्रम (spectrum) दिखायी पड़ता है, जबकि द्विभाषी समुदाय में स्थानीय भाषागत विशेषताएँ गहराई तक हैं।

एमेनो ने (1956:16) अपने लेख “India as a Linguistic Area” में भाषाई अभिसरण (Linguistic Convergence) की प्रक्रिया पर प्रकाश डालते हुए इसे दखिनी का महत्वपूर्ण लक्षण माना है जो दक्षिण की भाषाओं से प्रभावित है न कि उर्दू-हिंदी से।¹

सुब्बाराव और अरोड़ा (1989) ने अपने एक लेख “Some Aspects of the Syntax of Dakkhini” में तेलुगु वाक्यविन्यास का दखिनी पर प्रभाव और उसके कारण दखिनी के उर्दू से हटकर तेलुगु वाक्य-विन्यास को अपनाने के पक्ष में कई तर्क दिए हैं। उन्होंने उक्त लेख में दखिनी के रूप विन्यास और वाक्य-विन्यासात्मक (morphological and syntactic) अभिसरण की प्रक्रिया को तेलुगु और उर्दू के संदर्भ में विवेचित किया है।

इस आलेख में दखिनी, मानक हिंदी और तेलुगु के संदर्भ में कुछ प्रमुख बिंदुओं पर विचार किया जाएगा जो दखिनी के उक्त अभिसरण की पुष्टि करते हैं।

वाक्य रचना संबंधी विशेषताएँ

अकर्मक वाक्य – दखिनी में “ने” का प्रयोग नहीं मिलता। क्रिया अकर्मक वाक्य की तरह कर्ता के साथ अन्वित होती है। बाबूराम सक्सेना (1952:53) जी के अनुसार वैसे तो ‘ने’ दखिनी में नहीं प्रयुक्त होता पर जहाँ नहीं होना चाहिए कभी वहाँ होता है।

- (1) *उनीं बी बात को खोले है*
- (2) *खुदा के दोस्ताँ ने बोले हैं*
अन्य उदाहरण भी देखिए –

दखिनी	हिंदी
(3) <i>तुम फिल्म देखे क्या</i>	(क्या) तुमने फिल्म देखी ?
(4) <i>मैं हत्या नई करा।</i>	मैंने हत्या नहीं की
(5) <i>पुलिसवाला मारा</i>	पुलिसवाले ने मारा।

व्यवहार में प्रयुक्त न होने के कारण ‘ने’ का प्रयोग अर्जित व्यवहार के कारण या औपचारिक शिक्षण से ही संभव हो पाता है।

1 Linguistic Area is “an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families.” (Emeneau 1956:16 n.28)

कर्ता + को रचना

दखनी. में सर्वनाम + 'को, के' तिर्यक रूप मुझे, तुम्हें नहीं मिलते। तेलुगु की तरह सर्वनाम + को = ना + कु, मी + कु, आयन + कु = मेरे + कु, तेरे + कु, उन/ उना + कु का प्रयोग मिलता है

(6.1) तुमारे कु क्या होना? हिं. तुम्हें क्या चाहिए
ते. मीकू एम कावाले?

(6.2) मेरे कु एक साड़ी होना। हिं. (मुझे एक) साड़ी चाहिए
ते. (ना कु ओक) चीर कावाले

सुब्बाराव ने इसे इस तरह वर्गीकृत किया है- Degenitivization (loss of the genitive) and towards Dativization (replacing a genitive of Hindi-Urdu by a dative case marker. निम्नलिखित संबंध सूचक वाक्य Subbarao & Arora (1989:111) के शोध पत्र से उद्धृत हैं।

Possession/ संबंध कारक

a) Kinship

Hindi

(7) *sītā ke cār bacce haĩ*
Sita GEN four children are
'Sita has four children.'

Telugu (Dravidian)

(8) *sītā ki naluguru pillalu (unnaaru)*

b) Inalienable Possession

Hindi

(9) *kutt.e ke cār pāṅ hote haĩ*
dog.OBL GEN.PL four legs existing are
'A dog has four legs.'

Telugu

(10) *kukka ki nālugu kāḷḷu unṭāyi*
dog.OBL DAT four legs are
'A dog has four legs.'

Dakhini

- (11) *kutte ku cār pāvā raite*
dog.OBL GEN.PL four legs are

c) Concrete Possession

Hindi-Urdu

- (12) *us ke pās bahut paisā hai*
He OBL GEN near much money is
'He has a lot of money.'

Telugu

- (13) *āme ki cālā ḍabbu undi*
He DAT much money is
'He has a lot of money.'

Dakhini

- (14) *us ku bhot paisā ai*
he.OBL DAT much money is
'He has a lot of money.'

Agreement in gender and number / अन्विति – लिंग और वचन

दखनी में बहुवचन शब्द अंत में-याँ / -आँ लगाकर बनते हैं जैसे –

स्त्री. बात > बातों, पु. काम > कामों, बहू > बहुआँ, सूअर > सूआरों, रात > रातों, घर > घरों,
बच्ची > बच्चियाँ बैल > बैलों

अंग्रेज़ी शब्द के बहुवचन रूप – मार्का, क्वेश्चनों, नम्बरों, पैम्परों आदि।

Emphatic particles / बलाघात – ही, भी, तो

ही, भी का प्रयोग हिंदी में पूरा शब्द जोड़कर किया जाता है पर दखनी में ई/ई सर्वनाम में जुड़ जाता है – हमीं / तुमीं

हिंदी

दखनी

हम + ही, तुम + ही, आज + ही कल + ही हमीं, तुमीं, आजी, कली

बाबूराम सक्सेना (1952:53) ने 'च' के प्रयोग के कुछ उदाहरण दिए हैं –

भौतीच < बहुत + च हिंदी 'बहुत ही'; ऐसेच < ऐसे + च हिंदी 'ऐसे ही';

योंच < यों + च हिंदी 'यूं ही'; देखतेच < देखते + च हिंदी 'देखते ही' ।

हैदराबादी दखिनी

(15) कल से नई आती मैं पैलैच बोल दी

(16) झाड़ू पोंचा नई करती मैं पैलेच बोल दी

(17) मालूम इच है न सास कु डायबेटीज़ है जल्दी खाना होना मेरकु
हिंदी 'मालूम ही है न सास को डायबेटीज़ है मुझे जल्दी खाना चाहिए'
हिंदी 'भी' > बी / ब्बी – कब + ही > कब्बी, अब + ही > अब्बी

दखनी

हिंदी

(18) इत्ते दिनाँ में कब्बी नई फटकी
(फटकना = आना)

'इतने दिनों में कभी भी नहीं आई।'

(19) अब्बी सोके आई

'अभी सो कर आई हूँ।'

(20) किसको बी नको बोलो

'किसी को भी मत बताना।'

(22) उनाँ चुपके तो बी गल्ला फाड़ रैं

'वे लोग बेकार में चीख रहे हैं।'

(23) क्या तो बी बोल रैं

'पता नहीं क्या बोल रहे हैं'

सर्वनाम

अपना – दखनी में ‘अपना’ के स्थान पर ‘तुम्हारा’ का प्रयोग ही मिलता है ‘अपनी-अपनी’ और ‘अपने-अपने’ के प्रयोग भी देखिये।

दखिनी

- (24) तुम सब लोगाँ तुम्हारी किताबाँ खोलो
हिंदी ‘तुम सब अपनी किताबें खोलो।’
- (25) सारे जनाँ तम्हारी तुम्हारी किताबाँ खोल लेना
- (26) तुम सबों आज इच तुम्हारे-तुम्हारे घराँ कु चले जाव
हिंदी ‘आप सब आज ही अपने-अपने घर चले जाएँ’
- (27) मैं तुमन कु मेरा ई-मेल देताउँ
हम सब – हम लोगाँ / अपुन लोगाँ (वक्ता के साथ) = us (inclusive)
- (28) हम लोगाँ आनै इच वाले थे
हिंदी ‘हम लोग आने ही वाले थे।’
- (29) अपन लोग आज इच याद करै थे न
हिंदी ‘हम लोग आज ही याद कर रहे थे न।’

क्रिया – रहा ‘है’ का लोप

हिंदी ‘रहा + है’ = दखनी रा/ रय/ रै/ रई (वचन और लिंग की अन्विति)

वर्तमान काल – हिंदी की सहायक क्रिया “है, हूँ, हैं”

दखनी के क्रिया रूप – बोलतूँ, बोलतैं, बोलताउँ

भूत काल – सहायक क्रिया “था, थी, थे” का प्रयोग हिंदी और दखनी में भी होता है।

दखनी छिछोरे बाताँ करयथे

भविष्य काल – हिंदी के सहायक क्रिया-रूप “होगा, होगी, होंगे”

दखनी बोलतूँ फेल हुए सो

- (30) चार अंडे उबाल ले लेके लात्यूँ
हिंदी ‘चार अंडे उबाल कर ले आती हूँ।’
- (31) नई चला तो फेंक देंगे (रीमोट) अपना क्या जारा (जा रहा है)

इच्छार्थक क्रिया

दखनी में 'चाहना, चाहिए' का प्रयोग भी नहीं मिलता। संज्ञा + चाहिए की रचना में 'होना' का प्रयोग होता है जो मराठी के 'पाहिजे' और तेलुगु के 'कावाले/लि' के प्रभाव से है। कहीं कहीं 'मांगता 'पूछना' भी प्रयुक्त होता है

(32) तुम कु क्या होना

हिंदी	'तुम्हें क्या चाहिए'
मराठी	'तुला काय पाहिजे'
तेलुगु	'नीकु एम कावालि'
दखनी	एक्स्ट्रा पूछेंगे तो
हिंदी	'और अधिक / अतिरिक्त मांगेंगे तो...'

'चाहना, चाहिए' के अन्य संदर्भों में मूल वाक्य के अर्थ के अनुसार क्रियार्थक संज्ञा वाली रचनाएँ मिलती हैं। मध्यम पुरुष के वाक्य सुझाव, अप्रत्यक्ष विधि या निषेध की तरह आते हैं, अतः कर्ता प्रत्यक्ष होता है; उत्तम पुरुष में कर्ता + को की रचना मिलती है; अन्य पुरुष में इन दोनों का संयोजन है। कुछ उदाहरण देखें-

दखनी	हिंदी
(33) तुम जाना	तुम्हें जाना चाहिए
(34) तुम जाना नई	तुम्हें जाना नहीं चाहिए
(35) मेरे कु एक कमीज़ होना	मुझे एक कमीज़ चाहिए

निषेध सूचक— दखनी में 'मत' का प्रयोग नहीं मिलता, बल्कि 'नई' ही निषेध में आता है।

दखनी	हिंदी
(36) तुम नको जाओ	तुम मत जाओ
(37) तुम जाना नई	तुम्हें जाना नहीं चाहिए
(38) बहु नको खोलो बोली	बहु ने कहा कि मत खोलो

(39) मेरे कने नई आओ। मेरे पास मत आओ

(40) मिठाई नई खा। मिठाई मत खा

शब्द का क्रम काफी लचीला है। नहीं का रूप 'नई' वाक्य के अंत में ही नहीं मध्य में भी आ सकता है।

(41) क्या कि बोल नई हिंदी 'कुछ मत कह'

(42) नींद से नई उठा सकते मुझे दुनिया से उठाते कते
हिंदी 'जो नींद से उठा नहीं सकते वो कहते हैं मुझे दुनिया से उठा देंगे।'

“हाँ” और “नहीं”

“हौ”, “नको” दखनी का विशेष प्रयोग है, जो मराठी से आया है। सामान्य रूप से किसी बच्चे को किसी कार्य के लिए संकेत से मना करने के लिए जिस तरह “नहीं” कहते हैं, उस तरह यहाँ “नको” का प्रयोग मिलता है।

दखनी हिंदी

(43) तुम नको जाओ 'तुम मत जाओ'

(44) नको बोलो 'मत बोलो'

(45) लाल चटनी के शिगुफ़े नको करो 'लाल चटनी (मांगने) के नखरे मत करो'
'नको' दूसरी तरफ़ 'नहीं चाहिए' के अर्थ में भी आता है। कुछ प्रश्नोत्तर देखें -

(46) प्रश्न— तुमकू मिठाई होना? उत्तर— नको।

(47) प्रश्न— तुम खाना खाते क्या? उत्तर— अबी नको
दखनी “क्या कि” (हिंदी 'कुछ') प्रायः निषेधार्थक संदर्भों में ही आता है।

(48) प्रश्न— कौन आये थे? उत्तर— क्या कि 'नहीं मालूम।'
हिंदी के तकिया कलाम “पता नहीं” से इसकी तुलना कर सकते हैं।

(49) वो क्या कि बोल रा। 'पता नहीं, वह क्या बोल रहा है।'

इनमें इस प्रछन्न निषेध को देख सकते हैं। इसी दृष्टि से यह “क्या भी” ('कुछ भी') से भिन्न है। (48) और (49) में “क्या कि” का प्रयोग देखिये।

(50) क्या भी नई कर रा मैं 'मैं कुछ नहीं कर रहा।'

(51) क्या भी नको खाओ 'कुछ भी मत खाओ।'

प्रश्नवाचक शब्द अन्य अर्थ में

देखीये दखिनी के उदाहरण

(52) कैसा रैती कैसा चलती। '[कितने/स्टाइल से] रहती है, चलती है।'

(53) क्यों जोडना खाली-पीली। 'फिज़ूल में/बेवज़ह क्यों जोड़ें [नहीं]?'

(54) कायकु तो बी निकले तुम लॉकडाउन में। 'तुम लॉक डाउन में क्यों/बेवज़ह निकले?'

पूरक वाक्य की संरचनाएँ

पूरक वाक्य की संरचनाओं के अतिरिक्त अन्य स्थानों में सहायक क्रिया “है” का लोप होता है और “रह”, “रहा”, “रही” यह रूप लेता है रैं, रा, राउँ।

उच्चारण क्रम में “है” के लोप के कुछ अन्य उदाहरण है

(55) मैं चलतू। 'मैं चलता हूँ।'

(56) लाड़ँ कर रैई। 'लाड़ (बहुत) कर रही है।'

(57) माररा उन। 'वह मार रहा है।'

(58) मेरे को जाना (है)। 'मुझे जाना है।'

(59) तुमकू जाना (है) क्या? 'क्या तुम्हें जाना है?'

सहायक क्रिया “था” का प्रयोग सामान्य रूप से नहीं होता। क्रिया रचना के सरलीकरण के कारण वर्णन में इसका भी लोप देखा जा सकता है। किसी बच्चे के द्वारा एक कहानी उदाहरणार्थ प्रस्तुत है।

दखिनी में कहानी के दो रूप और हिंदी अनुवाद

दखनी – एक चूहा गलीच फिरता। उसके दुम में काँटा चुभता। हज्जाम मामू को बुलाता। बुलाके मेरा काँटा निकालो बोलता। बस्ता लेके भाग जाता।

हिंदी – ‘एक चूहा गली में ही फिरता था। उसके दुम में काँटा चुभ जाता है। नाई को बुलाता है। बुला कर कहता है, “मेरा काँटा निकालो”। बस्ता लेकर (बच्चा) भाग जाता है।’

बाबूराम सक्सेना (1952:53) भी अपनी पुस्तक *दक्खिनी हिंदी* में “हैगा” का उल्लेख करते हैं —
भौतीच हैगी।

आगरा और ब्रज प्रदेश में “है” की जगह “हैगा” का कथन शैली में प्रयोग देखिए –

एक चूहा गली में ही फिरता हैगा। उस के दुम में काँटा चुभ जाता हैगा। नाई को बुलाता हैगा। बुला कर कहता हैगा मेरा काँटा निकालो। बस्ता लेकर (बच्चा) भाग जाता हैगा।

यहाँ वास्तव में “है”, “था” के लोप का ही सवाल नहीं, बल्कि पक्ष व्यवस्था में भी सरलीकरण परिलक्षित होता है। वर्णन में प्रायः हर जगह अपूर्ण पक्ष की क्रिया देखते हैं। पूर्ण पक्ष में भी आया, आया है, आया था’ आदि रूपों में काल का अंतर सहायक क्रिया के लोप के कारण खत्म हो जाता है। भविष्य के संदर्भ में भी अपूर्ण पक्ष का प्रयोग है –

दखनी

हिंदी

(60) कपड़ा लाते ?

कपड़ा लाओगे ?

सरलीकरण के कारण तो यह है ही, अपूर्ण पक्ष के फुटकर खाते की तरह इस्तेमाल के कारण भी है।

हमने ऊपर उल्लेख किया कि पूरक वाक्यों में “है” आ सकता है। उल्लेखनीय है कि कई जगह इसके स्थान पर ‘होता’ का प्रयोग मिलता है।

(61) प्रश्न—

ये इसकी अम्मी है न ?

उत्तर— नई होती।

भूतकाल के लिए देखिए — हती/हता (हिंदी— थी/ था) एक राजा हता, एक रानी हती;
ब्रज— एक राम हते एक रावण ना, एक छत्ती हते एक बामन ना।

कई जगह अपूर्ण पक्ष की रचना के स्थान पर (बिना आवश्यकता के) निरंतरबोधक कृदंत का प्रयोग मिलता है। “यह काम होता है” की जगह काम होते रहता है इस प्रवृत्ति का द्योतक है। ऐसे संदर्भों में कृदंत तिर्यक, अविकारी रूप में आता है और केवल “रहता है” लिंग वचन के लिए अन्वित होता है।

(62) सलमाँ बाताँ करते रैता

‘सलमा बातें करती रहती है।’

(63) वो क्या कि बोलते रैता

‘न मालूम वह क्या कहता रहता है।’

(64) याँ एइच होता रैता

‘यहाँ यही होता रहता है।’

दखनी की विशेषता है “ला लेना”। सामान्य रूप से दखनी की रंजक क्रियाएँ मानक हिंदी के समान हैं। लेकिन रंजक क्रिया में पूर्वकालिक कृदंत का व्यापक प्रयोग (पकड़ लेके, कर लेके, खा लेके, ले लेके) यहाँ की विशेषता है।

“चला जाना” में ऊपर के उदाहरणों की तरह “चले जाना” रूप मिलता है (वह चले गया) और इसमें भी पूर्वकालिक कृदंत आता है (चले जाके)। इसी तरह रंजक क्रियाओं में “रहा” का प्रयोग भी व्यापक है।

मानक हिंदी में जिस तरह “जो... वह” से विशेषण उपवाक्य की रचना होती है, वैसी रचना का दखनी में अभाव है। यहाँ अस्वतंत्र उपवाक्य वाक्यांश स्तर पर अंतर्निहित संरचना के रूप में आता है, और यहाँ “सो” योजक शब्द हैं।

(65) आप बोले सो बात...

हिंदी— ‘आपने जो बात कही थी वह...’

विशेषण उपवाक्य के कर्ता, कर्म दोनों के लिए आता है। आगे के दो वाक्यों में क्रमशः कर्ता और कर्म का विशेषीकरण है। ध्यान दें कि दखनी में ये दोनों वाक्य सतही स्तर पर समान लगते हैं। कर्म-विशेषण वाक्य प्रायः अकर्तृक होता है।

(66) कोट पैने सो आदमी

‘वह आदमी जिसने कोट पहना है...’

(67) तेलुगू— कोट वेसुकुन्न मनिषि

‘*कोट पहना (हुआ) आदमी’

(68) यह देखे सोई फ़िल्म है

‘यह वही फ़िल्म है जो हमने देखी है’

(69) बोल्तू फेल हुए सो

‘तुम फेल हुए हो यह बात बोल दूँगा’

(70) ये बिस्तर अम्मा दिए सोई है याद रखो

हिंदी— ‘याद रहे कि यह बिस्तर हमारी अम्मा का ही दिया हुआ है।’

‘याद रखो कि यह वही बिस्तर है जो हमारी अम्मा ने दिया था।’

(71) फ़िज़ हमारे घर से आयसो है याद रखो

हिंदी— ‘याद रहे / रखो कि यह फ़िज़ हमारे घर से ही आया हुआ है।’

इन अकर्तृक वाक्यों में योजक से पहले का कृदंत अविकारी होता है² (अरोड़ा 2004)।

(72) फ़ोन पर बोले सो बातों याद रखो (* बोली सो बातों)। हिंदी— ‘बोली हुई बातें’

(73) इसमें रखे सो चिट्टी (* रखी सो चिट्टी), हिंदी— ‘रखी हुई चिट्टी’

2 “It may be noted that there is one to one correspondence between Dakhni and Telugu. In both of them the embedded clause syntactically occurs to the left of the complementizer and the matrix clause to the right. Further in both of them the complementizer is post-sentential. As opposed to this in Urdu the order is reverse; further the complementizer is pre-sentential.” (Arora 2004)

शायद इसका कारण है कि “बोले” (“ने” के अभाव में तथा कर्ता के उल्लेख के अभाव में) अंतर्निहित क्रिया वाक्यांश है, न कि विशेषण शब्द। तेलुगु में *चेप्पिन माट (बोले सो बात) / चेसिन पनी (करे सो काम)* का समान प्रयोग मिलता है।

कुछ रोचक वाक्य देखिए—

- (74) काला कोट पैंने सो है बोलके मालूम होता ।
हिंदी— ‘लगता है कि...वह है जो काला कोट पहने है।’
- (75) स्टैंड रखे सो ऊपर चढ़े देखो देखो । हिंदी— ‘जो स्टैंड रखा है उस पर चढ़ा हुआ है।’
- (76) तुम मेडिमेक्स से नहा लेंरें सोई बहुत है ।
हिंदी— ‘यही बहुत है कि तुम मेडिमेक्स से नहा पा रही हो।’

क्रियाविशेषण उपवाक्यों की रचना योजक शब्द “वैसा”, “तब”, “उधर” से होती है।

- (77) गन्ना खींच ले रैं/रँय वैसा कर रई ।
हिंदी— ‘यह ऐसा कर रही है, जैसे वे गन्ना खींच ले रहे हों।’
- (78) मैं आयी तब कौन भी नई था । हिंदी— ‘जब मैं आयी तब कोई भी नहीं था।’
- (79) लोगाँ बैठे हैं उधरीच । हिंदी— ‘जहाँ लोग बैठे हैं वहीं।’

इतिवृत्तात्मक वाक्य (Quotatives)

हिंदी में पुनःकथन “कि” से सूचित होता है, जबकि दखनी में बोलके/करके से या इनके अभाव में भी। हिंदी में पुनःकथन का उपवाक्य स्वतंत्र या मूल उपवाक्य के बाद “कि” से जोड़ा जाता है, जबकि दखनी में वह उपवाक्य के भीतर कर्म के स्थान में आता है। कुछ वाक्य देखिए—

- | दखनी | हिंदी |
|-----------------------------|-----------------------------------|
| (80) वह आतूँ (बोलके) बोला । | ‘उसने कहा था (कि) मैं आता हूँ।’ |
| (81) अम्मा मेरे को जाओ बोली | ‘माँ ने मुझसे कहा था कि तुम जाओ।’ |
| (82) आपको कौन जाओ बोले । | ‘आपसे किसने कहा कि आप जाएँ।’ |
| (83) क्या भी नको खाओ बोली । | ‘उसने कहा कि कुछ भी मत खाओ।’ |

- (84) नको जाओ बोले तो नई जातऊँ जी ।
हिंदी— ‘अगर तुम न जाने को कहो तो नहीं जाऊँगा जी ।’
उपरोक्त वाक्यों में लिंग, वचन की अन्विति स्पष्ट है।

“बोलके/करके” (अनि) का प्रयोग द्रविड भाषा परिवार की विशेषता का प्रभाव है (सुब्बाराव 2013)³। “बोलके” इतिवृत्तात्मक शब्द ही नहीं, प्रयोजन या उद्देश्य सूचित करने के लिए भी आता है। आगे के उदाहरण देखें—

- | दखनी | हिंदी |
|-------------------------------------|--|
| (85) मैं छोड़ दिया बच्चा है बोलके । | ‘मैंने इसलिए/यह सोचकर छोड़ दिया था कि बच्चा है।’ |
| (86) मैं जरूरी जाना बोलके गया । | ‘मैं इसी कारण गया था कि जाना जरूरी है।’ |
| (87) अच्छा है बोलके लिया था । | ‘यह सोच कर लिया था कि अच्छा है।’ |
| (88) मैं खाली बोलना बोलके बोला । | हिंदी—‘मैंने यों ही कह दिया था कि कुछ कहना है।’ |

किसी घटना के वर्णन में, खास कर दखनी भाषी बच्चों के द्वारा कहानी सुनाने में एक भिन्न प्रकार की इतिवृत्तात्मक रचना दिखाई पड़ती है जहाँ इतिवृत्तात्मक शब्द “कते” (< कहते) है।

- (89) एक राजा था कते ।
(90) वो लोग़ाँ शाम में आ रें कते ।

यहाँ “कहते” का प्रकार्य मानक हिंदी के “ऐसा है”, “बात (यों) है”, “कहा जाता है” (ऐसा है कि एक राजा था।.....) के समान है। स्मरणीय है कि दखनी में “कहना” का प्रयोग नहीं मिलता। यहाँ यह केवल इतिवृत्तात्मक शब्द है।

- (91) नींद से नई उठा सकते मुझे दुनिया से उठाते कते ।
हिंदी— ‘(जो)नींद से नहीं उठा सकते मुझे (वो) कहते हैं मुझे दुनिया से उठा देंगे।’
‘कि/या नहीं’ की जगह क्रिया की पुनरुक्ति देखी जा सकती है।
- (92) तुम सीता को देखे नहीं देखे ? हिंदी— ‘तुम ने सीता को देखा कि नहीं देखा?’

3 Interestingly, the calqued form *bol ke* in Dakkhini suggests that there are some abstract semantic properties associated with the abstract frame, since the fact that the Hindi/Urdu form *bol ke* is calqued from Telugu *-ani*, rather than *-ani* being employed in Dakkhini right away, indicates that Hindi/Urdu is the Embedded language already at stage 2 (Subbarao 2013).

(93) *उनों आये नई आये ?* हिंदी- ‘वे आए कि नहीं आए?’

“अगर करँ/करँगा” की रचना के स्थान पर “ने” रहित पूर्ण पक्ष की क्रिया का प्रयोग होता है तो “अगर” लुप्त रहता है।

(94) *तुम खाना नई खाये तो पीटूँगा।*

(95) *वो काम नई करे तो मारूँगा।*

(96) *पीली कमीज़ पहने खिल के आते।*

हिंदी- ‘पीली कमीज़ पहन लो तो निखर कर आते हो।’
बीते काल के संदर्भ में ‘अगर...जाते’ की रचना भिन्न होती है।

(97) *तुम जाते थे तो साब पैसे देते थे। तुम काय कू नई गये?*

‘अगर तुम जाते तो साहब पैसे देते। तुम क्यों नहीं गये?’

कृदंत- दखनी में पूर्वकालिक कृदंत की रचना “के” से होती है, जैसे- बोलके, करके, देखके, लाके, देख लेके, ला लेके, उतर जाके, पूछ लेके।

(98) *चार अंडे उबालके ले लेके आलूँ।* हिंदी- ‘चार अंडे उबाल कर लेकर आती हूँ।’

हिंदी के “बिना देखे” आदि के स्थान पर भी पूर्वकालिक कृदंत का प्रयोग होता है।

(99) *बिगर देखके, बिगर पूछके चला गया। (बिगर?)*

इसी तरह “उन्हें यहाँ आये दस दिन हो गये” आदि वाक्यों के कृदंत के स्थान पर पूर्वकालिक कृदंत का प्रयोग मिलता है।

(100) *उनको इधर आके दस रोज़ हो गये।* हिंदी- ‘उन्हें इधर आये दस दिन हो गए।’

कृदंत में चर्चा की गई है कि “आना” का व्यापार पूर्ण होने की स्थिति में ही “बाद” की सार्थकता है। दखनी के “आये बाद”, “गये बाद” आदि प्रयोग पूर्ण पक्ष के प्रयोग की इस विशेषता की ओर इंगित करते हैं।

प्रश्नवाचक वाक्य

दखनी में बहुधा प्रश्न बिना “क्या” के बनते हैं। बोलचाल की हिंदी में भी प्रश्न इस तरह संभव है।

(101) *मेरे को लेके जाते?* हिंदी- ‘मुझे ले जाओगे?’

(102) *कपड़ा लाते?* हिंदी- ‘कपड़ा लाओगे?’

हिंदी के विपरीत दखनी में अगर “क्या” का प्रयोग हो भी तो वह वाक्य के अंत में होता है। तेलुगु में अंत में “आ?” (= “क्या?”) का प्रयोग द्रविड भाषा परिवार की विशेषता है, जहाँ प्रश्नवाचक वाक्य निश्चयार्थक वाक्यों के अंत में कोई रूप जोड़ने से बनते हैं।

तेलुगु *वस्तारा* < *वस्ता* + *र* + *आ* ‘आओगे?’; *तिंटारा* < *तिंटा* + *र* + *आ* ‘खाओगे?’

दखनी में प्रश्नवाचक वाक्यों के अंत में “क्या” का प्रयोग होता है, जैसे —

दखनीहिंदी

- (103) कपड़े ला लूँ क्या? ‘क्या कपड़े ले आऊँ?’
 (104) छोकरा चले गया क्या? ‘क्या छोकरा चला गया?’
 (105) उनों भी आते क्या? ‘क्या वे भी भी आ रहे हैं?’
 (106) उनों आतुं बोले क्या? ‘क्या उन्होंने कहा था कि वे आएँगे?’

“क्यों” के स्थान पर कायकु का प्रयोग तेलुगु “एंदुकु” की तरह होता है, जैसे —

दखनीहिंदी

- (107) तुम काय कु नईं गये? ‘तुम क्यों नहीं गये?’
 (108) कायकू टेंशन लेरा? ‘टेंशन क्यों ले रहे हो?’

व्यंग्यात्मक प्रति प्रश्न

- (109) कामवाली हूँ नईं न?
 (110) खाना कौन बनातई? मैं?; साफ-सफाई कौन करतई? मैं?;
 कामाँ कौन करतई? मैं?

संबोधन सूचक वाक्य

दखनी भाषा भाषी के संबोधन *अम्मी*, *अब्बू*, *आपा* ‘बड़ी बहन’, *मामू* ‘मामा’, *खाला* ‘मौसी’, *खालू* ‘मौसा’ बहुत ही अपनत्व दर्शानेवाले संबोधन हैं। कहा जा सकता है कि पारिवारिक अपनत्व में -ऊ वाले रूप अधिक प्रचलित हैं।

“मामू” रिश्ते के अलावा “याराना” शब्द भी है—*मामू! एक बीड़ी तो पिलाओ।*

आदरार्थक एवं निकटता सूचक

“जान” आदरार्थक है जो सभी रिश्ते और संबोधन सूचक शब्दों के साथ प्रयुक्त होता है, जैसे—
अम्मीजान, अब्बाजान, खालाजान, आपाजान, मामूजान ।

“मियाँ”, “पाशा”, “जनाब” आदरार्थक संबोधन हैं। “पाशाजानी” का प्रयोग भी होता है।

(111) सुनाओ मियाँ! क्या चल राय ?

(112) हामिद मियाँ, कमाल पाशा काँ रैगै ! (रैगै = हिंदी ‘रह गए’)

“यारों” एकवचन और बहुवचन दोनों के लिए प्रयुक्त होता है –

(113) यारों/ यारों तुम खामोश रहो ।

(114) यारों! आज मीटिंग है । (पति पत्नी से)

स्त्रीवाचक प्रत्यय “अगे” संबोधन का शब्द है और सहेलियों के साथ “गे” निकटता दर्शाता है। छोटी बच्चियों को भी इस तरह संबोधित किया जाता है— *सुनगे/ नकोगे* ।

कहीं-कहीं पति-पत्नी आपस में भी “गे” का प्रयोग करते हैं। उदाहरणतः मरद कहेगा—

(115) कित्ता मारी गे!

(116) उठ गे! पानी गरम कर दूँ ।

“अई”, “उई”, “अब्बो”, “हाय”, “हाय अल्ला” औरतों में विस्मय प्रकट करने के लिए खूब प्रचलित हैं । “अब्बो”, “अब्बब्बो” का बोलचाल की दखनी में धड़ल्ले से प्रयोग होता है। “उई मा!”, “हाय अल्ला!” में लडकियों के शर्मने का भाव भी है।

“अम्मा” दक्षिण की भाषाओं में अपनी माँ ही नहीं किसी भी महिला या मालकिन के लिए आदरसूचक घनिष्ठता दर्शानेवाला संबोधन है। यहाँ तक कि महिला नामों के साथ “अम्मा” लगाने की प्रथा है— “नरसम्मा”, “बालम्मा” आदि।

“बाबू” भी नाम में या स्वतन्त्र रूप से संबोधन में पाया जाता है, जैसे— चंद्रबाबू नायडू, आंध्र प्रदेश के भूतपूर्व मुख्य मंत्री। यह उदाहरण भी देखिए— रिक्शेवाला – *बाबू! कहाँ जाना है?*

उसी तरह पुरुषों के लिए आदरार्थक “अय्या” का प्रयोग होता है जो तेलुगु में संस्कृत “आर्य” से आगत है— *नरसय्या, बालय्या* आदि।


विस्मय सूचक “अय्यो!”, “अम्मो!”। तेलुगु शब्द भी दखनी में बखूबी प्रचलित हैं।

संदर्भ-सूची

- Arora, Harbir 2004. *Syntactic Convergence: The Case of Dakkhini Hindi-Urdu*. Delhi: University of Delhi, Publication Division.
- Emeneau Murray B. 1956. "India as a Linguistic Area", *Language* 32,1 (Jan-March, 1956): 3-16.
- सक्सेना, बाबूराम 1952. *दक्खनी हिंदी*. हिंदुस्तानी एकादमी, उत्तर प्रदेश, इलाहाबाद.
- शर्मा वशिनी 1982. "हैदराबादी हिंदी", जगन्नाथन वी° आर° (संपादक): *प्रयोग और प्रयोग* में सम्मिलित। (*A Handbook of Modern Hindi Usage – Prayoga aur Prayoga*) December 24, 1982.
- Subbarao, K. V. 2009. "The Conjunctive Participle in Dakkhini Hindi-Urdu: Making the Best of Both Worlds", *Indian Linguistics* 70: 359-386.
- Subbarao, K. V. 2013. "Models for Language Contact: The Dakkhini Challenge: Formal Approaches to South Asian Languages" (FASAL) -3, Paper presented at the University of Southern California (USC).
- Subbarao, K. V. & Harbir Arora 1989. "Extreme Convergence: The Case of Dakkhini Hindi-Urdu", in: A. Mukherjee (ed.) *Language Change and Language Variation*. Hyderabad: Centre of Advanced Study, Osmania University, 105-122.

वीडियो संदर्भ – यूट्यूब

1. हैदराबाद डायरीज़
2. हैदराबाद कॉमेडीज़
3. वरंगल डायरीज़
4. घाबरामरद हद्दी (दखनी) कॉमेडी

Heinz Werner Wessler 

Cosmopolitan Hindustani Under Aurangzeb: Terminological Matters in François Marie de Tours' *Thesaurus Linguae Indianae*

हाइंस वरनर वेस्लर

औरंगज़ेब के ज़माने की भूमंडलीकृत हिंदुस्तानी: फ्रोंसुआ मारी दे तूर के शब्दकोष में पारिभाषिक शब्दावली पर एक टिप्पणी

Abstract The article explores terminological issues in a hitherto hardly explored chapter in early Hindi lexicography, namely in François Marie de Tours' extraordinary dictionary of 1703, the *Thesaurus Linguae Indianae*. The manuscript consists of 490 folios verso and recto organized in four columns: Latin headword, Hindi word in Devanagari, French rendering, and a phonological transcription with an advanced self-styled set of diacritics. Altogether, the language that the author presents in the dictionary and also in his grammar of Hindustani is astonishingly close to Modern Standard Hindi. Both extraordinary documents appear to be meant to go to the press, but for unclear reasons survived only as manuscripts. The dictionary is an important contribution to the early history of Hindi as well as of the early forms of the European encounter with South Asia. The following article is a provisional report on the project at Uppsala University and some of its findings.

Keywords Hindi, François Marie de Tours, Early Hindi, Surat, Hindi dictionaries.

सारांश यह लेख प्रारंभिक हिंदी शब्दकोषकार्य में पारिभाषिक मुद्दों की पड़ताल करता है। केंद्रबिंदु फ्रांसोआ-मरी दे तूर का असाधारण शब्दकोश 1703 है। पांडुलिपि में चार स्तंभों में संगठित 490 पन्नों पर दोनों तरफ सब मिलकर ११,००० लैटिन के शब्द हैं, हिंदी और फ्रेंच अनुवाद के साथ: लैटिन हैडवर्ड, देवनागरी में हिंदी शब्द, फ्रेंच का रूप, और एक स्वैच्छिक लिप्यंतरण के साथ एक उन्नत स्व-शैलीगत सेट रोमन लिपि में। कुल मिलाकर, लेखक जिस भाषा में भी कोश में प्रस्तुत करता है और हिंदुस्तानी के अपने व्याकरण में, वह आश्चर्यजनक रूप से मॉडर्न स्टैंडर्ड हिंदी के करीब है। दोनों असाधारण दस्तावेज प्रेस में जाने के लिए प्रतीत होते हैं, लेकिन अस्पष्ट कारणों से उस वक़्त छपा नहीं यानी केवल पांडुलिपियों के रूप में बच गया एक महत्वपूर्ण ऐतिहासिक शब्दकोष है। शब्दकोष का हिंदी के प्रारंभिक इतिहास के साथ-साथ दक्षिण एशिया के साथ यूरोपीय मुठभेड़ की शुरुआत में यह एक महत्वपूर्ण योगदान है। निम्नलिखित लेख उप्साला विश्वविद्यालय में परियोजना और इसके कुछ निष्कर्षों पर एक अनंतिम रिपोर्ट है।

मुख्य शब्द – हिंदी, फ्रांसोआ-मरी दे तूर, प्रारंभिक हिंदी, सूत्र, हिंदी शब्दकोश।

1 Introduction

The article explores the significance of an early Hindustani dictionary written by François Marie de Tours, bearing the date of 1703 (see Figure 1). Its digital version has been made available online some time ago.¹ This highly important dictionary as well as the grammar (see Figure 2) of this French Capuchin missionary, who had been in India probably since the 1680s based in Surat, had remained somehow forgotten by the research community until recently. It is only briefly mentioned by McGregor in his review of early Hindi lexicography (McGregor 2001: 9ff; compare McGregor 2003: 947ff).

The dictionary consists of about 11,000 headwords. Together with the grammar, it is an astonishing early masterpiece of missionary linguistics in South Asia and at the same time an important witness of Hindustani in a Gujarati-speaking environment. The lexicographical material allows a glimpse of the linguistic and cultural history of Early Modern Hindustani as a transregional dialect. At the same time, it opens a neglected chapter in the missionary history of the Capuchin Christian mission to the Indian Subcontinent and the history of the encounter between East and West in the early 18th century (Frykenberg 2003; Alam and Subrahmaniam 2007).

Manuscript 840 in the Indian manuscripts collection at the Bibliothèque nationale in Paris—i. e. de Tours' dictionary—appears to have been prepared for typesetting and printing, but it never went to the press. Manuscript 839 in the same collection is a copy of the manuscript from the hands of Anquetil-Duperron

1 <https://archivesetmanuscrits.bnf.fr/ark:/12148/cc96255g> (retrieved February 3, 2021)

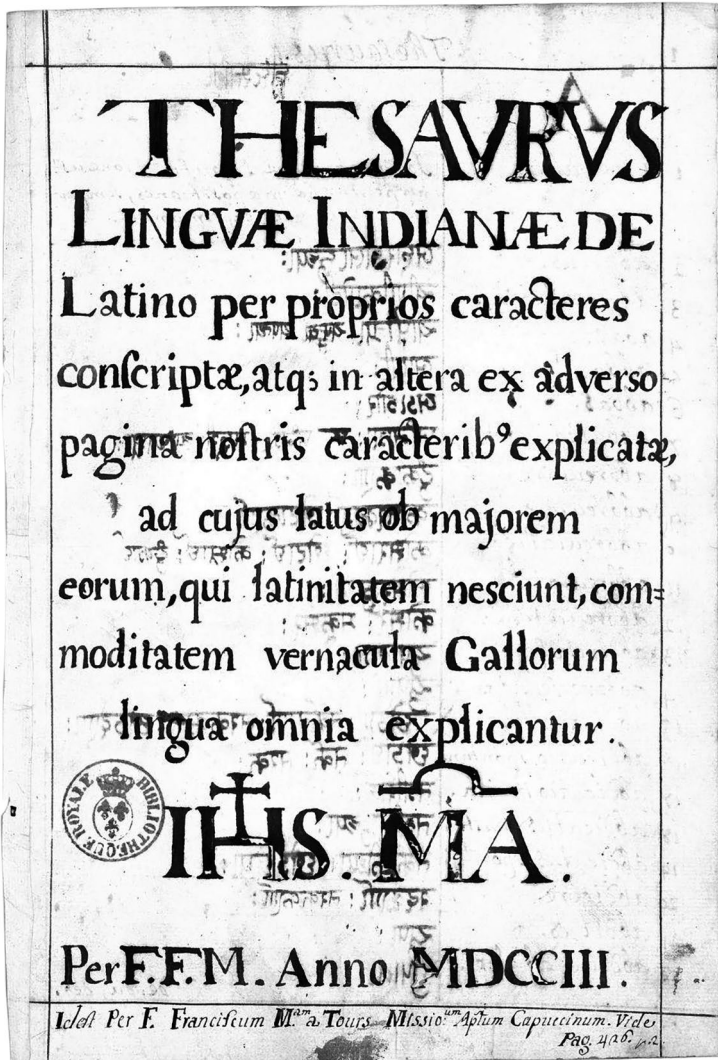


Figure 1. Title page of the dictionary MS 840, Bibliothèque nationale de France, Paris

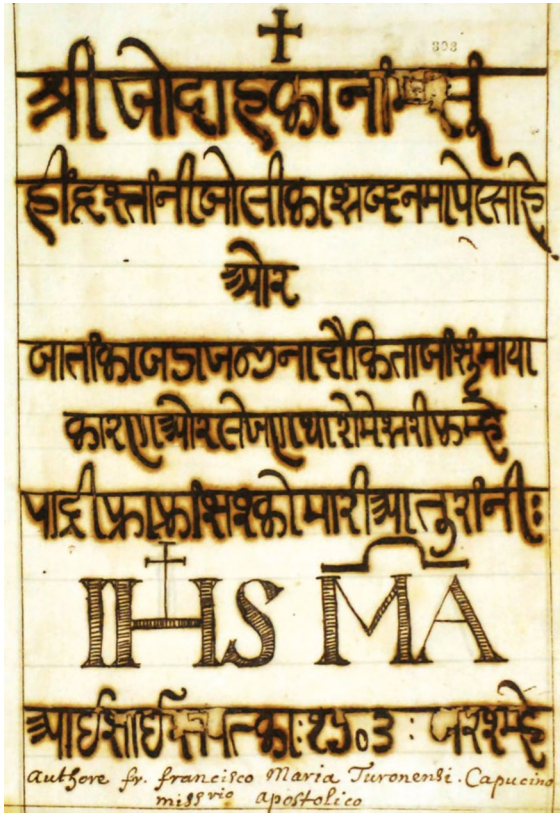


Figure 2. Title page of the grammar, Historical Archives of the Congregation for the Evangelization of Peoples or “de Propaganda Fide”, Urbaniana University, Rome

made in the 1780s.² The following article is a kind of report on a project based at the University of Uppsala and funded by the Swedish Research Council. It will give an overview of the project and a provisional insight into the terminological findings.

The planned “webonary” (digital online dictionary) based on François Marie de Tours’ dictionary is intended to be a starting point for an extended open dictionary of early Hindi based on word lists and dictionaries produced before the foundation of Fort William College in Kolkata (1800) and its impact on the development of Hindi/Hindustani.

2 The history of the original manuscript and the copy of Anquetil-Duperron is part of a research project of Gunilla Gren-Eklund (Uppsala) and will be published soon.

2 The historical and linguistic background of de Tours' Thesaurus

The French Capuchin friar François Marie de Tours came to India probably via the Levante and the Safavid empire—i. e., not via the sea route around the Cape of Good Hope—in the 1680s. The Capuchins had built up a wide network of houses in the Near East in the early 1600s. Since then, a steady flow of friars was sent out on missions towards the East. They would usually stay for some time in the Levante to study Arabic and get accustomed to the Orient before travelling further. The Capuchin order had been active in India since 1632 (Neill 1984; Frykenberg 2003).

Once in India, he got based in the Capuchin house in Surat founded about 1640 by Zénon de Beaugé (1603–1687), who had arrived in Goa in 1637—not much loved by the Portuguese and the Jesuits either—together with two other Capuchin friars sent by the Roman Congregatio de Propaganda Fide (founded 1622), the Pontifical office that tried to establish central coordination of worldwide missions of Catholic institutions—mostly religious orders. The Capuchin mission goes back to a request of Matheus de Castro (1594–1679), the first Indian bishop and Vicar Apostolic to the Kingdoms of Bijapur, Golconda, Abyssinia and Pegu (Rubiés 2001; Neill 1984). Surat (see Figure 3) was the then most important maritime trade hub on the Western coast in today's state of Gujarat, and starting point for the pilgrims to Mecca by ship (Malony 2003). The Capuchins established their central office in this place and therefore outside Portuguese hold territory. Surat had been part of the Mogul empire since the conquest of Gujarat in 1573 and hosted offices

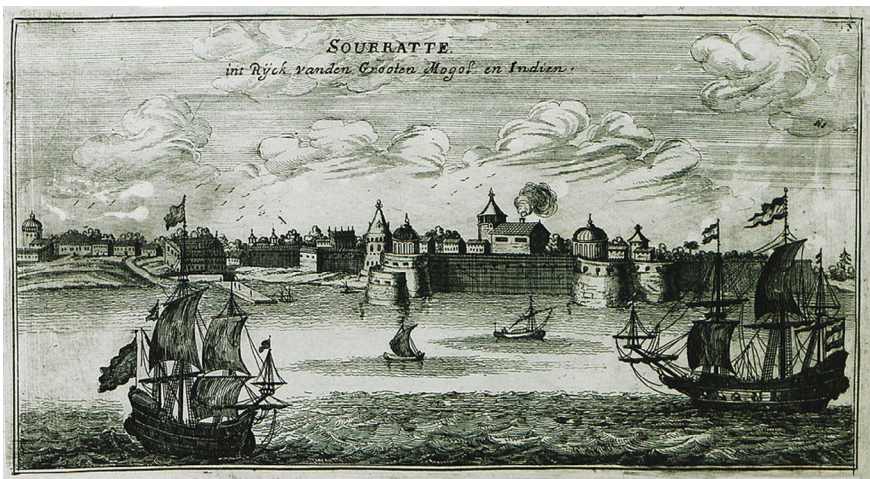


Figure 3. An image of Surat harbour at the beginning of the 18th century

of several European East India Companies in the 16th and 17th centuries before Mumbai (Bombay) took over the role as the central place in maritime trade with European companies.

While Surat remained the centre for Capuchin activities in South Asia, the friars soon opened branches in other places in the South as well as in the North of the empire. Much of de Tours' activities was travelling to the Capuchin houses in different places inside and outside the Mughal empire.

The oldest surviving grammar of Hindi, edited by Tej Bhatia and Kazuhiko Machida, was also written in Surat 1698. Its author was a civil officer of the Dutch East India Company based in Surat, Johan Josua Ketelaar (1659–1719).³ De Tours' dictionary manuscript is only five years younger, i. e. 1703, and was also composed in Surat, even though the surviving manuscript was put on paper during his prolonged stay in Rome (Aranha 2016). The dictionary dated 1703 on its cover page is today found in the Bibliothèque nationale in Paris, while the grammar composed around the same time is together with the complete archive of the Congregatio in the archives of the Urbaniana University in Rome.

De Tours himself uses the name “Hindustani” in the Devanagari title page as well as in the dictionary, where the Devanagari column has the title “Hīndustānī”—the first “ī” is long in his spelling. This is not necessarily astonishing, since in Marathi and partly in Gujarati, which have influenced his orthography as well as his grammar, the difference between long and short /i/ is blurred, particularly in the old versions of the languages. Otherwise, the language is called *lingua indiana* or *lingua mogolana* in the manuscript. Anyway, Hindustani or “the” Indian language is a term that is not astonishing. It appears, for example, also in Benjamin Schulze's famous early grammar (see below). It has, however, nothing to do with the split of “Hindi” and “Hindustani” which gained prominence after John Gilchrist and the early language primer writings in and around Fort William College in Kolkata in the first decade of the 19th century.

India is proud of the history of its grammatical and lexicographical traditions, starting with the linguistic interpretation of Vedic scriptures. The Sanskrit grammarian Panini (date unclear, possibly 4th–3rd century BC) is often regarded as the greatest grammarian in antiquity worldwide. The grammatical tradition, however, did not extend to the New Indo-Aryan languages. From the 16th century onwards, Europeans started to study Indian languages and write grammars and dictionaries. Joan Josua Ketelaar's grammar and dictionary of Hindi, completed in 1698, has been coined “the oldest grammar of Hīndustānī” by the famous Indian

3 A brief description of the complicated manuscript history and the digitized Utrecht manuscript are available on <https://www.uu.nl/en/special-collections/collections/manuscripts/modern-manuscripts/instructie-of-onderwijsinghe-der-hindoustanse-enspersiaanse-taalen-by-ketelaar> (retrieved July 18, 2020)

linguist Suniti Kumar Chatterjee in an article in 1933 and again by the editors of one of the editions of 2008 in three volumes (Bhatia & Machida 2008; see also Bhatia 1987: 21ff). Ketelaar's grammar and dictionary are almost half a century older than Benjamin Schulze's *Grammatica Hindostanica* of 1745 in Latin, which until 1893 was believed to be the earliest grammar of Hindustani resp. Hindi (Bhatia 1987: 50ff).

It is possible that the head of the first Jesuit mission at the court of Akbar (reigned 1556–1605), Jerónimo Xavier (1549–1617) or one of his successors may have produced a dictionary and a grammar of Hindustani (Maclagan 1932: 50ff, 193ff; Neill 1984; Onenkala 2015), which again may have survived unnoticed in some archive and may come to light at some point in the future. The word list of Hindustani preserved in the Marsden collection (MS 11952) in the library of the School of Oriental and African Studies (SOAS) in London under the name of Jerome (Jerónimo) Xavier appears to be much younger, according to a provisional estimate based on writing and paper. It may, however, eventually be a copy of an earlier word list from the Jesuit mission to the Mughal court. This issue needs further research, but for the time being the earliest grammar and dictionary (with less than 2000 headwords) goes to Ketelaar. Ketelaar produced word lists on different fields, all in Roman script in a rather vague transliteration into Roman script.

François Marie de Tours' dictionary is much more conclusive in many respects compared to Ketelaar. Both of them were based in Surat and must have known each other. The *Thesaurus Linguae Indiae*, dated 1703, consists of 427 pages. The layout of the dictionary is in four columns: Latin headword, Hindi equivalent in Devanagari, French rendering, and transcription with a self-styled set of diacritics. In many cases of Arabic or Persian loanwords, Arabic glosses in Arabic script are added to the last of the four columns in two or eventually three different handwritings in Naskh and Nastaliq script. Latin and Devanagari are rectos, while the French glosses and the transcription of each entry are versos just opposite. Altogether, the language that the author describes is astonishingly close to Modern Standard Hindi. Together with the dictionary, de Tours also composed a grammar of Hindustani.

Both documents—dictionary as well as grammar—appear to display a setup that was meant for the printer. In the case of the dictionary, this is particularly clear because of the numbering of each line on each page. Besides, there are occasional notes for the printer. However, neither the grammar nor the dictionary ever (yet) went into print. They survived in two manuscripts that are now being studied in a project at Uppsala University for the first time in detail.

During his visit to Rome in 1703–1704, as far as we know, nothing but a thin publication written by François Marie de Tours in form of a pamphlet was printed. This pamphlet containing 36 'doubts' (*dubia*) takes a position on the controversy on the so-called Malabar rites and was published in Liège (Belgium). It argues against certain forms of "accommodation" that the Jesuits advocated since the

start of the Madurai Mission at the beginning of the 17th century by the Italian Jesuit Roberto Nobili (1577–1656).⁴

The document takes the position in a controversy that was theological, but fought out between Capuchins and Jesuits, and may have been the reason why further publications of François Marie de Tours were blocked from within the Vatican, their value as important documents in missionary linguistics notwithstanding. However, the Capuchin position on the Malabar rites controversy found more and more support in the Vatican, which culminated in the papal constitution *Omnium sollicitudinum* by Benedict XIV from 1744, which restricted accommodative practices of the Catholic mission to a large extent.

The manuscript of the dictionary survived in the Bibliothèque nationale de France in Paris (MS 840), while the grammar is kept in the Archivio Storico of the Vatican.

Before its journey to Paris, the manuscript was preserved in the Archivio Storico di Propaganda Fide, which is today a part of the archives of the Pontifical Urbana University in Rome. It seems that the dictionary was separated from the grammar towards the end of the 18th century.⁵ However, it obviously had come to Paris in the 1780s, where the original and a copy from the hand of the famous orientalist Abraham Hyacinthe Anquetil-Duperron (1731–1805) are preserved in the Oriental manuscript collection of the Bibliothèque nationale (MS 839 and 840). The dictionary was identified as such by Ronald Stuart McGregor (McGregor 2000), but the issue needed further research.

In a brief French introduction to the manuscript, Anquetil-Duperron mentions that he had seen a version of the dictionary (in the form of a manuscript) during his stay in Surat in 1758 in the course of his study of the Zoroastrian textual traditions, and he regrets not having made a copy then. However, he prepared a handwritten copy of the dictionary after he had received the manuscript from Rome in 1784.

One of the interesting facts of the manuscript is that François Marie de Tours used Devanagari with some features of Modi and Gujarati scripts in the manuscript. As is well known, a century later and in Kolkata, John Gilchrist took it for granted that the language he classified as “Hindustani”—different from Hindi—was to be written in Nastaliq. Imre Bangha has however shown that the number of manuscripts from various parts of the Hindi/Hindustani speaking regions is much higher than expected (Bangha 2018). The Hindi/Hindustani binary and its focus on the script is a product of colonial linguistics. It is quite possible that de Tours had

4 More on the significance of François Marie de Tours in the Malabar rites controversy in Aranha 2016.

5 The details of the history of the manuscript is a research subject of Gunilla Gren-Eklund and will be published in due time.

employed an Indian scribe for that part, but in any case, this usage of the script is remarkable. At the same time, the manuscript's constant non-standard orthography is remarkable.

Beyond that, de Tours also developed a rather accurate transliteration with a self-developed system of diacritics that allow an insight into the pronunciation of the Early Modern Hindustani spoken in Surat around 1700. The comparison with Ketelaar shows that his dictionary and grammar were composed in the same epoch and at about the same time as Ketelaar's, but his grammar, as well as his dictionary, are much more extended and profound.

3 De Tours' Dictionary in the context of colonial linguistics

Colonialism did not only change political, administrative and judicial structures. It also had its deep impact on knowledge systems, perceptions and identity. As a kind of fallout of the discourse on orientalism that started with Edward Said's famous study from 1978 (Said 1978), the interaction between colonialism and indigenous knowledge systems has been studied from different angles in recent years.

Sheldon Pollock, with a perspicacious insight into the history of knowledge economies in early colonial South Asia, presents in several publications, particularly in a path-breaking edited volume *Literary Cultures in History: Reconstructions from South Asia* (Pollock 2003), the changing agency of the "vernaculars", i. e. modern languages, in the knowledge economies of the 18th century. Pollock's main focus is their exploration in the context of what is often termed "Early Modernity" or "Vernacular Modernity" in contemporary research.⁶ He sees similarities in the development of what he calls the "vernacular millennium" in Europe and India leading to a new form of cosmopolitanism.

Bernhard S. Cohn in his book on *Colonialism and Its Forms of Knowledge: The British in India* (Cohn 1996) argues that the British Orientalists' study of Indian languages was of primary importance to the colonial project of control and command. The seventeenth- and eighteenth-century travelling friars and missionaries have often been interpreted as agents of colonialism, but it has to be added that the anti-Orientalist and post-colonial gaze is more applicable to the later epoch of colonial domination after the drastic decline of the Mogul empire and the simultaneous rise of the East India Company Raj.

Sumit Guha (2011) explains how "lexical awareness", a central issue in missionary linguistics, has to be interpreted in a broader historical context that goes beyond the binary of the colonizer and the colonized. In recent years, the term

6 Compare also Pollock 2002 and Pollock 2011.

“cosmopolitanism” has been explored as an analytical tool in the study of cultural dynamics of Early Modern India (Lefèvre, Županov & Flores 2015). The focus is on the complexities of the interaction between different discourses, their languages and works of literature, and on the interest bearers in the interaction between South Asia, Europe, and Central Asia. Cosmopolitanism signals a shift from sociality to humanity, from primordial identities as terms of reference for group solidarity towards open discourse in a pluralist setup. The presence of Catholic missionaries in India before the rise of the British East India Company and particularly before the battle of Plassey 1757 in the Mogul empire has recently been discussed in the context of “Catholic Orientalism” (Xavier & Županov 2015).

Part of this endeavour was the linguistic and lexicographical research in New Indo-Aryan languages. This is the starting point of the Hindi grammatical tradition, which Tej Bhatia (1987: 15), therefore, calls “an alien tradition”. In the late 16th century, Christian missionaries started to study Indian languages, write grammars and dictionaries. It is astonishing that this endeavour appears not to have been extended to transregional Hindustani.

The first linguistic exploration of Hindi appears to have been made around the year 1700 and in Surat. It is, however, possible that a grammar of Braj, usually taken as a dialect of Hindi in the standard linguistic taxonomy, was composed in the Persian language about the same time as Ketelaar’s and François Marie de Tours’ grammars and dictionaries, or perhaps even earlier. This grammar constitutes part of the *Tuhfat ul-hind* by Mirzā Khān ibn-Fakhr ud-Dīn. The editor of the critical edition from 1935, M. Ziāuddīn, believes that Mirzā Khān’s grammar was written in or before 1676. Bhatia, however, persuasively suggests that the given evidence may rather refer to the year 1711, or later (Bhatia 1987: 19; McGregor 2003: 942ff).

Ketelaar, who was in the service of the Dutch East India Company, as well as François Marie de Tours, maintained relationships with higher dignitaries of the Mughal administration and even with the court itself. A manuscript kept in the University library in Uppsala from the Christopher Henrik Braad (1728–1781) collection, extracted from the French Capuchin diaries in Surat starting about 1650, mentions the name of Ketelaar, thus it is not impossible that Ketelaar and de Tours might have known each other. There is, however, no indication that the protestant Ketelaar and the Capuchin de Tours have entered a dialogue on linguistic or lexicographical issues.

Until the beginning of the 18th century, Surat was a much more important harbour and coastal town than Bombay on the Western coast, and the residence of a large number of European merchants and missionaries.

4 The dictionary

François Marie de Tours' stay in the region was interrupted by a return to Rome in 1703 as acting procurator of the French Capuchin missionaries in South Asia, and his grammar and dictionary were finished and ready for layout and printing. Several notes for the printer in the grammar manuscript demonstrate that it was meant to be handed over to the printing press for publication. Why this did not happen is not clear, but it might be related to the rivalries between the different orders engaged in India (Aranha 2016).

In his dictionary, François Marie freely translates Christian theological terms using Islamic or Hindu terminology (see Figures 4 and 5). The semantic study of these terms and the word patterns used are part of the project in Uppsala.

18	<i>amissio.</i>	जीव्यार्ष कसारः।
19	<i>amissus.</i>	जीया। जीव्या।
20	<i>amittere.</i>	जीयाः।
21	<i>amitti.</i>	जीयाः।
22	<i>amodo.</i>	आन्शः। आनीशः।
23	<i>amanitas.</i>	पुनी।
24	<i>amonus.</i>	पुनी। तमशकत्र।
25	<i>amor.</i>	देवताः। आरः। मोहिजनः।

Figure 4. A recto page from the dictionary

18	<i>perfe.</i>	Kóój, Kesára
19	<i>perdu.</i>	Kóóca, Kóógca.
20	<i>perdre.</i>	Kóónia.
21	<i>se perdre.</i>	Kóógána.
22	<i>dores en avant.</i>	dyson, labison,
23	<i>agréable, plaisante</i>	Kouchi.
24	<i>agréable, plaisant.</i>	Kouch, Kagescava.
25	<i>amour.</i>	doshi, pean. móhabet.

Figure 5. The next page (verso) with the French gloss and De Tours' transliteration

Deus *khodā, prameśvar*

Khudā is Persian (Sanskrit calque *svayaṃdatta-*, in the sense ‘he who determines upon himself’, compare also Skt. *svayaṃbhū* ‘he who has come into existence through himself’), while *prameśvar* is a corrupt spelling of the Sanskrit loanword *parameśvar*. Unorthodox spellings of Sanskrit loanwords appear quite regularly. Perso-Arabic and Sanskrit synonyms are very often mixed even in the sensitive realm of theological terminology.

Fidentia *vīsvās*

Viśvās is still a common Sanskrit gloss for ‘belief’ in modern Hindi, be it in Hindu or Christian contexts. The long “ī” instead of the grammatically correct short “i” is the common orthographical feature, even in the case of Sanskrit loanwords like this one. This relates to a common feature in spoken Gujarati and Marathi not to distinguish between /i/ and /ī/.

Fides divina *dīn, yīmān, mān*

“Divine belief”, different from “trust”, is where de Tours would semantically localize an Arabic loanword that is often used as equivalent to “religion”, *dīn*. The word *īmān* also is a borrowing from Arabic, while *mān* is Sanskrit. Interestingly, the dictionary does not use *dharm* as an equivalent. The words *dīn*, *īmāṃn* and *ansāf* (spelling of de Tours) appear again under the headword *religio*.

Fides humana *īnsāf, ītbār*

The dictionary has a different set of Arabic-origin glosses under the headword “human belief”, namely *ītbār* (typically with a long vowel “ī” in place of the Arabic short “i”).

Credere *mānṇā, atbārīrakhṇā, īmānrakhṇā*

“To believe” can be used in religious as well as in a non-religious context: *mānṇā*, *etbār rakhnā* and *īmān rakhnā* in Modern Standard Hindi (MSH). The treatment of the two parts of a conjunct verb as one word is a common feature in the dictionary, as well as the retroflexion of the infinitive suffix *-ṇā*.

Creare *paidekarṇā*

Creator *paidekarṇār*

“To create” is taken as a religious term: *paidā karṇā* in MSH,⁷ and the designation for “creator” corresponds to *paidā karne-vālā* in MSH, common suffixation forming an agent noun. In this case, the dictionary does not go back to a theological term from either the Islamic or the Hindu world. In some cases, this appears to be

7 The examples in this article are used in MSH and Urdu as well.

an ad hoc decision, but it could also be an effort to develop a Christian terminology, avoiding a terminology with a theological echo from either Hinduism or Islam.

Crucifigere *sūlīmhelaḡāvnā; sūlīkumārñā*

This would in MSH probably be *sūlī meḡ lagānā* ‘to put on the stake’, ‘to execute’ or ‘to put on the cross’. Similarly, *sūlī ko mārnā* or, more clearly, *sūlī par mārnā*. A Christian religious term (*śūlī*) goes back to Sanskrit *śūlikā*. Interestingly, Arabic *ṣalīb* is not used here, which de Tours must have known from his period in the Levante. This is a completely new effort to explain the meaning of crucifixion in Indian terms.

Salvator *khāles; mukhāles; salāmatīdār; coṭyā*

In this case, Arabic *khālis* and *mukhalliṣ* are used together with the ‘giver of peace’, i. e. Perso-Arabic *salāmatīdār*, and an indigenous nominal form related to MSH *chor-* ‘to release, set free’.

Salus *taslīm; dūā; mūjarā; saām; salāmat; khālasī; coṭ*

Religious ‘salvation; safety’ is glossed with six terms, among them five Arabic, and one being an indigenous (*coṭ*, corresponding to MSH *chūṭ*).

Salutare *dūaikarñā; mūjarekarñā; taslīmkarñā*

‘To save; to liberate; to greet’: Once again, three conjunct verb constructions with an Arabic word as first part plus MSH *karnā* ‘to do’: *dūā* ‘blessing, prayer’, *mujarā* (short “u”) ‘obeisance’ and *taslīm* ‘greeting’.

Sanctitas *pākījā*

MSH *pākīzā* is a Perso-Arabic loanword and an adjective, meaning ‘delicate, chaste’. In English, “holy” can be nominalised and adjective as well, but MSH *pākīzā* is clearly an adjective. This might be a mistake, since Latin *sanctitas* is clearly a noun.

Sanctuarium *devūḡā*

This must be a derivative of Sanskrit *deva* ‘god’ (originally one of many), similarly to MSH *devālā*.

Sanctus *pākbādhā; pīr; pīrjādā*

Under the keyword *sanctus* “holy” the Perso-Arabic adjective *pāk* ‘purified’ (cf. *pākīzā*), same as in MSH, appears combined with another adjective formation that goes back to Sanskrit *bādhaka/baddhaka* ‘bound, fixed’. The two following words are of Persian origin: *pīr* ‘old, to be respected’ and *pīrzādā* ‘related to the respected (person), son of a pīr’.

Reliquiae bākī

MSH *bāqī* ‘remains’ is a rather prosaic translation, even though it is Arabic in origin.

This short presentation of examples illustrates that the dictionary is rather pragmatic in the use of its theologically relevant terminology. Terminology without theological implications follows pragmatic patterns. It is unclear, who exactly the major informants of de Tours were, but it can be estimated that they belonged to the environment of converts.

5 Conclusion

Research in the semantics, orthography, phonetics and grammatical aspects in Early Modern Hindi lexicography is still in an initial stage. This involves historical linguistics of New Indo-Aryan languages, missionary linguistics, missionary history and cultural and social studies concerning the encounter between East and West in its formative phase.

In his famous Hindustani grammar of 1796, John Gilchrist argues that the language he describes could serve as a medium of administration in the East India Company’s territories better than the traditional Persian (Bhatia 1987: 79ff). This argument finds surprising support in de Tours’ statements on the functional capacity of the Hindustani of his epoch. Hindi—or, to be precise, Modern Standard Hindi (based on the dialect *Khaṛī bolī*)—is one of India’s 22 scheduled languages and the only official language of the Union of India, as stated in the Indian constitution from 1950 in its famous 8th Schedule in its amended form of 2002. Altogether, Hindi claims a literary tradition that goes back at least to the 12th century, but much of its earlier tradition is connected with dialects that are grammatically and lexicographically quite distinct from Modern Standard Hindi. The hitherto mostly unnoticed dictionary is of great importance for the early history of *Khaṛī bolī* as a “transregional idiom”, as McGregor calls it without referring to our lexicographer (McGregor 2003). De Tours’ dictionary also is an important trace of the pre-history of Hindi as an official language of India according to its constitution of 1950 and the modern question of the identity of Hindi (Wessler 2020).

The reference to Hindi/Hindustani as the *lingua mogolana* that is understood as lingua franca (*lingua vulgaris*) all over the Empire as well as on the coasts in the vicinity is a strong statement on the significance of Hindi long before the days of Fort William college’s *bhākhā munśīs*, and before Bhāratendu Hariścandra (1850–1885) and the growth of Hindi into its role as the official language of India according to the constitution of 1950.

François Marie de Tours’ dictionary “has the scope of a substantial reference work”, as McGregor rightly concludes (McGregor 2001: 11). As my preliminary presentation of terminological findings demonstrates, the approach of the

dictionary is pragmatic and cosmopolitan: it has a certain preference for Perso-Arabic terminology even when relating to theologically delicate matters, and at the same time it refers to Sanskrit terms as synonyms or in some cases solely to Sanskrit terms. Linguistic decisions on the use of terminology are hardly to be traceable, except that the dictionary does not much refer to Portuguese terms as loanwords.

The planned digital edition could form a basis for a conclusive webonary of Hindi in the Early Modern phase of Indian history, i. e. before the Battle of Plassey (1757) and the growth of the British East India Company Raj, before Lord Macauley's notorious "Minute on Education" of 1835 and its consequences for the status of the so-called vernaculars, and before the change from Persian to Urdu as the language of administration in the East India Company's North Indian territories in 1837.

ORCID®

Heinz Werner Wessler  <https://orcid.org/0000-0003-3053-4540>

Bibliography


- Alam, Muzaffar & Sanjay Subrahmaniam 2007. *Indo-Persian Travels in the Age of Discoveries 1400–1800*. Cambridge: Cambridge University Press.
- Aranha, Paolo 2016. "Vulgaris seu Universalis: Early Modern Missionary Representations of an Indian Cosmopolitan Space", *Puruṣārtha* 33: 331–360.
- Bangha, Imre 2018. "The Emergence of Hindi Literature: From Transregional Maru-Gurjar to Madhyadeśī Narratives", in: Williams, Tyler, Anshu Malhotra & John Stratton Hawley (eds): *Text and Tradition in Early Modern North India*. Oxford: Oxford University Press.
- Bhatia, Tej K. 1987. *A History of the Hindi Grammatical Tradition: Hindi-Hindustani Grammar, Grammarians, History and Problems*. (Handbook of Oriental Studies. Section 2 South Asia 4.) Leiden et al.: Brill.
- Bhatia, Tej & Kazuhiko Machida 2008. *The Oldest Grammar of Hindustānī: Contact, Communication and Colonial Legacy*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Cohn, Bernhard S. 1996. *Colonialism and Its Forms of Knowledge: The British in India*. Princeton: Princeton University Press.
- Frykenberg, Robert Eric (ed.) 2003. *Christians and Missionaries in India: Cross-Cultural Communication Since 1500*. (Studies in the History of Christian Missions.) Grand Rapids, Michigan: Eerdmans.

- Guha, Sumit 2011. “Bad Language and Good Language: Lexical Awareness in the Cultural Politics of Peninsular India, ca. 1300–1800”, in: Pollock, Sheldon (ed.): *Forms of Knowledge in Early Modern Asia: Explorations in the Intellectual History of India and Tibet, 1500–1800*. Durham; London: Duke University Press, 49–68.
- Lefèvre, Corinne, Ines G. Županov & Jorge Flores (éds) 2015. *Cosmopolitismes en Asie du Sud. Sources, itinéraires, langues (XVI^e–XVIII^e siècle)*. Paris: Éditions de l'École des hautes études en sciences sociales.
- Malony, Ruby 2003. *Surat: Port of the Mughal Empire*. Mumbai et al.: Himalaya Publishing House.
- McGregor, Ronald Stuart 2001. *The Formation of Modern Hindi as Demonstrated in Early 'Hindi' Dictionaries*. 2000 Gonda lecture. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- McGregor, Ronald Stuart 2003. “The Progress of Hindi, Part I: The Development of a Transregional Idiom”, in: Pollock, Sheldon (ed.): *Literary Cultures in History: Reconstructions from South Asia*. Berkeley: University of California Press, 912–957.
- Maclagan, Edward 1932. *The Jesuits and the Great Mogul*. London: Burns, Oates & Washborne.
- Neill, Stephen 1984. *A History of Christianity in India: The Beginnings to AD 1707*. Cambridge: Cambridge University Press.
- Onenkala 2015. “The Jesuit Perception of Later Mughals, Shah Jahan and Aurangzeb (1627–1707 AD)”, *Indian Church History Review* 49: 85–96.
- Pollock, Sheldon 2002. “Cosmopolitan and Vernacular in History”, in: Breckenridge, Carol A., Sheldon Pollock, Homi K. Bhabha & Dipesh Chakrabarty (eds): *Cosmopolitanism*. Durham et al.: Duke University Press, 15–53.
- Pollock, Sheldon (ed.) 2003. *Literary Cultures in History: Reconstructions from South Asia*. Berkeley: University of California Press.
- Pollock, Sheldon 2011. “The Languages of Science in Early Modern India”, in: Sheldon Pollock (ed.) *Forms of Knowledge in Early Modern Asia: Explorations in the Intellectual History of India and Tibet, 1500–1800*. Durham; London: Duke University Press, 19–48.
- Rubiés, Joan-Pau 2001. “The Jesuit Discovery of Hinduism: Antonio Rubino’s Account of the History and Religion of Vijayanagara (1608)”, *Archiv für Religionsgeschichte* 3: 210–256.
- Said, Edward W. 1978. *Orientalism*. London: Routledge & Kegan Paul.
- Wessler, Heinz Werner 2020. “Indian and Universal? The Significance of François Marie de Tours for the Case of Hindi”, in: Brandt, Carmen & Hans Harder (Hrsg.): *Wege durchs Labyrinth: Festschrift zu Ehren von*

- Rahul Peter Das*. Heidelberg; Berlin: CrossAsia-eBooks, 469–491
(DOI: <<https://doi.org/10.11588/xabooks.642>>, accessed: February 4, 2021).
- Xavier, Ângela Barreto & Ines G. Županov 2015. *Catholic Orientalism: Portuguese Empire, Indian Knowledge (16th–18th Centuries)*. New Delhi: Oxford University Press.

PART II

VARIATION IN THE GRAMMAR AND DISCOURSE OF STANDARD HINDI

Liudmila Khokhlova 

Conative: Compleitive Contrast in Hindi-Urdu Aorist Forms

ल्यूडमीला खोख्लोवा

हिंदी-उर्दू में सामान्य भूतकाल में प्रयत्न पक्ष और पूर्ण पक्ष का व्यतिरेक

Abstract Semantic contrast in the use of compound versus non-compound verbs in Hindi has been analysed by many scholars. Compound verbs (sequences of the main verb, conveying the basic meaning and the so-called ‘vector’ or ‘polar’ verbs, conveying attitudinal and aspectual specifications) are usually described as marked members in the opposition complete/incomplete action. It will be shown in this paper that compound verbs have compleitive meaning only in the case where the main verb is telic. If the main verb is not telic, the compound verb denotes the beginning of the action.

Almost all scholars (cf., e.g. Hook 1974; Nespital 1997; Montaut 2004) are of the opinion that the aorist forms generated from simple verbs are compatible with some arbitrary end point and may be followed by a culmination-cancelling clause. Contrary to this, the aorist forms produced from compound verbs imply that the event has culminated at its natural endpoint. However, in certain cases the compound verbs may also be compatible with an arbitrary endpoint (Kothari & Arunachalam 2009). The paper seeks to demonstrate that this dichotomy of compound verbs is based on the semantic properties of the theme object of the utterance: if it is incremental¹, the aorist form of a transitive compound verb may indicate only partial, but not complete change in the state of the theme.

1 I am using the term ‘incremental theme’ after Dowty (1991) who applied it to the argument of certain predicates involved in defining a homomorphism from its own physical extent to the temporal progress of the event it participates in. By this definition, verbs like ‘read’, ‘write’, ‘eat’, ‘sing’, ‘climb’, ‘clean’ are incremental (object) theme verbs. The term is used both for theme and patient, as their differentiation is not important for the present study.

Aorist forms of compound verbs are never used in the following cases: 1) the state of the object changes instantly; 2) the verb contains a conative component in its semantics; 3) the state of the object has not undergone any change.

Keywords Hindi aorist, compound verb, telic verb, completive meaning, incremental theme.

सारांश हिंदी की संयुक्त बनाम सरल क्रियाओं के अर्थपरक व्यतिरेक का विश्लेषण कई विद्वानों द्वारा किया जाता रहा है। कार्य-व्यापार के पूर्ण/अपूर्ण व्यतिरेक की दृष्टि से संयुक्त क्रियाओं को (संयुक्त क्रिया = मुख्य क्रिया + रंजक क्रिया जिसमें मुख्य क्रिया मूल अर्थ की वाचक होती है जबकि रंजक क्रिया विशिष्ट अभिवृत्तिपरक और पक्षपरक अर्थ को अभिव्यक्त करती है) आम तौर पर इस व्यतिरेक के चिह्नित सदस्य के रूप में देखा जाता है। इस आलेख में यह स्थापित किया गया है कि अगर मुख्य क्रिया प्रयत्नबोधक है तो संयुक्त क्रियाओं का अर्थ पूर्णताबोधक होगा। अगर मुख्य क्रिया प्रयत्नबोधक नहीं है तो संयुक्त क्रिया कार्य-व्यापार के आरंभ को दर्शाती है।

लगभग सभी विद्वानों (cf. Hook 1974; Nespital 1997; Montaut 2004, et al.) की यही राय है कि सरल क्रियाओं से उत्पन्न सामान्य भूतकाल रूप कतिपय मनमाने अंतिम छोर के अनुरूप होते हैं और उनके पीछे समापन-निष्प्रभाव करने वाले उपवाक्य आ सकते हैं। इसके विपरीत संयुक्त क्रियाओं से उत्पन्न सामान्य भूतकाल के रूपों का अर्थ यह होता है कि घटना का समापन उसके प्राकृतिक छोर पर हुआ है या होता रहा है, लेकिन कुछ मामलों में संयुक्त क्रियाएँ मनमाने अंतिम छोर (Kothari & Arunachalam 2009) के अनुरूप भी हो सकती हैं। इस आलेख में यह दर्शाने का प्रयास किया गया है कि संयुक्त क्रियाओं का यह द्विभाजन उक्ति के कथ्य वस्तु के आर्थी गुणों पर आधारित होता है: अगर यह वर्धमान है तो सकर्मक संयुक्त क्रिया का सामान्य भूतकाल रूप केवल आंशिक रूप में ही प्रकट हो सकता है, लेकिन कथ्य की अवस्था में पूर्ण परिवर्तन नहीं हो सकता।

संयुक्त क्रियाओं के सामान्य भूतकाल रूपों का प्रयोग निम्नलिखित परिस्थितियों में कभी नहीं होता है: 1) कर्म की अवस्था में तुरंत परिवर्तन होता है; 2) क्रिया के अपने अर्थ में क्रियात्मक (प्रयत्न पक्ष का) घटक निहित होता है; 3) कर्म की अवस्था में कोई परिवर्तन नहीं हुआ है।

मुख्य शब्द – हिंदी में सामान्य भूतकाल (aorist), संयुक्त क्रिया (compound verb), प्रयोजनमूलक क्रिया (telic verb), पूरक अर्थ (completive meaning), वर्धमान कथ्य (incremental theme)।

1 Introduction

This paper aims at throwing some additional light on the meanings of simple and compound aorist forms in Hindi and Urdu. The simple synthetic aorist is formed in Hindi by adding gender/number inflexions to the perfective verbal stem,² e. g. V-ā: *parh-ā* ‘read’, *gay-ā* ‘[he] went’. Compound verbal forms include the primary stem of a ‘major’ verb belonging to an unrestricted class and a finite form of a ‘light’ verb from a restricted verb class. The most frequently used light verbs are

- 2 The Hindi perfective suffix is -Ø- (zero) after consonant-final primary stems and -y- after vowel-final stems. For the sake of simplicity, they are not marked separately in the morphological glosses.

lenā ‘take’, *denā* ‘give’, *ānā* ‘come’, *jānā* ‘go’, *uṭhnā* ‘get up, rise’, *baiṭhnā* ‘sit down’, *paṛnā* ‘fall’, *ḍālnā* ‘throw’, and *rakhnā* ‘put’.

The main objective of this study was to discover and describe conditions under which compound verbs in the aorist form do not denote the completion of the action, thus deviating from their common distribution. It will be shown that a compound verb lacks the meaning of completeness if the main verb is not telic, and even telic verbs may denote incomplete action on condition that the object of the utterance is incremental.

The method used in the present work has been based on the eliciting of responses or reactions to utterances (see examples (23)–(31)) presented to speakers of Standard Hindi. The main respondents were students and teachers from Delhi University, Jawaharlal Nehru University in Delhi and Mahatma Gandhi Antarrashtriya Hindi Vishwavidyalaya in Wardha.³

I prepared small situations that were checked by two Hindi speakers and presented to the respondents in written form. The respondents were asked to read the text in my presence and correct the mistakes or write (+) if there were no mistakes. For example, sentence (9) below was put into the following situation:

*āj merī parīkṣā thī. māñ ne sāre savāl hal kiye, pāḍvāñ savāl thorā
muśkil thā, māñ ne hal to kiyā par hal nahī huā.*

In case all the respondents put (+), the sentence was considered correct: sentences (3), (8) – (12). As for utterances (23) – (31), each one was presented to the respondents in several situations. For example, sentence (25) ‘Gita sang a song but she forgot the last line (of it)’ was given in two situations: (1) Gita left without completing the song; (2) Gita remembered the last line and sang the song up till the end. The respondents were asked to choose either a simple or compound verb in each case. They were also asked to modify or extend the context to make the utterance more acceptable. Below I give the responses of five out of twenty-five informants that participated in this project. These students were particularly well versed in Hindi and Urdu prose and poetry. However, their responses did not differ significantly from those of other participants.

3 I am extremely grateful for guidance in my work to all faculty members of the Linguistic Department of Delhi University and the Centre for Linguistics in JNU University (Delhi), and especially to Professor Pradeep Kumar Das (JNU) as well as Professors Anil K. Pandey and Hanuman P. Shukla (Mahatma Gandhi Antarrashtriya Hindi Vishwavidyalaya, Wardha). I am deeply indebted to the students of these universities. My work would not have been possible without the help in particular of Shubha Srivastava, Jyoti Sharma, Nivedita Verma and Maaz Shaikh from Delhi University

2 Synthetic aorist forms

Classical Hindi grammarians like Kāmtā Prasād Guru (Guru 2017 V. S. = 1960 C. E.) and later linguists working in the same tradition tend to pay more attention to verb formation and the syntactic properties of verbal forms than to the meanings those forms are used to convey in different contexts. A typical description of the aorist in Hindi grammars may be found in Pande (2012: 180–181): *sāmānya bhūt ke dvārā vartmān se [...] pahle pūre hue vyāpār kā bodh hotā hai* ‘The simple past denotes action that took place before the present moment’. The author goes on to state that the action may take place both in the remote and more recent past (ibid.).

Yamuna Kachru termed this V-ā tense form (*kah-ā*) ‘simple perfect’—in opposition to present perfect (*kahā hai*) and past perfect (*kahā thā*). According to her, the simple perfect indicates a state resulting from the completion of an action or process. The present perfect indicates in addition the continued relevance of the completion of the action or process to the present moment, while the past perfect indicates completion in the past but without any continued relevance to the present time (Kachru 1980: 132).

Kachru’s description states the prototypical meanings of the three variants in denoting past action, but it does not suggest explanations for the simple past when used to mark the future (1) or for its usage in contexts where this tense indicates the continued relevance of the completion of the action or process to the present moment (2):

- (1) *abhī ā-yā*
just now come-AOR.M.SG
‘I am coming at once.’ (Montaut 2004: 106)

- (2) *are kitnā barā ho ga-yā*
interj how much tall be go-AOR.M.SG
‘O, he has grown so tall! / How tall he has grown!’ (exclamation of surprise) (ibid.: 107).

All those and many other non-trivial cases of usage of V-ā forms are described in Annie Montaut’s *Hindi Grammar* (Montaut 2004: 102–111)—the work which, among its other merits, is very useful for learning Hindi as a foreign language. In particular, unlike many Hindi manuals written in English, it teaches one to avoid the temptation to identify the usage of simple, present and past perfect in Hindi with that of the corresponding English forms.⁴ Following Montaut, I use for the

4 Many Hindi manuals written in English do this. For example, in a very good manual, *Teach Yourself Hindi* by Rupert Snell and Simon Weightman, which is one of the most

V-ā finite form the term ‘aorist’. She prefers this term because ‘...contrary to perfective known for instance from Slavonic languages, it is not used in infinitival or imperative as is usual with languages having a perfective /imperfective opposition, [...] besides, it does contrast with verbal vectors which partly convey perfective meanings’ (Montaut 2004: 103–104). According to Montaut, ‘...this form expresses a bound compact interval with no relation to the present time of speech: no adjacent state is taken into account, no internal divisions of the process, [...] its basic use, often the only one mentioned in grammars, refers to past events, [...] it is the most adequate form for the objective narrative where events are supposed to be told by nobody’ (Montaut 2004: 104–105). After giving this basic/prototypical definition of aorist, Montaut analyses quite a number of cases of non-trivial usage of the corresponding forms—including those of the types (1) and (2) mentioned above (Montaut 2004: 106–107). However, she does not pay enough attention to the ‘conative:completive’ contrast in Hindi aorist forms, maybe because this problem has been many times discussed by other Hindi grammarians (Hook 1974; Kachru 1980, etc.).

According to Kachru, the present, past and simple perfect indicate ‘the completed action or process’ (Kachru 1980: 132). She does not agree with Hook (1974) who argues that V-ā forms, when generated from simple verbal stems, are unmarked for aspect and thus may express “either perfective or imperfective view of the action”. In other words, they imply simultaneously an attempt to perform the action (the conative meaning), and also the completion of the action, that is, its successful performance to its logical end (the completive meaning) (Hook 1974). In an attempt to rebut Hook’s arguments Kachru gives a Hindi sentence which she considers to be ill-formed:

- (3) *rām ne tasvīr banā-ī par tasvīr nahī*
Ram ERG *picture.F.SG* *make-AOR.F.SG* *but* *picture.F.SG* *NEG*
ban-ī
 be.made-AOR.F.SG
 ‘Ram drew a picture but the picture did not get drawn.’ (Kachru 1980: 137)

However, all native speakers whom we consulted found this sentence to be correct.

popular in Europe, the perfect tenses are defined as ‘those which refer to an action that has been completed’, and, as illustrations, the following Hindi samples together with their English equivalents are suggested: *māī bolā* ‘I spoke’, *māī bolā hū* ‘I have spoken’, *māī bolā thā* ‘I had spoken’ (Snell & Weightman 1989: 122–123). Clearly, this type of description will not give students the necessary clues for translating correctly from English into Hindi.

Other Hindi grammarians (e. g., Nespital 1997; Pořízka 1967–1969; Liperovskij 1984) share the opinion that V-ā forms may have or not have the completive meaning, as examples (4) – (12) show:

- (4) *maĩ=ne rām ko apn-e pais-e di-ye*
 I=ERG Ram DAT self's-M.PL money-M.PL give-AOR.M.PL
 'I gave money to Ram.' (It is not clear if Ram took the money.)

- (5) *maĩ=ne parāṭh-ā khā-yā*
 I=ERG paratha-M.SG eat-AOR.M.SG
 'I ate the paratha⁵.' (Maybe only part of it) (Hook 1974: 167)

Such type of sentences may be followed by a culmination-cancelling clause:

- (6) *maĩ=ne us-e pais-e di-ye*
 I=ERG he-DAT money-M.PL give-AOR.M.PL
lekin us=ne nahĩ li-ye
 but he=ERG NEG take-AOR.M.PL
 'I tried to give money to Ram, but he did not take it.' (Hook 1974: 165)

- (7) *maĩ=ne parāṭh-ā khā-yā*
 I=ERG paratha-M.SG eat-AOR.M.SG
lekin sār-ā nahĩ
 but whole-M.SG NEG
 'I ate the paratha, but not all of it.' (Hook 1974: 167)

The transitive verbs are often used with conative meanings in sentences with their intransitive correlates generated from the same root:

- (8) *kisān ne khet jot-ā par vah nahĩ*
 peasant ERG field.M.SG plough-AOR.M.SG but it NEG
jut-ā
 become.ploughed-AOR.M.SG
 'The peasant ploughed (= 'tried to plough') the field but it did not become ploughed.'

5 A kind of flatbread.

- (9) *mohan ne savāl hal ki-yā par vah hal nahī hu-ā*
 Mohan ERG problem solution do-AOR.M.SG but it solution
 NEG be-AOR.M.SG
 ‘Mohan tried to solve the problem but it was not solved.’
- (10) *laṛk-e ne akhroṭ tor-ā par vah nahī ṭūṭ-ā*
 boy-OBL.SG ERG walnut.M.SG crack-AOR.M.SG but it NEG
 be.cracked-AOR.M.SG
 ‘The boy tried to crack the walnut but it did not crack.’
- (11) *naukar ne khiṛk-ī khol-ī par khiṛkī nahī khul-ī*
 servant ERG window-F.SG open-AOR.F.SG but window.F.SG NEG
 be.open-AOR.F.SG
 ‘The servant tried to open the window but it did not open.’
- (12) *maĩ=ne bahin ko bahut samjhā-yā par vah samjh-ī nahī*
 I=ERG sister DAT much explain-AOR.M.SG but she
 understand-AOR.F NEG
 ‘I many times explained (it) to (my) sister but she did not understand.’

3 The aorist forms of compound predicates

A number of scholars (e. g. Liperovskij 1984; Singh 1991; Butt 1995) have shown that compound verbs may signal either the inception or completion of a given action. The light verbs are analyzed as focusing either on the initial (the verbs *parṇā* ‘fall’, *denā* ‘give’, etc.) or on the final stage of the event (the verbs *lenā* ‘take’, *ḍālnā* ‘throw’, etc.). It seems that this statement is true only statistically, because the inceptive versus completive meaning of a compound verb depends on the telicity of the main verb. Compound verbs formed from non-telic stems usually bear inceptive meaning, while those formed from telic stems may have either conative or completive sense. Compare the following compound verbs with the same light verb *parṇā* ‘fall’ having inceptive meaning in (13), (14) and completive meaning in (15), (16):

- (13) *sītā hās paṛ-ī*
 Sita laugh fall-AOR.F.SG
 ‘Sita broke out laughing.’

- (14) *jo jahã th-ã vahĩ se daur paṛ-ã*
 who.REL where.REL be-PST.M.SG there.EMPH from run fall-AOR.M.SG
 ‘Everybody started running away from the place where they were.’
 (Liperovskij 1984: 178)
- (15) *maikrofon gir paṛ-ã*
 microphone fall fall-AOR.M.SG
 ‘The microphone fell [to the ground].’
- (16) *laṛk-ã ghoṛ-e se utar paṛ-ã*
 boy-M.SG horse-M.OBL.SG from get.off fall-AOR.M.SG
 ‘The boy got off the horse.’

The last example might contradict Butt’s statement that the light verb *paṛnā* may encode a lack of conscious choice (Butt 1995: 113).

Here we will concentrate on compounds formed only from telic verbs in order to analyze cases where these compounds may lose their meaning of ‘completeness’. There is a general belief that contrary to processes with simple verbs, the aorist forms featuring compound stems (that is, sequences of the main verb conveying the basic meaning and the so-called ‘vector’, ‘polar’ or ‘light’ verbs) signify that the event has culminated in its natural endpoint (Hook 1974; Liperovskij 1984; Nespital 1997; Pořízka 1967–1969 et al.):

- (17) *maĩ=ne rām ko kuch pais-e de di-ye*
 I=ERG Ram DAT some money-M.PL give give-AOR.M.PL
 ‘I gave some money to Ram (and he took the money).’ (Hook 1974: 165)
- (18) *maĩ=ne parāth-ã khā li-yā*
 I=ERG paratha-M.SG eat take-AOR.M.SG
 ‘I ate up the paratha.’ (Hook 1974: 167)

If a simple verb denotes a process, and in combination with a vector verb the result (e.g. *dhũṛhnā* ‘to look for’ and *dhũṛh lenā* ‘to find’), the aorist forms generated from simple stems imply the conative meaning, while those produced from compound stems express the completive meaning:

- (19) *ham=ne āp=kā patā dhũṛh-ã*
 we=ERG your address look.for-AOR.M.SG
 (*lekin nahĩ mil pā-yā*)
 (but NEG be.found be.able-AOR.M.PL)
 ‘We looked for your address (but could not find it).’

- (20) *ham=ne āp=kā patā dhū̃rh li-yā*
 we=ERG your address find take-AOR.M.SG
 ‘We have found your address.’ (Hook 1974: 164–165)

4 Conditions under which compound verbs do not denote the completion of the action

Kothari and Arunachalam (2009) studied simple (SV) and compound (CV) telic perfective predicates denoting partially completed and/or fully completed events. Each of 25 participants—adult Hindi speakers—viewed two different scenes for each of eight predicates (‘draw’, ‘eat’, ‘fill’, ‘extinguish’, ‘close’, ‘cover’, ‘pluck’ and ‘wake’). The first scene represented a full, and the second only a partial completion of the action. For example, in the first picture the door was tightly closed, while the second picture showed it still somewhat ajar. The participants were given two sentences—one with an SV, and the other with a CV predicate like (21) and (22) and their task was to choose the sentence that fits the scene:

- (21) *us=ne darvāz-e ko baṁd ki-yā*
 he=ERG door-OBL ACC closed do-AOR.M.SG
 ‘He closed the door.’
- (22) *us=ne darvāz-e ko baṁd kar li-yā*
 he=ERG door-OBL ACC closed do take-AOR.M.SG
 ‘He closed the door.’

The main results of the experiment were the following: 1) both SV and CV sentences were accepted for fully completed events; 2) SV sentences were accepted more often than CV sentences for partially completed events; 3) There was a lot of variation between predicates. The authors argue that “gradience between SV–CV and between predicates falls out from a pragmatic treatment: [...] the ending point for any event depends on what the individual believes should be the intended culmination” (Kothari & Arunachalam 2009: 26).

Below an attempt will be made to find other factors facilitating, impeding or blocking the usage of CV sentences for descriptions of incomplete action. I will concentrate here on the utterances featuring an incremental theme, as according to my observations the majority of CV sentences describing incomplete action have such a theme.⁶

6 The scope of this paper does not allow for detailed comments on Kothari and Arunachalam’s statement that there are no ‘systematic differences between incremental theme

The incremental theme may be defined as an argument whose properties determine the progress of the event it is involved in. In (5) and (7) above *the parāṭhā* is an incremental theme since every subpart of what is being eaten corresponds to a subpart of the event of eating that very *parāṭhā*.⁷ Three types of incremental theme verbs will be discussed below. When combined with the light verbs, they may denote an incomplete action.

1. The event's temporal terminus is 'achieved by progressing incrementally through the object' (Dowty 1991). The internal argument is created or consumed over time (Tenny 1994: 18): the verbs like *eat*, *build*. In many cases constructions with compound verbs signify that the speaker considers the goal achieved:

(23) *maĩ=ne yah seb khā li-yā bākī hiss-ā tumhār-ā*
 I=ERG this apple.M eat take-AOR.M.SG rest part-M.SG your-M.SG
h-ai
 be-PRS.3SG
 'I have partaken of the apple; the residue is yours.'

(24) *bāp ne imārat banā l-ī, ab is ko pūrā*
 father ERG building.F build take-AOR.F now he.OBL ACC full
karnā beṭ-e k-ī zimmedārī h-ai
 make son-M.OBL.SG GEN-F responsibility.F be-PRS.3SG
 'The father has built (his assigned part of) the house (to complete);
 his son's responsibility is to complete it (i. e. the house).'

In such cases the meaning of the compound verb seems to be intermediate between the conation and the completion: the action is not completed, but the speaker considers it as completed in the sense that the agent's duty has been fulfilled or part of the action has been completed. This is a typical case of 'successful result' described in Kothari and Arunachalam (2009). All the informants preferred CV construction in (23) and (24); SV was considered as possible, but worse.

Intransitive verbs with incremental theme show similar results:

and other predicate types' (2009). My results are different, and may be discussed in detail in another paper.

- 7 The analysis of this phenomenon provided in terms of 'Argument-to-Event Homomorphism' may be found in Dowty (1991); Krifka (1989, 1992); Ramchand (1997); Tenny (1992; 1994).

- (24a) *ghar ban ga-yā lekin pūrā nahī*
 house be.built go-AOR.M.SG but full NEG
 ‘The house is built, but not completely.’

One of the respondents said that both sentences have the same meaning, three preferred CV in the case where the result satisfies the speaker and there is very little to be done to complete the house, while one said that CV is marked for a good result and SV is unmarked: there may still be a lot to be done in order to complete building.

2. The event’s temporal terminus is achieved by progressing incrementally along the path object, but these objects ‘do not undergo change or motion during the event. Internal argument, although unchanged, provides a gradient along which the progress of the event may be measured’ (Tenny 1994: 17–18): the verb phrases like ‘climb the ladder’, ‘walk the trail’, ‘play the sonata’.

- (25) *gītā ne gān-ā gā li-yā par ākhirī lain bhūl*
 Gita ERG song-M.SG sing take-AOR.M.SG but last line.F.SG forget
ga-ī
 go-AOR.F.SG
 ‘Gita sang a song but she forgot the last line (of it).’

Three of the informants preferred CV for the situation when Gita left without singing the song to its end; SV was considered better in the case where she later remembered the last line and sang the song up to the end. For two informants CV and SV were equally possible in both situations.

- (26) *ve kitne kamzor haī ādhī sīrhī par caṛh ga-ye*
 they how.much weak are half staircase on climb go-AOR.M.PL
bāqī pāiri-yā caṛh-n-e k-ī tākat nahī h-ai
 left step-F.PL climb-INF-OBL GEN-F strength.F NEG be-PRS.3SG
 ‘How weak they are! They climbed half of the staircase (but) had no strength to climb the remaining steps.’

Three informants chose SV in the case where the agent is supposed to rest before going further and CV in the case where the future action is unknown or in the situation where the agent would never go further. One informant preferred CV in both cases, and the last one said that CV and SV may be equally used in both situations.

- (27) *laṛk-e ne kitab paṛh l-ī lekin do pann-e bāqī rah*
 boy-OBL ERG book.F read take-AOR.F but two page-M.PL left stay
ga-ye
 go-AOR.M.PL
 ‘The boy read the book, but two pages are left.’

Three informants preferred CV in the case where the boy would never read the remaining pages and SV if the situation is uncertain or the boy plans to read the book till the end later. One informant preferred CV in both situations and the other accepted both CV and SV irrespective of the continuation of the action.

3. The event’s temporal terminus is ‘achieved by progressing along measurable degrees of change in some property central to the verb’s meaning. Internal argument undergoes some change in a property over time’: change-of-state verbs like *ripen, redden* (Tenny 1994: 17–18).

- (28) *maĩ=ne ghar sāf kar di-yā par abhī kuch bākī*
 I=ERG house clean do give-AOR.M.SG but now something left
h-ai
 be-PRS.3SG
 ‘I cleaned the house but something is left (uncleaned).’

Three informants accepted CV for the situation when the agent does not plan to complete the cleaning and preferred SV for the case when he intends to do it, one preferred CV in both cases, and the last one did not express any preference.

In (28) the object undergoes only a partial but not complete change of state. In (29) and (30), cases where the object has not undergone any change, the construction with compound verb was not accepted by four Hindi speakers, and only one said that CV might be used in the situation when the agent is not supposed to do this work in future.

- (29) *maĩ=ne ghar sāf ki-yā (? kar di-yā) par vah*
 I=ERG house clean do-AOR.M.SG (? do give-AOR.M.SG) but it
sāf nahī hu-ā
 clean NEG be-AOR.M.SG
 ‘I cleaned the house but it did not become (sufficiently) clean.’

- (30) *maĩ=ne kapre dho-ye (? dho di-ye) lekin*
 I=ERG clothes.M.PL wash-AOR.M.PL (? wash give-AOR.M.PL) but
ve nahĩ dhul-e
 they NEG become.washed-AOR.M.PL
 ‘I washed the clothes but they were not (sufficiently) washed.’

The aorist forms of compound verbs cannot denote incomplete action in the following cases:

- a) The state of the object changes instantly:

- (31) *bacc-e ne akhroṭ tor-ā (*tor di-yā)*
 child-M.OBL.SG ERG walnut.SG crack-AOR.M.SG (*crack give-AOR.M.SG)
par vah nahĩ ṭūṭ-ā
 but it NEG be.cracked-AOR.M.SG
 ‘The child was cracking (= tried to crack) the walnut but it did not crack.’

- b) The verbal meaning contains a conative component:

- (32) *sohan ne bhāĩ ko bahut manā-yā*
 Sohan ERG brother ACC much persuade-AOR.M.SG
*(*man-ā di-yā) par vah nahĩ mān-ā*
 (*persuade give-AOR.M.SG) but he NEG be persuaded-AOR.M.SG
 ‘Sohan tried to persuade his brother many times, but he (= the latter) was not persuaded.’


- (33) *mohan ne apnĩ bĩvĩ ko bahut samjhā-yā*
 Mohan ERG self’s wife ACC much explain-AOR.M.SG
*(*samjhā di-yā) par vah nahĩ samjh-ĩ*
 explain give-AOR.M.SG but she NEG understand-AOR.F
 ‘Mohan many times tried to explain (something) to his wife, but she did not understand.’

5 Conclusion

It has been shown in this paper that compound verbs formed from non-telic stems usually convey inceptive meaning, while compound verbs formed from telic stems may have either a conative or completive sense. Compound verbs formed from telic stems do not necessarily denote the completion of the action when the theme of the utterance is incremental. Three types of incremental theme verbs denote

an incomplete action when combined with light verbs. In the case of an incremental theme, the event's temporal terminus may be achieved (1) by progressing incrementally through the object; (2) by progressing incrementally along the 'path object'; (3) by progressing along measurable degrees of change in some property central to the verb's meaning. In these respective cases, the internal argument: (1) is created or consumed over time; (2) remains unchanged; (3) undergoes some change in its properties over time. The most favourable contexts for using compound verbs are those when the action is not supposed to be completed at all or when the agent considers his goal achieved (in spite of the fact that the action has remained incomplete). The aorist forms from compound verbs cannot denote incomplete action if the state of the object changes instantly, if it does not change at all or there is a conative component in the verbal meaning.

ORCID®

Liudmila Khokhlova  <https://orcid.org/0000-0002-2152-6509>

Bibliography

- Butt, Miriam 1995. *The Structure of Complex Predicates in Urdu*. (Dissertations in Linguistics.) Stanford: Centre for the Study of Language and Information.
- Dowty, David 1991. "Thematic Proto-Roles and Argument Selection", *Language* 67: 547–619.
- Guru, Kāmtāprasād 2017 V. S. (1960 C. E. [1920]). *Hindī vyākaraṇ* (Hindi Grammar). Kāśī: Nāgarī Pracārinī Sabhā.
- Hook, Peter 1974. *The Compound Verb in Hindi*. (The Michigan Series in South and Southeast Asian Languages and Linguistics No.1.) Ann Arbor: The University of Michigan Center for South and Southeast Asian Studies.
- Kachru, Yamuna 1980. *Aspects of Hindi Grammar*. New Delhi: Manohar Publications.
- Kothari, Anubha & Sudha Arunachalam 2009. "Pragmatics and Gradience in Hindi Perfectives", in: *South Asian Languages Analysis Roundtable XXVIII: Denton, TX, October 9–11* (<https://blogs.bu.edu/sarunach/files/2011/05/KothariArunachalam2009_HindiPerfectives.pdf>, accessed: June 12, 2022).
- Krifka, Manfred 1989. "Nominal Reference, Temporal Constitution and Quantification in Event Semantics", in: Bartsch, Renate, J. F. A. K. van Benthem & P. van Emde Boas (eds): *Semantics and Contextual Expression*. Dordrecht: Foris Publications, 75–115.

- Krifka, Manfred 1992. “Thematic Relations as Links Between Nominal Reference and Temporal Constitution”, in: Sag, Ivan & Anna Szabolcsi (eds): *Lexical Matters*. (CSLI Publications.) Chicago: Chicago University Press, 29–53.
- Liperovskij, Vladimir 1984. *Glagol v jazyke xindi* (Verbs in the Hindi Language). Moscow: Nauka.
- Montaut, Annie 2004. *A Grammar of Hindi*. (LINCOS Studies in Indo-European Linguistics.) München: LINCOS.
- Nespital, Helmut 1997. *Dictionary of Hindi Verbs*. Allahabad: Lok Bharati Prakashan.
- Pãḍe, Hemacandra 2012. *Samasāmayik hindī vyākaraṇ* (Modern Hindi Grammar). Delhi: Granthlok.
- Pořízka, Vincent 1967–1969. “On the Perfective Verbal Aspect in Hindi. Some Features of Parallelism between New Indo-Aryan and Slavonic Languages”, *Archiv Orientální* 1967 (35): 64–68, 208–231; 1968 (36): 233–251; 1969 (37): 19–47, 345–364.
- Ramchand, Gillian 1997. *Aspect and Predication. The Semantics of Argument Structure*. Oxford: Clarendon Press.
- Singh, Mona 1991. “The Aspectual Content of Compound Verbs”, in: No, Yongkyoon & Mark Libucha (eds): *ESCOL’90: Proceedings of the Seventh Eastern States Conference on Linguistics. The Ohio State University, September 21–23*. Columbus, OH: Ohio State University Press, 260–271.
- Snell, Rupert & Simon Weightman 1989. *Teach Yourself Hindi*. London: Hodder & Stoughton.
- Tenny, Carol 1992. “The Aspectual Interface Hypothesis”, in: Sag, I.A. & A. Szabolcsi (eds): *Lexical Matters*. Stanford: CSLI, 1–27.
- Tenny, Carol 1994. *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

Saartje Verbeke 
and Aaricia Ponnet 

Animacy, Specificity, and Verb Semantics. What Drives Differential Object Marking in Hindi?

सार्त्य वरबेके और आरीसीया पोनेत्

सजीवता, विशिष्टता और क्रिया शब्दार्थ ।
डिफरेंशियल ऑब्जेक्ट मार्किंग “को” हिंदी में किन
कारकों से प्रभावित होता है?

Abstract This article aims to explore the dynamics of differential object marking (DOM) in Hindi. In Hindi, the *ko*-marking of the direct object (DO) is known to be influenced by two factors: animacy and definiteness or specificity. The marking of the inanimate object depends on the definiteness or specificity of the object. Unmarked inanimate objects can be interpreted as both definite and specific or indefinite and nonspecific, but when the inanimate object is marked with *ko*, it can only be interpreted as definite and/or specific. The definiteness/specificity opposition raises some questions: linguists do not seem to agree on which terminology to use. Some use definiteness and specificity as interchangeable concepts, others mention either definiteness or specificity, again others describe these factors as differentially influencing. This brings us to the following research question: do definiteness and specificity play different roles in the *ko*-marking of the DO? If so, which definitions of definiteness/specificity are pointedly applicable, and what are the different roles of these factors? Furthermore, how do they interact with animacy? Generally, animate arguments are marked with *ko*. However, Mohanan (1994) gives examples of animate nonspecific and indefinite DO's which are not *ko*-marked. Is this a rule or are there exceptions? Another influencing factor to be considered is the verb semantics. Several linguists describe that the *ko*-marking in Hindi also depends on the selectional properties of the verb. This asks for further investigation: does the DO of a certain verb always get *ko*-marked, and, vice versa, are there verbs of which the DO is never marked? To answer these questions, our

methodological approach will be twofold: firstly, we will perform a literature study of existing textbooks and grammars of Hindi to explore the different ways in which DOM in Hindi is being described and taught. Secondly, we will conduct a brief analysis of examples taken from the EMILLE corpus of Spoken Hindi¹ and the Corpus of Spoken Hindi² to find the frequencies of occurrence of *ko*-marked objects (as opposed to zero marking) and their correlation with the semantic parameters of animacy, definiteness, specificity and verb meaning.

Keywords Hindi, differential object marking, corpus linguistics.

सारांश इस लेख का उद्देश्य हिंदी में डिफरेंशियल ऑब्जेक्ट मार्किंग (डी ओ एम) अर्थात “प्रत्यक्ष कर्म को अलग-अलग ढंग से चिह्नित करना” की गतिशीलता का पता लगाना है। हिंदी में प्रत्यक्ष कर्म का “को” से चिह्नित होना दो कारणों से जाना जाता है: सजीवता और निश्चितता या विशिष्टता। निर्जीव वस्तु का अंकन वस्तु की निश्चितता या विशिष्टता पर निर्भर करता है। चिह्नित निर्जीव वस्तुओं की निश्चित और विशिष्ट या अनिश्चित और गैर-विशिष्ट दोनों के रूप में व्याख्या की जा सकती है, लेकिन जब निर्जीव वस्तु को “को” के साथ चिह्नित किया जाता है, तो इसे केवल निश्चित/विशिष्ट रूप में देखा जाता है। निश्चितता/विशिष्टता के अंतर से कुछ सवाल उठते हैं: भाषाविद इस बात से सहमत नहीं हैं कि किस शब्दावली का उपयोग करना है। जहाँ कुछ भाषाविद दोनों—निश्चितता और विशिष्टता—का उपयोग करते हैं, जैसे कि वे परस्पर परिवर्तनीय अवधारणाएँ हों, वहाँ अन्य या तो निश्चितता या फिर विशिष्टता का उल्लेख करते हैं और फिर कुछ इन कारकों के प्रभाव को अलग-अलग मानते हैं। यह बात हमें निम्नलिखित शोध प्रश्न पर ला खड़ा करती है क्या डी ओ की «को»-मार्किंग में निश्चितता और विशिष्टता अलग-अलग भूमिकाएँ निभाती हैं? यदि हाँ, तो निश्चितता/विशिष्टता की कौनसी परिभाषाएँ स्पष्ट रूप से लागू होती हैं, और इन कारकों की विभिन्न भूमिकाएँ क्या हैं? इसके अतिरिक्त उनका सजीवता से क्या लेना-देना है? आम तौर पर सजीव डी ओ “को” के साथ चिह्नित किया जाता है। हालाँकि एक भाषाविद (मोहनन 1994) सजीव गैर-विशिष्ट और अनिश्चित डी ओ के उदाहरण देती हैं जो “को” से चिह्नित नहीं हैं। क्या यह एक नियम है या क्या ये उदाहरण अपवाद हैं? एक और प्रभावशाली कारण क्रिया का शब्दार्थ है। कई भाषाविदों का मानना है कि हिंदी में “को”-चिह्नित क्रिया के चुनिंदा अर्थ संबंधी गुणों पर भी निर्भर करता है। प्रश्न यह उठता है कि क्या एक निश्चित क्रिया का डी ओ हमेशा “को” से चिह्नित किया जाता है, और, इसके विपरीत, क्या ऐसी क्रियाएँ हैं जिनके साथ डी ओ को कभी चिह्नित नहीं किया जाता है? इन सवालों का जवाब देने के लिए हमारी खोज दुहरी हो जाती है: सबसे पहले हम हिंदी की मौजूदा पाठ्यपुस्तकों और विभिन्न हिंदी व्याकरणों का अध्ययन करेंगे जिनमें हिंदी में डी ओ एम की मार्किंग का वर्णन है। दूसरा, ऐसे उदाहरण जो “को” से चिह्नित वस्तुओं और वे जो शून्य से अंकित हैं, अर्थात जहाँ “को” का लोप है, की तुलना आवृत्तियों के आंकड़ों के आधार पर करेंगे। इसके लिए “EMILLE कॉर्पस ऑफ़ स्पोकन हिंदी”¹ और “कॉर्पस ऑफ़ स्पोकन हिंदी”² से लिए गए उदाहरणों के सहारे हम इसका संक्षिप्त विश्लेषण करेंगे कि डी ओ की “को”-मार्किंग या “को” के लोप का सजीवता, निश्चितता, विशिष्टता एवं क्रियार्थ से क्या संबंध है।

मुख्य शब्द – विशिष्टता, डिफरेंशियल ऑब्जेक्ट मार्किंग, सजीवता, निश्चितता, क्रियार्थ।

1 <https://cqpweb.lancs.ac.uk/>, last accessed February 8, 2021.

2 <http://www.cosh.site/>, last accessed October 15, 2019.

1 Introduction

This article aims to explore the dynamics of differential object marking (DOM) in Hindi. In Hindi, the *ko*-marking of the direct object (DO) is known to be influenced by two factors: animacy and definiteness and/or specificity (Kachru 2006: 175). However, few studies (Bhatt & Anagnostopoulou 1996; Butt 1993; Self 2012) have exclusively focused on the interplay of these two factors under different circumstances and their impact on the formal marking of the patient.³ In this article, we intend to present the linguistic problems associated with Hindi objective marking and to compare the earlier linguistic analyses with the results of the present analysis of actual language usage. We will shed light on three issues pertaining to Hindi DOM marking: (a) the role of animacy, (b) the role of definiteness/specificity, and (c) the role of verb semantics.

First, it is traditionally assumed that an animate DO is marked with the objective marker *ko*.⁴ However, there are attestations of animate DOs which are not *ko*-marked. Is there a rule behind these unmarked DOs or are they simply exceptions? The unmarked animate DOs, as suggested by Mohanan (1994), are nonspecific. We discuss two possible hypotheses: either the lexical feature of animacy is overruled by the feature definiteness/specificity, or animacy does remain the primary factor determining the marking and we need to look at other syntactic properties to explain the unmarked forms.

Second, the specificity/definiteness opposition raises some questions: linguists do not seem to agree on which terminology to use. Some use ‘specificity’ and ‘definiteness’ as interchangeable terms; others mention either the role of definiteness or the role of specificity in DOM; again others describe these factors as each having a different impact on the marking. We will give an overview of the different approaches in the literature and focus on the following questions: Do specificity and definiteness play a different role in the *ko*-marking of the DO? If so, which definitions of definiteness/specificity should be used, and what are the different roles of these factors?

Third, another factor influencing DOM to be considered is the verb semantics. Several linguists (Mohanan 1994; de Hoop & Narasimhan 2005; Klein 2007; Geist et al. 2007) describe that the *ko*-marking in Hindi depends on the selectional semantic properties of the verb. We will look at several transitive verbs and investigate what kind of alignment structure they take.

3 The term is used in its broader meaning which includes also ‘theme’.

4 We prefer to use the term “objective” instead of “accusative” or “dative”. Hindi *ko* is used to mark the DO in a differential pattern, but it is also used to mark the IO. Neither the term “accusative” nor “dative” covers both uses. “Objective” refers to a case used to mark an object, and is as such more neutral than the two traditional terms.

To find an answer to these issues, our methodological approach will be two-fold: firstly, we will perform a literature study of existing textbooks and grammars of Hindi to explore the different ways in which DOM in Hindi is being described and taught. Secondly, we will conduct a brief analysis of 450 sentences extracted from the EMILLE Spoken Hindi corpus (Lancaster University, available on <<https://cqpweb.lancs.ac.uk/>>, created by Hardie 2012). We selected these sentences based on a search for constructions including a particular verb to find the frequencies of occurrence of *ko*-marked objects and their correlation with the semantic parameters of animacy, definiteness/specificity, and verb meaning. As such, this article aims, first, to offer an overview of the various linguistic analyses of differential object marking in Hindi, and second, to supply examples underpinning or negating these analyses using spoken language sentences documented from hearing.

Thus, this article uses two separate kinds of data, and this will, of course, influence the conclusions: the data taken from the textbooks and grammars is Hindi by prescription and is the closest to Standard Hindi one can find, whereas the sentences found in the corpora are renditions of spoken Hindi, and hence there may be dialectal variation or substandard grammar, thus breaking the grammatical rules of Standard Hindi, that will not be acceptable to every speaker of the language. However, compared to textbook material, corpus samples of Spoken Hindi are much closer to the actual language use, and a corpus is a representation of the changing and vibrant nature of language.

The structure of the article is as follows: section 2 briefly presents the basics of DOM in Hindi, based on traditional accounts found in grammar books. Section 3 focuses on the interaction between animacy, definiteness, and specificity. Section 4 gives an overview of the different accounts of definiteness and specificity determining the objective marking, and section 5 is a small-scale corpus study of the occurrences of marked and unmarked DOs with particular verbs.

2 Differential object marking

Traditionally, differential object marking in Hindi is said to be determined by the factors animacy and definiteness/specificity (Malchukov 2008; Klein & de Swart 2011; Aissen 2003; de Hoop & Narasimhan 2005; de Hoop & Malchukov 2008). The following examples from Mohanan (1994: 80) illustrate this pattern:

- (1) *ilā ne bacce ko uṭhāyā*
 Ila.F ERG child.M OBJ lift.AOR.M.SG
 ‘Ila lifted a/the child.’

- (2) *Ilā ne hār ko uṭhāyā*
 Ila.F ERG necklace.M OBJ lift.AOR.M.SG
 ‘Ila lifted the/*a necklace.’

- (3) *Ilā ne hār uṭhāyā*
 Ila.F ERG necklace.M lift.AOR.M.SG
 ‘Ila lifted a/the necklace.’

The DO *bacce ko* (DIR *baccā*) ‘child’ is animate and usually marked, whereas *hār* ‘necklace’ is inanimate and only marked when the interpretation is definite ‘the necklace’, marked by *ko*, as in (2).

The unmarked DO *hār* in (3) can be interpreted in three possible ways: 1) as a definite object ‘the necklace’, 2) as an indefinite specific object, which renders a translation with the indefinite determiner in English referring to a necklace known to the speaker (e.g. ‘Ila lifted a necklace [of those that were presented there]’), 3) as an indefinite nonspecific object ‘a necklace’.

Though this pattern of DOM might seem relatively straightforward, several complications in actual language usage occur. In the next section 3, we focus on the factor of animacy, and on the fact that not all animate DOs are always marked.

3 The role of animacy

Generally, animate arguments are marked with *ko*, as illustrated in (1) above. This is especially the case with animate arguments that are human (Kachru 2006: 175). For animate non-human arguments—one thinks in particular of animals—the rule is less strict. For instance, Montaut (2004: 170) gives the following example (4), in which *cuḥī* ‘mouse’ is not marked with *ko*.

- (4) *billī cuḥī khāegī*
 cat.F.SG mouse.F.SG eat.FUT.F.SG
 ‘The cat will eat the mouse.’

In general, though, animate arguments, even non-human, are much more likely to be marked with *ko* than inanimate arguments. However, some authors mention animate indefinite/nonspecific DOs that are not *ko*-marked. Montaut (2004: 171) gives the examples of *naukar rakhnā* ‘to keep a servant’ and *laṛkā dekhnā* ‘to look for a suitable boy’. In these instances, the animate argument is indefinite and nonspecific, features that seem to determine the absence of a case marking on the DO. Montaut (2004: 171) calls these DOs “no longer ... a human entity, but rather ... a general function”. In the same vein, Butt (1993) argues that *ko* can

only be omitted if the animate object is unambiguously indefinite and nonspecific. In a typological article on DOM, Klein and de Swart (2011) argue that animacy in Hindi is the trigger for the objective marking in the sense that animacy, as a lexical feature, immediately “causes” the objective marking of the DO. The instances of unmarked nonspecific animate DOs seem to be counterevidence to this hypothesis, as animacy as the trigger factor is here clearly overruled by the feature of non-specificity, which seems to determine the (absence of) marking.

Mohanani (1994: 109) offers another interpretation of these nonspecific animate arguments. Instead of taking the objects in their generic meaning, she interprets the whole construction as construction of incorporation. Her own example is the following (Mohanani 1994: 108):

- (5) *ilā bacce khojtī rahtī hai*
 Ila.F child.M.PL search.IPFV.F.SG DUR.PRS.F.SG AUX.PRS.3SG

Mohanani’s translation of this example is in accordance with a typical construction of incorporation: ‘Ila keeps performing the act of searching for children.’ In her opinion, the common translation that is given for such constructions, i. e. ‘Ila keeps searching for children’, is ungrammatical. She gives a number of syntactic properties of this construction, which seem to argue for an interpretation as construction of incorporation: *bacce* cannot take case marking, modification, conjoining, or be split from the verb by an adverb.⁵ The reasoning that the first property is not being allowed to take case marking is, of course, circular: the construction is considered as an instance of incorporation because there is no case marking – if there were case marking, it would not be considered as incorporation. With regards to the other properties, from the moment the DO would be determined by a modifier or conjoined with another NP, it loses its non-specificity, which is an argument for the hypothesis that the feature of specificity does indeed sometimes overrule animacy. The idea that an argument of a predicate can be incorporated into the predicate works well with inanimate nonspecific arguments – and to a certain extent with animate non-human arguments (Mohanani 1994: 106–108), but for animate nonspecific arguments, it leads to far-fetched interpretations of constructions.

From this discussion, it appears that animacy is the primary factor that determines the *ko*-marking; however, when a nonspecific meaning is intended, the marking can be absent. Specificity plays a different role than definiteness, which brings us to our next section. In this section, we discuss the issue of how definiteness and specificity relate to each other with respect to Hindi DOM.

5 A colleague pointed out that her respondents do accept the sentence with the DO split from the verb by an adverb of place or manner.

4 Definiteness and specificity: different roles?

In linguistics, two concepts have been deemed important with regard to ‘definiteness’, i. e. ‘familiarity’ or ‘identifiability’ and ‘uniqueness’ (Lyons 1999). The former refers to the role of definiteness as a way to mark the referent as already known to the hearer or not, the latter as a way to identify the referent as a unique entity in the discourse. The way definiteness is marked in languages cross-linguistically is highly diverse and has been extensively discussed and researched (Aguilar-Guevara, Pozas Loyo & Vázquez-Rojas Maldonado 2019). In many languages, there is a division in the way indefinite referents are marked. In this sense, indefinite arguments can be further subdivided into specific/nonspecific arguments (Fodor & Sag 1982). According to von Heusinger (2002), the function of definiteness in discourse is to mark familiarity, whereas specificity refers to the “certainty of the speaker about the identity of the referent”, i. e. uniqueness, which is the very basic definition of these concepts that will be followed here when we speak of definiteness vs. specificity.⁶ For this article, we follow the more mainstream assumption that specificity is relevant to determine indefinite arguments. As such, it is generally assumed that specificity is a subcategory of definiteness, in the sense that all definite arguments are by definition specific, and indefinite arguments are divided into specific and nonspecific arguments (Lyons 1999; von Heusinger 2002; Aissen 2003; Abbot 2004). Applied to DOM in Hindi, using the terminology of specificity is more precise than the broader categorization of definiteness. The following three subsections discuss the various accounts and analyses of definiteness and specificity and their relation with DOM in Hindi.

5 Definiteness

In languages with determiners, in particular articles, the definiteness of the argument is mostly taken to be indicated through these determiners. The accounts on Hindi which explain the use of *ko* with the concept of definiteness refer to the indefinite determiner *ek*, originally the numeral ‘one’, and to the indefinite demonstrative *koī* ‘some’ as indicators of indefiniteness (e. g. Montaut 2004: 54; Singh 1994: 219; Kachru 2006: 54–55).

According to Singh (1994: 227, followed by Aissen 2003), the indefinite marker *ek* cannot appear with a *ko*-marked inanimate object (hence the ungrammaticality of (6)). Singh and Aissen take this as evidence to state that *ko* marks definiteness with inanimate objects, in contrast to marking all animate objects, irrespective of

6 Note that von Heusinger (2002) himself argues to consider specificity not as a subcategory, but as a category that exists besides definiteness.

their definiteness (but see section 3 above). Singh (1994: 228) claims that the following example is ungrammatical:

- (6) *larke ne āj ek phūl *ko dekhā*
 boy.M.OBL ERG today one flower.M OBJ see.AOR.M.SG
 ‘The boy saw a flower today.’

Mohanani identifies definiteness as the feature which determines the marking of inanimate arguments. However, she mentions the complex interaction of definiteness with specificity in a footnote (Mohanani 1994: 80). Revealingly, one of her examples to illustrate this complexity is a similar construction to (6), argued to be ungrammatical by Singh (1994) and Aissen (2003). Mohanani (1994: 81) argues about this example (7): “[the object] may be either specific or nonspecific [...], but is indefinite because of the indefinite determiner.”

- (7) *ravī ek gāy ko kharīdnā cāhtā hai*
 Ravi.M one cow.F OBJ buy.INF wish.IPFV.M.SG AUX.PRS.3SG
 ‘Ravi wishes to buy a (particular) cow.’

Definiteness is the factor determining *ko*-marking for Mohanani, and indefiniteness is related to the determiner *ek*. She considers the DO of (7) as “ambiguous between specific and nonspecific”. In our opinion, however, *ek gāy* in (7) is indefinite but specific, and it is the specificity of the NP which causes the marking with *ko*.⁷

Furthermore, a perusal of the COSH (Nishioka 2016–2017) renders the following example (8), in which the same NP as in (6) *ek phūl ko* occurs.⁸ Though the combination with the genitive determiners hints at a numeral interpretation of *ek*, the indefiniteness of *phūl* remains. In other words, the use of the indefinite determiner does not exclude the objective marking of the DO – though such occurrences are rare.

7 Note that *gāy* ‘cow’, the DO of this construction is animate; however, as we discussed in section 3, for animate non-human DOs the general rule of *ko*-marking is less strict, and Mohanani finds the example without *ko*-marking also perfectly acceptable and common, and a true example of a nonspecific reading.

8 Another example is mentioned in Klein (2007: 5):
Adnan ne ek phūl kudi ko toḍā
 Adnan ERG one flower vase OBJ break.AOR.M.SG
 ‘Adnan broke a flower vase.’

A colleague pointed out that the word *kudi* is inscrutable. We agree this might be the case. Klein unfortunately does not give more details as to the origins of this example. Klein attributes the object marking in this example, which, unlike (6), is seen as grammatical, to the difference between the verbs *toḍnā* here and *dekhnā* in (6), but does not elaborate any further on what semantic aspects of the verbs in particular are relevant.

- (8) *is=i⁹kā¹⁰ ek phūl ko gur mē lapetkar [...]*
 this.OBL=GEN.M.SG one flower.M OBJ molasses in wrap.CVB [...]
 (COSH 33)
 ‘After dipping one of its flowers in molasses [...]’¹¹

In sum, at first sight, definiteness is an appropriate criterion to come to terms with the pattern of DOM in Hindi. However, definiteness alone is not sufficient to explain the markedness of inanimate indefinite DOs like example (8). Examples such as (8) bring us to accounts that rely on specificity instead.

6 Specificity

In another work, Singh (1994b) does consider specificity as playing a role in DOM, defining specificity as the knowledge that the object previously existed. She considers the following example as ungrammatical with the objective marker *ko*, because of the non-specificity of the DO cake (Singh 1994b: 80, also Bhatt & Anagnostopoulou 1996).

- (9) *rām ne cake¹²(*ko) banāyā*
 Ram.M ERG cake(*OBJ) make.AOR.M.SG
 ‘Ram made a cake.’

9 Ungrammatical forms and typos, as they appear in the sentences from the corpora, are marked with the low inverted exclamation mark.

10 The example phrase concerns a medical prescription. The genitive pronoun refers to the flower described in the preceding sentence:

वायगोला का दर्द:- सफेद अकौवा (मदार), इसे स्वेतार्क भी कहते हैं। इसका एक फूल को गुड़ में लपेटकर रोगी को खिलाकर पानी पिला दें, आधा घंटे में ही रोगी का दर्द सही हो जायेगा।

vāygolā kā dard: saphed akauvā (madār), ise svetārka bhī kahte hai. Iskā ek phūl ko guḍ mē lapetkar rogī ko khilākar pānī pilā dē, ādhā ghaṇṭe mē hī rogī kā dard sahī ho jāyegā.

‘Pain due to abdominal colic: white akauvā (swallowwort), also called *svetārka*. Wrap one flower of it in molasses and feed it to the patient, and give him/her water to drink, within half an hour the patient’s pain will be cured.’ A colleague points out that “iskā” in this example is ungrammatical, the genitive postposition being realised in the direct case instead of the oblique case. Since this is an example taken from a spoken web corpus, it might very well be that this reveals the process of building the sentence, in which the speaker is assumed to have decided to use a different way to express, in this particular case by means of using *ko*.

11 All the translations of the examples given from the corpora are own translations.

12 The transliteration was taken from the original source (in Roman script).

According to Butt (1993, also Montaut 2004: 170–171), who from the onset makes a clear distinction between definiteness and specificity, *ko* marks specificity with inanimates. The marked inanimate object can be both specific indefinite or definite. For Butt, an object being definite implies that the object is specific, but not the other way around. In the framework of Lexical-Functional Grammar, Butt & King (2004) consider only “specific” as a property, there is no “definite” property. Although de Hoop & Narasimhan (2005) and de Hoop & Malčukov (2008) base their account on Aissen (2003), who considers definiteness, they do not mention definiteness; for them, *ko* marks the specific object. Apart from these authors, others generally refer to both definiteness and specificity as determining the *ko*-marking, without making an explicit distinction or giving a clear definition of the two factors. Kachru (2006) mentions in her grammar that inanimate arguments are marked with *ko* if the NP is definite and specific (Kachru 2006: 173). Vasisht & Joseph (2008: 7) mention specificity/definiteness marking on the direct object without differentiating between the two.

From the above discussion, it is clear that relying on definiteness leads to some problems with interpreting certain occurrences of objective-marked inanimate indefinite arguments. Specificity brings the solution, as observed by the majority of the authors. As argued in section 3 above, specificity is also the criterion that might explain the unmarkedness of certain nonspecific animate arguments (see also Self 2012). We follow this view, and therefore we will use the term specificity in the remaining sections of this paper.

7 The role of verb semantics

Thus far, our review of the literature on the subject of the present article has suggested that animacy primarily determines the marking of the DO. In the case of nonspecific DOs, animacy might be overruled, and the DO might not get an objective marking because of its non-specificity. Inanimate DOs get no marking unless they are specific. However, some authors observe that there are deviations to these rules which seem to be related to the semantics of the verbs. For instance, de Hoop and Narasimhan (2005: 328, n. 3) observe that an animate noun like *bakrā* ‘goat’ is much more likely to receive *ko* object-marking with the verb ‘to kill’ than with the verb ‘to sell’. The marking of non-human animate nouns is somewhat a grey zone, and the hypothesis that the verb semantics determine in part the specificity and/or animacy of the non-human animate noun seems acceptable (compare section 3.). This observation of de Hoop and Narasimhan (2005) seems to refer primarily to a kind of semantic interaction between the verb and the DO, which ultimately has consequences for the interpretation and formal

marking of the DO.¹³ However, most accounts that argue for a clear relationship between the presence of *ko* and verb semantics go back to Mohanan (1994), who gives several specific verbs of which she argues they occur only with one type of DO – marked or unmarked.

Mohanan (1994: 81 and n. 34) champions the notion that the selectional restrictions of the verb on animacy restrict the use of *ko*. She argues that the choice of whether or not the DO receives *ko* “is available only to the objects of those verbs that are neutral to the animacy of their objects” (Mohanan 1994: 81). In the neutral class of verbs that permit variable use of *ko*, that is, those verbs the marking of whose DO is completely dependent on its properties of animacy/specificity, Mohanan includes for instance *khojnā* ‘search for’, *toṛnā* ‘break’, *kāṭnā* ‘cut’, and *lānā* ‘bring’. In her opinion, *likhnā* ‘write’, *parhnā* ‘read’, *pīnā* ‘drink’, *gānā* ‘sing’ and *banānā* ‘make’ belong to the class of verbs that require inanimate objects and do not allow *ko*. The verbs *bulānā* ‘call’ and *mārnā* ‘beat somebody’, ‘kill’, on the other hand, are among the verbs that require of their objects animacy and, accordingly, marking with *ko*. Mohanan’s hypothesis seems to be based on intuition and free language observation. In order to test for its validity, we have conducted a small-scale corpus investigation. As mentioned above, we used the EMILLE corpus of spoken Hindi (the University of Lancaster, available on <<https://cqpweb.lancs.ac.uk/>>, created by Hardie 2012), which means that the examples that we discuss in this article reflect free and natural language speech, and will thus not always reflect the grammatical rules of Standard Hindi. The data have an oral production source and have then been transcribed into Devanagari¹⁴ for the corpus¹⁵. No information on the sociocultural status of the speakers is available, although the deviations from Standard Hindi in grammar and register might imply that the data also include sentences by speakers from lower SES backgrounds. We selected 450 sentences based on the attestation of seven verbs, i. e. *banānā* ‘to make’; *parhnā* ‘to read’, ‘to study’; *gānā* ‘to sing’; *pīnā* ‘to drink’; *likhnā* ‘to write’; *mārnā* ‘to beat’, ‘to hit’, ‘to kill’; *bulānā* ‘to call’, ‘to invite’. All of these verbs have been argued to occur only with one type of DO, either marked or unmarked (Mohanan 1994: 81–82). We looked for all possible finite, active uses of these verbs, which

13 A colleague points out that this perspective is very different from Mohanan’s idea that an unmarked animate DO is incorporated. Indeed, De Hoop and Narasimhan consider the unmarked DO as a full-fledged argument in its own right, which determines the verb as much as it is determined by the verb.

14 All transliterations of the corpus examples are own transliterations. Wherever we give the context of the quoted examples, we have provided the original Devanagari text in footnote.

15 The corpus doesn’t provide further details on how and by whom the data were transcribed.

means that we did not take into account any passivized use.¹⁶ We also did not take into account constructions in which the DO is not overtly mentioned. Our investigation aimed to find out if the claim that certain verbs are exclusively associated with a DO marked with *ko* or with an unmarked DO is confirmed in actual language usage.¹⁷

8 *banānā* ‘to make’

The great majority of the patients of the verb *banānā* are in the unmarked case (121/125). These very frequent constructions are for instance of the type of (10), with a generic, nonspecific inanimate DO.

- (10) *bahut barhiyā khānā bhī banāte hai*
 very good food.M too make.IPFV.M.PL AUX.PRS.3PL
 (1ehinsp00a)
 ‘They make very good food too.’

Four constructions are with a marked objective patient. In all of these constructions, the patient is pronominal. In two of these four constructions, the DO is inanimate but specific, in the remaining two, the DO is animate and specific. In this discussion, we take both the categories animacy and specificity into account, although it must be clear from the discussion in section 3 above that animacy is the factor that primarily determines the *ko*-marking.

Inanimate specific

- (11) [...] **un=ko** *unhō=ne banāyā hai*
 they.OBL=OBJ they.OBL=ERG make.PFV.M.SG AUX.PRS.3SG
 (5ehinsp011)
 ‘They constructed them.’

The pronoun in this example is specific as it refers to the inanimate NP *amrīkī thikāne* ‘American places’ in the preceding sentence: *amrīkī kampaniyā̃ ihai islīe*

16 We realize that *ko*-marking also occurs with the patient in a passive construction, however, a discussion of when and where *ko* occurs in these passive constructions would transgress the scope of this article.

17 As Mohanan herself already pointed out: “when speaking of individual verbs and verb classes, we must expect there to be considerable speaker variation and grey areas in speaker judgments” (1994: 81).

*aur bhī jo amrīkī thīkāne hai unko unhōne banāyā hai*¹⁸ “[These] are American companies, this is why the other American residencies that are [there], they constructed them.”

- (12) *zindagī ikā jis=ko cāhā us=ko apnā*
 life.F GEN which.OBL=OBJ want.AOR 3SG.OBL=OBJ [his/her]own
banā na sake
 make not can.AOR.PL.HON/SBJV.3SG
 (63ehinsp00d)¹⁹
 ‘The one whom [s/he] loved in [his/her] life her/him [s/he] could not/probably will not get.’

Animate specific

- (13) *is=ko bartāniyā ke avām ne banāyā hai*
 he.OBL=OBJ Britain gen people ERG make.PFV.M.SG AUX.PRS.3SG
 (12ehinsp039)
 ‘The people of Britain have made him [the Prime Minister of his country].’

The pronoun in this example refers to the animate noun *prāīm ministar* ‘prime minister’ in the preceding sentence: *Prāīm ministar jo hai is mulk kā is ko bartāniyā ke avām ne banāyā hai*²⁰ ‘The people of Britain have made him the prime minister of this country.’

- (14) *bilkul us=ko chief minister ke jaisā banā rahe hāī*
 totally he.OBL=OBJ chief minister GEN such make PROG AUX.PRS.3PL
 (123ehinsp047)
 ‘They are completely making him [turning him into] the chief minister.’

The pronoun in this example refers to the noun *ādmī* ‘man’ in the preceding sentence: *matlab ek ādmī thā unkā jiskī śakal chief minister se kāphī miltī thī, to uskā iphūrā, matlab uske pūre mūh kā nakśā badal ke bilkul usko chief minister ke jaisā banā rahe hāī tākī vah usko chief minister kī jagah biṭhā sakē [...]*²¹ ‘I mean, among them was a man whose face approximated the chief minister, so his full,

18 अमरीकी कम्पनीयां है इसलीए और भी जो अमरीकी ठिकाने है उनको उन्होने बनाया है।

19 This sentence is not unconditionally acceptable to, at least some, native speakers.

20 प्राईम मिनिस्टर जो है इस मुल्क का इस को बर्तानिया के अवाम ने बनाया है।

21 मतलब एक आदमी था उनका जिसकी शकल chief minister से काफी मिलती थी, तो उसका फूरा, मतलब उसके पूरे मुँह का नक्शा बदल के बिलकुल उसको chief minister के जैसा बना रहे हैं ताकी वह उसको chief minister की जगह बिठा सकें [...]

I mean after changing the appearance of his entire face, [they] are making him really [look] like the chief minister [...].’

In general, *banānā* is constructed with an inanimate object. Animate DOs are rare: in our corpus, we found only eight instances. They were related to a pattern much observed in combination with *banānā* with the meaning ‘to make something/someone into something’. The following is an example.

Animate specific

- (15) *vahā pe us=ko kuttā banātā hai*
 there on he.OBL=OBJ dog.M make.IPFV.M.SG AUX.PRS.3SG
 (27ehinsp047)
 ‘There he makes him [his] dog.’

The pronoun is specific as it refers to the noun *pitājī* ‘father’ in the preceding sentence. The entire sentence refers to a movie scene where the villain makes the father of the hero act like a dog: *jahā pe uske pitājī ko mārā hotā hai usne. vahā pe usko kuttā banātā hai pahle [...]*²² ‘Where he constantly beat his father. There he first makes him [his] dog [...].’

- (16) *phir āp is=ko banāras kā pān vālā kyū banā rahe*
 then you he.OBL=OBJ Benares GEN betel-seller.M why make PROG.PL
haī?
 AUX.PRS.3PL
 (65ehinsp00f)
 ‘Then why are you making him a betel-seller of Benares?’

In this sentence, the speaker refers to a film character that he had played. The pronoun *usko* ‘him’ is hence specific as it refers to this film character: *To unhone saī parānjape ne mujhe ijāzat de dī ki ki lallan miyā pān vāle jo haī vo nām maine saješ kiyā thā unko maine kahā ki kahā ye śūtīng kar rahē haī kahne lage ki nizam ūddīn me to maine kahā kī phir āp isko banāras kā pān vālā kyū banā rahē haī*²³. ‘So she, Sai Paranjape, gave me permission to suggest the name of that Lallan Miyan pan seller, I asked her, where are you doing this shooting, she started to say in Nizamuddin, so I said “Then why are you making him a Benares pan seller?.”’

22 जहाँ पे उसके पिताजी को मारा होता है उसने। वहाँ पे उसको कुत्ता बनाता है पहले [...]

23 तो उन्होने सई परान्जपे ने मुझे इजाज़त दे दी कि कि लल्लन मियां पान वाले जो हैं वो नाम भी मैंने सजेस्ट किया था उनको मैंने कहा कि कहा ये शूटींग कर रहे हैं कहने लगे कि निज़ाम ऊद्दीन मे तो मैंने कहा की फिर आप इसको बनारस का पान वाला क्यु बना रहे हैं।

- (17) [...] *ivah*²⁴ *irākī śahriyō ko bambārī mē niśānā*
 [...] he Iraqi citizen.M.OBL.PL OBJ bombing.F in target.M
banā rahe hai
 make PROG.M.PL AUX.PRS.3SG
 (22ehinsp011)
 ‘[...] they are making the Iraqi citizens the target in the bombings.’

In other words, for *banānā*, it is certainly true that there is a tendency towards unmarked DOs, yet objective marked DOs are not strictly excluded.

9 *paṛhnā* ‘to read’

For *paṛhnā*, we needed to take into account that it is a polysemous verb, which means ‘to read’, as well as ‘to study’. In the latter meaning, it often occurs as an intransitive, without an explicit DO. As a transitive verb, surprisingly, the number of occurrences is rather limited. We only found 23 occurrences of *paṛhnā* in the EMILLE corpus. Of these, only one has an objective marked DO. In this example (18), the pronoun is specific as it refers to a sentence that is being read out by the speaker.

- (18) *ise*²⁵ *dobārā paṛhtā hū*
 this.OBJ twice read.IPFV.M.SG AUX.PRS.1SG
 (3ehinsp015)
 ‘I am reading this again.’

paṛhnā is similar to *banānā* in that it semantically takes a clear preference for an inanimate object. Again, however, if that inanimate object is definite, then it is perfectly possible that the object takes the objective case.

24 The pronoun *vah* ‘he, that’ is in the original sentence in the singular form, which occasionally is also used as plural. The agreement on the verb is ambiguous:
 उन्होने अमरीका और बर्तानिया पर इल्ज़ाम लगाया है ‘की वह इराकी शहरियों को बमबारी मे निशाना बना रहे है।
unhone amrīkā aur bartāniyā par ilzām lagāyā hai ki vah irākī śahariyō ko bambārī me niśānā banā rahe hai
 ‘He accused America and Britain of making the Iraqi citizens the target in the bombings.’

25 The analytic objective case form of the personal pronouns has a synthetic equivalent ending in *-e/ē*, e.g. in this example *ise* corresponds to *is=ko* (Kachru 2006: 64).

10 *gānā* ‘to sing’

We found 28 attestations of a finite form of the verb *gānā*. All of them were constructed with an unmarked case. In many cases, the patient was a form of the word ‘song’, either *gānā* (m.) or *gīt*. Remarkably, there were also a few specific inanimate objects which did occur unmarked, in contrast to the constructions of *banānā* and *paṛhnā*:

- (19) *yā ye gāyā thā* [...]

or this SING.PFV.M.SG AUX.PST.M.SG

(1ehinsp00d)

‘Or you sang this [...]’

The phrase comes from an interview with an artist about a certain song, the pronoun is hence specific as it refers to the song: *āp vo māyne rakh kar gāte hāi kyā? Yā ye gāyā thā jab māyne rakh kar vo log song karēge jo matlab lavī ḍavī hote hāi*²⁶ ‘Are you keeping that sense when you are singing? Or did you sing this when you were keeping the sense that people will sing a song that is lovey-dovey?’

In our corpus, we did not find any *ko*-marked DO. Note however that a corpus study is limited by the size of the corpus. Klein (2007: 5) found another example with *gānā*, in which the DO is marked:²⁷

- (20) *merā dost is gānā ko gātā hai*

my friend.M this.OBL song.M OBJ sing.IPFV.M.SG AUX.PRS.3SG

‘My friend sings this song.’

In this example, the DO has a modifier, which renders it definite/specific.

11 *pīnā* ‘to drink’

pīnā as a finite verb occurred 18 times in the corpus. It never occurred with a marked DO. However, most DOs were generic mass nouns, such as *dūdh* ‘milk’, *sigareṭ* ‘cigarette’²⁸, *cāy* ‘tea’, *śarāb* ‘alcohol’. All DOs were thus nonspecific and inanimate, which would make the objective marking quite unlikely.

26 आप वो मायने रख कर गाते हैं क्या? या ये गाया था जब मायने रख कर वो लोग संग करेगे जो मतलब लवी डवी होते हैं।

27 Klein does not mention his source of the example (note the ungrammatical *-ā* ending of *gānā*). A quick perusal on Google of *is gāne ko gānā* shows that this form does occur and is not unusual.

28 The word cigarette occurs in all cases we came upon in our study as a mass noun, meaning in connection with the verb—*sigareṭ pīnā*—‘to smoke’.

12 *likhnā* ‘to write’

In our corpus, we found *likhnā* with a DO 91 times. The majority of the patients are unmarked (89/91), none of them animate. 40/91 are specific objects. The majority of these are unmarked (38/40), even when they are pronominal:

Inanimate Specific

- (21) *'me=ne ye²⁹ nae dosto par likhā hai*
 I=ERG this new.OBL.PL friend.M.OBL.PL on write.PFV.M.SG AUX.PRS.3SG
 (78ehinsp0044)
 ‘I wrote this about new friends.’

Two of the attested DOs are in the objective case, both pronominal. The following example contains one such a marked pronominal DO.

- (22) *maī=ne ise 'likhī hai*
 I=ERG this.OBJ write.PFV.F.SG AUX.PRS.3SG
 (2ehinsp002)
 ‘I have written it.’

The pronoun is specific as it refers to the feminine noun *kitāb* ‘book’. This might also explain the verb form in the example sentence, even though the agreement with the anaphoric referent is ungrammatical (it should agree with *ise*, in which case the verb would have been in the default masculine, i. e. no agreement because of the objective case). This again reflects the free nature of the data, and stresses the value of researching variation in spoken Hindi: *uskā nām hai raph gāīḍ unhōne naī sirīz lagāī hai aur itihās ke lie aur unkī sirīz mē ye jo bhārat ke itihās kī kitāb jo maīne likhī hai. [...] uskā bhī phūl ṛāīṭal hai da raph gāīḍ kronikal iṃḍiyā eṃḍ ā maīne ise likhī hai* ‘Its title is *Rough Guide*, they launched a new series, on history, and in their series I wrote the book on the history of India [...] its full title is also the *Rough Guide Chronical India* and ah I wrote it.’

29 The pronoun *ye* refers to something that had been written before, the original Hindi version reads: आज से तीन साल पहले ‘मेने ये नए दोस्तो पर लिखा है ये आप को ही सुनाया था। *āj se tīn sāl pahle 'mene ye nae dosto par likhā hai ye āp ko hī sunāyā thā* ‘Three years ago from today I wrote this about new friends, this is what I made you listen to’.

The incorrect vowel in the 1.SG pronoun written as *me* instead of *maī* is most probably due to a typo in the corpus. The DO *ye* is either a literal transcription of its singular form *yah*, which is pronounced in the same way, or a grammatical mistake, in which case it is noteworthy that there is no plural agreement on the verb.

13 *mārnā*

There are two verbs which were mentioned as only taking objective marked DOs, *mārnā* and *bulānā*. Both of them have different meanings. The verb *mārnā* means ‘to beat’, ‘to hit’, and – as a rule in compound verbs – ‘to kill’. In the former meaning *mārnā* is constructed with an internal object, such as particularly frequent in our corpus, *golī* ‘bullet’. This is a conjunct verb. We found 107 constructions with *mārnā* in the corpus. 16 are constructions with an unmarked DO. 11 are conjunct verb constructions with an objective marked DO. Of the remaining sixteen, the object can either be considered as an internal object, or the construction is highly idiomatic. In those cases, no animate patient is present:

- (23) *āj maī=ne itnā baṛā hāth mārā hai*
 today I=ERG such big hand.M hit.PST.M.SG AUX.PRS.3SG
 (40ehinsp047)
 ‘Today I hit such a big hand.’ (i. e., I grabbed unlawfully plenty of money).
- (24) *ṭokrī pe lāt mārṭā hai*
 basket.F on kick.F hit.IPFV.M.SG AUX.PRS.3SG
 (23ehinsp048)
 ‘[He] kicks at the basket.’
- (25) *āmīr khaṇ ko bhī lāt mārṭī rahtī hai*
 Amir Khan OBJ too kick.F hit.IPFV.F.SG DUR.F.SG AUX.PRS.3SG
 (2ehinsp047)
 ‘She keeps on kicking Amir Khan too.’

14 *bulānā*

With *bulānā*, one has the meaning of ‘to call someone something’³⁰, ‘to invite’. The first meaning of course again implies a construction with a nominal predicative, which is unmarked, and a *ko*-marked object.

- (26) *ve bacce usī=ko pāpā bulāte hai*
 this.pl child.M.PL he.OBL.EMPH=OBJ daddy call.IPFV.M.PL AUX.PRS.3PL
 (1ehinsp047)
 ‘These children call him daddy.’

30 We didn’t find any example phrases with this meaning of *bulānā* which contain an inanimate DO.

36 of the 39 transitive constructions with *bulānā* in our corpus have an animate DO. All of these DOs are marked in the objective case. There are three occurrences of unmarked DOs, all of which are nonspecific and, interestingly, these three instances refer to a collective noun: *jalsā* ‘meeting’ in (27) and *mīṭiṅg* ‘meeting’ and *hartāl* ‘strike’ in the two other instances. Note that they are grammatically treated as inanimate, even though they semantically refer to animate entities.

- (27) *inhō=ne* *īrāk ke* *opposition ke* *tamām leaders³¹ kā*
 they.OBL=ERG Iraq GEN opposition GEN all leaders GEN
ek bahut barā *jalsā* *bulāyā* *thā*
 one very big meeting.M call.PFV.M.SG AUX.PST.M.SG
 (4ehinsp03b)
 ‘They had called a very big meeting of all the leaders of the Iraqi opposition.’


Our brief corpus investigation indicates that the verb semantics plays a role in preferences for one or the other marking of DO. However, it is unclear how strong this is related to the verb meaning rather than to the meaning of the object itself. Some verbs combine much more frequently with an animate or with an inanimate patient than others. With the verbs with an apparent predilection for unmarked DOs, we see that objective DOs are certainly not excluded, in particular, if they are pronominal, which concerns also inanimate nouns. The pronominal DOs in our examples are anaphoric, which implies that they are specific, and might require *ko*-marking, especially when they refer to an animate noun. The interesting thing here is that this shows that the verb semantics does not inhibit null-marking, on the contrary, these examples show that the speaker might want to stress the specificity by *ko*-marking the pronoun. It is only with *gānā* and *pīnā* that we found no marked DO in the corpus. However, both verbs take a much less specific patient than *banānā* and *paṛhnā*. The patients of *gānā* and *pīnā* are practically always mass nouns, or at least nonspecific. With *banānā* and *paṛhnā*, this is much less the case – here patient arguments can be specific, in the former case even animate. Concerning the verbs which only take objective patients, *mārnā* and *bulānā*, we find that they do occur with nonspecific DOs which do not take a marking (e.g. ‘meeting’ in example (27)). Apart from our findings on DOM, we also noted several ungrammaticalities in the examples, especially when a marked DO was present. Our corpus investigation thus suggests that more research into spoken Hindi is required. This might shed more light on substandard grammatical variation regarding agreement and its relation with, e.g. DOM.


31 “opposition” and “leaders” are in Roman script in the corpus.

15 Conclusion

In this article, we gave an overview of the different accounts of DOM in Hindi, focusing on the interaction between the different features argued to determine the objective marking of the DO. We saw that most accounts take animacy as the trigger factor, with animate human arguments generally taking *ko*-marking. In the group of inanimate DOs, we found that accounts that argue that the marking is determined by the feature of specificity are the most precise, based on attestations of inanimate, indefinite but specific DOs which take the objective marking. As such, we stand by our argument that a division of indefinite DOs into specific and nonspecific arguments for Hindi is more precise than the broader categorization of definiteness. The non-specificity of animate arguments is also sometimes related to the DO being unmarked, in particular, if the DO is animate but non-human. Finally, we presented some data from a small-scale corpus investigation of constructions with several verbs that had been reported to occur only with one type of DO. Through a fine-grained analysis of the examples, we observed that the meaning of the verb does not exclude a particular type of marking. Even though we would find many examples with a typical type of marking for this verb, we also found several examples of marked DOs that would be regarded atypical in light of the described restrictions. In each of these examples the important role of the feature of specificity was once more confirmed: the occurrence of an atypically marked DO in combination with such “restrictive” verbs could always be explained by the specificity of the DO. Additionally, we observed that interesting ungrammaticalities occur in this type of corpus data, which definitely asks for further research.

ORCID®

Saartje Verbeke  <https://orcid.org/0000-0002-9528-432X>

Aaricia Ponnet  <https://orcid.org/0000-0001-9708-3807>

Bibliography

- Abbott, Barbara 2004. “Definiteness and Indefiniteness”, in: Horn, Laurence R. & Gregory Ward (eds): *The Handbook of Pragmatics*. (Blackwell Handbooks in Linguistics 16.) Malden; Oxford; Carlton: Blackwell Publishing, 122–149.
- Aguilar-Guevara, Ana, Julia Pozas Loyo & Violeta Vázquez-Rojas Maldonado 2019. *Definiteness Across Languages*. Language Science Press (DOI: [10.5281/zenodo.3265959](https://doi.org/10.5281/zenodo.3265959)).

- Aissen, Judith 2003. “Differential Object Marking: Iconicity vs. Economy”, *Natural Language and Linguistic Theory* 21,3: 435–485.
- Bhatt, Rajesh & Elena Anagnostopoulou 1996. “Object Shift and Specificity: Evidence from ko-phrases in Hindi”, in: Dobrin, Lise M., Kora Singer & Lisa McNair (eds): *Papers from the Main Session, April 11–13, 1996, Chicago Linguistic Society* 32 (1). Chicago: Chicago Linguistic Society, 11–22.
- Butt, Miriam 1993. “Object Specificity and Agreement in Hindi/Urdu”, in: *29th Regional Meeting of the Chicago Linguistic Society* 1: 80–103.
- Butt, Miriam & Tracy Holloway King 2004. “The Status of Case”, in: Veneeta Dayal & Anoop Mahajan (eds): *Clause Structure in South Asian Languages*. Dordrecht: Kluwer, 153–198.
- de Hoop, Helen & Andrej L. Malčukov 2008. “Case-Marking Strategies”, *Linguistic Inquiry* 39, 4: 565–587.
- de Hoop, Helen & Bhuvana Narasimhan 2005. “Differential Case Marking in Hindi”, in: Mengistu Amberber & Helen de Hoop (eds): *Competition and Variation in Natural Languages: The Case for Case*. Amsterdam; Boston; Heidelberg et al.: Elsevier, 321–346.
- Fodor, Janet Dean & Ivan A. Sag 1982. “Referential and Quantificational Indefinites”, *Linguistics and Philosophy* 5: 355–398.
- Geist, Ljudmila, Dolgor Guntsetseg, Udo Klein, Edgar Onea & Klaus von Heusinger 2007. *Differential Object Marking and Verb Semantics*. Talk given at the SFB Seminar, University of Stuttgart.
- Hardie, Andrew 2012. “CQPweb - Combining Power, Flexibility and Usability in a Corpus Analysis Tool”, *International Journal of Corpus Linguistics* 17, 3: 380–409. (<<https://cqpweb.lancs.ac.uk/emillehinsp/>>, accessed: February 8, 2021).
- Kachru, Yamuna 1980. *Aspects of Hindi Grammar*. New Delhi: Manohar Publications.
- Kachru, Yamuna 2006. *Hindi*. Amsterdam: John Benjamins Publishing.
- Klein, Udo 2007. *Comparing Rule-Based and Constraint-Based Analyses of Differential Object Marking*. Paper presented at the workshop: Differential Case Marking, Stuttgart.
- Klein, Udo & Peter de Swart 2011. “Case and referential properties”, *Lingua* 121: 3–19.
- Lyons, Christopher 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Malčukov, Andrej L. 2008. “Animacy and Asymmetries in Differential Case Marking”, *Lingua* 118: 203–221.
- Mohanan, Tara 1994. *Argument Structure in Hindi*. Stanford: Center for the Study of Language.
- Montaut, Annie 2004. *A Grammar of Hindi*. München: LINCOM Europa.

- Nishioka, Miki & Lago Language Institute 2016–2017. Corpus of Spoken Hindi (COSH) and COSH Conc (< <http://www.cosh.site>>, accessed: October 15, 2019)
- Self, Stephen 2012. *Differential ko=marking: Case and Object Marking in Hindi*. Presentation at the GIAL Academic Forum, Dallas, 11 October 2012.
- Singh, Mona 1994. “Thematic Roles, Word Order and Definiteness”, in: Butt, Miriam, Tracy Holloway King & Gillian Ramchand (eds): *Theoretical Perspectives on Word Order in South Asian Languages*. (CSLI Lecture Notes. Stanford University Center for the Study of Language and Information 50.) Stanford: CSLI Publications, 217–236.
- Singh, Mona 1994. *Perfectivity, Definiteness and Specificity: A Classification of Verbal Predicates in Hindi*. Doctoral Dissertation (unpublished). The University of Texas at Austin.
- Spencer, Andrew 2005. “Case in Hindi”, in: Butt, Miriam & Tracy H. King (eds): *Proceedings of the LFG’05 Conference*. Bergen: CSLI Publications.
- Vasishth, Shravan & Brian D. Joseph 2008. “Constellations, Polysemy, and Hindi KO”, *Proceedings of the Berkeley Linguistics Society (BLS)* 28: 137–46.
- von Heusinger, Klaus 2002. “Specificity and Definiteness in Sentence and Discourse Structure”, *Journal of Semantics* 19, 3: 245–274.

Ekaterina Kostina 

To Link or Not to Link? Variation in Using Discourse Markers in Hindi

एकातेरीना कोस्तिना

लिंक किया जाए या नहीं? प्रोक्ति चिह्नों के प्रयोग में परिवर्तनशीलता

Abstract The paper attempts to understand some specific features and variability of discourse markers in Hindi through a linguistic experiment. Adopting the approach suggested by Vadim Kasevič (2014) to Hindi data, I will distinguish the following discourse markers (DMs): *bhī* ‘also’, ‘even’ and *lekin/par* ‘but’ as connectives, *hī* ‘only’, ‘particularly’, ‘just’, ‘exactly’ and *to* ‘indeed’, ‘sure’ as emphatic particles, and *to* ‘but’, ‘as for’ as the contrastive topic marker. Interjections, modal and parenthetical words, often treated in the literature as DMs, will not be discussed here.

In the experiment, the respondents received a fragment of a short story in Hindi with the DMs extracted.¹ The hypothesis was that the absence of DMs might stimulate the native speakers to reconstruct them. At the same time, substantial variability in reconstruction was also expected.

The results demonstrated that Hindi DMs were highly variable in terms of their frequency, location, and selection from the set of lexical items, which occasionally resulted in a change to the initial message or text structuring.

We may conclude that although the DMs are not autosemantic words, their role is not that of mere fillers, and they truly provide text integrity and coherence.

Keywords discourse markers, text coherence, Hindi, linguistic experiment, variability.

1 I thank all the participants of the experiment for providing valuable data and especially Mr. Gautam Kashyap (St. Petersburg State University) for his constant support and willingness to help. I am also most grateful to Dr. Joan Margaret Barry for meticulously checking the English of the paper.

सारांश प्रयुक्त शोध के द्वारा एक भाषाई प्रयोग के माध्यम से हिंदी में प्रोक्ति चिह्नों (डिस्कोर्स मार्कर या डी.एम.) की विशेषताओं और उनके कार्यों की परिवर्तनशीलता का विश्लेषण करने का प्रयास किया गया है। व. ब. कसेविच (Kasevič et al. 2014) के द्वारा सुझाए गए सिद्धांतों को हिंदी भाषा के लिए अपनाकर, कुछ प्रयुक्त प्रोक्ति चिह्नों/डी.एम. की चर्चा की जाएगी, उदाहरणस्वरूप – “भी” एवं “लेकिन” (पर) संयोजकों के रूप में, “ही” (बलाघात या बलसूचक (emphatic)) चिह्न के रूप में, “तो” व्यतिरेकी टॉपिक के मार्कर के रूप में। यहाँ विस्मयादिबोधक, मोडल और अभिभावक चिह्न की चर्चा अनपेक्षित है।

प्रयोग के दौरान कुछ उत्तरदाताओं द्वारा हिंदी की एक कहानी के अंश में से निकाले गए प्रोक्ति चिह्नों/डी.एम. को फिर जोड़ा जाना था। यह अनुमान लगाया गया कि मार्कर की अनुपस्थिति में पाठ अस्वभाविक लगता एवं हिंदी भाषी को इस कहानी में डी.एम. जोड़ने की अति उत्सुकता होती। साथ ही, चूँकि डी.एम. के उपयोग के कोई सुनिश्चित नियम नहीं हैं, इसलिए पुनर्निर्माण में परिवर्तनशीलता भी अपेक्षित थी।

प्रयोग के उपरांत परिणामों ने प्रदर्शित किया कि हिंदी के प्रोक्ति चिह्न बहुत ही परिवर्तनशील हैं। परिणामोपरांत यह पाया गया कि कुछ प्रतिभागियों (उत्तरदाताओं) ने कहानी के अंश में एक भी डी.एम. का उपयोग नहीं किया, अन्य प्रतिभागियों ने मूल पाठ की तुलना में तीन गुना अधिक डी.एम. का प्रयोग किया। तथापि, कुछ प्रतिभागियों ने एक ही स्थान के लिये अलग-अलग डी.एम. को जोड़ा, या डी.एम. को स्थानांतरित किया। जिसके परिणामस्वरूप कभी प्रारंभिक संदेश बदल जाता था तो कभी प्रारंभिक पाठ की संरचना प्रभावित हो जाती थी।

इस प्रयोग के आधार पर निष्कर्ष यह निकलता है कि यद्यपि प्रोक्ति चिह्न स्वतः शक्त शब्द नहीं है, उनकी भूमिका सामान्य पूरक की भूमिका से कहीं बढ़कर है और वास्तव में वे पाठ अखंडता और सुसंगतता का प्रावधान करते हैं।

मुख्य शब्द – प्रोक्ति चिह्न/डिस्कोर्स मार्कर (या डी.एम.), पाठ सुसंगतता, हिंदी, भाषाई प्रयोग, परिवर्तनशीलता।

1 Introduction

The concept of *discourse markers* (DMs) was introduced by D. Schiffrin in her dissertation *Discourse Markers: Semantic Resource for the Construction of Conversation* (Schiffrin 1982) and later developed in the book *Discourse Markers* (Schiffrin 1987) to describe peculiarities of a set of linguistic items that function in cognitive, expressive, social, and textual domains. Although the term has acquired significant popularity in studies on pragmatics (see Lenk 1998; Siepmann 2005; Ranger 2018), there is still no common definition of DMs or criteria to distinguish them from conjunctions, particles, modals or parenthetical words and expressions (see Fedriani and Sansó, 2017).

The concept of DMs has also been developed in a number of works on Hindi. Here again one may notice significant conceptual differences and mismatches in the lists of DMs. Thus, M. Shapiro (2003) defines them as “[l]inguistic devices by which sentences [...] are woven together in the formation of large chunks of discourse” and explains the discursive meaning of particles *to*, *hī* and *bhī*. He

also mentions *zarā*, *na*, *kyō*, *kyā* and *vaise to* without a detailed description of their semantics (ibid.: 275–276). A similar concept is represented in the works by A. Montaut (2004: 265, 278, 283 ff.; 2016). D. Sharma (1999: 5), focusing primarily on the issues of nominal clitics, discusses *hī*, *bhī*, *to*, *tak* and *bhar*. The collective paper “Towards an Annotated Corpus of Discourse Relations in Hindi” (Prasad et al. 2008: 75) focuses on discourse connectives, which, according to its authors, can be divided into five types: subordinating conjunctions, sentential relative pronouns, subordinators, coordinating conjunctions and discourse adverbials. Ya. Kachru (2006: 269) significantly widens the scope of the concept and defines it as “[a] number of linguistic items in Hindi which have no specific grammatical function, [such as] attitude markers, boundary-markers, confirmation-seekers, discourse connectives, pause fillers, hedging devices, hesitation markers, intimacy signals, prompters, repair markers, turn-taking signals, and topic-switch markers”.

This paper is based upon the concept of DMs worked out on the basis of evidence from various Asian and African languages within the framework of the project *Linear and Nonlinear Means of Ensuring the Coherence of the Eastern Text*² headed by V.B. Kasevič (Kasevič et al. 2014; 2018). In brief, distinctive characteristics of DMs in accordance with this concept are formulated as follows:

- DMs establish the coherence of discourse.
- DMs do not influence the meaning of an isolated phrase; if we extract a DM, the general idea of the sentence will remain the same.
- DMs do not deliver any grammatical meaning.
- DMs do not express the speaker’s emotions or attitude to what is being said.³

DMs can be classified into three major categories: connectives, emphatic markers and contrastive topic markers. Emphatic markers, in their turn, can express either strong or weak emphasis. If a word can act as both a DM and an auxiliary of another type, such usage shall be treated as manifestations of homonymy rather than polysemy.

2 Saint Petersburg University grant No 2.38.295.2014, 2014–2016.

3 It should be noted that DMs are very close to attitudinal and modal words and expressions, which are often viewed as one of the means of providing coherence of discourse (Kachru 2006: 269). For example, the Chinese language is reported to be rich in lexemes with modal semantics functioning as DMs (Kolpačková 2014: 91). According to Kasevič (2014: 87), allocation of DMs should not intervene in any generally recognized system in grammatical description, and modal words are integral to the complicated system of modality with its internal connections and contrapositions.

Applied to the Hindi data, these characteristics allow us to distinguish four groups of DMs:

Connectives: *lekin*, *par* and other lexemes with the basic meanings of ‘but’, ‘however’, *bhī* ‘also’, *aur* ‘and’, *phir* ‘then’.

Strong emphasis markers (emphasising the whole sentence): *kyā* ‘what’, *na* ‘is it not so’, ‘why not’.

Weak emphasis markers (emphasising words or phrases in a sentence): *hī* ‘particularly’, ‘just’, *bhī* ‘even’, *tak* ‘even’, *to* ‘certainly’, *na* ‘indeed’⁴.

Contrastive topic marker: *to* ‘as for...’, ‘indeed’.⁵

On the basis of ‘simple’ DMs, a great number of complex DMs can be formed, such as *phir bhī* ‘nevertheless’, *nahī to* ‘otherwise’, etc. So, the class of DMs can be perceived as an open one. Classification of DMs is for a great part contextual, as most of them can deliver a wide range of senses, sometimes spreading beyond the limits of pure connectivity or emphasis.

Emphatic pronouns in Hindi are the product of merging of a pronoun with the emphatic *hī*, e. g. *vahī* DIR, *usī* OBL ‘that very’. Such forms cannot be treated as DMs as they are pronouns, but since the emphatic part is inseparable from the pronominal one, it is necessary to consider them in a discussion on *hī*. Used after determiners, *hī* marks re-identification (‘the same’), while in combination with lexemes of other classes, including personal pronouns, it can be restrictive, centring the notion, denoting a high degree of feature manifestation or emphasising the predicate (Montaut 2004: 290). Following this observation, I distinguish the emphatic determiners as a special set.

Hindi DMs are usually discussed in the literature ‘in general’, with examples produced by the authors themselves or bearing no reference to the source. This paper demonstrates some specific features of these auxiliary lexemes, especially their variability, on the basis of a particular literary text and its reconstructions in the course of a linguistic experiment. The task of thirty-nine respondents was to reconstruct DMs in a written fragment of a short story in Hindi, from which the conductors of the experiment had extracted almost all DMs. The experiment will be described in detail in Section 2; the results and their interpretation will be presented in Sections 3 and 4.

Such an approach imposes considerable limitations on the scope of materials. I do not discuss DMs and contexts not available in the selected text or its reconstructions. This is at least partially compensated for by a strong focus on DM variations within a complete story, short though it might be. By virtue of the experiment, it becomes possible to see to what extent the discourse itself prompts the

4 *Na* expressing the weak emphasis usually follows the subject of the sentence.

5 The difference in semantics and usages of the coordinative, argumentative and topic-marking *to* has been thoroughly described by Annie Montaut (2016).

appearance of one or the other DM and what circumstances, including individual preference, are responsible for variation in DM usage. As far as I am aware, this is the first experiment aimed at DM reconstruction basing on Hindi material. Its results should be treated as preliminary, requiring further analysis on the basis of a wider corpus of texts and a larger number of respondents.

2 Description of the experiment

There are no fixed language rules determining DM usage, and they are subject to strong variation (Varma 2004: 70–74). It is only possible to determine the extent to which the variation is admissible by a set of experiments. A similar experiment was first conducted by a Sinologist, E. N. Kolpačkova. In the course of that experiment, native speakers of Chinese were asked to fill in gaps in the text (a cloze test) choosing from a list of DMs, so that grammatical correctness, as well as semantic and logical naturalness of the text, could be preserved (Kolpačkova 2015: 368). Representing the material in the form of a cloze test was, however, criticised by peers, for the conditions of the task seemed to prompt the decision to some extent. After an attempt to repeat the experiment with a cloze test using a Hindi text, it was decided that providing the respondents with a text without any gaps would deliver more interesting results.

The experiment was intended to provide the empirical data for answers to the following questions:

1. Is it possible for a native Hindi speaker to understand a text totally devoid of DMs?
2. Will the participants be able to put the extracted DMs back into their places?
3. Can the influence of social, gender or age factors be traced in DM usage pattern?
4. Is it possible to predict variation in DM usage from the phrase structure or context?
5. Can DMs alter and to what extent?
6. How does the usage of an alternative DM influence the coherence of the text?

The experiment was conducted in the following manner. An extract from the short story *Aṃtarāl* by Kamlā Kānt Tripāṭhī (2001), from which almost⁶ all the DMs had

- 6 Since there was only one *tak* meaning ‘even’ in the chosen part of the story, it was not extracted from the text and not included into the list of the words to be inserted, in order to avoid the possibility of reconstructions of the homophonous case marker *tak*. This proved to be the right decision, as even under the stated conditions there were cases of

been removed and in which the emphatic forms had been replaced by neutral ones, was offered to Hindi native speakers and Hindi experts who are not native speakers. This passage of text was selected for the experiment because it was a story within a story and seemed not to require acquaintance with the rest of the text to be understood properly. At the same time, it appeared quite rich in DMs of various types. The text deals with a hereditary tailor called Altaf, a Muslim friend of the narrator, who moved to Mumbai from their village in search of a better life. At first, he manages to find work and a place to live, but his life becomes endangered as the Hindu-Muslim riots begin, forcing Altaf to escape back home only to find their family nest and the traditional lifestyle pattern falling to decay. The genre of the text can be defined as the event-type narrative. The total number of words is 435, of which 16 are discourse markers: *hī* (4), *bhī* (5), *par* (1), *lekin* (2), and *to* (4); the text also contains four emphatic pronominal forms: *vahī* (1), *usī* (1), and *unhī* (2). The original text (transliterated) and its English translation are provided in the Appendix. All sentences are numbered with a combination of the letter T (for ‘text’) and their sequence numbers in the text, e. g. (T1). In this way, they are distinguished from the example numbers, which have no letter.

The participants received the “bare” text, a list of extracted words, and the following instructions (in Hindi): “In front of you is an extract from a short story in Hindi. Please read it carefully. For the purpose of the experiment, we have extracted the words *hī*, *bhī*, *lekin*, *par* (= *lekin*), and *to* and replaced *vahī* with *vah*, *usī* with *us* etc. If you feel something is missing, please insert it. If you feel something is incorrect, please correct it.” To ensure that respondents understood the task properly, they were provided with an explanation either in Hindi or in English. The pool of respondents included 39 native speakers of Hindi, the majority of whom are based in Delhi, Varanasi and Chandigarh. All of them are university graduates (philologists, whose speciality is Hindi or Russian language and literature, journalists, software developers, economists and engineers). In addition, four Russian and one Singhalese (from Sri Lanka) native speakers participated in the experiment. They are all Hindi teachers at different institutions. Information about the latter group is provided in Table 1a; in Tables 2 to 8 their results are marked with NH (non-Hindi). However, it was only the output from the data returned by the native speakers that was considered in the conclusions drawn.

Ideally, the experiment should have been conducted as one event, with respondents assembled on one day in one place and the experiment controlled

misinterpreted homophones: instead of the intended reconstruction of *par* ‘but’ the case marker *par* ‘on’ was reconstructed. Similarly, *aur* ‘and’ and *phir* ‘then’ (conjunctions homophonous with adverbs) were not extracted not to confuse the respondents.

while it was running. However, this kind of arrangement proved to be impossible, and the experiment was conducted distantly, in the form most comfortable for the participants. Most of them received the task as an email attachment after having agreed to participate. The others were provided with the printed version of the task text in one-off meetings in August and September 2019. In general, it took the participants of the experiment from half an hour to a couple of days to complete the task. The respondents returned the tasks by 30 December 2019. So, there was as good as no time limitation for the test nor control. It is almost certain that with an instant reaction, the individual results would have been somewhat different, but the overall difference would, most likely, have been negligible.

Since it may be presupposed that using the DMs lies beyond the sphere of pure logic and grammatical competence, this experiment can be treated as measuring the participants' unconscious behaviour and thus a departure from the principle of full disclosure⁷ (see, e.g. Blom & Unsworth 2010: 6) is possible. In most cases, providing little information on the actual purpose of the experiment proved its value. However, as nothing had been said about the origin of the text and also due to some specific features of the author's style, a number of participants decided they had to edit the story as if written by a foreigner, and instead of fulfilling the task offered their help as proofreaders. Two of the respondents obviously assumed that we were checking their knowledge of Hindi and, wishing to show better results, apparently found the original story on the internet and used it while working on the text.

Table 1 lists all the participants sorted by the number of successfully reconstructed DMs and includes information on their gender, age and occupation ("Hindi", "Russian" or "English" stand for professional occupation with the language). Non-native Hindi speaker respondents are additionally marked NH (non-Hindi), the participant who received the text with gaps in the places of DMs is marked with G (gaps), and those presumed to have used the original text while fulfilling the reconstruction task are marked D (doubtful respondents).

7 Following the principle of full disclosure is usually recommended for any research implicating quantitative data analysis. According to this principle, the participants "[s]hould be told what the purpose of the research is, what procedures will be followed, how data will be used, how it will be managed and stored, how their anonymity will be preserved, what benefits may be reasonably expected, as well as being informed of any risks or discomfort participation may involve" (Blom and Unsworth 2010: 6). However, as Blom and Unsworth further state, in some cases (including measuring the unconscious behavior) only partial disclosure or even opacity of the purpose of the experiment for the respondents is admissible (*ibid.*: 6).

Table 1 Participants sorted by their results

	Gender and age	Occupation's language	hī, 4 ⁸		vahī, usī, unhī, 4		bhī, 5		par + lekin, 3		to, 4		total, 20	
			ins	match	ins	match	ins	match	ins	match	ins	match	ins	match
1	F 34	Hindi	0	0	0	0	0	0	0	0	0	0	0	0
2	M 23	Other	0	0	0	0	0	0	0	0	0	0	0	0
3	M 24	Other	0	0	0	0	0	0	0	0	0	0	0	0
4	M 27	Other	0	0	0	0	0	0	0	0	0	0	0	0
5	M 39	Journalist, Hindi and English	0	0	0	0	0	0	0	0	0	0	0	0
6	M 40	Other	0	0	0	0	0	0	0	0	0	0	0	0
7	M 45	Other	0	0	0	0	0	0	0	0	0	0	0	0
8	M 49	Hindi	0	0	0	0	0	0	0	0	0	0	0	0
9	M 32	Other	0	0	1	0	0	0	1	0	0	0	2	0
10	M 63	Other	0		1	1	0	0	0	0	0	0	1	1
11	F 34	Other	0	0	1	1	1	0	0	0	0	0	2	1
12	M 65	Other	0	0	1	1	2	0	0	0	0	0	3	1
13	F 60	Hindi	0	0	0	0	3	0	1	1	0	0	4	1
14	M 39	Other	1	0	0	0	3	0	1	1	0	0	5	1
15	F 30	Other	0	0	1	1	5	1	0	0	0	0	6	1
16	F 32	Other	2	0	1	1	2	0	1	0	0	0	6	1
17	M 35	Russian	3	0	2	1	4	0	0	0	0	0	9	1
18	M 49	Other	5	0	0	0	5	1	0	0	0	0	10	1
19	M 26	Other	0	0	1	0	1	0	1 ⁹	1	0	0	3	2
20	M 46	Journalist, Hindi and English	0	0	2	1	0	0	1	1	0	0	3	2
21	F 67	Hindi	2	0	0	0	1	0	3	2	0	0	6	2
22	M 48	Other	2	0	1	1	6	2	1	0	1	0	11	3

8 The digit after the comma indicates the number of DM occurrences in the original text. The column “ins” (inserted) contains the number of DMs inserted by a respondent, and the column “match” contains the number of exact matches with the original DMs.

9 The respondent used *paramtu* instead of *lekin* or *par* that were listed for insertion.

	Gender and age	Occupation's language	hī, 4 ⁸		vahī, usī, unhī, 4		bhī, 5		par + lekin, 3		to, 4		total, 20	
			ins	match	ins	match	ins	match	ins	match	ins	match	ins	match
23	M 42	Other	4	0	1	1	6	2	2	0	0	0	13	3
24	M 40	Hindi	2	1	1	1	2	1	2	1	0	0	7	4
25	M 32	Other	1	1	2	1	5	1	3	1	0	0	11	4
26	M 24	Russian	2	1	0	0	4	0	9	3	3	0	18	4
27	F 32	Other	7	0	3	1	9	2	4	1	0	0	23	4
28	F 57	Journalist, Russian, Hindi, English	7	1	2	2	10	2	0	0	0	0	19	5
29	M 28	Other	7	0	3	2	9	0	5	2	2	1	26	5
30	M 55	Other	4	2	1	1	1	1	3	2	1	0	10	6
31	M 77	Russian, Hindi	5	0	1	1	2	1	3	3	4	1	15	6
32	F 28	Other	3	1	0	0	10	3	5	1	4	1	22	6
33	M 40 (second attempt)	Other	4	0	0	0	10	2	7	3	6	1	27	6
34	M 50	Hindi	3	1	1	1	9	2	3	1	4	2	20	7
35	F 64	Hindi	4	1	2	2	9	2	2	2	3	1	20	8
36	M 28	Russian	3	2	2	2	4	3	2	2	3	2	14	11
37	M 70 G	Russian	3	2	1	1	6	4	3	3	3	3	16	14
38	M 35 D	Other	2	2	3	3	5	5	3	3	3	2	16	15
39	M 40 D	Other	4	4	4	4	6	5	3	3	4	4	21	20

Table 1a Non-Hindi native participants sorted by their results

1	F 25NH	Hindi	1	0	2	2	1	0	3	0	1	0	8	2
2	F 39NH	Hindi, English	1	0	3	3	3	0	1	0	0	0	8	3
3	F 45NH	Hindi	1	0	1	1	4	0	2	1	2	1	10	3
4	F 40NH	Hindi	5	0	1	1	3	1	3	1	1	1	13	4
5	F 41NH	Hindi	1	1	2	1	3	1	2	1	2	1	10	5

3 Discussion of the results of the experiment

Eight participants obviously misunderstood the instructions and did not insert any DMs. This does not mean that they were unscrupulous while doing the task: none of them returned blank pages. They corrected (or miscorrected) everything they considered a grammatical or orthographic mistake or stylistic failure. Their corrections and suggestions included, among others:

- spelling mistakes (*puštainī* ‘hereditary’ for *puštaunī*, *khetihar* ‘farmer’ for *khetihār*) or variants (*uskā* for *us kā* ‘his’, *vo* for *vah* ‘this, he’);
- grammar mistakes (*damge kī capeṭ mē* ‘in the blow.f of the riot’ for *damge ke capeṭ mē* ‘in the blow.m of the riot’);
- Indo-Aryan words instead of Arabic-Persian loanwords;
- simple past forms instead of the past perfect forms;
- changes in the degree of politeness/subordination.

At least two participants stated overtly that the text did not need any changes except for some spelling corrections: “I realise that the story has been written by one of the old-time authors. The expression is like that. Hence the grammar is correct in accordance with the time frame of the story. Although I found a couple of spelling mistakes, I would not exactly call them mistakes because each author has a different style of writing”¹⁰ (M 39). One of the participants (M40) did not insert any DMs at first, but then he re-read the text keeping in mind the necessity to use DMs and literally “tried to use as many of them as possible”. These facts prove that the chosen words truly fulfil the requirements for DMs, as they do not influence the comprehensibility of the text although their absence may make it look a bit awkward from the stylistic point of view.

The preliminary observations from the experiment are summarised as follows:

1. Usage of DMs appears to be highly dependent on the personal preferences of the speaker. Most full matches were demonstrated by those participants who, irrespective of their gender, age or occupation, managed to “catch the tune” of the author. None of the participants (including those who presumably broke the experiment guidelines by using the original text) reconstructed all the DMs. Even the person who was given the paper with gaps in place of extracted DMs interchanged *hī* and *bhī* in some cases (this observation is developed further in 4). Since the number of male participants considerably exceeds that of females, it is difficult to trace gender-specific trends in DM reconstruction except for the fact that in general women tried harder to follow the requirements of the task and use all extracted words in the list and did not

10 In the quotations from the participants’ remarks to the text their spelling is retained.

attempt to copy the original text. Although some of the participants noticed that the style of the text appeared to be quite old-fashioned, age does not seem to be an influencing factor: it cannot be stated that older participants were more successful, and respondents of the same age group demonstrated totally different results. Participants with a background in linguistics (especially non-native Hindi speakers) tended to use more DMs of all types, as they obviously looked at the text from the point of view of a researcher. However, such a thoughtful attitude did not make their reconstructions any closer to the original text.

2. Despite being provided with the list of extracted words, 21 of the respondents did not use one or more of the DMs on the list. 18 respondents missed *to*, 8 did not use any *hī*, 7 papers lack emphatic pronouns and only 3 papers demonstrate a total absence of *bhī*. This leads us to the tentative conclusion that DMs with connective meaning are the easiest to reconstruct and their absence is the most discernible, while the necessity to mark the contrastive topic is the weakest.
3. In most cases variation in DM usage observed in the experiment papers does not lead to the emergence of incorrect structures. In this context, the postposition *ke bāvajūd* ‘in spite of’, followed by *bhī* ‘even’, deserves a special mention. Strict stylistic rules of Standard Hindi do not allow this sequence (Maharotā 2004: 62). The original phrase contains no DM: *tamām dar aur dahśat ke bāvajūd ve apne gāv-deś kā nātā nahī bhūle the* ‘**In spite of** all fear and anxiety, they did not forget the ties with their native village.’ Seven respondents put *bhī* after the postposition (*ke bāvajūd bhī*), thus doubling the concessive meaning of the phrase and in this way producing a stylistically disputable structure. However, this frequently occurs not only in the spoken language but also in literary and scholarly works (see, e. g. Simh 2007: 206). Other cases of variable DM usage by the participants of the experiment are not questionable from either a grammatical or stylistic point of view, although they often result in a shift of semantic accents or change in text structuring.

4 Variations in DM reconstruction

All variations studied fall into the following types:

1. DM vs. \emptyset — a DM of the original text is not reconstructed in any form.
2. \emptyset vs. DM' — a DM of any type is inserted at a place containing no DM in the original.
3. DM vs. DM' — a DM of another type is reconstructed either exactly at the original DM location or in the same phrase.

A shift in DM positioning, when a DM is reconstructed at a place very close to, yet different from the initial one, although technically a combination of variation types DM vs. Ø and Ø vs. DM', in some cases deserves special attention from the point of view of text coherence.

Variation of the first type is overwhelming: not a single DM of the original text was reconstructed by at least half of the participants. It may be interesting to observe the “reconstruction rate” of each of the DMs and attempt to determine the reasons behind it. At the same time, a number of structures, by contrast, appear to be very “attractive” for DM insertion, although they lack a DM in the original text. These cases also deserve special attention.

Tables 2, 4 and 6 show the phrases with DMs from the returned documents with their English translations and glossing lines. As the agreement is of no importance for the current study, gender, number and case, except for the ergative, are not glossed unless it is needed to avoid ambiguity. The number of papers in which the DMs reconstructions exactly match the original phrases is presented in the rightmost column. Table 3 illustrates the cases of excessive – as compared to the original text – usage of connectors. Table 5 demonstrates different ways of structuring the text resulting from variations in DM usage. Table 7 deals with one particular phrase which appeared to demonstrate the largest number of DM variations, while Table 8 demonstrates multiple DM variants suggested by the respondents in the phrases initially containing no DMs. Within the tables (except for Table 7), phrases are listed in the order of their appearance in the text¹¹ and are sometimes abridged. The numbers after the original sentences refer to the complete text represented in the Appendix. The results of the two participants defined as “doubtful” are provided in the form “+2D” or “+1D”, while “+1G” stands for the responses from the paper with gaps (M70G). Below the sentences exactly reproducing the original, the same phrases containing alternative DMs suggested by the participants are provided (if any). Such phrases are marked with (A) for “alternative”. If the original sentence contains two or more DMs of different types, figures in the column “Number of respondents” are provided only for the DM in question in the corresponding paragraph, irrespective of whether the respondents reconstructed the other DMs.

11 For more clarity and description convenience, in Table 2 phrases with *lekin* or *par* are grouped together in the top and are followed by phrases with *bhī*, and in Table 4 emphatic pronouns are separated from ptcls. In Table 8 sentences with similar DM variations are grouped.

4.1 Variation in expressing connectivity

Table 2 Reconstruction of connectives

Phrase, glossing line and translation	DM	Number of respondents
(1) <i>use</i> [...] <i>faikṭarī mē</i> [...] <i>kām mil gayā thā</i> . to.him [...] factory in [...] work become.available go.PFV was. par <i>unhī</i> <i>dinō vahāḍ damṅe</i> <i>bharaḥ uṭhe</i> <i>the</i> (T5, T6) but those.EMPH days there riots erupt rise.PFV were 'He had got a job at the factory. But riots had erupted there on the very same days.'	<i>par</i> ¹² / <i>lekin</i>	7+1NH+ 2D+1G
(A) <i>un</i> <i>dinō vahāḍ damṅe</i> <i>bhī</i> <i>bharaḥ uṭhe</i> <i>the</i> those days there riots also erupt rise.PFV were 'Riots also had erupted there those days.'	<i>bhī</i>	1
(2) <i>do-ek bār kasbe mē jākar sūī lagvāī</i> <i>thī</i> . two-one time town in having.gone needle cause.to.be.applied.PFV was. lekin <i>koī fāyadā nahī huā</i> (T34, T35) but any result not was '[Father] had gone to the town a couple of times to be given injections. But to no avail.'	<i>lekin</i> / <i>par</i> / <i>parantu</i>	15+1NH+ 2D+1G
(3) <i>usne</i> [...] <i>ādmī se bāt kī</i> <i>hai jo saūdī mē rahtā</i> <i>hai</i> . he.ERG [...] man with talk do.PFV is who Saudi in live.IPFV is. lekin <i>vīsā vagairah me baṛā kharc hai</i> (T39, T40) but visa etc. in big cost is 'He has had a talk with a man who lives in Saudi [Arabia]. But the visa etc. cost much.'	<i>lekin</i> / <i>par</i> / <i>magar</i>	6+1NH+ 2D
(4) <i>kuch log use sṭeśan tak choṛne</i> <i>bhī</i> <i>āe</i> <i>the</i> (T18) some people him station up.to to.see.off also come.PFV were 'Some people also had come to the station to see him off.'	<i>bhī</i>	5+1NH+ 2D+1G
(A) <i>sṭeśan tak</i> <i>bhī</i> <i>choṛne</i> <i>āe</i> <i>the</i> station up.to even to.see.off come.PFV were 'Some people even had come to the station to see him off.'	<i>bhī</i>	2
(5) <i>ab to abbā ko</i> <i>bhī</i> <i>kām nahī miltā</i> ¹³ (T25) now but father to also work not become.available.PRS 'But now father also does not get work.'	<i>bhī</i>	3+2D+1G

12 The slash here indicates that any of the DMs listed in the table cell can be found in the respondents' papers, but all of them have been calculated together: in this experiment they proved to be interchangeable. Some participants even used the combination "*par/lekin*" themselves instead of choosing one word.

13 The experiment papers contained so many variants of DM usages for this phrase, that I found it reasonable to discuss all of them in a separate section, see Table 7.

Phrase, glossing line and translation	DM	Number of respondents
(6) <i>jise silānā bhī hotā hai vah</i> [...] (T27) to.whom sewing even be.IPFV is he [...] ‘[Even if] someone needs [clothes] to be sewn [...]	<i>bhī</i>	0+2D
(A) <i>jise silānā hī hotā hai vah</i> [...] (4A) to.whom sewing EMPH be.IPFV is he [...] ‘The one who does need [his/her clothes] be sewn [...]	<i>hī</i>	1NH+1G
(7) <i>dhīmā-dhīmā bukhār bhī rahtā hai</i> (T32) slow-slow fever also stay.IPFV is ‘[He] also [has] a slight fever.’	<i>bhī</i>	11+2D+1G
(8) <i>umar bhī to ho gā hai</i> (T36) age also after.all be go.PFV is ‘After all, [his] age is also high.’	<i>bhī</i>	5+1NH+2D+1G

DMs with the adversative connective meaning are the easiest to reconstruct, the most variable in terms of mutual interchangeability and the least variable in terms of interchangeability with other markers. Despite the fact that the list of extracted words was provided, some of the respondents preferred other adversatives, such as *magar*, and the conjunction *paraṃtu* (without taking into consideration the Urdu-oriented style of the text). The high reconstruction rate for *lekin* in (2) is caused by the prototypical character of the situation of “failed expectations”, a basic meaning of adversative connectors.

The contraposition in (1) and (3) is indirect, and the intermediated character of contraposition makes the relation between the clauses less clear. In (1) it is not the events described in the phrases connected with *par* (Altaf’s finding a job and the beginning of riots) but the consequences of these events (i. e., ‘peaceful life’ and ‘troubles’) that are juxtaposed. Example (3) contraposes the first step in Altaf’s plan to move abroad (talk with a man living in Saudi Arabia) to his realisation of his inability to do so. In these cases, the grounds for reconstructing an adversative connector are still less obvious.

As for *bhī*, the most easily reconstructed are the cases with purely connective meaning, such as (4), (7) or (8), presenting “complex situations”. In the phrase (4), however, *bhī* can also be interpreted as an emphatic marker. As a Hindi native speaker explained, “under the circumstances described (riots), the final event [i. e., going to see Altaf off — *E. K.*] is the least likely one as measured by the scale of common suppositions, therefore, [it means] ‘even’”. Such interpretation is supported by the even more “emphatic” variation (4A) suggested by two respondents, who shifted *bhī* to the place after *steśan tak* ‘up to the station’. In this position, *bhī* cannot be interpreted as marking the complex situation (there is no listing of

other places besides the station to which Altaf could have been taken), and the DM clearly bears the emphatic meaning ‘even’.

The least evident is the usage of *bhī* ‘also’ in the connective-emphatic sense (6). Although some participants felt the necessity to place a DM in this sentence, *bhī* was reconstructed only by the two doubtful respondents. Even the participant provided with the paper with gaps preferred to insert *hī* ‘exactly’. Two native speakers (philologists) were asked to explain the difference between the phrases (6) and (6A), and both agreed that the general meaning was the same, although they confirmed that there was a subtle difference between the original sentence and the reconstructed one. This makes it clear that complex semantic issues deserve further study and cannot be discussed in detail at this stage of research.

The number of connectors suggested by the respondents but absent in the original text is high; see Table 3. It may be noted that some descriptions of complex situations have no connectors in *Āmtarāl*, while in other cases, the connector *bhī* marks the last link in a chain of events or a list of attributes. Similarly, the author does not use *bhī* in the description of a complex situation if it contains conjunction. In such cases, quite a number of respondents suggested using a connector, as in (9) and (10) in Table 3 or shifted *bhī* to the space after the second component of the complex description, as in (11) and (12)¹⁴ in Table 3. The doubtful respondent M40 even provided an extra connector *bhī*, although he presumably was aware of its absence in the original text (12).

The absence of *bhī* at the places where it is most expected to appear may be explained as a feature of the style, which is quite different in the selected extract compared to the rest of the story. Although it is not a first-person narrative, the author attempts to imitate spontaneous narration by a village tailor of his misfortunes rather than relate the story artistically from the distant point of view of a writer. This results in short sentences, sometimes incomplete or lacking the subject (this has been noticed by some participants of the experiment who suggested subject reconstruction), devoid of connectors (including the adversative ones, see (13), (14) in Table 3) or connective-emphatic markers.

14 The complex situation in the first case includes not only feeding Altaf and booking him the ticket but also going to the station to see him off, and the sentence (4) (T18), which follows (10) (T17) in the text, contains *bhī* in the original. In the second case, the situation also includes the cough that won’t stop and, finally, a slight fever (the last component of the description is marked with *bhī* in the original text (T32), as we have seen in the example (7) above.

Table 3 Connectors reconstructed at “wrong” places

Phrase, glossing line and translation	DM	Number of respondents
(9) <i>Altāf ke [...] bāl khicī ho gae the.</i> Altāf GEN [...] hair salt-and-pepper be go.PFV were. <i>gāl bhī picak gae the (cf. T1, T2)</i> cheeks also shrivel go.PFV were 'Altāf's [...] hair had begun to turn gray. His cheeks had also shriveled.'	<i>bhī</i>	15+1NH
(10) [...] <i>uske dost ko cākū ghōp diyā gayā thā.</i> [...] his friend DAT knife pierce give.PFV AUX.PASS.PFV was. <i>ṭāksī bhī jalā dī gāī thī (cf. T8, T9)</i> taxi also burn give.PFV AUX.PASS.PFV was '[...] his friend had been pierced by a knife. Also the taxi had been burnt down.'	<i>bhī</i>	7+3NH
(11) [...] <i>ṭāksī jalā dī gāī thī.</i> [...] taxi burn give.PFV AUX.PASS.PFV was. <i>par vah kisī tarah bhāg niklā thā [...]</i> (cf. T9, T10) but he some way run come.out was [...] '[...] the taxi had been burnt down. But he had run away somehow [...]	<i>par</i>	3
(12) <i>unhōne [...] use khilāyā-pilāyā thā aur [...]</i> they.ERG [...] him feed.PFV-give.drink.PFV was and [...] <i>ṭikaṭ bhī nikalvākar diyā thā (cf. T17)</i> ticket also having.caused.to.be.taken.out give.PFV was 'They were the ones who had fed him [...] and booked him a ticket [...]	<i>bhī</i>	7+1NH
(13) [...] <i>bīmār hī cal rahe hai.</i> [...] ill EMPH go PROG AUX.PRS.3HON. <i>khāṭ bhī pakāṭ lī hai (cf. T29, T30)</i> bed also catch take.PFV is '[...] [father] has been running constantly ill. [He] is also bedridden.'	<i>bhī</i>	7+1D
(14) <i>dhīmā-dhīmā bukhār bhī rahtā hai.</i> slow-slow fever also stay.IPFV is. <i>par ab kisko dikhāē?</i> (cf. T32, T33) but now to.whom should.show '[Father] also [has] a slight fever. But to whom should [we] show [him] now?'	<i>par, lekin</i>	2+1NH

4.2 Variation in expressing emphasis

Table 4. Reconstruction of emphatic pronominal forms and emphatic DMs

Phrase, glossing line and translation	DM, pronominal form	Number of responders
(15) <i>par unhī dīnō vahāṅ dāṃge bhāṛak uṭhe the (T6)</i> but those.OBL.EMPH days there riots erupt rise.PFV were 'But riots had erupted there on the very same days.'	<i>unhī</i>	2+1D+2NH
(A) <i>par un dīnō hī vahāṅ dāṃge bhāṛak uṭhe the</i> but those days EMPH there riots erupt rise.PFV were 'But riots had erupted there on the very same days.'	<i>hī</i>	1

Phrase, glossing line and translation	DM, pronoun form	Number of responders
<p>(16) <i>usī kī ākhō ke sāmne uske dost ko</i> he.OBL.EMPH GEN eyes GEN before his friend DAT <i>cākū ghōp diyā gayā thā</i> (T8) knife pierce give.PFV AUX.PASS.PFV was ‘His friend had been pierced by a knife before his own eyes.’</p>	<i>usī</i>	3+1NH+2D
<p>(A) <i>us kī ākhō ke sāmne hī</i> [...] he.OBL GEN eyes GEN before EMPH [...] ‘[...] right before his eyes.’</p>	<i>hī</i>	4+1NH
<p>(17) <i>unhī ke yahā [...] vah [...] ek-ek</i> they.OBL.EMPH GEN here [...] he [...] one-one <i>rāt kāt rahā thā</i> (T15) night spend PROG was ‘At their places he was spending night by night.’</p>	<i>unhī</i>	1+1NH+2D
<p>(A) <i>un ke yahā bhī</i> [...] they.OBL GEN here also [...] ‘also in their location [...].’</p>	<i>bhī</i>	1
<p>(18) <i>lauṭkar phir vahī dhāk ke tīn pāt</i> (T23) having.returned again that.EMPH butea GEN three leaves ‘On coming back, it was all the same again.’</p>	<i>vahī</i>	16+3NH+1G+2D
<p>(19) <i>ilāke ke himdū mazdūro aur ṭaiksivālō</i> region GEN Hindu laborers and taximen <i>kā hī use sahārā milā thā</i> (T14) GEN EMPH him help become.available.PFV was ‘He had been supported by none other than Hindu laborers and taximen from [his home] region.’</p>	<i>hī</i>	4+2D
<p>(A1) <i>ilāke ke himdū mazdūrō aur ṭaiksivālō kā</i> region GEN Hindu laborers and taximen GEN <i>bhī use sahārā milā thā</i> also him help become.available.PFV was ‘He had been also supported by Hindu laborers and taximen from [his home] region.’</p>	<i>bhī</i>	1+1G
<p>(A2) <i>lekin/par/paraṅtu</i> [...] <i>mazdūrō aur ṭaiksivālō</i> but [...] workers and taximen <i>kā use sahārā milā thā</i> GEN him help become.available.PFV was ‘But he had been supported by [...] laborers and taximen.’</p>	<i>lekin/ par/ paraṅtu</i>	6+2NH
<p>(20) <i>unhōne hī karakī me use khilāyā-pilāyā thā</i> (T17) they.ERG EMPH distress in him feed.PFV-give.drink.PFV was ‘They were the ones who had fed him in times of distress.’</p>	<i>hī</i>	3+1NH+1G+1D
<p>(A) <i>unhōne karakī me bhī use khilāyā-pilāyā thā</i> they.ERG distress in also him feed.PFV-give.drink.PFV was ‘They had fed him even in times of distress.’</p>	<i>bhī</i>	6+1NH

Phrase, glossing line and translation	DM, pronoun form	Number of responders
(21) [...] <i>bahut dinō se bīmār hī cal rahe hāi</i> (T29) [...] many days since ill EMPH go PROG AUX.PRS.3HON '[...] [Father] has been running constantly ill for a long time.'	<i>hī</i>	1+1D
(A) [...] <i>bahut dinō se bīmār bhī cal rahe hāi</i> [...] many days since ill also go PROG AUX.PRS.3HON '[...] has also been running ill for a long time'	<i>bhī</i>	5+1NH+1G
(22) <i>ek hī bār mē itnā karzā-kuām ho gayā</i> (T37) one EMPH time in so.much debt be go.AOR '[I] got into so much debt in one single go.'	<i>hī</i>	1+2D
<i>ek bār mē hī itnā karzā-kuām ho gayā</i> one time in EMPH so.much debt be go.AOR '[I] got into so much debt in one single go.'	<i>hī</i>	3
(23) <i>umar bhī to ho gāi hai</i> (T36) age also after.all be go.PFV is 'After all, [his] age is also high.'	<i>to</i>	1NH+1G+2D
<i>umar jo ho gāi hai</i> age as be go.PFV is 'As [his] age is high.'	<i>jo</i> ¹⁵	1

It may be noticed that a great number of respondents reconstructed the emphatic determiner *vahī* (18) in the idiom *vahī dhāk ke tīn pāt* (lit. 'the same three leaves of *dhāk*,' which is the *Butea monosperma* tree famous for its trifoliolate leaves), meaning "all the same". The emphatic form here conveys the idea of equity through the similarity of attributes.

Usually, this saying comes with *vahī* 'that very' and if it is preceded by the non-emphatic *vah*, the word is used as a personal pronoun 'he/she/it' rather than a determiner 'that'. It acts then as the sentence subject and requires a verbal form to complete the sentence as in *partāl karne par jo tasvīr milī, vah dhāk ke tīn pāt jaisī hai* (Jāgaraṅ 2013) 'The picture that was found on conducting the search is like the three leaves of *dhāk*' (i. e. absolutely the same). Some of the participants who did not know the idiom and understood the *vahī* as a personal pronoun, stated that the sentence made no sense and was incomplete. This could potentially be interpreted in such a way as to say that the means of emphasis (including the emphatic pronouns, see Introduction) in proverbs and sayings should not be treated as DMs at all as they are integral to fixed expressions and their extraction prevents one from understanding the text correctly.

Contrary to the case with *vahī*, the reconstruction rate for other emphatic pronouns is average. In (15), *unhī dinō*—the emphatic pronoun conveys the idea of

15 The respondent used *jo* despite the fact it was not listed in the task, thus emphasising the causative interpretation of *to*.

re-identification. In other cases—(16) and (17)—the restrictive-emphatic meaning is implied. The same restrictive-emphatic meaning is typical of combinations of *hī* with substantival parts of speech (19), including those pronominal forms which do not take the emphatic marker (20). In combination with the numeral *ek* ‘one’ *hī* emphasises the idea of uniqueness (22) and in (21) *hī* emphasises the high degree of feature manifestation (‘really’, ‘totally’)—as Montaut (2004: 290) describes the meaning of the particle.

In (23) we can see *to* in the emphatic meaning. Here it makes “the clause a crucial discursive argument” (Montaut 2016: 276) and adds value to what is being said rather than changing the topic, which appeared quite difficult to reconstruct.

The analysis of connectors demonstrated a discrepancy between the author’s preference of not using DMs at certain places and the intuition of most respondents who inserted DMs at these places. The same kind of divergence is also evident in the case of emphatic DMs. For example, 14+1NH respondents inserted *hī* to make the phrase *khāṣī baṃd hone kā nām (hī) nahī leī* (T31) ‘The cough won’t stop’ more expressive. Against the background of the average reconstruction rate for this text, the number appears quite high and needs to be explained. According to several native Hindi speakers, the phrase with *hī* sounds more natural, and when the expression *nām na(hī) lenā*, literally meaning ‘not to take [one’s] name’, i. e., ‘not to mention’ is used metaphorically (‘to refrain entirely’), the word *nām* ‘name’ should be accompanied with *hī* or *tak* ‘even’. However, literary sources do not confirm this observation unambiguously. For example, numbers of exact matches in Google Book search for the phrases *baṃd hone kā nām nahī leī* and *baṃd hone kā nām hī nahī leī* in similar contexts are almost the same.¹⁶

Compared to the connectives, the emphatic DMs demonstrate quite naturally a clearer trend to shift within a phrase. In (15) and (16) the participants of the experiment preferred to insert the particle *hī* after the noun rather than change the form of its pronominal attribute. In (22) three respondents decided to emphasise the whole syntagma rather than the single numeral *ek*, which, however, did not lead to any considerable change in text coherence.

In (19) we can see restrictive usage of *hī*. Replacing it with *bhī* ‘also’ leads to the crucial change in the phrase semantics: instead of emphasising the fact that Altaf’s Hindu fellow-countrymen were the ones who helped him (despite the anti-Muslim riots), the reconstructed phrase implies that a number of people helped

16 Ten matches for the non-emphatic phrase (including the one from *Aṃtarāl*) and eight matches for the emphatic one. Changing the spelling (*band* instead of *baṃd*) or the morphological features of the verb does not influence the ratio. Both the non-emphatic and the emphatic variations appear in similar contexts, such as continuous laughing, crying, or raining.

him, including Hindu labourers and taxi drivers. This implication contradicts the context and the reasons behind the two respondents' choice of *bhī* in this sentence remain unclear.

By contrast, usage of a connector of the *lekin* 'but' type is predictable in this sentence due to the contraposition of the behaviour of local labourers and taxi drivers (who ran away from Mumbai) and that of the Hindu labourers from Altaf's region. However, earlier in the text we saw a notice that Altaf had found shelter at the place of his Hindu fellow villagers: *vah [...] apne ilāke ke hinduō kī jhuggī mẽ pahūc gayā thā* 'He [...] had reached the slum where his Hindu fellow-countrymen lived' (T10), and this fact forms the background to further events. Unlike *lekin*, which usually introduces the new alternative information, the restrictive ('none other than X') emphasiser *hī* refers to already known facts: its usage presupposes that we already know that someone does (or did) something and now we learn that nobody else does this. Therefore, its usage here seems absolutely reasonable and preferable.

In (21) most respondents used *bhī*, whereas the original has *hī*. Here — as in (6) — the difference in semantics between the original and the reconstructed phrases is very subtle. Some native speakers formulated it in this way: the emphatic *hī* in the original indicates the gravity of the father's condition (he is seriously ill) and the connective DM in the reconstruction marks the complex situation (not only is the father unable to get a job, but also, he is ill). Due to the presence of one more DM at the beginning of this sentence (abridged here), we will return to DM variations within it in the next section (example 26).

Table 5 Original and reconstructed semantic blocks

Original text, (T5) – (T7)	Semantic blocks	Reconstruction, 5+2NH respondents	Semantic blocks
<i>[...] use [...] faiktārī mẽ [...] kām mil gayā thā</i> 'He had got a [...] job at the [...] factory.'	1. The background for further events.	<i>use [...] faiktārī mẽ [...] kām mil gayā thā</i> 'He had got a job at the factory.'	1. Altaf's first days in Mumbai.
<i>par unhī</i> <i>dinō vahā</i> but those.OBL.EMPH days there <i>damge bharak uṭhe the</i> riots erupt rise.PFV were 'But riots had erupted there on the very same days.'		<i>Ø unØ dinō vahā damge bharak uṭhe the</i> 'Riots had erupted there those days.'	2. Riots begin and Altaf suffers an attack.
<i>us din vah [...] kahī jā rahā thā</i> 'On that day he was going somewhere [...].'	2. Altaf suffers an attack.	<i>usī</i> <i>din vah [...]</i> that.OBL.EMPH day he [...] <i>kahī jā rahā thā</i> somewhere go PROG was 'On that very day he was going somewhere [...].'	

The experiment demonstrated that DMs help divide the text into semantic blocks, and due to the absence of DMs in the task the respondents were not always able to define the borders between these blocks and presumably divided the text differently from the author, see Table 5.

In the original text the emphatic *unhī* ‘those very’ inextricably links the phrase ‘The riots had erupted in those days’ with the previous sentence and affiliates it with Block 1. The next sentence, beginning with the neutral *us* ‘that’, introduces Block 2. In the bare text the absence of the connective *par* ‘but’ results in the fading of the sense of contraposition, and substitution of the emphatic pronoun with the neutral one leads to a loosening in the time relation. As a result, 5 + 2 NH respondents moved the border between the semantic blocks. Having retained the unemphatic *un* ‘those’, they failed to reconstruct the time identity between the first two sentences and having introduced the emphatic *usī* ‘that very’, they established this identity between the second and the third phrases.

4.3 Variation in contrastive topic marking

Table 6 Reconstruction of the contrastive topic marker

Phrase, glossing line and translation	DM	Number of respondents
(24) <i>ab to abbā ko bhī kām nahī milā</i> ¹⁷ (T25) now but father to also work not become.available.PRS ‘But now father also does not get work.’	<i>to</i>	5+1NH+1G+1D
(25) <i>yah to reḍīmed kā zamānā hai</i> (T26) this indeed readymade GEN epoch is ‘This indeed is the epoch of readymade [clothes].’	<i>to</i>	3+1G+2D
(26) <i>idhar to bahut dinō se bīmār hī cal rahe hāī</i> (T29) here indeed many days since ill EMPH go PROG are ‘Moreover, [he] has been running constantly ill for a long time.’	<i>to</i>	0+1G+2D
(A) <i>idhar vah bhī bahut dinō se bīmār cal rahe hāī</i> Here he also many days since ill go PROG AUX.3PRS.HON ‘Recently he also has been running ill for a long time.’	<i>bhī</i>	1

It has already been noted that many respondents did not use *to* in their papers. Correct reconstructions of *to* are quite rare, and cases of its placement in positions other than in the original are, for the most part, individual.

In (26) *to* was not reconstructed by anyone apart from the two respondents presumed to have cheated and the one who had a paper with gaps and therefore knew something had to be inserted in that place. This may result from

17 This sentence will be discussed in detail in Table 7.

understanding *idhar* as ‘recently’, while in combination with *to* it can be translated as ‘moreover’. If *idhar* is translated as ‘recently’, usage of the contrastive topic marker *to* after it requires that the current state of the father’s health be contraposed to his condition previously. The selected extract does not contain enough information on earlier days, and so the respondents did not feel the necessity to put the contrastive topic mark after the time reference. Here we can see that usage of a DM can sometimes be determined by rather distant parts of discourse. For some of the respondents the phrase seemed unnatural and they suggested saying simply *in dinō vah bahut bīmār hai* ‘These days he is very ill’ or corrected it by adding the missing subject *vah* ‘he’ or replacing the discomfoting *idhar* with the honorific form *ve* ‘they’ making it agree in number with the predicate.

Table 7 Variations in DM usages in the phrase “Now father does not get work”

Characteristics of the text	Phrase, glossing line and translation	Number of respondents
Bare phrase	<i>ab abbā ko kām nahī miltā</i> ‘Now father does not get work.’	25
Original text	<i>ab to abbā ko bhī kām nahī miltā</i> now but father to also work not become.available.PRS ‘But now father also does not get work.’	3+1D+1G
Reconstructions by participants of the experiment	<i>ab to abbā ko kām nahī miltā</i> now but father to work not become.available.PRS ‘But now father does not get work.’	1NH
	<i>ab to abbā ko kām bhī nahī miltā</i> now but father to work even not become.available.PRS ‘But now father does not get even work.’	1+1NH
	<i>ab bhī abbā ko kām nahī miltā</i> now also father to work not become.available.PRS ‘Also now father does not get work.’	1
	<i>ab abbā ko bhī kām nahī miltā</i> now father to also work not become.available.PRS ‘Now father also does not get work.’	+1D
	<i>par ab abbā ko kām nahī miltā</i> but now father to work not become.available.PRS ‘But now father does not get work.’	1
	<i>ab abbā ko kām bhī nahī miltā</i> now father to work also not become.available.PRS ‘Now father does not get work too.’	5
	<i>ab abbā ko kām hī nahī miltā</i> now father to work EMPH not become.available.PRS ‘Now father absolutely does not get work’	1+1NH
	<i>ab abbā ko to kām nahī miltā</i> now father to but work not become.available.PRS ‘Now as for father, he does not get work.’	1

One structure (T25) deserves special attention, as its interpretation demonstrates the full variety of DM usages, see Table 7. In order to explain such variability, it seems necessary to provide the context in which the phrase under discussion (put in bold) appears (T23–T26). The places in which DMs were used either by the author or by the respondents are left blank and underlined.

Lauṅkar phir vahī dhāk ke tīn pāt. gāv ke puṣṭainī peṣe mẽ ab kyā rakhā hai? _____ ab _____ abbā ko _____ kām _____ nahī miltā. yah to reḍīmeḍ kā zamānā hai. ‘On returning, everything was the same. Now what has remained of the traditional occupation of the village? **Now father does not get work.** This indeed is the epoch of readymade [clothes].’

Although the absence of DMs or their occurrence in other places than those of the original did not prevent the respondents from understanding the text, it is clear that without the missing part of the short story it is difficult (if not impossible) to reconstruct the author’s idea. In the original text the combination of *to* and *bhī* marks the juxtaposition of the new situation with the previous one: the father used to be the only tailor in the neighbourhood who had a sewing machine and to whom people from all the surrounding villages (including the narrator and his family) used to come to have their clothes sewn.¹⁸ Otherwise, it can mark the complex situation (both the whole village and now also the father cannot get work).

Two respondents emphasised the sense of difficulty in finding work in general by adding the emphatic *hī* after *kām* ‘work’, and seven respondents placed *bhī* ‘also’ not after *abbā ko* ‘to father’ but after *kām* ‘work’. With this DM the phrase became connected directly with the neighbouring sentences adding yet another feature to the description of a complex situation (the father and son’s and also the whole village’s misfortune).

The majority of respondents did not topicalise *ab* ‘now’. This word is included in the rheme of the previous sentence and its functioning as the topic of this phrase appears natural even without *to*. One of them decided (supposedly based on the phrase *vahī dhāk ke tīn pāt* ‘everything was the same’ and having no idea of the family’s relatively prosperous life when Altaf was a child) that the sentence was to illustrate the fact that at present nothing was changing for the better and so provided *ab* ‘now’ with *bhī* ‘also’. At the same time, eight respondents worked out that the miserable situation of the present day was being contraposed to the good old times and expressed this idea with the help of the contrastive topic marker.

One of the respondents preferred to topicalise with *to* not the present time (*ab*), but the father himself (apparently basing this on the presence of *ab* ‘now’ in the previous sentence and the implication that the father belonged to the traditional

18 This information is provided in the passage preceding the selected episode.

occupation mentioned in it). This example demonstrates that the connective power of a DM can sometimes stretch far beyond the immediate surroundings.

Finally, in a number of structures participants suggested different means of providing discourse coherence, none of which had been used by the author.

Table 8 Variation of Ø vs. DM'

Characteristics of the text	Phrase, glossing line and translation	Number of respondents
(27) Original (bare) phrase	<i>uske pās kirāe kī kaun kahe, khāne tak¹⁹ ko paise nahī the</i> (T13) 'As for the money to pay for rent, what to say about it, he had no money even to buy food.'	
Reconstruction	<i>uske pās kirāe kī to kaun kahe</i> [...] he.OBL.GEN at rent GEN even who will.say [...] 'As for the money to pay for rent, what to say about it [...]'	2
	<i>lekin uske pās kirāe kī kaun kahe</i> [...] but he.OBL.GEN at rent GEN who will.say [...] 'But what to say about the money to pay for rent [...]'	2+1NH
	[...] <i>khāne tak ko bhī paise nahī the</i> [...] food even for also money not were ' [...] [he] had also no money even for food.'	5
	[...] <i>khāne tak ko paise bhī nahī the</i> [...] food even for money also not were ' [...] [he] had no even money even to buy food.'	2+2NH
(28) Original (bare) phrase	<i>do-ek bār kasbe mē jākar sūī lagvāī thī</i> (T34) '[Father] had gone to the town a couple of times to be given injections.'	
Reconstruction	<i>do-ek bār kasbe mē jākar sūī</i> two-one times town in having.gone needle <i>to lagvāī thī</i> but cause.to.be.applied.PFV was 'As for injections, [father] had gone to the town a couple of times to be given them.'	5
	<i>do-ek bār kasbe mē jākar sūī</i> two-one times town in having.gone needle <i>bhī lagvāī thī</i> also cause.to.be.applied.PFV was '[Father] also had gone to the town a couple of times to be given injections.'	2

19 The DM *tak* was not extracted from the text, see Note 6.

Characteristics of the text	Phrase, glossing line and translation	Number of respondents
(29) Original (bare) phrase	<i>usne paṛos ke gāv̄ ke ek ādmī se bāt kī hai</i> [...] (T39) 'He has had a talk with a man from the near village [...].'	
Reconstruction	lekin <i>usne paṛos ke gāv̄ ke ek ādmī</i> but he.ERG nearby GEN village GEN one man <i>se bāt kī hai</i> with talk do.PFV is 'But he has had a talk with a man from the near village.'	2
	<i>usne paṛos ke gāv̄ ke ek ādmī se</i> he.ERG nearby GEN village GEN one man with bhī <i>bāt kī hai</i> also talk do.PFV is 'He has had a talk with a man from the near village too.'	2
	<i>usne paṛos ke gāv̄ ke ek ādmī se bāt</i> he.ERG nearby GEN village GEN one man with talk bhī <i>kī hai</i> also do.PFV is 'He has also had a talk with a man from the near village.'	1
	<i>usne paṛos ke gāv̄ ke ek ādmī se bāt</i> he.ERG nearby GEN village GEN one man with talk to <i>kī hai</i> indeed do.PFV is 'He has indeed had a talk with a man from the near village.'	1
(30) Original (bare) phrase	[...] <i>apne bac nikalne kā bharosā nahī huā thā</i> (T22) '[he] had had no belief in escaping.'	
Reconstruction	[...] <i>apne bac nikalne kā bharosā</i> [...] own escape get.away GEN belief hī <i>nahī huā thā</i> (T22) EMPH not be.PFV was '[he] had had absolutely no belief in escaping.'	5
	[...] <i>apne bac nikalne kā bharosā</i> [...] own escape get.away GEN belief bhī <i>nahī huā thā</i> (T22) even not be.PFV was '[he] had had not even belief in escaping.'	1+1NH
(31) Original (bare) phrase	[...] <i>mumbaī jāne kī himmat nahī hotī</i> (T37) '[...] there's no courage left to go to Mumbai.'	
Reconstruction	[...] <i>mumbaī jāne kī himmat hī nahī hotī</i> (T37) [...] Mumbai going GEN bravery EMPH not be.PRS '[...] there's absolutely no courage left to go to Mumbai.'	9+2NH
	[...] <i>mumbaī jāne kī himmat bhī nahī hotī</i> (T37) [...] Mumbai going GEN bravery even not be.PRS '[...] there's not even courage left to go to Mumbai.'	1

In (27) usage of the structure [...] *kī kaun kahe* with the emphatic DM *tak* was emphatic enough for the author, while the respondents suggested adding contrastive and emphatic-connective DMs to the phrase. The contrastive topic marker can deliver a meaning very close to that of a connector, and both DMs were added by some respondents to emphasise the contraposition expressed by the adversative connector in (28) and (29).

Sentences (30) and (31) demonstrate variation in *hī/bhī* usage. Although both DMs deliver the emphatic meaning, there is a significant difference between them here. In negative sentences *bhī* marks the expected event that does not occur and its usage denotes failed expectations (previously had bravery/belief but no longer has). On the other hand, *hī* in such cases denotes maximal realisation. Since bravery or belief are not the expected feelings here, usage of *bhī* seems less relevant, and this observation is supported by the notable prevalence of reconstructions of *hī*.

5 Conclusions

Despite the limited volume of primary data and limited range of examined DMs the experiment has enabled some conclusions to be made about the scope and types of DM variation in Hindi, which can lay a foundation for further analyses of the phenomenon.

1. Variability is a typical characteristic of DMs. Lack of variation in the usage of a subsidiary lexical unit points with high probability to the fact that it is not a DM in the proper sense of the term or that it is being used in a proverb or fixed expression.
2. Usage of DMs strongly depends on personal preference and sense of style. While some of the participants did not insert a single DM, others used up to three times more DMs than there were in the original text.
3. Age, gender or the level or type of education of the respondent do not appear significantly to influence his /her use of DMs. Data for further socio-cultural groups have to be investigated.
4. The use of DMs also obviously depends on the sentence syntax and lexis. Some sentences in the text are neutral to DMs: a DM can be inserted or not, while other sentences appear to provoke a strong inclination to use a DM of one or another type.

5. DMs of all types not only help to bind the text together and highlight the key points, but also mark borders between semantic blocks of which the discourse is built, and which shift with the changing of a DM.
6. The linking power of a DM can stretch beyond the limits of adjoining sentences or paragraphs: sometimes just knowing the wider context can ensure the usage of the authentic DM. Otherwise, people may (as is mostly the case) insert a range of DMs at one and the same place, which results in a shift of emphasis. For the most part, the preferability of the author's choice of a DM can be explained, but only as a posteriori truth.
7. Not unexpectedly, DMs conveying prototypical and simple meanings are the easiest to reconstruct, while DMs with complex meanings are subject to greater variation.


Bibliography

- Blom, Elma & Sharon Unsworth (eds) 2010. *Experimental Methods in Language Acquisition Research*. (Language Learning and Language Teaching Series 27.) Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Fedriani, Chiara & Andrea Sansó (eds) 2017. *Pragmatic Markers, Discourse Markers and Modal Particles: New Perspectives*. (Studies in Language Companion Series 186.) Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Kachru, Yamuna 2006. *Hindi*. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Kasevič, Vadim B., Olga A. Bernikova, Pavel L. Grohovskij, Evgenija I. Injuškina, Elena N. Kolpačkova, Anna V. Kuryševa, Dar'ja I. Šbnaja & Natalija S. Jahontova 2014. "Diskursivnye slova vostočnogo teksta" (Discourse Markers in the Oriental Text), in: *Jazyki stran Dal'nego Vostoka, Jugo-Vostočnoj Azii i Zapadnoj Afriki: Materialy XI Meždunarodnoj Naučnoj Konferencii. Sankt-Peterburg, 30.09.2014–02.10.2014*. (Languages of Far East, Southeast Asia and West Africa: Materials of the 11th International Scientific Conference. Saint Petersburg, 30.09.2014–02.10.2014). Sankt-Peterburg: Studija NP-Print, 86–90.
- Kasevič, Vadim B., Evgenija I. Injuškina, Ekaterina A. Kostina, Anna V. Mazarčuk & Natalija S. Jahontova 2018. "Celostnost' i svjaznost' teksta: nekotorye predvaritel'nye soobraženija" (Text Integrity and Coherence: Some Preliminary Considerations), in: *Jazyki stran Dal'nego Vostoka*,

- Jugo-Vostočnoj Azii i Zapadnoj Afriki: Materialy XIII Meždunarodnoj Naučnoj Konferencii: Sankt-Peterburg, 30.10.2018–01.11.2018* (Languages of Far East, Southeast Asia and West Africa: Materials of the 13th International Scientific Conference. Saint Petersburg, 30.10.2018–01.11.2018). Sankt-Peterburg: Izdatel'stvo CSO, 92–97.
- Kolpačkova, Elena N. 2014. “Modal'nye i diskursivnye slova v sovremennom kitajskom jazyke” (Modal Words and Discourse Markers in Contemporary Chinese), in: *Jazyki stran Dal'nego Vostoka, Jugo-Vostočnoj Azii i Zapadnoj Afriki: Materialy XI Meždunarodnoj Naučnoj Konferencii. Sankt-Peterburg, 30.09.2014–02.10.2014* (Languages of Far East, Southeast Asia and West Africa: Materials of the 11th International Scientific Conference. Saint Petersburg, 30.09.2014–02.10.2014). Sankt-Peterburg: Studija NP-Print, 91–94.
- Kolpačkova, Elena N. 2015. “Vlijanie gendernogo faktora na vybor sredstv obespečenija svjaznosti kitajskogo teksta (na osnove lingvističeskogo eksperimenta) (Influence of the Gender Factor on the Choice of Means of Provision of the Chinese Text Coherence (based on evidence of a linguistic experiment))”, in: *Azija i Afrika v menjajuščemsja mire. XXVIII Meždunarodnaja Konferencija po istočnikovedeniju i istoriografii stran Azii i Afriki, Sankt-Peterburg, 22–24 aprolja 2015* (Asia and Africa in the Changing World. The XXVIII International Conference on Historiography and Source Studies of Asia and Africa, Saint Petersburg, 22–24 April 2015). Sankt-Peterburg: VF SPBGU, 368.
- Lenk, Uta 1998. *Marking Discourse Coherence: Functions of Discourse Markers in Spoken English*. (Language in Performance Series 15.) Tübingen: Narr.
- Maharotrā, Rameś Candra 2004. *Mānak hindī ke śuddh prayog* (Correct Usage of Standard Hindi), Vol. I. Dillī: Rādhākṛṣṇa Prakāśan.
- Montaut, Annie 2004. *A Grammar of Hindi*. München: LINCOM Europa.
- Montaut, Annie 2016. “The Discourse Particle *to* and Word Ordering in Hindi”, in: Fernandez-Vest, Jocelyne & Van Valin, Robert (eds): *Information Structuring and Spoken Language from a Cross-linguistic Perspective*. Berlin; Boston: De Gruyter Mouton, 263–282.
- Prasad, Rashmi, Samar Husain, Dipti Sharma & Aravind Joshi 2008. “Towards an Annotated Corpus of Discourse Relations in Hindi”, in: *Proceedings of the 6th Workshop on Asian Language Resources*: 73–80. (<https://www.aclweb.org/anthology/I08-7010.pdf>), accessed: 15 March 2019).
- Ranger, Graham 2018. *Discourse Markers: An Enunciative Approach*. Cham: Palgrave Macmillan.
- Schiffrin, Deborah 1982. *Discourse Markers: Semantic Resource for the Construction of Conversation*. Dissertations available from ProQuest. AAI8307360. (<https://repository.upenn.edu/dissertations/AAI8307360>), accessed: 28 July 2020).

- Schiffirin, Deborah 1987. *Discourse Markers*. (Studies in Interactional Sociolinguistics Series 5.) Cambridge University Press.
- Shapiro, Michael 2003. “Hindi”, in: Cardona, George & Jain, Dhanesh (eds): *The Indo-Aryan Languages*. (Routledge Language Family Series.) London; New York: Routledge, 250–285.
- Sharma, Devyani 1999. “Nominal Clitics and Constructive Morphology in Hindi”, in: Butt, Miriam & Tracy Holloway King (eds): *Proceedings of the LFG99 Conference*. CSLI Publications, 5–16.
- Siepmann, Dirk 2005. *Discourse Markers Across Languages: A Contrastive Study of Second-level Discourse Markers in Native and Non-native Text with Implications for General and Pedagogic Lexicography*. (Routledge Advances in Corpus Linguistics.) London; New York: Routledge.
- Siṃh, Mālṭī 2007. *Ādhunīk hindī kāvyā aur purāṇ kathā* (Contemporary Hindi Poetry and Puranic Tales). Ilāhabād; Naī Dillī; Pāṭnā: Lokbhārtī Prakāśan.
- “Tīn daṛe beasar, cauthe par nazar” (The three ran away headless, keeping an eye on the fourth), *Jāgaraṇ* (05.12.2013) (<<https://www.jagran.com/uttar-pradesh/kanpur-city-10913553.html>>, accessed: 2 July 2020).
- Tripāṭhī, Kamlā Kānt 2001. *Aṃṭarāl aur anyā kahāniyā* (The Interval and Other Stories). Naī Dillī: Ārya Buk Ḍīpo.
- Varma, Terry 2004. “Scalarity and Let Alone Constructions in Hindi”, in: Singh, Rajendra (ed.): *The Yearbook of South Asian Languages and Linguistics 2003*. Berlin: De Gruyter Mouton, 71–95.

ORCID®

Ekaterina Kostina  <https://orcid.org/0000-0002-5742-9516>

Appendix

The analysed extract from the short story *Aṃṭarāl* by Kamlā Kānt Tripāṭhī

Transliteration

altāf ke sar aur dārḥī ke bāl khicṛī ho gae the. (T1) gāl picak gae the. (T2) ākhō mē barbād zimdagī aur tūṭe bhaviṣya kī nīrāsā thī. (T3) kuch sālō pahle māmā ke choṭe laṛke ke sāth vah mumbaī bhāg gayā thā. (T4) kuch din idhar-udhar bhaṭakne ke bād use ek gārmentī faikṭarī mē silāī kā kām mil gayā thā. (T5) par unḥī dinō vahñ damge bharak uṭhe the. (T6) us din vah apne ek musalmān dost kī ṭaiksī mē uske sāth kahñ jā rahā thā ki acānak damge ke capeṭ mē ā gayā thā. (T7) usī kī ākhō ke sāmne uske dost ko cākū ghōp diyā gayā thā. (T8) ṭaiksī jalā dī gaī thī. (T9) vah kisī tarah bhāg niklā thā aur chipte-chipāte apne ilāke ke himḍuō kī jhuggī mē pahūc gayā thā. (T10) ek hafte vahñ chipā rahā thā. (T11) idhar ke sāre mazdūr, kārīgar aur ṭaiksīvāle ṭaigar ke ḍar se sāse hue the aur ek-ek kar mumbaī choṛ rahe the. (T12) uske pās kirāe kī kaun kahe, khāne tak ko paise nahñ the. (T13) ilāke ke himḍū mazdūrō aur ṭaiksīvālō kā hī use saharā milā thā. (T14) unḥī ke yahñ jagah badal-badalkar vah kisī tarah ek-ek rāt kāt rahā thā. (T15) tamām ḍar aur dahaśat ke bāvajūd ve apne gāv-deś kā nātā nahñ bhūle the. (T16) unḥōne hī kaṛakī mē use khilāya-pilāyā thā aur uske lie muluk kā ṭikaṭ nikalvākar diyā thā. (T17) kuch log use ṣeśan tak choṛne bhī āe the. (T18) kaisī kālī-ḍarāvānī rāt thī vah! (T19) har pal maut sāmne nāc rahī thī. (T20) har ādmī dūsre par śak kar rahā thā. (T21) jab tak gārī sahar se bāhar nahñ ā gaī, apne bac nikalne kā bharosā nahñ huā thā. (T22)

lauṭkar phir vahī dhāk ke tīn pāt. (T23) gāv ke puṣṭainī peśe mē ab kyā rakhā hai? (T24) ab to abbā ko bhī kām nahñ miltā. (T25) yah to reḍīmeḍ kā zamānā hai. (T26) jise silānā bhī hotā hai vah kasbe mē jākar silvātā hai. (T27) abbā ko nae faiśan kī cīzē silnī kahā ātī hāī! (T28) idhar to bahut dinō se bīmār hī cal rahe hāī. (T29) khāṭ pakaṛ lī hai. (T30) khāśī baṃd hone kā nām nahñ letī. (T31) dhīmā-dhīmā bukhār bhī rahtā hai. (T32) ab kisko dikhaē? (T33) do-ek bār kasbe mē jākar suī lagvaī thī. (T34) lekin koī fāyadā nahñ huā... (T35) umar bhī to ho gaī hai... (T36) ek hī bār mē itnā karzā-kuām ho gayā aur jān ke aise lāle pāe ki ab dubārā mumbaī jāne kī himmat nahñ hotī. (T37) ab vah kārīgar se khetihār mazdūr ban gayā hai aur apanī aur apāne kunbe kī gārī jaise-taise khñc rahā hai. (T38) usne paṛos ke gāv ke ek ādmī se bāt kī hai jo saūdī mē rahtā hai. (T39) lekin vīsā vagairah mē baṛā kharc hai. (T40) allāh ko jaisā manzūr hoga. (T41)

In most cases the translation provided here is almost entirely congruent with the translations of individual sentences in the paper. However, the Past Perfect used in the examples to reflect the original structure of a Hindi sentence is replaced by the Past Simple. The translation of (T6) in the examples and Table 5 also aims at reflecting the original structure and illustrating the usage of the emphatic pronoun. The structure of this sentence has been changed to make it sound more natural.

Translation

Altaf's hair and beard had begun to grow gray. (T1) His cheeks were shrivelled. (T2) There was disappointment in his eyes: his life was ruined, and his future broken. (T3) Some years ago, he ran away to Mumbai with his uncle's younger son. (T4) He wandered here and there for a while, and, in a couple of days he got a job as a tailor at a garment factory. (T5) But riots erupted there on the very same days. (T6) That day he was driving somewhere with his Muslim friend in his taxi when, all of a sudden, they got gripped by rioters. (T7) His friend was pierced by a knife before his own eyes. (T8) The taxi was burnt down. (T9) He himself somehow managed to run away and covertly reached the slum where his Hindu fellow-countrymen lived. (T10) There, he stayed hiding for a week. (T11) All local labourers, artisans and taxi drivers were leaving Mumbai one by one, scared of Tiger²⁰. (T12) As for the money to pay for rent, what to say about it, he had no money even to buy food. (T13) He was supported by none other than Hindu labourers and taximen from [his home] region. (T14) At their places he was staying, moving from one to another every night. (T15) Despite all fear and panic, they didn't break the link with their native village and land. (T16) They were the ones who fed him in times of distress and booked him a ticket to the village. (T17) Some people even came to the station to see him off. (T18) Oh, how dark and scary that night was! (T19) Death was dancing in front of him every moment. (T20) All people were suspicious of one another. (T21) Before the train came out of the city, [he] had no belief in escaping. (T22)

On coming back, it was all the same. (T23) What has remained of the hereditary occupation of the village these days? (T24) Now even father doesn't get work. (T25) This indeed is the epoch of readymade clothes. (T26) Even if someone needs clothes to be sewn, he goes to the town and orders them there. (T27) Little father knows about sewing new fashion clothes! (T28) Moreover, he has been constantly ill for a long time. (T29) [He] is bedridden. (T30) The cough won't stop. (T31) [He] also [has] a slight fever. (T32) To whom should we show him now? (T33) [Father] went to the town a couple of times to be given injections. (T34) But to no avail... (T35) After all, his age is also high... (T36) [I] got into so much debt in one single go, and so many times life was in danger that now there's no courage left to go to Mumbai once again. (T37) From an artisan, he turned into a field worker, and now somehow manages to carry the van of his and his family's life. (T38) He had a talk with a neighbour who lives in Saudi Arabia. (T39) But visa and all other staff cost so much. (T40) If Allah wills. (T41)

20 Supposedly the author refers to Tiger Memon (nickname of Ibrahim Mushtaq Abdul Razzaq Memon), a Muslim mafia don and terrorist who was standing behind the Mumbai bomb blasts of 1993.

Tatiana Oranskaia 

Hindi Clause Strings with Adverbial Clauses: A Tentative Formalisation

तत्याना ओरांस्कया

हिंदी की उपवाक्य शृंखलाएँ – अंतरिम सूत्रबद्ध प्रणाली

Abstract The article presents a tentative formalisation system for multicomponent Hindi clause strings which include at least one adverbial clause. The aim of the study is to create a tagged database of such constructions. The relations between the members of “classical” biclausal constructions, subordinate and coordinate, serve as the benchmarks for investigating longer clause sequences occupying a level intermediary between complex, compound and compound-complex sentences on the one hand and larger text units on the other hand. The special interest in the subordinate modifier clauses is caused by their semantic variety and relatively loose—as compared to complement clauses—connection with the main clause. Due to the latter characteristic adverbial clauses offer more points of syntactic similarity with simple sentences, and so a better possibility for testing syntactic relations between members of a clause string as well as their semantic grounds. Only finite clauses are considered here components of clause strings. The formalisation is expected to on the one hand explicate the results of the preliminary analysis and on the other hand facilitate further exploration on the topic.

Keywords subordination, coordination, formalisation system, gradience, structural variance.

सारांश प्रस्तुत लेख हिंदी की ऐसी बहुअवयवी उपवाक्य शृंखलाओं (multicomponent clause strings) के अध्ययन का प्रयास है जिनमें कम-से-कम एक क्रियाविशेषण उपवाक्य विद्यमान हो। इस तरह के उपवाक्यों की एक अंतरिम सूत्रबद्ध प्रणाली (tentative formalisation system) यहाँ प्रस्तुत की गई है। हमारा उद्देश्य इस तरह की वाक्य-रचनाओं का चिह्नित डाटाबेस तैयार करना है। प्रधान और आश्रित उपवाक्यों से बनी 'पारंपरिक' (classical) वाक्य-रचनाओं के अंगों (members) के पारस्परिक संबंधों को आधार बनाकर ऐसी उपवाक्य शृंखलाओं का अलग से अध्ययन किया जा सकता है जो, एक ओर तो मिश्र, संयुक्त तथा संयुक्त-मिश्र वाक्यों की दृष्टि से तथा दूसरी ओर संपूर्ण पाठ जैसी अधिक बड़ी इकाई की दृष्टि से „मध्यवर्ती” (intermediary) संरचनाएँ मानी जा सकती हैं। आश्रित विशेषक उपवाक्य (subordinate modifier clauses) हमारा ध्यान इसलिये आकृष्ट करते हैं क्योंकि उनमें अर्थगत विविधता होती है तथा प्रधान उपवाक्य के साथ, पूरक उपवाक्यों की तुलना में, उनका संबंध अपेक्षाकृत शिथिल (loose) होता है। प्रधान उपवाक्य के साथ इस अपेक्षाकृत शिथिल संबंध के कारण क्रियाविशेषण उपवाक्यों की वाक्यरचनागत समानता सरल वाक्यों के साथ अधिक दिखाई देती है और इसलिये ऐसे उपवाक्यों के आधार पर उपवाक्य शृंखलाओं के अंगों के बीच विद्यमान वाक्यगत संबंधों और उनकी विशेषताओं की जाँच करना संभव प्रतीत होता है। यहाँ पर केवल समापिका क्रियायुक्त उपवाक्यों (finite clauses) को ही उपवाक्य शृंखलाओं का अवयव (component) माना गया है। उपवाक्य शृंखलाओं को सूत्रबद्ध करने (formalisation) के इस प्रयास के द्वारा प्रारंभिक विश्लेषण के परिणाम कुछ अधिक स्पष्ट होने के साथ-साथ इस अध्ययन को आगे बढ़ाने में भी सहायता मिल सकेगी।

मुख्य शब्द – आश्रितता/अधीनता, समानाधिकरण, सूत्रबद्धीकरण प्रणाली, ग्रेडिएन्स (gradience), संरचनात्मक अंतर।

1 Introduction

The article presents a tentative version of a system to formalise clause strings in Modern Standard Hindi and the analysis underlying the formalisation. It is the first published outcome of my research project “Hypotaxis in Spoken and Literary Hindi: A Comparative Analysis of Complex Sentences with Adverbial Clauses” at the Mahatma Gandhi Antarrashtriya Hindi Vishwavidyalaya (Mahatma Gandhi International Hindi University) in Wardha, Maharashtra.¹ Its eventual objective is to create a tagged database for complex syntax in Hindi focusing on adverbial clauses.

- 1 I am deeply grateful to the faculty and staff of this—in spite of its young age—renowned university for the wonderful opportunity to work there and benefit from kind assistance of a number of colleagues as well as from the vibrant atmosphere of the university. I am greatly obliged to the Indian Council of Cultural Relations whose fellowship under the programme “Fellowship to Foreign Nationals for Research in India” enabled my research stays in India in 2016–2018. My heartfelt thanks are due to ex-Vice Chancellor Prof. Girishwar Misra, ex-Deputy Vice Chancellor Prof. Anand Vardhan Sharma, Dr Usha Sharma, Dean of the School of Language Prof. Hanuman Prasad Shukla and Prof. Anil Kumar Pandey for their kind attention to my research efforts and for facilitat-

The article is split into two parts. The first part is subordinate, the second part is the main one, to use two of the key concepts in the present discussion. The first part contains an introductory review of issues relevant to the suggested formalisation system. The main part presents a tentative formalisation of clause strings and the underlying analysis. It is a linear tectogrammatical presentation of interclausal relations, refraining from an analysis of the interior syntax of the clauses. No hierarchical trees are built.

This is a result of the first stage of the project, in which the data are elicited from written sources. The next stage of the study will be based on spoken language data.

The definition “clause strings” refers to clause sequences consisting of more than two clauses. Such long strings occupy a level intermediary between complex and compound sentences on the one hand and larger text units on the other hand. Interpretation of these structures coincides with the problem of information hierarchy in human language, which includes the issue of subordinating devices.

We can imagine a long clausal string as a sequence of links, each consisting of two clauses and a binding link carrying the meaning of the semantic connection between them. The dependency distance between the constituent clauses is a variable related to the semantic type of their interconnection. The binding link is not necessarily formally explicit: the connection may rely on implicit meanings binding the clauses together.

“We construe the same situation in alternative ways”, to cite Langacker (2010: 55). The more complex the situation is, the more propositions participate in portraying it. Accordingly, the number of clauses framing them is higher, which allows a higher variation in their linkage leading, in its turn, to a variation increase in perspectives on the situation.

Multicomponent-sequences, which include at least one adverbial clause, are the focus of the project. The adverbial interclausal relations belong principally to the domain of subordination.

As an umbrella notion, complex syntax shares part of its domain with the syntactic organisation of larger text chunks where sentences may have no other form

ing them. Very special thanks are due to Dr Shamim Fatma and Dr Dhanjee Prasad from whose scholarly knowledge, skills and active interest in my topic I have immensely benefited. I cannot thank enough the bright and enthusiastic students who participated in the project. It is with a heavy heart that, due to lack of space, I have to leave them anonymous and also exclude some other parts of my gratitude list, such as my heartfelt thanks to the staff of the Guest House. It is my pleasant duty to thank Dr Joan M. Barry and Mr Philip H. Pierce profoundly for their insightful corrections of my English and Dr Hem Chandra Pande for patiently answering my innumerable questions on linguistic terminology in Hindi and other issues.

to express their semantic connections but the adjoining position—and sometimes not even that.

With the best will in the world complex syntax cannot be considered a new research subject. However, the discussion on clause complexing remains profuse and intense, gaining new force since the 1960es. The subject has been studied in different theoretical frameworks using various data types and formats (see, e.g. Haiman & Thompson 1988; Shopen 2007). As a resource for constructing our experience of the “flow of events”, it is especially significant for the understanding of human cognitive activities, discourse structuring and its hierarchy. Especially promising in this respect are studies on spoken language syntax based on substantial data collections made possible by modern technical devices (Auer et al. 2009).

Adverbial clauses have been attracting considerable attention from syntacticians (see Thompson, Longacre & Hwang 2007). One of the reasons is their relatively loose link to the matrix clause, which is a significant fact for the debate on gradience in grammar (Fanselow et al. 2006). The cross-linguistic semantic and syntactic heterogeneity of adverbial modifiers in general (see, e.g. Eifring 1995: 54; Ricca 2010) is the reason why adverbial clauses span over the subordination–coordination axis.

Against this rich and multifaceted background, research on complex syntax in Hindi is rather scarce. Studies on subordination in Hindi deal for the most part with complement and noun-modifying relative clauses (e.g. Ananthanarayana 1996; Bhatt 2003; Kachru 1978; Dayal 1996; Kothari 2010). Adverbial interclausal relations are to a considerable extent neglected. Accordingly Hindi data are only occasionally used in typological research and on the periphery of scholarly discussion. Thus, Hindi is not considered in the generalising works on syntax mentioned above (Haiman & Thompson 1988; Shopen 2007), nor is it among the 60 languages involved in the crosslinguistic study on causal clauses (Diessel & Hetterle 2011).² Some attention has been given to conditional constructions in Hindi (Oranskaya 2005; Sharma 2010; 2012).

Naturally, analyses of complex syntax underlie implicitly or explicitly annotations in Hindi databanks, such as for example in the Hindi Discourse Relation Bank (Umangi et al. 2009) or the annotation in The Hindi/Urdu Treebank Project (Bhat et al. 2017). Another example is the annotated corpus data on relative clauses contained in the Appendix A to the PhD thesis of Anubha Kothari (Kothari 2010).

This article is a step towards a tentative version of a formalisation system for complex syntax in Hindi. It concentrates on interclausal relations with the adverbial semantics. The formalisation system is expected to enable a closer look into

2 Of Indian languages only Santali and Kannada are considered in this study.

variations in the ways to connect clauses, the placement of adverbial clauses in clause sequences, and the scope of the latter.

2 Basic concepts and terms

The analysis and, accordingly, the formalisations of clause strings in 5.2 use traditional concepts. The major concept is clause: it is a structure including at least a subject and a predicate. Predicate is understood here as a finite predication phrase. This is different from the usual approach, which relates also infinite verb forms to clausal predicates (cf., e.g. Lehmann 1985). Homogeneous subjects and homogeneous predicates are considered to belong to the same clause.

Clause is a relative notion determined by syntactic context. A structure of any syntactic complexity is a clause if it is itself a part of a composite unit whose integrity is based on semantic and syntactic relations within it. This interpretation of the term follows the definition of sentence as a combination of clauses (Longacre 1970; 2007). A sentence may also contain just one clause.

Further concepts belong to the sphere of complex syntax. They are listed here according to their complexity, starting with the simplest one.

C o m p o u n d s e n t e n c e is a composite syntactic unit whose constituent clauses are in a structural equivalence relation, that is **c o o r d i n a t i o n**.

C o m p l e x s e n t e n c e is a composite syntactic unit whose constituent clauses are in a structurally hierarchical relation, that is **s u b o r d i n a t i o n**.

C o m b i n e d s e n t e n c e: the term is reserved here for clause strings that conjoin more than two clauses using explicit devices of both subordinating and coordinating types. The phrase “explicit devices” refers almost exclusively to formal lexico-grammatical markers and to a few semantically based common types of clause binding, for example, attitude verbs. Otherwise, a semantically motivated clause string with an unmarked clause adjunction does not qualify as a sentence.

C l a u s e s t r i n g may be a sentence of one of the types characterised above or a sequence of clauses whose semantic interrelations do not necessarily receive an expression through lexico-grammatical means.

The notion of clause string raises the question of its right boundary. The question of boundaries is typically a tricky one. Most difficult to overcome in research on the syntax of spoken language (Auer 1992: 41), it also presents enough difficulties in syntactic exploration of written texts.

In functional linguistics, a sentence of a written text is defined on the basis of graphological features. As a rule, in texts written in such scripts as Roman or Cyrillic, the beginning of a sentence is marked by an uppercase letter and punctuation marks are generally conceived as signs corresponding to the prosodic

signs of intonation units in oral communication (Chafe 1984). However, the Devanagari script (used by Hindi) does not distinguish between uppercase and lowercase letters. Punctuation in Hindi is sparser than in texts written in the most widely-used letter scripts. Although all punctuation marks of these scripts are also used in Devanagari the sign for a full stop—a short vertical line (।)—often also occurs in syntactic positions, where a European text has a comma or another sign for marking a syntactic unit. Just as such punctuation usage complicates defining the sentence boundaries, so it may produce a yet stronger variation in interclausal parsing. If a clausal string does not fit into any definition of a sentence, this study draws on the semantic interrelations characterising a multiple event sequence as the ground for delimitation of strings. The criterion is, of course, anything but accurate. However, on the whole it seems to work. Propositional relations build a by and large usable foundation for clause-by-clause parsing.

Coming to *c o o r d i n a t i o n* and *s u b o r d i n a t i o n*, the broadest concepts relevant for the study, we find that their opposition is somewhat problematic. Not only do constructions show mixtures of subordination and coordination (Haspelmath 2004: 37), but both types of clause-linkage may overlap in expressing the same meaning. Owing to the semantic and formal multifariousness of interclausal bonds, the status of subordination as a grammatical category in its own right is placed in doubt (Cristofaro 2014; Herlin et al. 2014).

A more persuasive approach views subordination as gradient along the opposition axis whose other pole is coordination. It is an established idea that grammatical categories are gradient (Fanselow et al. 2006; Traugott & Trousdale 2010) and there is no reason why subordination would differ from other syntactic phenomena. For the weaker grade of the hierarchical dependence in a sentence, i. e. not embedded subordination, Foley & Van Valin (1984, ch.6) use the term “cosubordination”.

The array of phrasal and adpositional modifiers to the terms “coordination” and “subordination” speaks clearly of their terminological insufficiency. Compare, for example, “syndetic” and “asyndetic” applied to both the terms (Jucker 1991), “pseudo-coordination” (Ross 2016), “genuine coordination” (Ledgeway 2016: 157) or the term “insubordination”, which appeared at the break of the 21st century and immediately gained strength in linguistic theory (Heine et al. 2016).

For all its imprecision, subordination is an unavoidable concept, convenient as a contrast to “coordination” (among other things). There is no other term that can be deployed when discussing the hierarchical organisation of clausal units within clause strings.

3 Adverbial clauses

Adverbial clauses share specific characteristics with lexical adverbials. Most important is that semantic factors dominate over grammatical in determining the position of adverbial clauses in complex syntactic hierarchies. It has been demonstrated for English that the semantic content of interclausal relations plays the major role in the ordering of main and adverbial clauses (Diessel 2005).

The grade of variation is in direct proportion to the scope of grammatical freedom. The syntactic heterogeneity of clause-linkages with adverbial semantics goes along with a relatively loose connection between adverbial and main clauses (Chafe 1984), or “loose subordination” (Givón 1990). Adverbial clauses are characterised by a “low degree of integration into the matrix clause... and a low degree of interlacing”, as it is the case in the core languages of Europe (Kortmann 1997: 241). We can add also Hindi, insofar as it concerns conjoined clauses with finite verb forms.

The dependence distance between an adverbial and the superordinate clause is variable. Another specificity is that an adverbial clause’s governor may be a complex multiclausal structure of which it itself is a part. Semantic variety of adverbial clauses is combined with diverse syntactic marking. In other words, adverbial clauses show a gradience in the degree of subordination. This makes the task of presenting their relations within long clause strings through a formalised coding system look like a promising method of capturing their basic syntactic features and idiosyncrasies.

Various perspectives on characteristics of adverbials have been summarised by Ernst (2020). Among the subjects of discussion are their location, distribution, correlation between semantic and syntactic factors, etc. Formalisation, especially in the form of a database, makes it easier to capture such issues and the variations in framing various communicative strategies.

4 Data and principles of analysis

The formalisation presented in this article is based on data from written sources. The major part of the processed data stems from essays of modern Hindi writers accessible on the Internet (<http://hindisamay.com>).³ The choice of the genre of essay is due to its closer connection to the reality lived by the author. Its fragmented, predominantly monologic form and incoherent composition show an unconventionally strong personal influence of the author on the form and content

3 As the texts placed on the website had been typed out from books and magazines, diversions from the originals are possible.

(Wang & Jan 2018: 296). Because of these and some other, less relevant, features essay is closer to the spoken language than other literary genres. This is especially strongly felt in the syntactic characteristics of these texts. In them, utterances recorded in the written form are characterised by a comparatively free form of assemblage of clauses and a strong tendency to build long clause sequences, thus encompassing on the average a higher number of mutually related micro-situations than a sentence in a text of a higher literary level. Syntactic and semantic connections within a string can spread over distant clauses, resulting in a portrayal of a multidimensional complex situation beyond the scope of complex and compound sentences.

As of this stage of the study, all respondents have been educated Hindi native speakers from various dialect backgrounds. Students of the School of Language, Mahatma Gandhi Antarrashtriya Hindi VishwaVidyalaya, were assisting in data collection and in preparing a database. Further respondents have given occasional assistance in the former task. Many of the students participated in a six-day training workshop *Hindī miśr vākyō mē ṭaigḍ ḍeṭābes nirmāṇ* (Creation of Tagged Database for Hindi Complex Sentences; 12–17 February 2018). Working together with the students opened to me new vistas and resulted in a number of changes introduced since then in the formalisation system. At this point I heartily thank the students once again. The opportunity to enjoy this collaboration I mostly owe to two of their teachers, Dr Shamim Fatma and Dr Dhanjee Prasad. It was they who came up with the idea of organising such a workshop and did a superb job of bringing it to life. Needless to say, their contribution to the project was not limited to the organisational aspect.

The obtained data bring to light strong structural variations in expressing the same logico-semantic structures. Hypotactic and paratactic clause complexes alternate in their nexus meanings with each other and with sequences of simple sentences. Such alternations can hardly be free considering that the broadest scope of logico-semantic relations obtains at the level of clause complexes (Halliday 1985).

Tectogrammatical representations of clause strings have been developed with a view to creating an interactive database. For representations see 5.2 below.

At the initial stage the formalisation system is based on an analysis of a limited amount of data and is being developed by involving larger data. At the same time, its development serves to elaborate and correct the methods and results of the analysis.

The analysis underlying the formalisation proceeds from biclausal to multi-clausal sequences. The benchmark in the analysis is clausal units which incorporate two clauses and a semantic link between them. Of their two basic types—compound sentence and complex sentence—the latter, comprising a main clause and a subordinate adverbial clause, is more significant for the discussion here.

However, the adverbial semantics of the link between two clauses can in a number of cases be expressed by compounding them, exposing a partial synonymy of hypotactic and paratactic constructions. Moreover, the semantic relation can also exist between juxtaposed sentences which are otherwise syntactically independent (Aguar & Barbosa 2016: 12).

The variety of expression of interclausal adverbial relations reveals their strongly gradient character.

5 Formalisation system

A string is a hierarchical structure. Nevertheless, formalisations are structured horizontally in order to reflect the unfolding of strings along the time axis in spoken and written language forms and in accordance with graphic presentation of a language in a left-to-right script, which is the case with Hindi written in Devanagari.

The preliminary variant of the formalisation presented here is being developed and expanded with new data. Search for an optimal formalisation facilitates the analysis. The procedure accepts the standpoint that strings are conveniently analysed as a linear structure (Longacre 1960).

The tagging procedure follows a major principle of Natural Language Processing, according to which annotations should not alter the underlying corpus in any way; that is, tags are separated from the data using them. This is known as the principle of *stand-off annotation* (Ide & Romary 2004).

Four steps precede the tagging procedure:

- (a) Parsing into clauses
- (b) Disambiguation of interclausal meanings
- (c) Disambiguation of intersentential meanings
- (d) Establishing types of clause combining.

In the schematic presentation all subordinating connectors are placed in the subordinate clauses and coordinating connectors are placed in the matrix clause. This principle is also deployed in the clause-by-clause parsing of the strings. This strategy is tentative, adopted for mere convenience. Clausal affiliation of various connecting devices needs to be explored.

A range of clause binding means corresponds to the semantic heterogeneity of adverbials, from marked subordination through compounding to juxtaposing clauses which are formally independent sentences.

In order to capture the formal variety of interclausal relations five types of brackets are used along with other tags. It is a peculiarity of the system. Other linear presentations of syntactic constructions use two or at most three types: round,

square and angled brackets. So, Langacker (2014) uses three types of brackets and additionally slashes and double slashes alternately for presenting clausal and phrasal structure of sentences linearly and the greater-than sign for establishing interclausal hierarchy in asyndeton. In linear tectogrammatical schemes he combines round and square brackets. Three types are also used in the Transcription System of Spoken Language (Auer et al. 2009). The annotations in the Hindi Discourse Relation Bank use square and curly brackets to mark the ordering of clauses (Umangi et al. 2009). Kothari (2010) uses round brackets to demarcate clauses and square brackets for morphosyntactic tags. Tree-form annotation is superfluous to this review.

The tools used in the tectogrammatical formalisation described below are presented in Table 1. Most syntactic tags are common for linear syntactic annotations. Some tags occur here for the first time. In any case, I never came across them in the literature on the topic. Some semantic tags used here are mine, four are borrowed from English Propbank (Bonial et al. 2015).

5.1 Formalisation tools and tagging guidelines

Table 1 Brackets

<p>Angle brackets < ></p> <p>Tags for each clause (also when a clause is a sentence) are given in angle brackets.</p>
<p>Round brackets ()</p> <p>Tags for all composite (non-simple) sentences are taken in round brackets when a composite sentence occurs in the string for the first time (unlike U-brackets, see below). A left round bracket introduces each compound, complex and combined sentence. The number of round brackets on the right boundary of a tectogrammatical scheme corresponds to the number of conjoined sentences.</p>
<p>Square brackets []</p> <p>Sequences of clauses forming a compound or complex sentence are given in square brackets.</p>
<p>Curly brackets { }</p> <p>All tags with associated lexical semantics, including terminology, are given in curly brackets. These are conjunctions and all functional expressions used to combine clauses, or else are denotations of semantic relations between clauses. The conjunctions and other lexical items with syntactic functions are italicised in the schemes.</p>
<p>U-brackets ⊂ ⊃</p> <p>Tags for composite sentences and connectors participating in more than one structural relation are given in U-brackets, when not manifest in the surface structure. Repeatedly used sentences are indicated by the same subscript letter.</p>

Table 2 Terms, tags and tagging examples

<p>Terms and their tags</p>	<p>Tagging examples</p> <p>The examples here are from the essay of Ajayendranāth Trivedī <i>Bar̥kā jāmun</i> <www.hindisamay.com/contentDetail.aspx?id=5514&pageno=1>. In other cases reference is given to the formalisations in 5.2.</p>
<p>Clause <C> Main clause <MC> Subordinate clause <SC> Complex sentence <CxS></p>	<p>(<CxS {tmp}>[<i>tez havā caltī</i> <SC> Strong wind blew <SC> <i>to kathjāmun jamīn par bich jāte</i> <{to} MC>]) then java.plums ground on spread went<{then}MC>]) '[When] strong wind was blowing the java plums were raining to the ground.'</p>
<p>Compound Sentence Clauses in a source text passage are numerically indexed according to their sequence in the source text passage: C1, C2 ... Cn</p>	<p><CpS> (<CpS {aur} {tmp}> [subah hotī <C1> {aur} [morning came <C1> {and} <i>usī bar̥kā jāmun ke tale dhān kī</i> that.very big java.plum GEN under rice GEN <i>piñī śurū hotī</i> <C2>]) threshing beginning was <C2>]) 'Morning came and under the same big java plum tree began threshing of rice.'</p>
<p>Combined sentence</p>	<p><CdS> See for an example 5.2 (D).</p>
<p>Semantics of the relations between clauses; the conjunctions are italicised in the schemes.</p>	<p>temporal {tmp}, location {loc}, cause {cau}, effect {eff}, {cau-eff}, conditional {cnd}, resultative {res}, purpose {prp}, concessive {cnces}, consecutive {cnsect}, complement {cmpl}, quotation {quot}, consequence {cnseq}, manner {mnr}, restriction {rstr}, comparison {cmpr}, attributive {attr}</p> <p>Meaning concretisation is expressed by a colon (:) before the tag extension, e.g. {tmp: immediate sequence}</p>
<p>When conjunctions are shifted from their initial position to a position inside the clause, they are marked in the scheme by hyphens on both sides of the tag, (as -{yadi}- in the example here).</p>	<p>(<CxS {cnd-res} {yadi-to}[<SC {cnd} -{yadi}>-> <i>apnā patā-ṭhikānā yadi kisī=ko</i> own address-living.place if somebody=DAT <i>batānā ho</i> <{to}MC {res}>]) <i>to bas</i> tell.INF be.CONJ.3SG then just <i>itnā hī kahnā hamāre lie kāfī thā</i> ... that.much only say.INF us for enough was ... 'If we had to explain somebody where we live, it was enough to say just...'</p>
<p>Conjunctions in their usual position at the beginning of a clause</p>	<p><{to} MC>, <{tab} MC> <{agar} SC>, <{jab} SC></p>
<p>Clause order</p>	<p>[<SC> < MC>] or [<MC> <SC>] or [<MC -<SC>->] The latter scheme presents embedded SC.</p>

Terms and their tags	<p>Tagging examples</p> <p>The examples here are from the essay of Ajayendranāth Trivedī <i>Bar̥kā jāmun</i> www.hindisamay.com/contentDetail.aspx?id=5514&pageno=1.</p> <p>In other cases reference is given to the formalisations in 5.2.</p>
<p>Clause valence – the term denotes the number of clauses with which a clause is syntactically connected. If a clause is involved in syntactic relations with more than one clause, it is assigned a valence corresponding to the number of clauses connecting to it. The number of the valence tag is the sequence number of the valence realisation in the string linear structure. The valence tag is parenthesised together with its clause tag and separated from it by a colon.</p>	<p>val1, val2, val3 [<MC:val1><{ki}SC>] See for an example 5.2 (A).</p>
<p>Level – the letter “L” with a following number denotes the level in the hierarchical structure of the strings.</p>	<p>[{<jab/cause}SC-L3><{to}MC-L2:val1>] See for an example 5.2 (C).</p>
<p>An elucidating sentence depicting a situation reflected in the analysed clause(s) is given in double slashes //.</p>	<p>//Hamāre bar̥kā jāmun ko kaun nahī jāntā// //“Who doesn’t know our big java plum tree?”//</p>

- Clause tags are positioned after each clause, i. e. after its number and before punctuation signs.
- If an embedded clause is placed inside the matrix clause, the initial part bears the clause number with the postposed number sign (#), whereas the bare number appears at the end of the clause.
- In order to distinguish main clauses and subordinate clauses of different levels in a string hierarchy, each clause is indexed with its level number, e.g. MC-L2 means “main clause of the 2nd level”.
- The subscript numbers in the data refer to the clauses. Along with the numbering, they mark the right boundaries of the clauses. In the text of the article the subscript numbers are substituted with bracketed numbers.⁴
- Tectogrammatical structures are positioned after the clause strings.

4 In the continuation of the studies the clause count will be used in a quantitative analysis.

5.2 An analysis and a tentative formalisation of clause strings

This part includes an analysis of strings in a passage from an essay by Buddhināth Miśra (Miśra s. a.) *Phūl āe haī kanerō mē* ('Oleanders are Blossoming'). I selected this passage because it includes four clause strings which build an almost uninterrupted sequence and thus present a convenient opportunity to explore the transitional level between sentence syntax and text syntax. As stated above, the style characteristics of the essay genre are, as a rule, fairly close to those of oral narrative, which is the case in the analysed extract. In this part of the text, the key figure is not the author, an uncommon characteristic of essays. This is about a person who found himself in an unknown village and asked for shelter for the night in a house which, like his own house, had a jujube tree in front of it. The host on his return home from the fields found an unknown person there who had introduced himself to the family as their relative. Next morning the host asked the guest about his place of residence and their relationship. Clause strings (A)–(C) are parts of the conversation corresponding to its timeline. Clause string (D) precedes the conversation in narrative time. However, it is placed last in the analysis in order to separate it from the strings which are considered sentences.

- (A) *ham jānnā cāhte haī₁ ki āp kis gāv ke sam̐bandhī haī₂,*
 We know want AUX₁ that you what village of relation are₂
kyōki āj tak hamne kabhī āpko dekhā nahī₃.
 because today until we somewhen you.OBJ saw not₃.

(<CxS{cmpl}{cause}>(<CxS_i{cmpl}>[<MC₁:val1><{ki}SC₂>])(<CxS{cause}>
 [<<CxS_i[<MC₁:val2><{ki}SC₂>>>]<{kyōki}SC₃>]))

'We want to know₁, from what village are you our relation₂,
 because until today we had never seen you₃.'

String (A) is an exemplary case of an adverbial connection between two syntactic units, the first of which is a biclausal complex sentence and the second a clause carrying the adverbial meaning. In traditional terminology it is a complex sentence with two subordinate clauses, whereas the deeper structural relations need further comments. Both subordinate clauses depend on the same main clause. The dependencies within the string are asymmetrical. Clause (2), a complement of the verb *jānnā* 'to know', is embedded in clause (1), the connection being marked by the complementiser *ki* (from Persian, lit. 'who', 'which', 'why'), approximately corresponding to 'that' but, unlike the English conjunction, with an interrogative "pedigree" and with a broader set of subsidiary syntactic functions. The verb 'to know' is the immediate governor of clause (2). The clause could have the pronoun

yah ‘it’, ‘this’ in the position before the infinitive ‘know’ serving as a cataphoric prop for the subordinate clause (2).

The verb ‘to know’ governing clause (2) is, in its turn, the object complement of the finite predicate ‘want’. This verb is sub-classified as a verb of mental attitude, a semantic sub-category with a range of idiosyncratic features, the capacity to take a clausal complement along with one or two nominal arguments (Pearson 2021). The nominal arguments may be of predicative nature: infinite verbal forms extending the combining capacity of the clause. The bi-verbal character of the finite VP in the main clause determines its double syntactic valency, that is its capacity to subordinate two clauses — (2) and (3). The former has one infinite verb form as its governor, whereas the syntactic governor of the latter is the whole bi-verbal VP. The semantic scope of the clause (3) dependency is still broader—it is the whole complex sentence. The connecting device *kyōki* ‘because’ (*kyō-ki* ‘why-[subordinator] that’) is a fully-fledged conjunction, as is to be expected in adverbial clauses (cf. Lehmann 1988). The dependency distance of the adverbial clause is longer than that of the object clause. This also points to its rather loose formal connection within the string, which is obviously compensated for semantically.

A more interesting, although predictable, aspect of the syntactic asymmetry is the inverse proportionality between the strength of the syntactic connection and the semantic dependency scope: the object complement clause, which is firmly embedded in the main clause, is semantically connected to its predicate, whereas the juxtaposed adverbial clause is semantically linked to the whole preceding complex sentence.

In the following excerpt the structures incorporating clauses (4) to (7) can be ignored in the discussion (hence they are marked with double slashes on both sides). They help us to understand the context of the sequence (8) to (11) which builds to a longer string—string (B).

//*Yah sunkar atithi muskurāe aur bole₄ – hamāre āpke bīc Bādrāyaṅ sambamdh hai₅.*
 this having.heard guest smiled and said₄ – me you between Bādrāyaṅ’s⁵ relation is₅.
Jaise āpke darvāze par ber (badrī phal) kā peṛ hai₆, vaise hī mere darvāze par bhī ber
 Like your door at jujube (badrī fruit) of tree is₆, so exactly my door at too jujube
kā peṛ hai₇.
 GEN tree is₇.//

(B) *cūki rāt ho rahī thī₈ aur pūrā gāv mere lie aparicit thā₉, islie*
 as night falling was₈ and whole village me for unknown was₉, therefore
maīne yah sambamdh nikālā₁₀ ki kuch to samāntā hai mere āp=mē₁₁.
 I this relation thought.up₁₀ that some at.least similarity is me you=in₁₁.

5 The idiomatic expression ‘Bādarāyaṅ’s relation’ means ‘a far-fetched, just a nominal relation’.

(<CxS {cause-eff} {cml}>[(<CxS {cause-eff}>[<{cũki} SC₈> <{aur}> <{cũki}> SC₉> <{islie} MC_{10i}:val1>]](<CxS {cml}>[< <MC_{10i}>:val2> <{ki} SC₁₁>]]))

//‘Having heard this the guest smiled and said₄, “We are distantly related₅. Just as there is a jujube tree in front of your house₆, so there is also a jujube tree in front of my house₇.”’//

(B) ‘As night was falling₈ and I didn’t know anybody in the village₉, I therefore thought up this relationship₁₀, so that at least there is something in common between us₁₁.’

String (B) is a complex sentence. It consists of four units with subordinate and coordinate interclausal relations. Although it includes both relations types it is not considered a combined sentence because the coordinate link is located not on the highest level of the string. Subordinate relations prevail in (B), whereas the only coordinate bond expressed by the conjunction *aur* ‘and’ connects two collateral subordinate clauses (8) and (9). They build a sequence and share the causal conjunction *cũki* ‘as’, which connects them to the nucleus⁶ of the string—the main clause (10) introduced by the adverbial connector *islie* ‘therefore’.

All clauses are introduced by connectors: three of them by conjunctions—*cũki* ‘as’ (8), *cũki* ‘as’... *aur* ‘and’ (9), *ki* ‘that’ (11)—and one by the adverbial connector *islie* ‘therefore’ (10). The conjunctions *cũki* and *ki* are borrowings from Persian. Both are functional derivations of interrogative-relative pronouns. The former is a combination of *cũ* ‘where’, ‘why’, ‘how’ and *ki* (< ‘who’, ‘what’), which can introduce almost any subordinate clause and which appears as a separate conjunction in the final clause of the sentence. The coordinating conjunction *aur* (< Skt. *apara-* ‘other’, see Turner 1966: 20) conjoins two causal subordinate clauses. The connector *islie* gives the clause the meaning of effect and marks it as the main clause. It consists of the oblique form of the deictic pronoun *is* ‘it’ and the deverbal marker *lie* (< *lenā* ‘to take’), the meaning on the whole being approximately ‘this taken’. The pronominal anaphor refers to the situation rendered by the preceding proposition of the subordinate constituent. Here we have a case where each part of the cause-effect relation is marked with its own device making explicit the meaning of the interclausal link. However, this tandem is not absolutely necessary and some language purists even consider the double marking stylistic negligence. Each of the markers alone serves the same semantic effect, but the syntactic connection is then realised differently in each case. If the causal

6 The term ‘nucleus’ is understood here as the sentence core, not in the sense in which it is used by Longacre (1970; 1985) and in subsequent studies adopting the same or a similar approach.

marker *cūki* has zero correspondence in the effect clause, which is the main clause, the clauses build a sentence. If there is explicit marking only in the effect clause, the syntactic integrity is weakened and both clauses may be considered separate sentences. According to my preliminary observations, the latter way of mapping the causal relation is the most frequent one among several marked types of Hindi multiclausal causal constructions.

The second valency of the main (effect) clause is induced by the object of the finite verb: *sambamdh* ‘relation’. It is determined by the descriptive relative clause with the conjunction *ki* ‘that’ specifying the noun and correlated with the preposed deictic pronominal attribute *yah* ‘this’.

- (C) *yah sunkar sabne zor kā thahākā lagāyā aur atithi ko*
 This having.heard all strength of laughter laid.out and guest OBJ
sādar vidā kiyā, yah kahkar₁₂ ki₁₃ jab sambamdh sthāpit
 respectfully see.off did, this having.said₁₂ that_{13#} when relation established
ho hī gayā hai₁₄, to_{15#} jab kabhī idhar se guzrē₁₆,
 be really gone is₁₄, then_{15#} when sometime here through would.pass.by₁₆,
yahī rātri-vīsrām karē_{15;13}.
 here.only night-rest would.do_{15;13}.

(<CxS {cmpl} {cause-eff} {tmp}> [<MC₁₂-L4<< {ki} SC₁₃-L3>] (<CxS {cause-eff}> [{jab/cause} SC₁₄-L2<< {to} MC₁₅-L3:val1>] (CxS {tmp}> [< {jab} SC₁₆-L1> <<MC₁₅-L3:val2 >>]))))

‘Having heard this [they] all guffawed and bade the guest a respectful farewell saying₁₂ that_{13#}, “As the relationship has been established₁₄, so_{15#} whenever you pass by here₁₆, stay only here [in this house] for the night₁₅”₁₃.’

String C is a four-level complex sentence with stepwise subordination. The fourth level main clause (12) joins the postposed subordinate clause (13–16) due to the valency of a verb of saying (used in the converb form *kahkar* ‘having said’). With regard to its syntactic structure the subordinate clause is a tripartite complex sentence conjoined to the fourth-level main clause by the complementiser *ki*. Its main clause (15), located on the third level, has two valences. One of them is filled by a preposed clause (14) and the other one by an embedded clause (16). Both clauses are introduced by the conjunction *jab* ‘when’ but they differ in relation to the main clause (15) with regard to both semantics and syntax. In the subordinate clause (14) *jab* is used as a cause marker and correlated with the conjunction *to* introducing the main clause. (The conjunction is marked as a distant clause part.) In Hindi, markers of main clauses are generally more significant in establishing interclausal relations than markers of subordinate

clauses; as a rule, these can easily be omitted. The embedded subordinate clause (16) uses the time conjunction in its basic meaning of time. This marker has no correlative in the main clause of the complex sentence, as an embedded clause does not need any further tie to the matrix clause. The connection is clear from the clause location and the subordinating conjunction *jab*. Thus, two subordinate clauses occupy different levels: (14) is on the second level, whereas (16) demonstrating the strongest bond with the main clause (15) is on the first level. The positions of the subordinate clauses (14) and (16) in the syntactic hierarchy of the string may be correlated with their semantics. It has been suggested that causal clauses' connection to the main clause is the loosest among all other semantic types of adverbial clauses (Diessel & Hetterle 2011). It is conceivable that adverbial clauses with the basic adverbial meanings of time and space, which are expressed in clauses by lexical adverbial modifiers, enjoy a closer relation to the main clause.

- (D) *saṃyog se, us din khetō mē kām zyādā thā₁₇, aur koī sarkārī*
 Chance by that day fields in work much was₁₇ and any government
naukrī tō thī nahī₁₈ ki kām pūrā ho na ho₁₉, gharī
 job indeed was not₁₈ that work finished be.SBJV not be.SBJV₁₉ clock
dēkhkar log ghar bhāgē₂₀. so, unke āte-āte kāfī der
 look.CVB people home run₂₀. So their coming-coming enough tardiness
ho gāī₂₁.
 be went₂₁.

String{cause-eff}(<CdS>([<CpS{aur}>[<C₁₇-L3>(<{aur}>CxS{attr}>[<MC₁₈-L3>
 <{ki}>SC₂₀-L2>])<CxS{cnces}>(<SC₁₉-L1><MC₂₀-L2>]<{so/eff}>MC₂₁-L4>))))

‘By chance that day there was much work in the fields₁₇, and it wasn’t a government job₁₈, in which whether the work is finished or not₁₉, the people look at the clock and run home₂₀. So when they [the menfolk of the family] came it was [already] late₂₁.’

Unlike the clause strings (A)–(C), string (D) does not fit into any definition of a sentence. In establishing the boundaries of clause strings which do not fit into any definition of a sentence this study draws on the semantic interrelations characterising a multiple event sequence as the ground for delimitation of strings.

(D) is a five-part string which includes two syntactic segments graphically framed as sentences. The concluding simple sentence is separated from the preceding combined sentence of four clauses by the Devanagari full stop sign. Nevertheless, it belongs semantically to the same string, presenting the final event of the whole situation and building a clear logical transition to the following complex

situation. The fifth constituent is introduced by the pronominal conjunction *so* ‘so’, which may introduce a new graphical sentence, as in this case. In similar contexts it may be separated from the previous part of a sentence by a comma or not separated at all (the same as in English).⁷ This variation in punctuation shows the possible transitions between syntactic independence and a tighter formal bond to the semantic correlate within a clausal string. This kind of alternating syntactic framing seems to be typical of consecutive clauses.

The string has a four-level hierarchical structure. At its highest point (fourth level) is the final clause (21), which is related to the whole preceding clause sequence as effect to cause. The third level is formed by a simple (17) and a complex (18–20) sentence connected by the coordinating conjunction *aur*. The complex sentence exhibits stepwise subordination forming the second level: the rightward valency of the main clause (18) is filled by a continuative relative clause (19–20), which depends on the main clause subject and expands its content. Finally, the unmarked conditional concessive relation between the subordinate (19) and the main (20) clause is located on the lowest (first) level.

6 In place of conclusion. Future directions of the data analysis and formalisation: a view

The intended study needs a variety of numerically reliable information in the form of a database. Such a database built on written and oral sources will be useful for relational research in Hindi linguistics and in typology. The generally estimated workable database volume is set at 5 million words with a desirable (but in our case unrealistic) expansion up to a sample of 20 million words (Matthiessen 2002: 252). Currently, the primal data universe comprises somewhat more than 100 data units.

The perspective on the general issues needs to be broadened to cover data elicited not only from written texts but also oral discourse. Inclusion in large quantities of discourse material in data will shed a new light on Hindi complex communication structures and build a solid basis for exploring the cognitive characteristics of its syntactic complexity.

Complex syntax is tightly related to variation. The more complex the situation, the more propositions participate in its portrayal. Accordingly, the number of clauses framing them is higher which allows a higher variation in their linkage. This in turn leads to an increase in variation in perspectives on the situation. Variance in

7 Compare: *gāv ke sāre log aśikṣit haī, so mazdūrī par zimdā rahte haī, [...]* ‘All villagers are uneducated, so they live from their manual labour, [...]’ (Simh 2018) and *ārthik sthiti thīk nahī so mazdūrī kī majbūrī hai*. ‘[They are] not well financially so manual work is a necessity.’ (“Dūr nahī...”) / *Dainik Bhāskar*.

the interplay between meanings and formal tools brings up the issue of “choice” as propounded by Halliday (Halliday & Matthiessen 2014), in the sense of a speaker’s primary choice of meaning and its correlation with multiple forms.

Processing original Hindi texts aims, among other things, at compiling count data on the total number of clauses, in particular, of clauses with adverbial meanings, as well as on variations in them. These have to do with the semantics of the interclausal bonds. Further variations have to do with the information structure and can impact the placing of conjunctions. The feature [+/- focused] can also be responsible for the clausal order. Such phenomena as foregroundedness and backgroundedness also come into play.

A productive and, in terms of text generation, promising analysis would involve contrasting bipartite complex components. Two types of opposition between ways of framing interclausal adverbial relations should be considered: 1. various adverbial relations within the same clause string and 2. same adverbial relations in different clause strings. The major analytic perspective would start from the semantic vantage point and integrate contextually conditioned choices of structural-syntactic and auxiliary lexico-grammatical means of expression.

In terms of information structure, the relation between parts of a minimal—bipartite—hypotactic string is interpreted in the following way: a governing/main clause profiles a process that is foregrounded, while the sub-clause profiles a backgrounded process, be it causal, conditional or circumstantial. Various types of conceptual and functional subordination underlie the subordinating structures, which are shaped with the help of grammatical and lexico-grammatical means.

The exploration is underpinned by the general observation that some types of intersentential relations may be explicated formally or else be expressed without using any special formal means. The latter type of connection results from the order of sentences and the semantics of their key terms; the former type uses semantically specific markers.


Complex sentences in Hindi belong to the marked type. The markers in a number of cases may be dependent on the meaning of the interclausal relation modified by the content of the protasis and apodosis. A large formalised database should provide a reliable foundation for establishing the marking rate, positional and scope variations of the markers as well as semantic and logical grounds of their overlapping. Further tasks deal with information structure and communicative functions of the formal varieties (structuring types of composite sentences) and amorphous types of clausal linkage.

The part of the database planned to formalise oral data offers much more significant challenges on each step of the process—from collecting data to parsing the speech stream and stratifying the relations between the syntactic segments. It is axiomatic that dia- and polylogues present the most challenging problem. This

lies in the unpredictable turns a syntactic trajectory may take at any moment in the time resulting in clausal structures within the speaker's turn domain or across the speakers' turns (Auer 2005).

Higher syntax in Hindi still awaits a thorough exploration of its written and especially oral discourse incarnations. The formalised corpus is expected to eventually provide a substantial basis for an insightful conceptualisation of clause ordering, text structure and the varied relationships between semantics of inter-clausal connections and the language means marking them. Enhanced data and new foci, especially a cognitive one, will bring a new dimension also to typological research involving Hindi. The results of the future work can be used to enhance modelling of probabilistic sequences of clauses and sentences, also in descriptive, analytic and computational linguistics.

ORCID®

Tatiana Oranskaia  <https://orcid.org/0000-0001-7207-9190>

Bibliography

- Aguiar, Joana & Pilar Barbosa 2016. "Establishing Causal Relations in European Portuguese: Coordination and Subordination in a Stratified Corpus", in: Pratas, Fernanda, Sandra Pereira and Clara Pinto (eds): *Coordination and Subordination: Form and Meaning—Selected Papers from CSI Lisbon 2014*. Newcastle upon Tyne: Cambridge Scholars Publishing, 1–22.
- Ananthanarayana, H. S. 1996. "Relative Clause in Indo-Aryan", in: Swarajya, Lakshmi & Aditi Mukherjee (eds): *Word Order in Indian Languages*. Hyderabad: Centre of Advanced Study in Linguistics, Osmania University, Hyderabad; Booklinks Corporation, 31–46.
- Auer, Peter 1992. "The Neverending Sentence: Rightward Expansion in Spoken Language", in: Kontra, Miklós (ed.): *Studies in Spoken Languages: English, German, Finno-Ugric*. (Sonderdrucke aus der Albert-Ludwigs-Universität Freiburg. Budapest: Hungarian Academy of Sciences, Linguistics Inst.), 41–59.
- Auer, Peter 2005. "Projection in Interaction and Projection in Grammar", *Text & Talk. An Interdisciplinary Journal of Language, Discourse & Communication Studies* 25,1: 7–36. (<<https://doi.org/10.1515/text.2005.25.1.7>>, accessed: September 6, 2020).
- Auer, Peter et al. 2009. "Gesprächsanalytisches Transkriptionssystem 2 (GAT 2)", *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 10: 353–402 (<<http://www.gespraechsforschung-ozs.de>>, accessed: February 1, 2021).

- Bhatt, Rajesh 2003. “Locality in Correlatives”, *Natural Language & Linguistic Theory* 21,3: 485–541. (<<http://www.jstor.org/stable/4048041>>, accessed: March 5, 2021).
- Bhat Riyaz Ahmad et al. 2017. “The Hindi/Urdu Treebank Project”, in: Ide, Nancy & James Pustejovsky (eds): *Handbook of Linguistic Annotation*. Dordrecht: Springer (<https://doi.org/10.1007/978-94-024-0881-2_24>, accessed: September 12, 2020)
- Bonial, Claire et al. 2015. *English PropBank Annotation Guidelines*. Center for Computational Language and Education, Research Institute of Cognitive Science, University of Colorado at Boulder. (<<http://verbs.colorado.edu/propbank/EPB-Annotation-Guidelines.pdf>>, accessed: May 5, 2019).
- Chafe, Wallace 1984. “How People Use Adverbial Clauses”, in: *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 437–449. (<<http://journals.linguisticsociety.org/proceedings/index.php/BLS/article/viewFile/1936/1708>>, accessed: November 13, 2019).
- Christofaro, Sonia 2014. “Is There Really a Syntactic Category of Subordination?”, in: Visapää, Laura, Jyrki Kalliokoski & Helena Sorva (eds): *Contexts of Subordination: Cognitive, Typological and Discourse Perspectives*. (Pragmatics & Beyond New Series, vol. 249.) Amsterdam; Philadelphia: John Benjamins Publishing Company, 73–91.
- Dayal, Veneeta 1996. *Locality in WH Quantification: Questions and Relative Clauses in Hindi*. (Studies in Linguistics and Philosophy 62.) Dordrecht; London: Kluwer.
- Diessel, Holger 2005. “Competing Motivations for the Ordering of Main and Adverbial Clauses”, *Linguistics* 43,3: 449–470. (<<https://doi.org/10.1515/ling.2005.43.3.449>>, accessed: February 5, 2018).
- Diessel, Holger & Katja Hetterle 2011. “Causal Clauses: A Crosslinguistic Investigation of Their Structure, Meaning, and Use”, in: Siemund, Peter (ed.): *Linguistic Universals and Language Variation*. (Trends in Linguistics. Studies and Monographs 231.), 23–54.
- “Dūr nahī – rājdhānī Patnā se Sipārā pul ke nīce roz”, *Dainik Bhāskar* (date cannot be retrieved) (<<https://www.bhaskar.com/latest-patna-news-043521-2483803.html>>, accessed: March 29, 2021).
- Eifring, Halvor 1995 [1993]. *Clause Combination in Chinese*. (Sinica Leidensia 32.) Leiden; New Yourk; Köln: E. J. Brill.
- Ernst, Thomas 2020. “The Syntax of Adverbials”, *Annual Review of Linguistics* 6,1: 89–109 (<<https://www.annualreviews.org/doi/10.1146/annurev-linguistics-011619-030334>> accessed: March 15, 2021).

- Fanselow, Gisbert, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds) 2006. *Gradience in Grammar: Generative Perspectives*. Oxford; New York: Oxford University Press.
- Foley, William A. & Van Valin, Robert D., Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Givón, Talmy 1990. *Syntax. A Functional-Typological Introduction*, vol. II. Amsterdam: John Benjamins.
- Haiman, John & Sandra A. Thompson (eds) 1988. *Clause Combining in Grammar and Discourse*. (Typological Studies in Language 18.) Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Halliday, Michael A. K. 1985. *An Introduction to Functional Grammar*. London: Arnold.
- Halliday, Michael A. K. & Christian Matthiessen 2014 [1985; 1994; 2004]. *Halliday's Introduction to Functional Grammar*. New York: Routledge.
- Haspelmath, Martin 2004. “Coordinating Constructions: An Overview”, in: Haspelmath, Martin (ed.): *Coordinating Constructions*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 3–39.
- Heine, Bernd, Gunther Kaltenböck & Tania Kuteva 2016. “On Insubordination and Cooptation”, in: Watanabe, Honoré & Nicholas Evans (eds): *Insubordination*. Amsterdam; Philadelphia: John Benjamins, 39–64.
- Herlin, Ilona, Jyrki Kalliokoski, and Lana Visapää 2014. “Introduction”, in: Visapää, Laura, Jyrki Kalliokoski, and Helena Sorva (eds): *Contexts of Subordination: Cognitive, Typological and Discourse Perspectives* (Pragmatics & Beyond New Series, vol. 249.) Amsterdam; Philadelphia: John Benjamins Publishing Company, 1–16.
- Ide, Nancy M. & Laurent Romary 2004. “International Standard for a Linguistic Annotation Framework”, *Natural Language Engineering* 10,3–4: 211–225. (<<https://doi.org/10.1017/S135132490400350X>>, accessed: January 9, 2017).
- Jucker, Andreas H. 1991. “Between Hypotaxis and Parataxis. Clauses of Reason in Ancene Wisse”, in: Kastovsky, Dieter (ed.): *Historical English Syntax*. (Topics in English Linguistics.) Berlin; New York: Mouton de Gruyter, 203–220.
- Kachru, Yamuna 1978. “On Relative Clause Formation in Hindi-Urdu”, *Linguistics* 207: 5–26.
- Kortmann, Bernd 1997. *Adverbial Subordination: A Typology and History of Adverbial Subordinators Based on European Languages*. (Empirical Approaches to Language Typology 18.) Berlin; New York: Mouton de Gruyter.
- Kothari, Anubha 2010. *Processing Constraints and Word Order Variation in Hindi Relative Clauses*. Stanford University, ProQuest Dissertations Publishing, 28168220. (<<http://purl.stanford.edu/vh402hk4761>>, accessed: October 17, 2020).

- Langacker, Robert 2010 [1998]. “Conceptualization, Symbolization and Grammar”, *International Journal of Cognitive Linguistics* 1: 31–63.
- Langacker, Ronald W. 2014. “Subordination in a Dynamic Account of Grammar”, in: Visapää, Laura, Jyrki Kalliokoski & Helena Sorva (eds). *Contexts of Subordination: Cognitive, Typological and Discourse Perspectives* (Pragmatics & Beyond New Series, vol. 249). Amsterdam; Philadelphia: John Benjamins Publishing Company, 17–72.
- Ledgeway, Adam 2016. “From Coordination to Subordination: the Grammaticalisation of Progressive and Andative Aspect in the Dialects of Salento”, in: Pratas, Fernanda, Sandra Pereira & Clara Pinto (eds): *Coordination and Subordination. Form and Meaning—Selected Papers from CSI Lisbon 2014*. Newcastle upon Tyne: Cambridge Scholars Publishing, 157–184.
- Lehmann, Christian 1985. “On Grammatical Relationality”, *Folia Linguistica* 19: 67–109. (<<https://www.degruyter.com/document/doi/10.1515/flin.1985.19.1-2.67/html>>, accessed: October 17, 2020).
- Lehmann, Christian 1988. “Towards a Typology of Clause-Linkage”, in: Haimann, John & Sandra Thompson (eds): *Clause Combining in Grammar and Discourse*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 181–225. (<<https://www.christianlehmann.eu/publ/linkage.pdf>>, accessed: July 15, 2020)
- Longacre, Robert E. 1960. “String Constituent Analysis”, *Language* 36,1(1): 63–88.
- Longacre, Robert E. 1970. “Sentence Structure as a Statement Calculus”, *Language* 46,4: 783–815.
- Longacre, Robert E. 2007 [1985]. “Sentences as Combinations of Clauses”, in: Shopen, Timothy (ed.): *Language Typology and Syntactic Description*. Vol. II: *Complex Constructions*. Cambridge: Cambridge University Press.
- Matthiessen, Christian M. I.M. 2002. “Combining Clauses into Clause Complexes: A Multi-faceted View”, in: Bybee, Joan L. & Michael Noonan (eds): *Complex Sentences in Grammar and Discourse: Essays in Honor of Sandra A. Thompson*. Amsterdam; Philadelphia: John Benjamins Publishing Company, 235–320.
- Miśra, Buddhināth s. a. *Phūl āe haī kanerō mē* (Oleanders are Blossoming), (<<http://www.hindisamay.com>>, accessed: January 20, 2018).
- Oranskaya, Tatyana I. 2005. “Conditional Constructions in Hindi”, in: Xrakovskij, Victor S. (ed.): *Typology of Conditional Constructions*. (LINCOM Studies in Theoretical Linguistics 25.) Munich: Lincom Europa, 218–245.
- Pearson, Hazel 2021. “Attitude Verbs”, in: Gutzmann, Daniel et al. (eds): *Companion to Semantics*. Hoboken: Wiley Blackwell, 2021. (<<https://semanticsarchive.net/Archive/mIwMTc1Z/Attitude%20verbs%20Pearson%20to%20appear.pdf>>, accessed: January 30, 2021).

- Ricca, Davide 2010. “Adverbs”, in: Baldi, Philip & Pierluigi Cuzzolin (eds): *New Perspectives in Historical Latin Syntax*. Vol. 2. *Constituent Syntax: Adverbial Phrases, Adverbs, Mood, Tense*. (Trends in Linguistics. Studies and Monographs 180,2.) Berlin; New York: De Gruyter Mouton, 109–191.
- Ross, Daniel 2016. “Between Coordination and Subordination: Typological, Structural and Diachronic Perspectives on Pseudocoordination”, in: Pratas, Fernanda, Sandra Pereira & Clara Pinto (eds): *Coordination and Subordination. Form and Meaning—Selected Papers from CSI Lisbon 2014*. Newcastle upon Tyne: Cambridge Scholars Publishing, 209–244.
- Sharma, Ghanshyam 2010. “On Hindi Conditionals”, in: *Annual Review of South Asian Languages and Linguistics*, 107–134.
- Sharma, Ghanshyam 2012. “On the Role of Protases in Conditional Statements: Some Evidence from Hindi”, in: *Annual Review of South Asian Languages and Linguistics 2011*, 49–78.
- Shopen, Timothy (ed.) 2007 [1985]. *Language Typology and Syntactic Description*. Vol. 2: *Complex Constructions*. Cambridge: Cambridge University Press.
- Simh, Tavlin 23 Jan 2018. Dalitō kī āvāz bhī sunnī paregī. *Amar Ujālā* (<<https://www.amarujala.com/columns/opinion/the-voice-of-the-dalits-will-also-be-heard>>, accessed: March 29, 2021).
- Thompson, Sandra A., Robert E. Longacre & Shin Ja J. Hwang 2007 [1985]. “Adverbial Clauses”, in: Shopen, Timothy (ed.): *Language Typology and Syntactic Description*. Vol. 2. *Complex Constructions*. Cambridge: Cambridge University Press, 237–300.
- Traugott, Elizabeth C. & Graeme Trousdale (eds) 2010. *Gradience, Gradualness & Grammaticalization*. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Umangi, Oza et al. 2009. “Experiments with Annotating Discourse Relations in the Hindi Discourse Relation Bank”. (<<https://www.researchgate.net/publication/228574312>> accessed: April 11, 2020).
- Wang, Yaqin & Jianwei Jan 2018. “A Quantitative Analysis on a Literary Genre Essay’s Syntactic Features”, in: Jiang, Jingyang & Haitao Liu (eds): *Quantitative Analysis of Dependency Structures*. (Quantitative Linguistics 72.) Berlin; Boston: De Gruyter Mouton, 295–314. (<https://www.researchgate.net/profile/Yaqin-Wang-3/publication/328147823_A_Quantitative_Analysis_on_a_Literary_Genre_Essay%27s_Syntactic_Features/links/5d6a0f20a6fdcc547d6d195e/A-Quantitative-Analysis-on-a-Literary-Genre-Essays-Syntactic-Features.pdf>, accessed: October 8, 2018)

PART III

VARIATION ISSUES IN HINDI TEACHING

Rama Kant Agnihotri 

Multilinguality and Language Teaching

रमा कान्त अग्निहोत्री

बहुभाषिता एवं भाषा शिक्षण

Abstract In this paper I argue that it is essential to be sensitive to language for language teaching. It is important to address such questions as: What is the nature of language? What is its structure like? How does it change? These are some of the questions about which there must be a minimal understanding. The question of language acquisition by children is also associated with these questions. This does not, however, mean that every teacher must also be a linguist. It only means that any person who claims to intervene in education must know something about language; what kind of education or knowledge can we have without language? The real nature of language is “multilinguality” and not a “homogeneous language”; diversity rather than uniformity is the defining feature of language. One more thing: issues of language teaching must be addressed. But what is special about Hindi language teaching? What is so special about English language teaching? English Language Teaching (ELT) is a big industry. It seems in the absence of the British Empire we have to have the empire of English. Our response to this socio-political pressure is not to start manufacturing as it were another brand of soap or oil. The response should be to locate the teaching of languages in the context of multilinguality.

Keywords multilinguality, language teaching, nature of language, phonology, morphology, syntax.

सारांश इस लेख में मेरा यह प्रयास है कि हम यह बात समझ सकें कि भाषा शिक्षण के लिए भाषा के प्रति सामान्य रूप से संवेदनशील होना आवश्यक है। भाषा की क्या प्रकृति है, भाषा की संरचना कैसी होती है, भाषा किस तरह बदलती है आदि कुछ ऐसे प्रश्न हैं जिनके बारे में कुछ समझ होनी ज़रूरी है। इन प्रश्नों से यह प्रश्न जुड़ा है कि बच्चे भाषा सीखते कैसे हैं। इसका अर्थ यह कदापि नहीं है कि हर भाषा शिक्षक को भाषा वैज्ञानिक बनना होगा। केवल इतना ही कि वह हर व्यक्ति जिसका शिक्षा से कुछ भी लेना देना है उसे भाषा के बारे में कुछ समझ होनी चाहिए; भाषा के बिना भला कैसी शिक्षा और कैसा ज्ञान। भाषा की वास्तविक प्रकृति बहुभाषिता है न कि एकरूपता; विविधता

है न कि समांगता। एक अन्य बात : भाषा शिक्षण तो समझ में आता है पर यह हिंदी भाषा शिक्षण अलग से क्या? बहुत जोर रहा है अंग्रेज़ी भाषा शिक्षण का; आज भी है। बहुत विशाल उद्योग है। अंग्रेज़ी साम्राज्य नहीं तो अंग्रेज़ी भाषा का साम्राज्य ही सही। इसका उत्तर यह नहीं कि हम एक और साबुन या तेल की तरह एक नया उद्योग शुरू कर दें। इसका उत्तर यह है कि बहुभाषिता के संदर्भ में भाषा शिक्षण को समझें।

मुख्य शब्द – बहुभाषिता, भाषा शिक्षण, भाषा प्रकृति, ध्वनि संरचना, शब्द संरचना, वाक्य संरचना।

१ भाषा

भाषा को एक उचित परिभाषा देना अपने आप में एक विशेष समस्या है। आम लोग, भाषा वैज्ञानिक एवं दार्शनिक भाषा को समझने का सदियों से प्रयास कर रहे हैं। आम लोग भाषा को केवल एक संप्रेषण का माध्यम भर समझते हैं। कई भाषा वैज्ञानिक भाषा को वाक्यगत संरचना एवं एक शब्दकोष का नियमबद्ध मिलन मानते हैं। अनेक दार्शनिकों को तो लगा कि दर्शन को समझना अंततः भाषा को ही समझना है। असल में भाषा की वास्तविक प्रकृति बहुभाषिता है और बहुभाषिता इंसान होने की पहचान है (Agnihotri 2009)। चाहे वह भाषा हिंदी हो या फिर कोई अन्य भाषा, भाषा शिक्षण निरर्थक ही सिद्ध होगा जब तक हम भाषा की सही प्रकृति एवं संरचना को नहीं समझ लेते। इसलिए यह कोई हैरानी की बात नहीं कि सालों भाषा शिक्षण होने के बाद भी १२ वर्ष की शिक्षा के बाद भी विद्यार्थी न तो भाषा में ही कोई दक्षता हासिल कर पाते हैं और न ही उनकी बौद्धिक क्षमता का कोई विशेष विकास हो पाता है। भाषा शिक्षण की कक्षा में बहुभाषिता एक बहुत समृद्ध स्रोत होता है। बहुभाषिता का सरलता से भाषा शिक्षण एवं बौद्धिक विकास के लिए उपयोग किया जा सकता है। इस प्रक्रिया से भाषा के प्रति संवेदना भी बढ़ेगी और विद्यार्थी कक्षा में एक दूसरे की भाषा की इज्जत करना भी सीखेंगे। हिंदी (या किसी भी और भाषा) की किसी भी कक्षा में कई अन्य भाषाएँ मौजूद रहती हैं। लेकिन उन पर कोई ध्यान नहीं दिया जाता, चाहे वह उत्तरी भारत की हो या दक्षिणी या फिर कोई विदेशी भाषा हो। केवल 'शुद्ध भाषा' पर जोर रहता है। भाषा और बहुभाषिता में क्या रिश्ता है और कैसे एक सार्वभौमिक व्याकरण भाषाओं के आपसी संवाद में सहयोग देता है इस बात पर गौर नहीं किया जाता। जब तक भाषा पर संवाद शिक्षण-प्रशिक्षण का एक आवश्यक अंग नहीं बन जाता, तब तक भाषा शिक्षण में कोई मूल बदलाव संभव नहीं।

इसमें कोई दो राय नहीं कि एक स्तर पर भाषा एक नियमबद्ध व्यवस्था है जिसके सिद्धांत ध्वनि के, शब्दों के, वाक्य के या फिर संवाद के – यानी हर स्तर पर – कार्य करते हैं। सार्वभौमिक स्तर के ये सिद्धांत आम तौर पर कई या अनेक भाषाओं के लिए सामान्य होते हैं। उदाहरण के लिए ध्वनि के स्तर पर कुछ साधारण सिद्धांत देखें। हिंदी को ही लें। हिंदी में 'ड' की ध्वनि शब्द के शुरू में नहीं आती। वास्तव में यह बात कई भारतीय भाषाओं के बारे में भी सच है।

यही बात अंग्रेज़ी के बारे में भी सच है। आप अंग्रेज़ी में किङ (king), सिङ (sing) आदि तो कह सकते हैं लेकिन *डकि या *डसि नहीं। इसी तरह आप हिंदी में 'लडका' तो कह सकते हैं पर /ड/ ध्वनि का शब्द के शुरू में प्रयोग नहीं कर सकते। *डकाल जैसा कोई शब्द हिंदी में संभव नहीं। संसार की अधिकतर भाषाओं की शब्दावली में ध्वनि के स्तर पर व्यंजन व स्वर ध्वनि का क्रम रहता है (Hyman 2008: 101)। यानी (व्यंजन)-स्वर-(व्यंजन) आदि। इस संरचना का अर्थ यह हुआ कि स्वर के बिना शब्द नहीं बन सकता; व्यंजन के बिना बन सकता है। /आ/ हिंदी की ध्वनि है। लेकिन यह हिंदी का एक शब्द भी है और वाक्य भी। 'आ' का अर्थ है केवल 'तू आ'।

बहुत ही कम शब्द ऐसे होंगे जिनमें दो या दो से अधिक व्यंजन साथ-साथ सुनाई दे रहे हों। कुछ शब्दों में ऐसे व्यंजन समूह वर्तनी के कारण दिखाई दे रहे हैं, जैसे school, cycle, psychology आदि। Psychology में आपको शुरु में पाँच व्यंजन दिख रहे हैं लिपि के आधार पर। लेकिन आप ध्यान से बोलें तो केवल एक ही व्यंजन ध्वनि है /स्/। हिंदी हो चाहे अंग्रेज़ी, शब्द के आरम्भ में केवल तीन ही व्यंजन ध्वनियाँ संभव हैं। हिंदी में शायद 'स्त्री' जिसमें /स्/, /त्/, /र/ साथ-साथ आते हैं। इसी प्रकार अंग्रेज़ी के निम्न शब्दों को देखिये:

spray, street, screw, splash, squash.

इन शब्दों के शुरु में तीन व्यंजन ध्वनियाँ अवश्य हैं। लेकिन क्या व्यंजन ध्वनियों को साथ-साथ बोलने में हमें कोई कीमत अदा करनी पड़ी? बहुत अधिक। यदि हम (व्यंजन)-स्वर-(व्यंजन)-स्वर की सामान्य संरचना के आधार पर चलते हैं तो हमें कोई रोक टोक नहीं होती। हम 'काला', 'खाला', 'नाला', 'पाला', 'बाला' आदि जैसे लाखों शब्द बना सकते हैं। लेकिन आप यदि यह चाहते हैं कि हिंदी या अंग्रेज़ी शब्द के आरंभ में कई व्यंजन हों तो आप के ऊपर कई बंधन हैं। कई नियमों का पालन करना पड़ेगा। जैसे:

१. तीन से अधिक व्यंजन ध्वनियाँ शब्द के आरम्भ में नहीं आ सकतीं।
२. यदि तीन हैं तो पहला केवल /स्/ हो सकता है।
३. दूसरा केवल /प्/, /त्/, /ट्/, /क्/ ही हो सकता है।
४. और तीसरा व्यंजन केवल /य्/, /र्य्/, /ल्य्/, /व्य्/ ही हो सकता है।

यह बात कई भारतीय भाषाओं के लिए भी सच है (Pandey 2014)। यही कारण है कि अधिकतर भाषाएँ सीखते समय यह एक स्वाभाविक प्रक्रिया होती है कि भारतीय बच्चे या बड़े संयुक्त व्यंजनों को पहले या फिर बीच में स्वर लगाकर तोड़ देते हैं। उदाहरण के लिए: 'स्ट्रीट' (street) को इस्-ट्रीट (is-treet) या सट्-रीट (sat-reet)।

बोलना एक स्वाभाविक बात है। समय के चलते सब बच्चे और बड़े सीख जाते हैं यदि उचित मात्रा में सही प्रयोग से सामना होता रहे तो। एक शिक्षक के लिए चाहे वह हिंदी पढ़ा रहा हो या अंग्रेज़ी ऐसी बातों की समझ होना ज़रूरी है। मोटी बात यह कि यदि आपको यह समझ नहीं कि भाषा सीखने में मुख्य मुकाम कौन कौन से होते हैं तो आप कैसे एक अच्छे भाषा शिक्षक बन सकते हैं। और यह समझना भी उतना ही ज़रूरी है कि अंग्रेज़ी या हिंदी के कई ऐसे रूप होंगे जहाँ इन नियमों का पालन इस प्रकार से नहीं होगा। उदाहरण के लिए हिंदी बोलनेवाले कई लोग 'स्त्री' को /सतरी/ या /इस्-ती/ बोलते हैं। राजस्थान के कई हिंदीभाषियों के लिए 'चाय' का रूप /साय/ है।

शब्दों के स्तर पर देखिये। यह एक मौलिक बात है कि भाषा की संरचना जैसी ध्वनि व वाक्य के स्तर पर होती है वैसी शब्द के स्तर पर नहीं। ध्वनि व वाक्य के नियम हर जगह लागू होते हैं। ध्वनि के क्षेत्र में हमने अभी ऊपर देखा। वाक्य संरचना में हम अभी नीचे देखेंगे। ध्वनि व वाक्य के नियम हम एक बार सीख लेते हैं। बस वह काफी होता है जीवन भर भाषा का उपयोग करने के लिए। लेकिन नए शब्द, मुहावरे, उपमाएँ आदि सीखने की कोई सीमा नहीं। भाषा के इस पहलू पर

स्कूल के शिक्षक विशेष ध्यान दे सकते हैं। लेकिन इसके लिए शब्द संरचना की कुछ समझ होनी ज़रूरी है। वह इसलिए भी कि सदियों से चले आ रहे पूर्वाग्रहों को हटाना भी ज़रूरी है। उदाहरण के लिए हिंदी और अंग्रेज़ी में लिंग व्यवस्था को लीजिये। दोनों में नर और मादा होते हैं। अंग्रेज़ी में मादा होने की परिस्थिति में 'she' सर्वनाम का उपयोग होता है। लेकिन इसका अर्थ यह नहीं कि अंग्रेज़ी में भी व्याकरण के स्तर पर हिंदी की तरह कोई लिंग व्यवस्था है। व्याकरण की दृष्टि से हिंदी लिंग-युक्त है और अंग्रेज़ी लिंग-रहित। लेकिन अंग्रेज़ी और हिंदी में लिंग व्यवस्था स्कूलों में एक ही तरह से पढ़ाई जाती है। हिंदी व अंग्रेज़ी पढ़ानेवाले अध्यापक अक्सर विद्यार्थियों को हिंदी व अंग्रेज़ी की लिंग व्यवस्था सिखाते हैं बिना इस बात पर ध्यान दिए कि अंग्रेज़ी के वाक्यों को लिंग के बदलने से कोई फर्क नहीं पड़ता। हिंदी व अंग्रेज़ी के निम्न वाक्य देखिये:

(१) राम सेब खाता है।

(२) गीता सेब खाती है।

क्रिया में कर्ता के लिंग के अनुसार बदलाव आता है। लेकिन अंग्रेज़ी में ऐसा नहीं है:

(३) Ram eats an apple.

(४) Geeta eats an apple.

वाक्य (३) और (४) में क्रिया में कोई अंतर नहीं आया; क्रिया को कोई फर्क नहीं पड़ता चाहे राम खाए, चाहे गीता। इन वाक्यों से वाक्य संरचना आधारित नियमों की पहचान भी होती है। हिंदी में शब्द-क्रम मुख्यतः 'कर्ता-कर्म-क्रिया' होगा और अंग्रेज़ी में 'कर्ता-क्रिया-कर्म'। यह भी देखें कि केवल कर्ता के वर्तमान काल में अन्य पुरुष एकवचन होने पर ही क्रिया के साथ '-स्' का प्रयोग होगा अन्यत्र कहीं नहीं। इससे स्कूलों में अंग्रेज़ी पढ़ाने वाले शिक्षकों को यह भी साफ़ हो जाना चाहिए कि बच्चे अंग्रेज़ी सीखते वक्त क्यों स्वाभाविक रूप से बोलते व लिखते हैं:

(४क) * Geeta eat an apple.

शब्दों को लेकर हिंदी में अक्सर यह पढ़ाया जाता है कि नर होगा तो पुल्लिंग और मादा होगी तो स्त्रीलिंग। सही भी है लेकिन कुछ हद तक ही। 'घोड़ा' पुल्लिंग है तो 'घोड़ी' स्त्रीलिंग। पर 'कोयल' सदा स्त्रीलिंग और 'चीता' सदा पुल्लिंग! समूहवाचक संज्ञाओं में देखें: 'भीड़' चाहे आदमियों की हो या औरतों की, सदा स्त्रीलिंग ही होगी। दूसरी तरफ 'झुण्ड' चाहे आदमियों का हो या औरतों का, पुल्लिंग ही होगा। बच्चों को यह भी पढ़ाया जाता है कि व्यंजनांत ('घर', 'बालक', 'संदूक' आदि) एवं आकारांत ('कमरा', 'गाना', 'कपड़ा' आदि) शब्द पुल्लिंग होते हैं। पर 'किताब', 'बात', 'चमक' और 'माला', 'हवा' आदि तो स्त्रीलिंग हैं। शब्दों की दुनिया ध्वनि व वाक्यों की दुनिया से बहुत अलग होती है। शब्दों की दुनिया में एक ही व्याकरणिक रिश्ते को दिखाने के अक्सर कई तरीके होते हैं (Singh & Agnihotri 1997; 2007)। हमने ऊपर देखा कि संरचना के आधार पर यह कहना संभव नहीं कि शब्द पुल्लिंग होगा या कि स्त्रीलिंग। वास्तव में शब्दों की दुनिया में एक ही व्याकरणिक रिश्ते को कई तरीकों से दिखाया जा सकता है। 'लेख' में केवल '-अक' लगाकर 'लेखक' बनता है लेकिन 'शिक्षा' में पहले '-आ' तो हटाना पड़ता है उसके बाद '-अक' लगता है

जैसे 'शिक्षक'। कुछ शब्दों में यह काम '-कार' लगाने से होगा यथा 'कला' से 'कलाकार'। यदि शिक्षक इन बातों के प्रति संवेदनशील नहीं होगा तो वह हिंदी या अंग्रेजी या फिर किसी भी भाषा का एक प्रेरणादायक शिक्षक कैसे बनेगा?

२ बहुभाषिता

भाषा को बहुभाषिता समझ कर ही हम उसे सही ढंग से परिभाषित कर सकते हैं (Agnihotri 2006, 2007a, 2007b, 2009, 2012)। यह सत्य है कि भाषा की एक सार्वभौमिक जन्मजात छवि हम सबके मानस में रहती है। वह न हो तो हम इस बात का कोई उत्तर नहीं दे सकते कि बच्चे कैसे चार वर्ष की आयु में ही भाषा की दृष्टि से व्यस्क हो जाते हैं। वे सहजभाव से किसी भी भाषा की, जो उनके पर्यावरण का हिस्सा हो, जटिल व विषम ध्वनि व वाक्य संरचना को आत्मसात कर लेते हैं। कभी कोई बच्चा अपनी मातृभाषाओं में पाँच छह साल की उम्र के बाद कोई लुटि नहीं करता। भाषा को एकरूपता के ढाँचे में ढालना एक भाषावैज्ञानिक, राजनैतिक या साहित्यिक मजबूरी हो सकती है पर इन सबसे भाषा कि वास्तविक बहुभाषिता छुपती नहीं। केवल शब्दों का ही नहीं बल्कि ध्वनियों व वाक्य संरचनाओं का भी लेन देन चलता रहता है उन सब के बीच जिन्हें हम 'एकरूप भाषा' का नाम देते हैं। उस 'एकरूप भाषा' का अपना वास्तविक स्वरूप बहुभाषिता है। काफी साल पहले जब मैंने बहुभाषिता (Achmat 1992; Agnihotri 1992) की बात की और उसे multilinguality कहना शुरू किया तो उसके बाद कई भाषा वैज्ञानिकों व भाषा शिक्षकों ने कई ऐसी परिकल्पनाएं व शब्द इस्तेमाल करने शुरू किये जिससे इसी बात का आभास होता है: superdiversity, translanguaging, polylingualism, hyperlingualism, interlanguaging.

भाषा को बहुभाषिता (multilinguality) कहने से हमारे क्या तात्पर्य हैं और उस परिप्रेक्ष्य में भाषा/ हिंदी भाषा शिक्षण से हम क्या समझते हैं? पहली बात तो यह कि एक ही व्यक्ति के मानस व व्यवहार में कई भाषाएँ साथ-साथ रहती हैं, उस व्यक्ति की उनमें प्रवीणता के आयाम अलग-अलग हो सकते हैं और उनमें निरंतर एक 'तरलता' का संबंध बना रहता है। यह नहीं है कि बहुत सी भाषाएँ हैं तो बहुभाषिता है। यह कि बहुभाषिता है उसके सिवा और कुछ नहीं। अभिव्यक्ति के सम्पूर्ण साधनों को मिलाकर बहुभाषिता बनती है व्यक्तिगत व सामाजिक स्तर पर। जितने समृद्ध किसी व्यक्ति के भाषागत संसाधन होंगे उतनी ही समृद्ध उसकी संज्ञानात्मक क्षमता होगी और उतनी ही गहरी होगी उसकी अन्य भाषाओं व भाषा बोलनेवालों के प्रति संवेदनशीलता। हर समाज बहुभाषी होता है और इसलिए हर कक्ष भी बहुभाषी होती है। इस तरह की बहुभाषी कक्षाओं में किस तरह के अध्यापक होने चाहिये, किस तरह की पाठ्य सामग्री का प्रयोग होना चाहिए और किस तरह की शैक्षणिक प्रक्रिया व गतिविधियाँ होनी चाहिए इस पर विचार करना आवश्यक है।

३ भाषा शिक्षण

यदि हम भाषा शिक्षण में कुछ भी सार्थक नवाचार करना चाहते हैं तो इस प्रक्रिया में पहला कदम होगा शिक्षक प्रशिक्षण। इसकी शुरुआत एक छोटे स्तर से करनी होगी। बाद में इसका विस्तार हो सकता है यदि हमारे परिणाम अच्छे होंगे। यानी कुछ भाषाओं में प्रवीणता, कक्षा में उपलब्ध सभी भाषाओं में रूचि, संज्ञानात्मक विकास व सामाजिक सहिष्णुता- यह सब संभव होने चाहिए। शिक्षक

प्रशिक्षण का मुख्य हिस्सा होगा भाषा के चार पक्ष: भाषा सीखने की प्रक्रिया, भाषा की संरचना, भाषा में ऐतिहासिक व सामाजिक स्तर पर बदलाव और भाषा के सामाजिक व मानसिक पहलू। भाषा शिक्षक के प्रशिक्षण में अधिकतर जोर दिया जाता है पाठ्य सामग्री व शैक्षणिक प्रक्रिया पर। वास्तविकता यह है कि न तो पढ़ाने का कोई एक आदर्श तरीका हो सकता है और न ही कोई ऐसी पाठ्य सामग्री होती है जिसे आदर्श मान लिया जाए। इसके विपरीत हम यह जानते हैं कि जिन अध्यापकों को भाषा के उपरोक्त पहलूओं का आभास होता है और जो स्वयं भाषा में प्रवीण होते हैं वे स्वयं उचित सामग्री व पढ़ाने के उचित तरीके खोज लेते हैं।

३.१ भाषा शिक्षण के तरीके

भाषा किन तरीकों से पढ़ाई जाए इसका इतिहास काफी लंबा व जटिल है। इन तरीकों का अपने समय में प्रचलित भाषावैज्ञानिक व मनोवैज्ञानिक सिद्धांतों व कार्यप्रणालियों से गहरा संबंध रहा है। एक सबसे पुराना तरीका जो आज भी हम पर काफी हावी है वह है मुखर रूप से व्याकरण पढ़ाना व मातृभाषा व सीखी जा रही दूसरी भाषा में निरंतर अनुवाद करते रहना। यह माना जाता था कि यदि विद्यार्थी अपनी भाषा से टारगेट भाषा में व टारगेट भाषा से अपनी भाषा में अनुवाद कर लें तो मान लेना चाहिए कि उसे टारगेट भाषा आ गई। इस मंजिल तक पहुँचने के लिए व्याकरण में संज्ञा व क्रिया के जटिल रूपों को कंठस्थ करना आवश्यक माना जाता था। यह तरीका आज भी काफी प्रचलित है। इसमें यह चिंता नहीं रहती कि शायद बच्चे को टारगेट भाषा में कुछ भी बोलना न आए। और न यह कि वह शायद कभी भी अपने विचारों को नई भाषा में व्यक्त न कर पाए। Lado (1957) से लेकर Kumaravadivelu (2006) तक भाषा शिक्षण के तरीकों में कई उतार-चढ़ाव आये। संरचनात्मक भाषावैज्ञानिक (Bloomfield 1933) व व्यवहारवादी मनोविज्ञान के दिनों में तो अनुकरण व अभ्यास का खूब बोल-बाला रहा (Skinner 1957; Thorndike 1931)। भाषा शिक्षण मुख्यतः इस बात पर आधारित था कि विद्यार्थियों को एक शब्दावली दे दी जाए और उन्हें वाक्य संरचना के साँचे सिखा दिए जाएँ जिनमें वे शब्दों को भरते जाएँ। इस बात की कोई समझ नहीं थी कि बच्चे चार वर्ष की आयु से पूर्व ही भाषा की दृष्टि से व्यक्त हो जाते हैं (Chomsky 1957; 1959; 1965)। मेग्रेटिक टेप के आने से भाषा शिक्षण के तरीकों में कई बदलाव हुए। इसके साथ ही दोनों विश्व युद्धों में जासूसी करने के लिए बहुत ही कम समय में ही दुश्मन की भाषा सीखने की ज़रूरत सामने आई। टेप रिकॉर्डर व उसका विकसित रूप लैंग्विज लैब सामने आये। अब यह संभव था कि एक ही संरचना को आसानी से बार बार सुना जा सकता था और सीखनेवाले अपनी इच्छा व समय के अनुसार बिना अध्यापक के भी अभ्यास कर सकते थे और अपनी अशुद्धियों को समझ कर ठीक कर सकते थे। इस ऑडियो-लिंगुअल तरीके के मुख्य आधार थे:

- भाषा सही मायने में बोलना व सुनकर समझना है।
- भाषा एक आदत है। अभ्यास से यह स्वतः एक आदत सी बन जाती है।
- भाषा पढ़ाओ न कि भाषा के बारे में। यानी व्याकरण अलग विषय के रूप में पढ़ाने की कोई आवश्यकता नहीं।
- भाषा वही है जो उसके मातृभाषी बोलते हैं। न कि कोई ऐसा रूप जो आदर्श माना जाता हो।

ऑडियो-लिंगुअल तरीका अनुवाद व व्याकरण-आधारित तरीके के विरोध में एक आवाज़ भी था। सैद्धांतिक भाषा विज्ञान में फिर एक आवाज़ उठी डेल हायमज़ (Hymes 1966; 1974) की। उनका

कहना था कि भाषा केवल संरचना की ही बात नहीं; उसमें संदर्भ व समाज का भी एक आधारभूत पक्ष है। हम जो भी कहते हैं वह इस बात पर निर्भर करता है कि हम किसके साथ, कहाँ और किस विषय पर बात कर रहे हैं। इस संप्रेषण आधारित (Brumfit & Johnson 1978) भाषा शिक्षण के तरीके ने पहले इस्तेमाल हो रहे सभी तरीकों पर प्रश्न चिह्न लगा दिया। यह समझ आने लगा कि केवल व्याकरण युक्त सही भाषा बोलना ही काफी नहीं। क्या सामाजिक संदर्भ है और आप भाषा के साथ क्या करना चाहते हैं यह भी ध्यान में रखना आवश्यक है। आप केवल कोई सूचना देना चाहते हैं या आज्ञा देना चाहते हैं या फिर किसी से कोई प्रार्थना करना चाहते हैं; किसी को खुश करना चाहते हैं या उससे दूरी दिखाना चाहते हैं – इन सभी परिस्थितियों में भाषा का स्वरूप बदलेगा। हिंदी का वाक्य देखें या साथ ही उसका अंग्रेज़ी अनुवाद –

दरवाज़ा खुला है। 'The door is open.'

इन वाक्यों के शब्दिक अर्थ के बिलकुल इतर संदर्भ के आधार पर एक ही वाक्य के अलग अलग मायने हो सकते हैं।

१. (दरवाज़े पर कोई दस्तक दे रहा है; अर्थ होगा): आप अंदर आ जाँ।
२. (इंटरव्यू में नाराज़ अधिकारी यदि उम्मीदवार से कहे तो): आप बाहर जा सकते हैं।
३. (बहुत सर्दी हो और कोई भी बिस्तर से न निकलना चाहे तो): तुम (कोई भी बिस्तर में लेटा दूसरा व्यक्ति) दरवाज़ा बंद कर दो।

संप्रेषण-आधारित भाषा शिक्षण में अधिक ज़ोर ख़ासकर शुरू के दिनों में धाराप्रवाह बोलने पर था न कि शुद्धता पर। जब संरचनात्मक भाषाविज्ञान व व्यवहारवादिता (Bloomfield 1933; Skinner 1957) का बोलबाला था तो यह माना जाता था कि विद्यार्थी की अपनी भाषा उनके अन्य भाषा सीखने में बाधा बनेगी। स्कूल में बच्चों को नई भाषा सीखने में क्या परेशानियाँ आयेंगी यह समझने के लिए भाषा वैज्ञानिक व अध्यापक घर की भाषा व टारगेट भाषा का तुलनात्मक अध्ययन करते हैं।

पर आधुनिक शोध में यह जल्दी ही साफ़ हो गया कि ध्वनि के स्तर को छोड़ कर टारगेट भाषा सीखनेवाले की मातृभाषा किसी भी स्तर पर बाधक नहीं होती है। और ध्वनि के स्तर पर तो मातृभाषाओं में भी खूब विविधता होती है। वास्तव में दुले, बर्ट व क्राशन (Dulay, Burt & Krashen 1982) ने साफ़ दिखा दिया कि किसी भी भाषा को अर्जन करने की राह एक ही है चाहे उसे पहली भाषा के रूप में सीखो या दूसरी। क्राशन (Krashen 1987) का कहना था कि दूसरी भाषा सीखने के लिए उपयुक्त माला में समझ में आने वाला चुनौतीपूर्ण इनपुट होना चाहिए। हम सब में भाषा सीखने की क्षमता रहती है और एक ऐसा 'मॉनिटर' भी रहता है जो समय समय पर खुद भाषा को सुधारता रहता है, चाहे वह कोई भी भाषा हो, पहली, दूसरी या फिर तीसरी। कुमारवादिवेलु (Kumaravadivelu 2006) तो 'पोस्ट-मेथड' की बात करते हैं—यह एक ऐसा मुकाम है जहाँ पढ़ाने का कोई भी एक सही तरीका नहीं है। हर अध्यापक अपना अलग तरीका तलाश करता है यह समझते हुए कि उसके विद्यार्थी अलग अलग सांस्कृतिक पृष्ठभूमियों से आते हैं, उनकी पहचान अलग-अलग है और जिन सामाजिक, राजनैतिक व आर्थिक परिस्थितियों में वे विद्यार्थी पढ़ा रहे हैं वे अलग हैं। वही तरीका सफल होगा जिसमें अध्यापक अपनी परिस्थिति की

विशिष्टता को पहचानेगा, यह समझेगा कि उसे अपनी परिस्थिति से ही अपने सिद्धांत बनाने हैं और यह भी समझना है कि उसके संदर्भ में क्या संभावनाएँ हैं।

३.२ बहुभाषिता एवं भाषा शिक्षण

भाषा शिक्षण की कोई भी विधि प्रचलित रही हो किसी ने यह सवाल कभी नहीं पूछा कि उन भाषाओं का क्या होगा जो बच्चे अपने साथ लेकर स्कूल आते हैं। क्या यह मानना उचित होगा कि सभी बच्चे एक ही भाषा बोलते हैं जबकि हम जानते हैं कि कक्षाएँ सामान्यतः बहुभाषी होती हैं? क्या यह सोचना आवश्यक नहीं कि बच्चों की भाषाएँ नई भाषाएँ सीखने के लिए एक स्रोत हो सकती हैं? यही नहीं ये भाषाएँ बच्चों के संज्ञानात्मक विकास का एक समर्थ साधन बन सकती हैं। अंग्रेज़ी पढ़ाने के संदर्भ में इस तरह की सोच केन हेल व ओनील और उनके कुछ साथियों ने एम् आइ टी में शुरू की। होंडा, ओनील व पिपिन (Honda, O'Neil & Pipin 2007) ने दिखाया कि किस प्रकार भाषाविज्ञान का उपयोग अंग्रेज़ी की कक्षा में हो सकता है। बहुभाषिता के संदर्भ में यह प्रयोग १९९२ के आस पास केप टाउन में हुई कुछ कौन्सिलिंग व कार्यशालाओं में हुआ जिन पर अद्वारित कुछ वीडियो (Achmat 1992) भी बने और कई लेख भी लिखे गए (Agnihotri 1992; 1995; 1997)। बहुभाषिता के संदर्भ में भाषा शिक्षण—हिंदी भी को लेकर—पर काफी काम हुआ (Agnihotri 2006; 2007a; 2007b; 2009; 2010; 2012; Garcia 2009; Heugh 2010; Canagarajah 2011 आदि)। सैद्धांतिक तौर पर इस दृष्टिकोण की मुख्य बातें हैं:

- यह बिलकुल ज़रूरी नहीं है कि हम सदा एक भाषा व एक पाठ्यपुस्तक के जाल में रहकर ही शिक्षा व भाषा शिक्षण के बारे में सोचें। बहुभाषिता के संदर्भ में अनेक तरह की सामग्री का प्रयोग हो सकता है। उदाहरणतः जो कहानी व गीत बच्चे अपने साथ लाते हैं वह सभी शिक्षा की सामग्री बन सकते हैं।
- हर बच्चा एक समृद्ध व जटिल भाषा लेकर कक्षा में आता है। उसकी भाषा का किसी भी कीमत पर निरादर नहीं होना चाहिए।
- हमारा प्रयास होना चाहिए कि हर बच्चे की भाषा कक्षा में हो रही गतिविधियों का हिस्सा बने। जो कहानी, गीत, मुहावरे वह साथ लेकर आता है वह उसकी विरासत का हिस्सा हैं। उनका निरादर होगा तो बच्चों का मन शिक्षा में नहीं लगेगा।
- इस तरह के तरीके के लिए यह कदापि आवश्यक नहीं कि अध्यापक को बच्चों की हर भाषा आती हो। हाँ यह ज़रूर हो सकता है कि शिक्षक बच्चों के द्वारा लाई हुई सामग्री से खुद बराबर सीखता रहे। इस तरह की सामग्री लाने के लिए बच्चों को निरंतर प्रेरित करना चाहिए, वे घर से ला सकते हैं या गली कूचों से। पोस्टरों से या इश्तिहारों से।
- बहुभाषी शिक्षण के तरीके अपनाने का यह अर्थ नहीं है कि बच्चे दो या तीन भाषाओं में प्रवीण नहीं होंगे। इन भाषाओं में प्रवीणता के साथ-साथ उनका परिचय उन भाषाओं से भी होगा जो कक्षा में उपलब्ध होंगी और वे समझ पायेंगे कि हर बच्चे की भाषा राजनैतिक व सामाजिक स्तर पर न सही पर संरचनात्मक स्तर पर बराबर है।
- बच्चों की भाषाओं का इस्तेमाल करते हुए यह सहज ही संभव हो सकता है कि निम्न उद्देश्यों की पूर्ति हो: हर बच्चे की भाषा को आवाज़ मिले और उसका आत्मविश्वास बढ़े; व्याकरण के किसी एक बिंदु को लेकर (यथा लिंग, वचन आदि), विभिन्न भाषाओं से डेटा एकलित किया जाए और उसका नियमबद्ध तरीके से वर्गीकरण हो। समूहों में बैठकर बच्चे इनका विश्लेषण करेंगे तथा

नियम बनायेंगे। इन नियमों को फिर से नए डेटा के साथ जाँचेंगे। उदाहरण के लिए शिक्षक किसी भी एक भाषा को लेकर बोर्ड पर कुछ शब्द लिख देगा। बच्चों के साथ मिलकर उनके बहुवचन भी लिखेगा। बस इसके बाद सब काम विद्यार्थी खुद कर सकते हैं। उन्हीं शब्दों को लेकर ३ या ४ भाषाओं से एकवचन व बहुवचन लिखे जायेंगे। यह काम विद्यार्थी खुद करेंगे। इसके बाद छोटे छोटे समूहों में बैठकर वे उस डेटा का निरीक्षण करेंगे, उसे अपने अवलोकनों के आधार पर अलग-अलग श्रेणियों में वर्गीकृत करेंगे और बहुवचन बनाने के नियम बनायेंगे हर भाषा में। यहाँ सवाल सही या गलत का उतना नहीं है जितना कि एक संज्ञानात्मक वैज्ञानिक प्रक्रिया का।

इस तरह की गतिविधियों से यह सुनिश्चित होगा कि बच्चों की भाषा के प्रति संवेदनशीलता बढ़ेगी और उनका संज्ञानात्मक विकास होगा।

३.३ कक्षा में: शिक्षण विधि का उदाहरण

कबीरा का एक दोहा है। यह अक्सर स्कूलों में पढ़ाया जाता है। हम तीसरी कक्षा से लेकर दसवीं कक्षा तक कहीं भी इसका प्रयोग कर सकते हैं। हर स्तर पर ज़ाहिर है कि गतिविधियाँ अलग अलग होंगी।

कबीरा खड़ा बज़ार में, माँगे सबकी खैर,
ना काहू से दोस्ती, ना काहू से बैर।

इस एक छोटे से दोहे से हम क्या क्या नहीं कर सकते। अच्छा रहे यदि बच्चे इसे याद कर लें और रोमन में भी लिख लें और इसके अर्थ समझ लें। जैसे:

*kabiiraa khaRaa bazaar men, maange sab kii khair,
naa kaahuu se dostii, naa kaahuu se bair .*

(Roman letters have their normal phonetic value. Vowel length is marked by doubling, such as /aa/, /ii/ etc., /n/ after vowels is used to mark nasalization.)

‘कबीरा’ को ही लें। रोमन में लिखें या देवनागरी में कोई फर्क नहीं पड़ता। मुद्दा यह है कि संरचना व्यंजन-स्वर क्रम की है। एक साधारण सी गतिविधि से शुरू कर सकते हैं। यह गतिविधि तीसरी या चौथी कक्षा में हो सकती है। हर बच्चा अपना नाम लिख सकता है देवनागरी में भी व रोमन में भी। भारत के सरकारी स्कूलों में भी यह दोनों लिपियाँ पहली-दूसरी कक्षा से ही अक्सर शुरू हो जाती हैं। कुछ नाम हो सकते हैं: विराट / Viraat, सुरेश / Suresh, कविता / Kavita, होमना / Homnaa, सविता / Savitaa आदि। यह समझने में अधिक समय नहीं लगेगा कि इन सभी शब्दों में व्यंजन-स्वर-व्यंजन का क्रम रहता है। रोमन में लिखने से यह बात अधिक साफ़ हो जाती है। आप पूरे दोहे को देखें यही क्रम है केवल ‘दोस्ती’ को छोड़कर। उसमें भी आप ‘स्’ के बाद रुकते हैं। भाषा के बारे में इस प्रकार की चेतना से बहुत लाभ होता है। सभी बच्चे समझने लगते हैं कि विभिन्न भाषाओं में शब्दों की संरचना अधिकतर एक तरह की ही होती है। थोड़े से प्रयास से यह संभव है कि कक्षा में विद्यार्थी अपने आस पास की दुनिया के शब्दों पर एक नज़र डालें। जैसे: ‘कोका कोला’, ‘पानी’, ‘दूध’, ‘चाय’, ‘चीनी’, ‘कलम’ आदि।

यदि कोई भारतीयभाषी दो या दो से अधिक व्यंजन एक साथ बोलने का प्रयास करें तो कई बंधन होते हैं और वह कोई स्वाभाविक प्रक्रिया भी नहीं लगती। अक्सर बच्चे व अन्य भाषाएँ सीखनेवाले लोग व्यंजनों को तोड़कर ही बोलते हैं यथा स्कूल को इसकूल या फिर सकूल। जैसा कि हमने ऊपर देखा तीन से अधिक व्यंजनों को बोलने पर तो भारतीय भाषाओं और अंग्रेज़ी में बहुत ही सख्त अंकुश है।

इस तरह की सब बातें बच्चे स्वयं तलाश कर लेते हैं। जैसा कि मैंने ऊपर कहा अध्यापक की भाषा विज्ञान के कुछ आधारभूत पहलुओं पर पकड़ होना ज़रूरी है। अक्सर यह प्रश्न उठाया जाता है कि बहुभाषी कक्षा में इस तरह की गतिविधियाँ करते हुए अध्यापक किस भाषा का प्रयोग करे। इस तरह की आशंकाओं के पीछे यही मानसिकता छुपी रहती है कि किसी भी कक्षा में शिक्षण का ढांचा 'एक किताब-एक भाषा-एक अध्यापक' का होना चाहिए। इससे दूर हटना आवश्यक है। तभी अध्यापक व विद्यार्थियों की सृजनशीलता को जगह मिलेगी। हमें सोचना चाहिए कि अलग-अलग भाषा बोलने वाले विद्यार्थी जो कक्षा में चुप-चाप बैठे रहते हैं वही खेल के मैदान में एक दूसरे से खूब बातें करते हैं और कई बार तो वह सब भी एक दूसरे को समझाते हैं जो कक्षा में अध्यापक ने पढ़ाया। यह तो साफ़ है कि कक्षा में अध्यापक उन्हीं भाषाओं का प्रयोग कर सकता है जो उसे आती हैं। सवाल इस बात का है कि वह विद्यार्थियों की भाषाओं को कितनी जगह देता है।

शब्दों के स्तर पर देखिये। बच्चों के साथ संज्ञा पर बातचीत शुरू करें। कुछ ही देर में बच्चे यह शब्द दोहे में से निकाल लेंगे: 'बज़ार', 'दोस्ती', 'खैर', 'बैर'। बस अब अध्यापक का काम ख़त्म। इसके बाद जब ज़रूरत होगी तो अध्यापक मदद करेगा अन्यथा वह विद्यार्थियों को समूहों में काम करते हुए देखेगा। छोटे छोटे समूहों में बैठकर बच्चे इन शब्दों के लिए अपनी अपनी भाषा के शब्द लिखें। रोमन लिपि में भी व देवनागरी में भी। इनके लिए अंग्रेज़ी के शब्द भी बताये जा सकते हैं। कक्षा पर निर्भर करता है। इसके बाद इन शब्दों के बहुवचन बनाएँ। 'बज़ार', 'खैर' व 'बैर' का तो कुछ न होगा पर 'दोस्ती' का 'दोस्तियाँ'। लेकिन कुछ ही देर में यह साफ़ होगा कि 'बाज़ारों' तो होता है। आप कक्षा में इस तरह की गतिविधि करें और देखें क्या समां बंधता है। बच्चे अपने आस पास के शब्दों को लिखकर उनके बहुवचन बनाते हैं और हिंदीभाषी बच्चे हिंदी में बहुवचन बनाने के जटिल नियम खुद तलाश कर लेते हैं। अन्य समूह अन्य भाषाओं में ऐसे ही नियम तलाश करेंगे। बात नियम तलाश करने की ही नहीं। जैसा कि मैंने पहले कहा बात है उस प्रक्रिया से गुज़रने की जो वैज्ञानिक है और जो केवल भाषा के माध्यम से ही संभव है खासकर ७-८ साल की छोटी उम्र में। इस उम्र में आप भौतिक या रसायन शास्त्र का प्रयोग अभी नहीं कर सकते वैज्ञानिक प्रक्रिया से परिचित कराने के लिए। बच्चे स्वयं इस बात को समझ लेते हैं कि हिंदी में शब्द का बहुवचन बनाने के लिए यह जानना आवश्यक है कि शब्द की अंतिम ध्वनि क्या है और उसका लिंग क्या है। 'मकान' और 'घर' व्यंजनांत हैं पर पुल्लिंग भी। साधारण बहुवचन में इनमें कोई परिवर्तन नहीं होगा। 'बड़ा मकान' / 'बड़े मकान' आदि। लेकिन 'किताब' और 'पुस्तक' भी व्यंजनांत हैं लेकिन स्त्रीलिंग। इनका बहुवचन बिलकुल अलग 'किताबें', 'पुस्तकें'। परसर्ग (यानी 'में', 'ने' आदि) आने से पहले सभी बहुवचन बदलेंगे अलग अलग नियमों से परंतु परसर्ग के साथ एक ही सामान्य नियम के अनुसार: *मकानों / घरों / किताबों / पुस्तकों में*—इसी प्रकार बच्चे अपनी अपनी भाषा के नियम बना सकते हैं। यदि यह गतिविधि प्राइमरी कक्षाओं में हो रही है तो 'विशेषण' या 'परसर्ग' जैसे तकनीकी शब्द इस्तेमाल करने की आवश्यकता नहीं।

अंग्रेज़ी में लिंग नहीं है। बहुवचन बनाने के लिए शब्दांत ध्वनि को ही सुनना है ध्यान से। उदाहरण के लिए बच्चे जल्द ही यह नियम पकड़ लेते हैं कि यदि शब्द का अंत /स्/, /श/, /च्/ या /ज्/ से हो तो बहुवचन बनाने के लिए अंग्रेज़ी में 'इज़' लगाना होता है यथा 'बसिज़', 'बुशिज़' आदि।

प्रश्न नियम सीखने का नहीं है। प्रश्न है वैज्ञानिक प्रक्रिया से गुजरने का; भाषा के प्रति संवेदनशीलता का; संज्ञानात्मक विकास का व सामाजिक सहिष्णुता का। प्रश्न भाषा में प्रवीणता का भी है। फिर से कबीर के दोहे को देखें। इसे हिंदी कहें, अवधि या ब्रज या उस समय में आम लोगों में बोली जाने वाली हिन्दुस्तानी। (यह सवाल हर भाषा के बारे में पूछा जाना चाहिए।) किसी बड़ी कक्षा में इस बात पर भी खूब चर्चा हो सकती है। इतिहास, राजनीति व भाषा के संघर्ष का यह एक बहुत ही जटिल मुकाम है। तीसरी-चौथी कक्षा के बच्चे भी बहुत कुछ कर सकते हैं। बहुभाषिता के संदर्भ में भाषागत प्रवीणता के लिए बहुत जगह रहती है। अनुवाद करना व अलग अलग भाषाओं में सरलार्थ व टिप्पणी करना कुछ सक्शत तरीके बहुभाषिता का उपयोग करने के हैं। सवाल आरंभ में शुद्धता का नहीं। सवाल है बेझिझक बोलने व लिखने का; समझकर पढ़ने का। मान लो कक्षा में हिंदी व अंग्रेज़ी के साथ साथ मेवाड़ी, भीली, मंडियाली व गुजराती भी हैं। कक्षा में खुलकर बोलने, अपनी भाषाओं का प्रयोग करने में व भाषा के नियम खोजने की प्रक्रिया में कोई अंतर्विरोध नहीं। बात है विद्यार्थियों के आत्मविश्वास का एवं उनके संज्ञानात्मक विकास का। विद्यार्थी छोटे छोटे समूहों में बैठकर इस दोहे को गद्य में लिखते हैं। अलग भाषाओं में अनुवाद कर सकते हैं। जैसे:

कबीरा बाज़ार में खड़ा है, वह सबकी खैर मांगता है
उसकी न किसीसे दोस्ती है और न ही किसी से बैर।

मानिये अंग्रेजी में कुछ ऐसा अनुवाद हो:

‘Kabira stands in the market place and prays for the
welfare of one and all.
He is neither anybody’s friend nor anybody’s foe.’

हो सकता है कक्षा में कोई विद्यार्थी हिमाचल के मंडी शहर का हो। वह लिख सकता है:

कबीर खड़ी रा बजारा, सब री माँगया करौँ खैर
न केसी कन्ने तेसरी दोस्ती न केसी कन्ने बैर।

यदि कोई विद्यार्थी मेवाड़ी जानता होगा तो कुछ ऐसा:

कबीरा उबो बजार वचे, मांगे सबरी खैर,
ना कणी सु यारी, ना कणी सु वेर

हम ने ऊपर देखा कि हिंदी व अन्य भारतीय भाषाएँ (खासी को छोड़ दें) क्रियांत हैं। हिंदी के वाक्य (१) (राम सेब खाता है) व (२) (गीता सेब खाती है) में चार-चार शब्द हैं। यानी इनसे शब्दों को इधर उधर कर अधिक से अधिक चार! जो कि $४ \times ३ \times २ \times १ = २४$ वाक्य बन सकते हैं। हिंदी में लचीलापन है इसलिए कुछ इस तरह के वाक्य संभव हैं:

(१क) सेब राम खाता है।
(२क) खाता है राम सेब।

पर अंग्रेज़ी में ऐसा लचीलापन संभव नहीं। क्रिया मध्य में है, यथा (३) Ram eats an apple। यहाँ भी चार शब्द हैं और २४ वाक्य बनना संभव है पर वाक्य (३) के अलावा एक भी व्याकरणिक तौर पर मान्य नहीं होगा और अर्थ के स्तर पर भी यह अटपटा:

(३क) *An apple eats Ram.


इन सब बातों को लेकर १३-१४ साल के विद्यार्थी खूब चर्चा कर सकते हैं। इन अनुवादों से साफ़ होगा कि हिंदी में काफी लचीलापन है। क्रिया को हम कुछ इधर उधर कर सकते हैं। अंग्रेज़ी में यह संभव नहीं। और भी कई आयामों पर चर्चा संभव है। हिंदी में 'में' 'बज़ार' के बाद आता है और अंग्रेज़ी में पहले। इस तरह के कई अंतर सामने आयेंगे। यह सब कक्षा ८-९ तक आते-आते हो सकता है।

साहित्य, भाषा व भाषा विश्लेषण की प्रक्रिया को हम जितना साथ साथ रखेंगे उतना ही लाभ होगा।

४ निष्कर्ष

कबीर के दोहों को देखें या प्राचीन भारत की मणिप्रवाल प्रथा को या फिर आम आदमी की भाषा को देखें, हमें चारों ओर भाषा का रूप बहुभाषिता जैसा ही दिखाई देता है। इसी बहुभाषिता को सामने रखते हुए ही हमें भाषा शिक्षण के बारे में सोच विचार करना चाहिए। हर विद्यार्थी भाषागत क्षमता व कुछ भाषाएँ लेकर ही कक्षा में आता है। बहुत से विद्यार्थी, बहुत सी भाषागत विविधता। हर बच्चे की आवाज़ को जगह मिलनी चाहिए। इसी तरीके से ऐसा भाषा शिक्षण संभव होगा जिसकी मानवता को ज़रूरत है।

ORCID®

Rama Kant Agnihotri  <https://orcid.org/0000-0003-0878-0913>

संदर्भ-सूची

Achmat, Z. 1992. "Yo dude, cosa wena kyk a? *The Multilingual Classroom*", Salt River, South Africa: National Language Project. (<<http://www.youtube.com/watch?v=174ULxuBM3EAgniho>>, accessed: April 14, 2014).

Agnihotri, Rama Kant 1991. "India: Multilingual Perspectives". Paper presented at the International Conference on Democratic Approaches to Language Planning and Standardization. National Language Project and University of Cape Town, Cape Town, September 11-13, 1991.

- Agnihotri, Rama Kant 1992. "India: Multilingual Perspectives", in: Crawhall, N. T. (ed.): *Democratically Speaking: International Perspectives on Language Planning*. Salt River, South Africa: National Language Project, 46–55.
- Agnihotri, Rama Kant 1995. "Multilingualism as a Classroom Resource", in: Heugh, K., A. Siegrühn & P. Plüddemann (eds): *Multilingual Education for South Africa*. Johannesburg: Heinemann, 3–7.
- Agnihotri, Rama Kant 1997. "Multilingualism, colonialism and translation", in: Ramakrishna, S. (ed.): *Translation and Multilingualism: Post-colonial Contexts*. Delhi: Pencraft International, 34–46.
- Agnihotri, Rama Kant 2006. "Identity and Multilinguality: The Case of India", in: Tsui, A. B. M. & J. W. Tollefson (eds): *Language Policy, Culture, and Identity in Asian Contexts*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 185–204.
- Agnihotri, Rama Kant 2007a. "Towards a Pedagogical Paradigm Rooted in Multilinguality", *International Multilingual Research Journal* 1: 79–88.
- Agnihotri, Rama Kant 2007b. *Hindi: An Essential Grammar*. London: Routledge.
- Agnihotri, Rama Kant 2009. "Multilinguality and a New World Order", in: Mohanty, A. K., M. Panda, R. Phillipson & T. Skutnabb-Kangas (eds): *Multilingual Education for Social Justice: Globalizing the Local*. New Delhi: Orient BlackSwan, 268–277.
- Agnihotri, Rama Kant 2010. "Multilinguality and the Teaching of English in India", *The EFL Journal* 1: 1–14.
- Agnihotri, Rama Kant 2012. "Multilinguality, Marginality and Social Change", in: Khanna, A.L & A. S. Gupta (eds): *Essential Readings for Teachers of English: From Research Insights to Classroom Practices*. New Delhi: Orient BlackSwan, 13–24.
- Bloomfield, Leonard 1933. *Language*. London: George Allen and Unwin.
- Brumfit, C. J. & K. Johnson 1978. *The Communicative Approach to Language Teaching*. Oxford: Oxford University Press.
- Canagarajah, A. S. (2011). "Codemeshing in Academic Writing: Identifying Teachable Strategies of the Translanguaging", *Modern Language Journal* 95: 401–417.
- Chomsky, Noam 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam 1959. "Review of BF Skinner's Verbal Behaviour", *Language* 35: 26–58.
- Chomsky, Noam 1965. *Aspects of Theory of Syntax*. Cambridge, Mass: MIT Press.
- Dulay, H., M. Burt & S. Krashen 1982. *Language Two*. New York: Oxford University Press.
- Garcia, O. 2009. "Education, Multilingualism and Translanguaging in the 21st Century", in: Mohanty, A. K., M. Panda, R. Phillipson & T. Skutnabb-Kangas

- (eds): *Multilingual Education for Social Justice: Globalizing the Local*. New Delhi: Orient BlackSwan, 128–145.
- Hymes, D. (ed.) 1966. *Language in Culture and Society*. New York: Harper and Row.
- Hymes, D. 1974. *Foundations in Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press.
- Heugh, K. 2010. “Productive Engagement with Linguistic Diversity in Tension with Globalised Discourses in Ethiopia”, *Current Issues in Language Planning* 11, 378–396.
- Honda, Maya, Wayne O’Neil, and David Pippin 2007. “On Promoting Linguistics Literacy: Bringing Language Science to the English Classroom”, in: Denham, Kristin and Anne Lobeck (eds): *Linguistics at School: Language Awareness in Primary and Secondary Education*. Cambridge: Cambridge University Press, 175–188.
- Hyman, Larry M. 2008. “Universals in Phonology”, *The Review* 25(1–2): 83–137. (<<https://doi.org/10.1515/TLIR.2008.003>>, accessed: March 25, 2019).
- Krashen, S. 1987 [1982]. *Principles and Practice in Second Language Acquisition*. New York (et al.): Prentice Hall International.
- Kumaravadivelu, B. 2006. *Understanding Language Teaching: From Method to Post-Method*. Marwah, N.J.: Lawrence Erlbaum.
- Lado, R. 1957. *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.
- Pandey, Pramod 2014. *Sounds and Their Patterns in Indic Languages*. Vol. 1. *Sound Patterns*. New Delhi: Cambridge University Press.
- Singh, Rajendra & Rama Kant Agnihotri 1997. *Hindi Morphology: A Word-based Description*. Delhi: Motilal Banarsidass.
- Skinner, B.F. 1957. *Verbal Behaviour*. New York: Appleton Century Crofts.
- Thorndike, Edward Lee 1931. *Human Learning*. New York: The Century.

Anil Kumar Pandey

How to Determine Grammatical Gender in Hindi

अनिल कुमार पाण्डेय

हिंदी में लिंग निर्धारण¹

Abstract Hindi has two grammatical genders: masculine and feminine. Students of Hindi, especially those who learn it as a second language, have difficulties in determining the nouns' gender. This creates syntactic problems in constructing and comprehending sentences. Nominal suffixes—both inflexional and derivational—may serve as a criterium in determining the noun gender and thus help to a certain extent to solve this problem. The present article approaches this issue from a pragmatic perspective. It provides a broad overview of the Hindi nominal suffixes to be used by Hindi learners and in the classroom.

Keywords nominal gender, determining the gender, inflexional suffixes, derivational, suffixes, lexical semantics, learners of Hindi as a second language.

सारांश हिंदी में दो ही लिंग हैं – पुल्लिंग और स्त्रीलिंग। हिंदी सीखने वाले गैर-हिंदी भाषियों, भारतीय एवं विदेशियों के लिए संज्ञाओं का लिंग निर्धारित करना कठिन होता है। यह कठिनाई वाक्यविन्यास में समस्या पैदा करती है। हिंदी में काफ़ी प्रत्ययों के आधार पर लिंग का निर्धारण किया जा सकता है चाहे वे रूपसाधक (Inflexional) हों अथवा व्युत्पादक (Derivational)। हिंदी में लिंग निर्धारण संबंधी समस्याओं को इन प्रत्ययों के आधार पर कुछ हद तक दूर किया जा सकता है। प्रस्तुत आलेख के द्वारा व्यावहारिक रूप से इस समस्या का समाधान ढूँढ़ने का प्रयास किया गया है।

मुख्य शब्द – संज्ञाओं का लिंग, लिंग का निर्धारण, रूपसाधक प्रत्यय, व्युत्पादक प्रत्यय, शब्दार्थ, हिंदी, सीखनेवाले हिंदीतर भाषी।

1 प्रस्तुत आलेख “हिंदी में प्रत्यय विचार (लिंग निर्धारण के विशेष संदर्भ में)” (पाण्डेय 2010) आलेख का संवर्धित-परिवर्धित किया हुआ नया संस्करण है।

१ परिचय

अधिकांश भारतीय-आर्य भाषाओं में तीन लिंग होते हैं परंतु हिंदी में दो ही लिंग हैं, पुल्लिंग और स्त्रीलिंग। हिंदीतर भाषियों की हिंदी सीखते समय एक बड़ी समस्या है कि किन शब्दों को स्त्रीलिंग कहें, किन शब्दों को पुल्लिंग और इसका आधार क्या है? इस संदर्भ में प्रसिद्ध वैयाकरणों ने भी अपनी विवशता प्रकट की है। हिंदी के मूर्धन्य वैयाकरण कामता प्रसाद गुरु के शब्दों में निम्नलिखित कथन उद्धृत है, “हिंदी में अप्राणिवाचक शब्दों का लिंग जानना विशेष कठिन है; क्योंकि यह बात अधिकांश व्यवहार के अधीन है। अर्थ और रूप दोनों ही साधनों से इन शब्दों के लिंग जानने में कठिनाई होती है।” (गुरु 1920: 162)।

समस्याएँ अधिकतर वहाँ आती हैं जहाँ निर्जीव अथवा अमूर्त शब्दों के लिंग का निर्धारण करना हो। इसके अलावा मानव अथवा पशु-पक्षियों के लिंग निर्धारण की जब बात हो तो वहाँ भी समस्याएँ खासकर अहिंदी भाषी के लिए आती हैं। प्रस्तुत आलेख का उद्देश्य है कि अहिंदी भाषी शिक्षार्थियों के साथ हिंदी भाषी शिक्षार्थियों के लिए लिंग निर्धारण के कुछ व्यावहारिक पक्ष प्रस्तुत किए जाएँ जिससे कुछ हद तक उस संदर्भ में आने वाली समस्याओं से छुटकारा मिल सके।

हिंदी में लिंग का निर्धारण कुछ हद तक प्रत्ययों के आधार पर किया जा सकता है। ये प्रत्यय रूपसाधक (Inflectional) एवं व्युत्पादक (Derivational) हैं। इन्हीं दोनों प्रकार के प्रत्ययों के आधार पर लिंग का निर्धारण किया जा सकता है। जब किसी शब्द में कोई व्युत्पादक प्रत्यय जुड़ता है तो मूल शब्द संज्ञा, विशेषण अथवा कृदंत होते हैं। यदि संज्ञा शब्द है तो वह स्त्रीलिंग होगा अथवा पुल्लिंग। विशेषण शब्दों का अपना कोई लिंग नहीं होता। जिन संज्ञा शब्दों की वे विशेषता बताते हैं, उन्हीं संज्ञा शब्द के साथ उनकी अन्विति होती है।

२ हिंदी में लिंग के सूचक

२.१ आर्थी लिंग

वाक्यगत संरचनाओं में लिंग का अपना महत्व होता है। अधिकांश भारतीय-आर्य भाषाओं में वाक्यविन्यासीय स्तर पर लिंग का प्रभाव स्पष्ट रूप से परिलक्षित होता है। प्रत्येक भाषा में लिंगों की संख्या भिन्न-भिन्न हैं। अब सवाल यह है कि संज्ञा को लिंग की कौनसी उपकोटि—स्त्रीलिंग या पुल्लिंग—में रखा गया है। उसका अपना कोई सिद्धांत भी नहीं है कि अमुक सिद्धांत के आधार पर इन्हें अलग-अलग रखा जा सके। सजीव शब्दों का तो लिंग निर्धारण आसानी से शब्दार्थ के आधार पर ही किया जा सकता है, जैसे: पुरुष – महिला, माता – पिता, गाय – बैल, परंतु लिंग अक्सर शब्द-रूप से भी स्पष्ट होता है, जैसे: राजा – रानी, हाथी – हथिनी, घोड़ा – घोड़ी आदि (हर जोड़े में पहले पुल्लिंग शब्द आता है)। हिंदी में पशु, पक्षी व कीटवाचक कुछ संज्ञा शब्दों को पुल्लिंग व कुछ स्त्रीलिंग की कोटि में रखा गया है जैसे – पुल्लिंग: कौआ, खटमल, सारस, चीता, उल्लू, केंचुआ, भेड़िया आदि। स्त्रीलिंग: कोयल, छिपकली, लोमड़ी, दीमक, चील, मैना, गिलहरी, तितली, मक्खी आदि।

पुरुषवाचक एवं निश्चयवाचक सर्वनामों में लिंग का निर्धारण संदर्भ से किया जाता है – वह, यह, हम, मैं, तुम, आप आदि। उदाहरण के लिए: वह आ रहा है; वह आ रही है। मैं आ रहा हूँ; मैं आ रही हूँ। तुम कब आओगे; तुम कब आओगी? यहाँ व्याकरणिक संदर्भ से निर्धारित किया जा सकता है कि आने वाला पुल्लिंग है अथवा स्त्रीलिंग, यह क्रिया में प्रतिबिंबित होता है। इसी

तरह—वह लड़का है; वह लड़की है—जैसे वाक्यों में लिंग विधेय में द्रष्टव्य है। अब ये उदाहरण देखिए: तुम आ जाओ। आप कौन हैं? इन उदाहरणों में लिंग का कोई व्याकरणिक सूचक नहीं है। कर्ता का लिंग अधिभाषिक संदर्भ से ही पहचाना जा सकता है। इन सभी उदाहरणों में सर्वनामों का प्रयोग दोनों लिंगों में किया जा सकता है परंतु इनके स्थान पर यदि संज्ञा शब्द का प्रयोग करेंगे तब वहाँ पुरुषवाची संज्ञा व स्त्रीवाची संज्ञा का प्रयोग होगा एवं उसके कारण क्रिया का प्रयोग भी स्त्रीवाची और पुरुषवाची के रूप में किया जाएगा। जैसे: लड़का आ रहा है (पुल्लिंग); लड़की आ रही है (स्त्रीलिंग)।

सजीव शब्दों में भी कुछ ऐसे शब्दों का प्रयोग किया जाता है जो पुरुष व स्त्री दोनों के वाचक होते हैं परंतु उन्हें या तो पुल्लिंग की कोटि में रखा गया है अथवा स्त्रीलिंग की कोटि में, जैसे:

पुल्लिंग: मनुष्य, मानव, पशु, अभिभावक, समाज, पक्षी, विद्यार्थी।
स्त्रीलिंग: संतान, भीड़, औलाद, चिड़िया, मक्खी।

पशु वर्ग में अधिकांश शब्द के पुल्लिंग व स्त्रीलिंग रूप मिलते हैं यथा: घोड़ा – घोड़ी, हाथी – हाथिनी, ऊँट – ऊँटनी, कुत्ता – कुतिया, बाघ – बाघिन, शेर – शेरनी आदि। कभी-कभी पशु का लिंग अलग-अलग शब्द प्रकट करते हैं, जैसे: गाय – बैल।

२.२ व्याकरणिक लिंग के सूचक – प्रत्यय

लिंग अधिकतर प्रत्ययों से स्पष्ट होता है। पहले रूपसाधक (Inflectional) प्रत्यय देखें, जो संज्ञाओं के रूप निर्धारण में उपकारी होते हैं, इसके बाद व्युत्पादक (Derivational)।

नोट: सामासिक शब्द के द्वितीय घटक के आधार पर ही समस्त पद का लिंग निर्धारण होता है। यदि द्वितीय घटक स्त्रीलिंग है तो समस्तपद स्त्रीलिंग होगा, इसके विपरीत पुल्लिंग। उदाहरणतः

– सहन (पु°) + शक्ति (स्त्री°) = सहनशक्ति (स्त्री°), मनः (पु°) + स्थिति (स्त्री°) = मनः स्थिति (स्त्री°);

भू (स्त्री°) + विज्ञान (पु°) = भूविज्ञान (पु°), प्रीति (स्त्री°) + भोज (पु°) = प्रीतिभोज (पु°)।

२.२.१ रूपसाधक प्रत्यय

हिंदी में पाँच रूपसाधक प्रत्यय सुस्पष्ट लिंग-सूचक हैं। ये प्रत्यक्ष रूप बहुवचन प्रत्यय -एँ, -याँ (- (य)आँ)^२, -ए एवं शून्य (-Ø)^३ तथा तिर्यक रूप एकवचन प्रत्यय -ए^४ होते हैं। बहुवचन प्रत्ययों में -एँ, -याँ प्रत्यय स्त्रीवाची प्रत्यय हैं। ये जिस किसी शब्द में जुड़ते हैं, वे सभी शब्द स्त्रीलिंग होते

२ प्रत्यय है -आँ। ई- एवं इ-कारांत स्त्रीलिंग शब्दों में -आँ जुड़ने से उत्पन्न -य प्रातिपदिक में शामिल है। सुविधा के लिए नीचे प्रत्यय के रूप में -याँ लिखा जाता है।

३ शून्य प्रत्यय को “शून्य विभक्ति” (Zero Morph) कहा गया है।

४ ये दो समस्वन रूपिम हैं।

हैं तथा -ए एवं शून्य प्रत्यय जिस शब्द के साथ जुड़ते हैं, वे पुल्लिंग होते हैं। यहाँ शून्य से तात्पर्य है, प्रत्यय रहित। जिन शब्द रूपों में आकार ध्वन्यात्मक तौर पर अभिव्यक्त नहीं होता है या फिर व्यंजन में अंत होता है वहाँ शून्य प्रत्यय माना जाता है जैसे कि 'घर [-Ø]'। एकवचन में शून्य प्रत्यय वाले शब्द स्त्रीलिंग एवं पुल्लिंग हो सकते हैं। परंतु -ए प्रत्यय वाले शब्द सिर्फ पुल्लिंग होते हैं।

नीचे तालिकाओं १, २, ३ व ४ में बहुवचन संज्ञाओं में लगे इन प्रत्ययों के उदाहरण द्रष्टव्य हैं।

तालिका १ व्यंजनांत, आकारांत एवं ऊकारांत स्त्रीलिंग संज्ञाओं के बहुवचन

एकवचन शब्द	प्रत्यय -एँ	बहुवचन शब्द	लिंग
मेज़	+एँ	मेज़ें	स्त्रीलिंग
पुस्तक	+एँ	पुस्तकें	स्त्रीलिंग
सीमा	+एँ	सीमाएँ	स्त्रीलिंग
कथा	+एँ	कथाएँ	स्त्रीलिंग
बहू	+एँ	बहुएँ	स्त्रीलिंग

इसी संरचना के कुछ और स्त्रीलिंग संज्ञाओं के उदाहरण हैं:

आदतें, सड़कें, क्रियाएँ, भावनाएँ, दिशाएँ, पत्रिकाएँ, रेखाएँ, संरचनाएँ, वासनाएँ, आशाएँ आदि।

तालिका २ ईकारांत स्त्रीलिंग संज्ञाओं के बहुवचन

एकवचन शब्द	प्रत्यय -याँ (-आँ)	बहुवचन शब्द	लिंग
कुर्सी	+याँ	कुर्सीयाँ	स्त्रीलिंग
लकड़ी	+याँ	लकड़ियाँ	स्त्रीलिंग
बकरी	+याँ	बकरियाँ	स्त्रीलिंग
सवारी	+याँ	सवारियाँ	स्त्रीलिंग
चुनौती	+याँ	चुनौतियाँ	स्त्रीलिंग
गाड़ी	+याँ	गाड़ियाँ	स्त्रीलिंग

ईकारांत शब्दों के बहुवचन रूप के कुछ उदाहरण और दिए जा रहे हैं। जैसे:

परी – परियाँ, कमी – कमियाँ आदि।

तालिका ३ आकारांत पुल्लिंग संज्ञाओं के बहुवचन

एकवचन शब्द	प्रत्यय -ए	बहुवचन शब्द	लिंग
फोड़ा	+ए	फोड़े	पुल्लिंग
लड़का	+ए	लड़के	पुल्लिंग
कपड़ा	+ए	कपड़े	पुल्लिंग
दरवाज़ा	+ए	दरवाज़े	पुल्लिंग
हादसा	+ए	हादसे	पुल्लिंग

तालिका ४ व्यंजनांत पुल्लिंग संज्ञाओं के बहुवचन – शून्य प्रत्यय (-०)

एकवचन शब्द	प्रत्यय शून्य -०	बहुवचन शब्द	लिंग
प्रेम	+०	प्रेम	पुल्लिंग
प्यार	+०	प्यार	पुल्लिंग
घर	+०	घर	पुल्लिंग
समाज	+०	समाज	पुल्लिंग
ग्रंथ	+०	ग्रंथ	पुल्लिंग

हिंदी में किसी शब्द के वचन के स्तर पर परिवर्तन करने पर जो परिवर्तन होते हैं, वे प्रत्यय के स्तर पर ही होते हैं। ऐसे प्रत्यय को भाषावैज्ञानिक दृष्टि से बद्धरूपिम (Bound Morpheme) की संज्ञा दी गई है।

उपर्युक्त उदाहरणों में शब्द के साथ -ए, -(य)आँ, -एँ तथा -शून्य बहुवचन प्रत्यय जुड़े हैं। बहुवचन प्रत्यय के भी दो रूप मिलते हैं। पहला रूप वह प्रत्यय जो शब्दों के बहुवचन रूप बनने पर प्रत्यक्ष रूप (Direct Form) में दिखे, जैसा कि उपर्युक्त उदाहरणों में दिया गया है। दूसरा रूप वह जो तिर्यक रूप (Oblique Form) में दिखे, जैसे: लड़कों (को), कुर्सियों (को), मेजों (को), समाजों (को), मनुष्यों (को), जानवरों (को) आदि में -ओं प्रत्यय में। (यह रूप सामान्यतः परसर्गों से पहले आता है।) दोनों बहुवचन प्रत्ययों में व्याकरणिक दृष्टि से वाक्यगत जो अंतर है वह यह कि प्रत्यक्ष रूप के साथ कोई परसर्ग (विभक्ति चिह्न) नहीं जुड़ता। हाँ, आकारांत शब्दों में -ए प्रत्यय जुड़ने से उनके बहुवचन रूप बनते हैं। परसर्ग (विभक्ति चिह्न) के जुड़ने पर -ए युक्त रूप बहुवचन न होकर तिर्यक एकवचन रूप होते हैं। उदाहरण के लिए:

कुर्सियाँ टूट गईं।

* कुर्सियों को ले आओ। (अस्वाभाविक प्रयोग)

मेजें मजबूत बनी हैं। (बहुवचन रूप)

* मेजें को कमरे में लगाओ। (अस्वाभाविक प्रयोग)

कपड़े फट गए। (बहुवचन रूप)

कपड़े पर दाग है। (तिर्यक एकवचन रूप)

लड़के खेल रहे हैं। (बहुवचन रूप)

लड़के ने खेला है। (तिर्यक एकवचन रूप)

उपर्युक्त उदाहरणों से स्पष्ट है कि शब्द में प्रत्यक्ष रूप बहुवचन प्रत्यय लगने पर कोई परसर्ग नहीं लगेगा। आकारांत शब्दों में बहुवचन प्रत्यय -ए लगाने पर उसके साथ परसर्ग तो लग सकते हैं परंतु परसर्ग लगाने पर बहुवचन न होकर वे तिर्यक एकवचन रूप हो जाते हैं। पुल्लिंग आकारांत शब्दों में बहुवचन प्रत्यय -ए केवल जातिवाचक संबंधवाचक (पापा, मामा, चाचा, राजा, काका आदि) एवं संस्कृत कर्तृवाचक (श्रोता, वक्ता आदि) संज्ञाओं में नहीं लगता।

बहुवचन का तिर्यक रूप (Oblique Form)

जैसा कि ऊपर कहा गया है बहुवचन प्रत्यय का तिर्यक रूप सभी संज्ञाओं में -ओं है। अतः इसके आधार पर लिंग का निर्धारण करना असंभव है।

२.२.२ व्युत्पादक प्रत्यय (Derivational Suffixes)

वे प्रत्यय जो किसी शब्द के साथ जुड़कर दूसरे शब्द निर्मित करते हैं उन्हें व्युत्पादक प्रत्यय कहते हैं। जैसे: सुंदर (विशेषण) + ता = सुंदरता (संज्ञा)।

प्रत्यय जुड़ने से मौलिक प्रातिपदिकों में रूपस्वनिमिक (morphophonological) परिवर्तन हो सकते हैं। आम तौर पर ऐसे परिवर्तन संस्कृत तत्सम तथा उन्हीं संरचनाओं के अनुसार बने नवसंस्कृत शब्दों की विशेषता होते हैं, परंतु हिंदी और दूसरे मूल (फ़ारसी, अरबी) शब्दों में भी आते हैं। यहाँ इन परिवर्तनों पर कोई विशेष टिप्पणी नहीं की जा रही है।

स्त्रीवाची प्रत्यय

-ता स्त्रीवाची प्रत्यय

तालिका ५ में कुछ स्त्रीवाची प्रत्ययों एवं उनसे बनने वाले शब्दों के उदाहरण दिए जा रहे हैं।

तालिका ५ - ता स्त्रीवाची प्रत्यय एवं भाववाचक संज्ञाएँ

मूल शब्द	- ता स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
मानव	+ ता	मानवता
स्वतंत्र	+ ता	स्वतंत्रता
पति	+ ता	पतिता
महत्	+ ता	महत्ता
लघु	+ ता	लघुता
दुष्ट	+ ता	दुष्टता
मम	+ ता	ममता
आधुनिक	+ ता	आधुनिकता

इसी नमूने पर निम्न संज्ञाएँ निर्मित हैं:

दीर्घता, मूर्खता, मधुरता, कटुता, गुरुता, विशेषता, शूरता, जड़ता, एकरूपता, आत्मीयता।

इस तद्धित प्रत्यय के शब्द में जुड़ने से जो शब्द बनते हैं, वे भाववाचक स्त्रीलिंग संज्ञा शब्द होते हैं।

- ता प्रत्यय जिन शब्दों में लगता है, वे शब्द बहुधा विशेषण पद होते हैं।

- ता प्रत्यय से कुछ समूहवाचक अथवा अन्य शब्द भी निर्मित होते हैं, जो स्त्रीलिंग ही होते हैं:

जन + ता = जनता, कवि + ता = कविता

- इमा स्त्रीवाची प्रत्यय

- इमा तद्धित प्रत्यय से युक्त सभी शब्द भाववाचक स्त्रीलिंग संज्ञाएँ होते हैं (देखिए तालिका ६)। उपर्युक्त प्रत्यय बहुधा विशेषण शब्दों में जुड़ता है।

तालिका ६ - इमा प्रत्यय एवं भाववाचक स्त्रीलिंग शब्द

मूल शब्द	-इमा स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
लाल	+ इमा	लालिमा
काला	+ इमा	कालिमा
रक्त	+ इमा	रक्तिमा
लघु	+ इमा	लघिमा

निम्न संज्ञाएँ भी इसी नमूने पर निर्मित हैं:

नीलिमा, अरुणिमा, हरीतिमा, श्वेतिमा, धवलिमा, भंगिमा, महिमा, पूर्णिमा, मंदिमा, बंकिमा, म्लानिमा, मधुरिमा ।

- आवट स्त्रीवाची प्रत्यय

- आवट प्रत्यय कृदंत प्रत्यय है। इस प्रत्यय से निर्मित शब्द भाववाचक स्त्रीलंग संज्ञाएँ होते हैं (देखिए तालिका ७)।

तालिका ७ - आवट प्रत्यय एवं स्त्रीलिंग शब्द

मूल शब्द	-आवट स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
लिख (ना)	+ आवट	लिखावट
रुक (ना)	+ आवट	रुकावट
मिल (ना)	+ आवट	मिलावट
गिर (ना)	+ आवट	गिरावट

- आवट प्रत्यय लगे हुए कुछ और संज्ञा शब्द द्रष्टव्य हैं:

थकावट, बनावट, बुनावट, सजावट, कसावट, खिंचावट, चुनावट, जमावट, फुलावट, फैलावट, दिखावट।

- आहट स्त्रीवाची प्रत्यय

- आहट बहुधा कृदंत प्रत्यय है। इसके विशेषण व धातु या प्रातिपदिक के साथ जुड़ने से भाववाचक स्त्रीलंग संज्ञा शब्द बनते हैं (देखिए तालिके ८)। -आहट प्रत्यय से जो शब्द बनते हैं, वे अक्सर ध्वनि अनुकरणात्मक शब्द होते हैं।

तालिका ८ - आहट प्रत्यय एवं स्त्रीलिंग शब्द

मूल शब्द	-(आ)हट स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
कड़वा	+ (आ)हट	कड़वाहट
चिकना	+ (आ)हट	चिकनाहट
गरमा	+ (आ)हट	गरमाहट
छटपटा	+ (आ)हट	छटपटाहट
चिल्ला (-ना)	+ (आ)हट	चिल्लाहट

इसी प्रकार के अन्य उदाहरण द्रष्टव्य हैं:

अचकचाहट, उकसाहट, किचकिचाहट, खिसियाहट, गड़गड़ाहट, जगमगाहट, झूँझलाहट, तड़तड़ाहट, तुतलाहट, धकधकाहट, फुसफुसाहट, बिलबिलाहट, बौखलाहट, मिनमिनाहट, मुस्कुराहट, लड़खड़ाहट, सनसनाहट, सरसराहट ।

-अना स्त्रीवाची प्रत्यय

इस कृदंत प्रत्यय से निर्मित शब्द भाववाचक स्त्रीलिंग संज्ञाएँ होते हैं। यथा:

अर्च+अना= अर्चना, अवहेल्+अना= अवहेलना, कल्प्+अना= कल्पना, घट्+अना= घटना, काम्+अना= कामना ।

इसी प्रकार के संज्ञा शब्द देखिए, जैसे:

आराधना, गर्जना, तर्जना, भावना, रचना, वर्जना, गणना, यातना, याचना, वेदना, सूचना, स्थापना, वंचना, तुलना, प्रेरणा, आदि ।

-अना प्रत्यय संस्कृत धातुओं में जुड़ता है। जिन संज्ञाओं में यह प्रत्यय लगा है, वे अधिकतर तत्सम शब्द होते हैं।

-आ स्त्रीवाची प्रत्यय

-आ कृदंत प्रत्यय है। इससे युक्त स्त्रीलिंग शब्द भाववाचक संज्ञाएँ होते हैं।

कथ्+आ=कथा, चिन्त्+आ = चिन्ता, पूज् +आ= पूजा, परीक्ष् +आ=परीक्षा, कृ+आ= क्रिया । निम्न इसी प्रकार के संज्ञा शब्द द्रष्टव्य हैं, जैसे:

सज्जा, चर्चा, उषा, प्रच्छा, बाधा, भाष्+आ= भाषा, भिक्षा, शोभा, रक्षा, सेवा, कुंठा, लज्जा, घृण्+आ= घृणा, प्रथा, व्यथा, कृपा, स्पृध्+आ= स्पर्धा आदि ।

-ती स्त्रीवाची प्रत्यय

यह प्रत्यय तद्धित और कृदंत दोनों रूपों में प्रयुक्त हो सकता है। इससे बनने वाले शब्द स्त्रीलिंग होते हैं (देखिए तालिका ९)।

तालिका ९ क्रिया से बने संज्ञा शब्द

मूल शब्द	-ती स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
पाव/पा(ना)	+ ती	पावती
चढ़(ना)	+ ती	चढ़ती
गिन(ना)	+ ती	गिनती
भर(ना)	+ ती	भरती
बोल	+ ती	बोलती
चल	+ ती	चलती

-इ, -ति स्त्रीवाची प्रत्यय

-इ अथवा -ति प्रत्यय से युक्त भाववाचक शब्द स्त्रीलिंग होते हैं (देखिए तालिका १०)। परंतु ऐसे ही प्रत्ययों के साथ युक्त प्राणिवाचक शब्द पुल्लिंग होते हैं (नीचे भारतीय-आर्य पुरुषवाची प्रत्ययों के परिच्छेद में देखिए)। स्त्रीवाची -ति प्रत्यय कृदंत प्रत्यय है। वह संस्कृत क्रिया धातुओं के साथ संयुक्त होता है। स्त्रीवाची -इ प्रत्यय बहुधा कृदंत प्रत्यय परंतु तद्धित प्रत्यय के रूप में भी प्रयुक्त होता है।

तालिका १० संस्कृत क्रिया धातुओं से बने स्त्रीलिंग संज्ञा शब्द

मूल शब्द	-ति स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)	मूल शब्द	-इ स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
कृ	+ ति	कृति	कृष्	+इ	कृषि
प्री	+ ति	प्रीति	हन्	+इ	हानि
शक्	+ ति	शक्ति			
स्था	+ ति	स्थिति			
स्मृ	+ ति	स्मृति			

इसी प्रकार -ति प्रत्यय वाले संज्ञा शब्दों के और उदाहरण द्रष्टव्य हैं:

जाति, दृष्टि, क्षति, ख्याति, भीति, भ्रान्ति, सृष्टि, श्रुति, नीति, दीप्ति, तृप्ति, भक्ति, क्रांति, नियुक्ति, जाति, ज्योति, शांति, गति, रीति ।

-आई स्त्रीवाची प्रत्यय

-आई प्रत्यय कृदंत और तद्धित दोनों में प्रयुक्त होता है (देखिए तालिका ११) ।

तालिका ११ विशेषण (ऊपर) एवं क्रिया मूल से (नीचे) बने भाववाचक संज्ञा शब्द

मूल शब्द	- आई स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
सच्चा	+आई	सचाई,
ढीठ	+आई	ढिठाई
एका	+आई	एकाई / इकाई
रूखा	+आई	रूखाई
खट्टा	+आई	खटाई
लड़(-ना)	+आई	लड़ाई
पीट(-ना)	+आई	पिटाई

इसी प्रकार – महँगाई, पढ़ाई, लिखाई, पिसाई, बुराई, चिकनाई, सिलाई, ठकुराई, चराई, पंडिताई, कमाई आदि ।

-ई स्त्रीवाची प्रत्यय

-ई प्रत्यय कृदंत और तद्धित दोनों में प्रयुक्त होता है । ईकारांत भाववाचक शब्द स्त्रीलिंग ही होते हैं (देखिए तालिका १२) ।

इस ग्रूप में भारतीय-आर्य एवं फ़ारसी-अरबी, यहाँ तक कि अंग्रेज़ी भी -ई प्रत्यय युक्त संज्ञाएँ हैं ।

तालिका १२ -ई प्रत्यय से बने स्त्रीलिंग संज्ञा शब्द

मूल शब्द	-ई स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
सावधान	+ ई	सावधानी
खेत	+ ई	खेती
डॉक्टर	+ ई	डॉक्टरी
गृहस्थ	+ ई	गृहस्थी
जवान	+ ई	जवानी
हँस(-ना)	+ ई	हँसी
धमक	+ ई	धमकी
दोस्त	+ ई	दोस्ती
गरीब	+ ई	गरीबी
जुदा	+ ई	जुदाई
खुदा	+ ई	खुदाई

इसी प्रकार: अमीरी, चोरी, टोली, दलाली, फाँसी, बुद्धिमानी, बोली, बदी, मनमानी महाजनी, सुस्ती आदि।

-औती स्त्रीवाची प्रत्यय

-औती बहुधा कृदंत प्रत्यय है, तद्धित के रूप में भी प्रयुक्त होता है। उससे भाववाचक स्त्रीलिंग शब्द बनते हैं (देखिए तालिका १३)।

तालिका १३ -औती प्रत्यय से बने स्त्रीलिंग शब्द

मूल शब्द	-औती स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
बूढ़ा	+ औती	बुढ़ौती
मन	+ औती	मनौती
फिर(ना)	+ औती	फिरौती
चुन(ना)	+ औती	चुनौती
बाप	+ औती	बपौती
कट(ना)	+ औती	कटौती
चुक(ना)	+ औती	चुकौती

-आनी स्त्रीवाची प्रत्यय

यह प्रत्यय सजीव पुल्लिंग शब्द में जुड़कर सजीव स्त्रीलिंग संज्ञा बनाता है (देखिए तालिका १४)।

तालिका १४ - आनी प्रत्यय से बने स्त्रीलिंग शब्द

मूल शब्द (पुल्लिंग मनुष्य संज्ञाएँ)	-आनी स्त्रीवाची प्रत्यय	निर्मित शब्द (स्त्रीलिंग मनुष्य संज्ञाएँ)
पंडित	+आनी	पंडितानी
मास्टर	+आनी	मास्टरानी
जेठ	+आनी	जेठानी
चौधरी	+आनी	चौधरानी
नौकर	+आनी	नौकरानी
देवर	+आनी	देवरानी
मेहतर	+आनी	मेहतरानी

कुछ अपवाद हैं, जैसे क्षत्र+आनी = क्षत्राणी । इसमें -आनी/-आणी प्रत्यय समूहवाचक शब्द में लगा है ।

-इन स्त्रीवाची प्रत्यय

यह प्रत्यय -आनी प्रत्यय की भाँति सजीव पुल्लिंग शब्दों में लगकर स्त्रीलिंग संज्ञाएँ बनाता है (देखिए तालिका १५) ।

तालिका १५ -इन प्रत्यय से बने स्त्रीलिंग संज्ञा शब्द

शब्द (सजीव पुल्लिंग संज्ञाएँ)	-इन स्त्रीवाची प्रत्यय	निर्मित शब्द (सजीव स्त्रीलिंग संज्ञाएँ)
सुनार	+ इन	सुनारिन
नाती	+ इन	नातिन
लुहार	+ इन	लुहारिन
तेली	+ इन	तेलिन
अहीर	+ इन	अहिरिन
बाघ	+ इन	बाघिन
साँप	+ इन	साँपिन

-शून्य(-Ø) स्त्रीवाची प्रत्यय

यह कृदंत प्रत्यय है, जो क्रिया मूल में जुड़कर उसके समान रूप की भाववाचक संज्ञाएँ बनाता है (देखिए तालिका १६) ।

तालिका १६ -शून्य प्रत्यय से बने स्त्रीलिंग भाववाचक संज्ञा शब्द

मूल शब्द	-ना प्रत्यय का लोप+ -Ø स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
छूटना	Ø	छूट
चमकना	Ø	चमक

इस संरचना की पूरी व्युत्पत्ति प्रक्रिया नीचे देखिए: -ना पुरुषवाची एवं -० स्त्रीवाची प्रत्यय।

पुरुषवाची प्रत्यय

स्त्रीवाची प्रत्ययों की भाँति पुरुषवाची प्रत्ययों की संख्या अधिक है। नीचे कुछ पुरुषवाची प्रत्ययों के उदाहरण दिए गए हैं।

-त्व पुरुषवाची प्रत्यय

यह तद्धित प्रत्यय है। शब्दों में जुड़कर यह प्रत्यय भाववाचक पुल्लिंग संज्ञाएँ बनाता है (देखिए तालिका १७)

तालिका १७ संस्कृत -त्व प्रत्यय से बने संज्ञा शब्द

मूल शब्द	-त्व पुरुषवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
मम	+ त्व	ममत्व
ब्राह्मण	+ त्व	ब्राह्मणत्व
बंधु	+ त्व	बंधुत्व
एक	+ त्व	एकत्व
मातृ	+ त्व	मातृत्व
धन	+ त्व	धनत्व
देव	+ त्व	देवत्व
महत्	+ त्व	महत्त्व

इसी प्रकार: सतीत्व, स्त्रीत्व, राजत्व, पितृत्व, प्रभुत्व, चुंबकत्व, द्वित्व, अमरत्व, नेतृत्व, पुरुषत्व, लघुत्व।

-अक पुरुषवाची प्रत्यय

-अक कृदंत प्रत्यय है जो संस्कृत धातुओं के साथ संयुक्त होकर जातिवाचक—प्राणिवाचक व वस्तुवाचक—पुल्लिंग संज्ञा शब्द बनाता है (देखिए तालिका १८)।

तालिका १८ -अक प्रत्यय से बने जातिवाचक संज्ञा शब्द

मूल शब्द	-अक पुरुषवाची प्रत्यय	निर्मित शब्द जातिवाचकसंज्ञाएँ)
आलुंच्	+ अक	आलोचक
ग्रह्	+ अक	ग्राहक
गै	+ अक	गायक
दीप्	+ अक	दीपक
दा	+ अक	दायक
दृश्	+ अक	दर्शक
पठ्	+ अक	पाठक
सूच्	+ अक	सूचक

इसी प्रकार: जनक, चिंतक, तारक, दोहक, धावक, नाशक, प्रेषक, रक्षक, भक्षक, वेधक, वंचक, संचारक, वादक, संपादक, लेखक, साधक, द्योतक, याचक, वाचक ।

-अन पुरुषवाची प्रत्यय
(देखिए तालिका १९) ।

तालिका १९ -अन पुरुषवाची प्रत्यय से बने संज्ञा शब्द

मूल शब्द	-अन पुरुषवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
हन्	+ अन	हनन
पाल्	+ अन	पालन
अंक्	+ अन	अंकन
वच्	+ अन	वचन
जीव्	+ अन	जीवन
आकृष्	+ अन	आकर्षण
मृ	+ अन	मरण
शी	+ अन	शयन
वृ	+ अन	वरण

कुछ स्वनात्मक संदर्भों में -अन का -अण हो जाता है, जैसे: मरण, आकर्षण ।

-आव पुरुषवाची प्रत्यय यह एक कृदंत प्रत्यय है। इससे बनने वाले शब्द भाववाचक संज्ञाएँ होते हैं (देखिए तालिका २०) ।

तालिका २० -आव प्रत्यय से बने संज्ञा शब्द

मूल शब्द	-आव पुरुषवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
लग(-ना)	+आव	लगाव
बच(-ना)	+आव	बचाव
छिड़क(-ना)	+आव	छिड़काव
तन(-ना)	+आव	तनाव
घेर(-ना)	+आव	घेराव
खींच(-ना)	+आव	खिंचाव

इसी प्रकार — कटाव, गलाव, ठहराव, चढ़ाव, चुनाव, जमाव, झुकाव, पड़ाव, पहनाव, बरताव, बहाव, भटकाव, भराव, रिझाव, सड़ाव ।

-आवा पुरुषवाची प्रत्यय यह प्रत्यय भी-आव प्रत्यय की भांति कृदंत प्रत्यय है। इससे बनने वाले शब्द भी भाववाचक पुल्लिंग संज्ञाएँ होते हैं (देखिए तालिका २१) ।

तालिका २१ - आवा प्रत्यय से बने पुल्लिंग संज्ञा शब्द

मूल शब्द	- आवा पुरुषवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
बुला(-ना)	+ आवा	बुलावा
भूल(-ना)	+ आवा	भूलावा
बहक(-ना)	+ आवा	बहकावा
पहन(-ना)	+ आवा	पहनावा
पछता(-ना)	+ आवा	पछतावा

इसी प्रकार: चलावा, छलावा, चढ़ावा, दिखावा ।

-पन एवं -पा पुरुषवाची प्रत्यय

इस भाववाचक तद्धित प्रत्यय से भाववाचक पुल्लिंग शब्द बनते हैं (देखिए तालिका २२) ।

तालिका २२ - पन प्रत्यय से बने पुल्लिंग संज्ञा शब्द

शब्द	-पन एवं -पा पुरुषवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
काला	+पन	कालापन
बाल	+पन	बालपन
बच्चा	+पन	बचपन
पागल	+पन	पागलपन
एकाकी	+पन	एकाकीपन
लचीला	+पन	लचीलापन
बूढ़ा	+ पा	बुढ़ापा
मोटा	+ पा	मोटापा

इसी प्रकार -पन प्रत्यय जुड़े हुए संज्ञा शब्दों के अन्य उदाहरण द्रष्टव्य हैं:

अंधापन, अजनबीपन, अक्खड़पन, अल्हड़पन, उपजाऊपन, ओछापन, कच्चापन, गंजापन, धुंधलापन, बाँझपन, लड़कपन, सयानापन आदि ।

-इ, -ति पुरुषवाची प्रत्यय

यह कृदंत प्रत्यय हैं जो संस्कृत धातुओं में जुड़कर तत्सम प्राणिवाचक शब्द बनाते हैं । इनकी संख्या कम है (देखिए तालिका २३) ।

तालिका २३ - इ कृदंत प्रत्यय से बने पुल्लिंग संज्ञा शब्द

मूल शब्द (संस्कृत धातु)	-इ पुरुषवाची प्रत्यय	निर्मित शब्द (प्राणिवाचक संज्ञाएँ)
कप्	+इ	कपि
कू	+इ	कवि
मुन्	+इ	मुनि
पा	+ति	पति

व्याकरणिक पर्यायवाची पुरुषवाची एवं स्त्रीवाची शब्द

-ना पुरुषवाची एवं -Ø (शून्य) स्त्रीवाची प्रत्यय

-ना कृदंत प्रत्यय से युक्त क्रियार्थक संज्ञा (Infinitive) पुल्लिंग होती है। इन्हीं क्रियार्थक संज्ञाओं में जुड़े -ना प्रत्यय की जगह -Ø प्रत्यय लगने से अर्थात् क्रिया के धातु रूप के समान जो भाववाचक संज्ञाएँ बनती हैं, वे कुछ थोड़े अपवादों को छोड़कर (जैसे जोड़-तोड़) सभी स्त्रीलिंग होती हैं। उदाहरण के लिए: जाँचना, लूटना, महकना, डाँटना आदि क्रियार्थक संज्ञाएँ पुल्लिंग हैं, जबकि लूट, महक, डाँट भाववाचक संज्ञाएँ स्त्रीलिंग हैं।

नीचे कुछ क्रियार्थक संज्ञाओं एवं -ना प्रत्यय लोप के उदाहरण दिए जा रहे हैं। जहाँ क्रियार्थक संज्ञा पुल्लिंग है, वहीं -ना का लोप होकर -शून्य (-Ø) लगी भाववाचक संज्ञाएँ स्त्रीलिंग हैं (देखिए तालिका २४)।

तालिका २४ प्रत्यय -ना के लोप से बने संज्ञा शब्द

-ना युक्त मूल शब्द (क्रियार्थक संज्ञा)	लिंग	-ना लोप एवं -Ø जुड़ने से निर्मित संज्ञा शब्द	परिवर्तित लिंग
जाँचना	पुल्लिंग	जाँच	स्त्रीलिंग
लूटना	पुल्लिंग	लूट	स्त्रीलिंग
पहुँचना	पुल्लिंग	पहुँच	स्त्रीलिंग
उछलना-कूदना	पुल्लिंग	उछल-कूद	स्त्रीलिंग
फटकारना	पुल्लिंग	फटकार	स्त्रीलिंग
डाँटना-डपटना	पुल्लिंग	डाँट-डपट	स्त्रीलिंग

इसी प्रकार के निम्नलिखित शब्दों के जोड़ों में पहले क्रियार्थक संज्ञा (पुल्लिंग) आती है, इसके बाद इससे निर्मित -Ø प्रत्यय वाली संज्ञा (स्त्रीलिंग):

चमकना – चमक, समझना – समझ, छींकना – छींक, पूछना – पूछ, फाँसना – फाँस, छापना – छाप, रगड़ना – रगड़, दौड़ना – दौड़, महकना – महक, चहकना – चहक।

२.३ मूलतः अरबी-फ़ारसी के स्त्रीवाची और पुरुषवाची प्रत्यय

हिंदी में कुछ अरबी-फ़ारसी प्रत्ययों का प्रयोग हो रहा है, जो स्त्रीवाची और पुरुषवाची प्रत्यय के रूप में प्रयुक्त हो रहे हैं। इनके आधार पर भी लिंग निर्धारण संभव है।

अरबी-फ़ारसी के कुछ स्त्रीवाची प्रत्यय

-इश स्त्रीवाची प्रत्यय

-इश प्रत्यय लगे शब्द भाववाचक स्त्रीलिंग संज्ञाएँ हैं (देखिए तालिका २५)।

तालिका २५ -इश प्रत्यय से बने भाववाचक संज्ञा शब्द

मूल शब्द	-इश स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
कोश	+इश	कोशिश
गुज़ार	+इश	गुज़ारिश
नाला	+इश	नालिश
पैदा	+इश	पैदाइश
साज़	+इश	साज़िश

इसी प्रकार: गरदिश, नुमाइश, आज़माइश, बंदिश, मालिश ।

-ई स्त्रीवाची प्रत्यय

-ई प्रत्यय से युक्त शब्द भाववाचक संज्ञाएँ हैं (देखिए तालिका २६)। थोड़े अपवाद हैं, जैसे समूहवाचक संज्ञा आबादी (< आबाद+ई)।

तालिका २६ -ई प्रत्यय से बने भाववाचक स्त्रीलिंग संज्ञा शब्द

मूल शब्द	-ई स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
दूर	+ई	दूरी
ईमानदार	+ई	ईमानदारी
चालाक	+ई	चालाकी
गुलाम	+ई	गुलामी
ज़बरदस्त	+ई	ज़बरदस्ती

इसी प्रकार— चापलूसी, मजबूरी, मंजूरी, माफ़ी, वीरानी, शायरी, शैतानी, होशियारी, सख्ती, हाज़िरी, शौक़ीनी।

-गी स्त्रीवाची प्रत्यय

-गी प्रत्यय स्त्रीवाची भाववाचक प्रत्यय है। यह प्रत्यय बहुधा विशेषण शब्दों में जुड़ता है (देखिए तालिका २७)।

तालिका २७ -गी प्रत्यय से बने भाववाचक स्त्रीलिंग संज्ञा शब्द

मूल शब्द	-गी स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
जिन्दा	+गी	जिंदगी
नाराज	+गी	नाराजगी
नापसंद	+गी	नापसंदगी
संजीदा	+गी	संजीदगी

इसी प्रकार: आवारगी, बंदगी, शर्मिंदगी, मरदानगी, ख़ानगी, सादगी ।

-गीरी स्त्रीवाची प्रत्यय (निम्नलिखित पुरुषवाची प्रत्यय -गीर देखिए।)
इस संयुक्त प्रत्यय के दो भाग हैं: -गीर एवं -ई। -ई वही प्रत्यय है जो ऊपर दिखाया जा चुका है, जैसे दूर-दूरी में।

-गीर कृदंत प्रत्यय है जो फ़ारसी क्रिया से बना हुआ है। इसका शाब्दिक अर्थ है 'लेता हुआ', 'लेने वाला'। इससे भारत-आर्य एवं फ़ारसी मूल के संज्ञा शब्दों से भाववाचक (व्यवसायवाचक) संज्ञाएँ बनती हैं (देखिए तालिका २८)।

तालिका २८ -गीरी प्रत्यय से बने भाववाचक स्त्रीलिंग संज्ञा शब्द

शब्द	-गीरी (< -गीर+ -ई) स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक व्यवसायवाचक संज्ञाएँ)
बाबू	+गीरी	बाबूगीरी
गुंडा	+गीरी	गुंडागीरी
नेता	+गीरी	नेतागीरी
राह	+गीरी	राहगीरी
मुंशी	+गीरी	मुंशीगीरी

इसी प्रकार: दाईगीरी, बढईगीरी, राजगीरी।

-अत स्त्रीवाची प्रत्यय
यह प्रत्यय भाववाचक स्त्रीलिंग संज्ञाएँ बनाता है, यथा:

शब्द	-अत स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक संज्ञाएँ)
खिलाफ़	+ अत	खिलाफ़त

-अत प्रत्यय मूलतः एक अरबी व्युत्पादक संरचना का अंश है। हिंदी में यह व्युत्पादक प्रक्रिया स्पष्ट रूप से दिखाना अधिकतर शब्दों को लेकर मुश्किल है। इस लिए नीचे दिए गए शब्दों में -अत प्रत्यय जुड़ा न लिखकर इसको सिर्फ़ चिह्नित किया गया है। सदाकत, शराफ़त, हिमाकत, लियाकत, शरारत आदि।

-इयत स्त्रीवाची प्रत्यय
यह अरबी प्रत्यय शब्द में जुड़कर भाववाचक स्त्रीलिंग संज्ञा बनाता है। अधिकतर यह अरबी मूल के विशेषणों में जुड़ता है (देखिए तालिका २९)।

तालिका २९ -इयत प्रत्यय से बने भाववाचक स्त्रीलिंग संज्ञा शब्द

शब्द	-इयत स्त्रीवाची प्रत्यय	निर्मित शब्द (भाववाचक एवं वस्तुवाचक संज्ञाएँ)
असल	+ इयत	असलियत
खास	+ इयत	खासियत
मनहूस	+ इयत	मनहूसियत
मासूम	+ इयत	मासूमियत
महरूम	+ इयत	महरूमियत
ईसान	+ इयत	ईसानियत
अंग्रेज	+ इयत	अंग्रेजियत
क्रब्ज	+ इयत	क्रब्जियत

अरबी-फ़ारसी के कुछ पुरुषवाची प्रत्यय

-आक पुरुषवाची प्रत्यय

इस प्रत्यय से भाववाचक एवं वस्तुवाचक स्त्रीलिंग संज्ञा शब्द बनते हैं (देखिए तालिका ३०)।

तालिका ३० -आक प्रत्यय से बने स्त्रीलिंग संज्ञा शब्द

शब्द	-आक पुरुषवाची प्रत्यय	निर्मित शब्द (वस्तुवाचक संज्ञाएँ)
खूर	+ आक	खुराक
पोश	+ आक	पोशाक

-आना पुरुषवाची प्रत्यय

यह प्रत्यय संज्ञा शब्दों में जुड़कर अधिकतर भाववाचक एवं कभी-कभी समूहवाचक संज्ञाएँ बनाता है (देखिए तालिका ३१)।

तालिका ३१ -आना प्रत्यय से बने स्त्रीलिंग संज्ञा शब्द

मूल शब्द	-आना पुरुषवाची प्रत्यय	निर्मित शब्द (भाववाचक एवं समूहवाचक संज्ञाएँ)
घर	+ आना	घराना
नज़र	+ आना	नज़राना
मेहनत	+ आना	मेहनताना
हरज	+ आना	हरजाना
यार	+ आना	याराना

-गार पुरुषवाची प्रत्यय

यह प्रत्यय संज्ञा शब्दों में जुड़कर भाववाचक स्त्रीलिंग संज्ञाएँ बनाता है (देखिए तालिका ३२)।

तालिका ३२ - गार प्रत्यय से बने भाववाचक संज्ञा शब्द

मूल शब्द	- गार पुरुषवाची प्रत्यय	निर्मित शब्द (जातिवाचक एवं भाववाचक संज्ञाएँ)
याद	+ गार	यादगार
रोज़	+ गार	रोज़गार

- गीर पुरुषवाची प्रत्यय (ऊपर लिखित स्त्रीवाची प्रत्यय - गीरी देखिए)
इस प्रत्यय से पुल्लिंग जातिवाचक संज्ञाएँ बनती हैं (देखिए तालिका ३३) ।

तालिका ३३ - गीर प्रत्यय से बने संज्ञा शब्द

मूल शब्द	पुरुषवाची प्रत्यय	निर्मित शब्द (पुल्लिंग जातिवाचक संज्ञाएँ)
उठाई	+ गीर	उठाईगीर
आलम	+ गीर	आलमगीर
जहान	+ गीर	जहाँगीर
राज	+ गीर	राजगीर
राह	+ गीर	राहगीर

- ची पुरुषवाची प्रत्यय

यह प्रत्यय संज्ञाओं में लगकर जातिवाचक-कर्तृवाचक संज्ञाएँ बनाता है । - ची मूलतः तुर्की प्रत्यय है ।
हिंदी में यह फ़ारसी से होकर आया (तिवारी 2016: 329) (देखिए तालिका ३४) ।

तालिका ३४ - ची प्रत्यय से बने संज्ञा शब्द

मूल शब्द	- ची पुरुषवाची प्रत्यय	निर्मित शब्द (जातिवाचक—कर्तृवाचक—संज्ञाएँ)
अफ़ीम	+ ची	अफ़ीमची
बंदूक	+ ची	बंदूकची
खज़ाना	+ ची	खज़ानची
मशाल	+ ची	मशालची
नक़ल	+ ची	नक़लची
तोप	+ ची	तोपची

३ निष्कर्ष


प्रस्तुत आलेख के द्वारा व्यावहारिक रूप से संज्ञाओं का व्याकरणिक लिंग पहचानने का प्रयास किया गया है।

व्यावहारिक रूप इसलिए कहना उपयुक्त है कि इसे सिद्धांत नहीं कह सकते क्योंकि सिद्धांत तो सिद्ध होता है परंतु मेरा प्रयास हिंदी के प्रयोक्ताओं के लिंग निर्धारण संबंधी समस्याओं को कुछ हद तक कम करना है।

उपर्युक्त प्रत्ययों में कुछ महत्वपूर्ण स्त्रीवाची एवं पुरुषवाची प्रत्ययों के उद्धरण दिए गए हैं। इनमें कुछ कृदंत प्रत्यय हैं, कुछ तद्धित प्रत्यय तथा कुछ कृदंत एवं तद्धित दोनों रूपों में प्रयुक्त प्रत्यय। इनमें भारतीय-आर्य भाषाओं एवं फार्सी-अरबी मूल के प्रत्यय हैं।

संदर्भ सूची

- गुरु, कामता प्रसाद 1920. *हिंदी व्याकरण*. काशी: नागरी प्रचारिणी सभा.
तिवारी, उदयनारायण 2016 [1950]. *हिंदी भाषा का उद्गम और विकास*. दिल्ली: लोकभारती प्रकाशन.
पाण्डेय, अनिल कुमार 2010. “हिंदी में प्रत्यय विचार (लिंग निर्धारण के विशेष संदर्भ में)”, *भाषा* 49,3. नई दिल्ली: केंद्रीय हिंदी निदेशालय.

Sandhya Singh 

Attitude Towards Learning Hindi in Singapore Universities: A Study

संध्या सिंह

सिंगापुर के विश्वविद्यालयों में हिंदी सीखने का रुझान: एक अध्ययन

Abstract Hindi is taught as a foreign language in two Singaporean universities. This reflects the interest in India and South Asia, which has increased in universities and the academic world over the last two-three decades. This interest has stemmed from India's growth especially in the rise of the purchasing power of her growing middle class. There are also various untapped opportunities in India. In Singapore, the students who take Hindi are diverse. Some students are second/third generations Indians. Native Singaporean ethnic Chinese and Malays also do take up the course. The motivations for doing so range from a desire to understand Bollywood films without subtitles to gaining a marketable skill set.

This paper seeks to explain this phenomenon by understanding the broad mechanisms and reasons behind the uptake of Hindi as a subject. The variation amongst students in terms of linguistic and cultural differences is analysed as a variable influencing their motivation to learn Hindi. Both primary and secondary data are utilized in doing so, ranging from existing literature reviews to interviews and surveys.

Keywords Indian diaspora, Hindi teaching in foreign countries, Hindi learning, culture, script.

सारांश सिंगापुर के दो विश्वविद्यालयों में हिंदी विदेशी भाषा के रूप में सिखाई जा रही है और कहीं न कहीं यह रुचि भारत और दक्षिण एशिया के प्रति उस रुझान का प्रतिबिंब है, जिसका प्रचलन विश्व भर के विश्वविद्यालयों और अकादमिक संसार में पिछले दो-तीन दशकों में बढ़ा है। भारत को एक उदय होती शक्ति के रूप में देखा जा रहा है और बाज़ारीकरण के इस युग में हिन्दुस्तान व हिन्दुस्तानियों के प्रभुत्व को नहीं नकारा जा सकता है। इन विश्वविद्यालयों में हिंदी लेने वाले कुछ दूसरी या तीसरी पीढ़ी के भारतीय भी होते हैं पर, एक बहुत बड़ा वर्ग बॉलीवुड प्रेमियों का भी है। साथ ही भारत की उभरती अर्थ-व्यवस्था, भारत की आर्थिक वृद्धि को भी रोज़गार या व्यापार की संभावनाओं के रूप में देखा जा रहा है।

अतः इस लेख में सिंगापुर के विश्वविद्यालयों में हिंदी सीखने वाले छात्रों के भिन्न पक्षों का विश्लेषण एवं हिंदी के प्रति उनके रुझान पर टिपण्णी की गई है। एक ओर जहाँ हिंदी सीखने वाले छात्रों के बीच भाषाई सांस्कृतिक भिन्नता है वहीं दूसरी ओर हिंदी पढ़ने की उनकी प्रेरणा में भिन्नता के बीच संबंध “परिवर्तनशीलता” का रूप है। साथ ही बहुत ही संक्षेप में हिंदी अध्यापन के कुछ पहलू रखे गए हैं। इस शोध के लिए प्राइमरी और सेकेंडरी दोनों तरह के आंकड़ों का इस्तेमाल किया गया है। साहित्य समीक्षा, साक्षात्कार, सर्वेक्षण, भिन्न हिंदी विभागों से बातचीत, व्यावहारिक परिप्रेक्ष्य में हिंदी का अध्ययन, संवादात्मक समारोहों का अवलोकन, समीक्षा, आदि माध्यमों का सहारा लिया गया है।

मुख्य शब्द – भारतीय डायस्पोरा, विदेशों में हिंदी शिक्षण, हिंदी सीखना, संस्कृति व लिपि।

१ परिचय

सिंगापुर को स्वतन्त्र राष्ट्र की संज्ञा सन् १९६५ में मिली। सन् १९५९ में सिंगापुर ब्रितानी साम्राज्य के अधीन एक स्वतंत्र राज्य बन गया था। कुआलालम्पुर सरकार के साथ बढ़ते मतभेद को शान्त करने के लिए सिंगापुर ने संपूर्ण स्वराज्य की ओर कदम बढ़ाया और ९ अगस्त १९६५ को पूर्ण स्वराज्य हासिल कर लिया (Chew & Lee 1991)।

सिंगापुर कुल ७१६.१ वर्ग किलोमीटर में फैला हुआ क्षेत्र है। जातीयता में अनोखापन सिंगापुर की एक महत्वपूर्ण खूबी है। सिंगापुर में कई धर्मों में विश्वास रखने वाले, विभिन्न देशों की संस्कृति, इतिहास तथा भाषा के लोग रहते हैं। मलय प्रायद्वीप होने के कारण मलय लोगों को यहाँ का मूल निवासी कहा जाता है लेकिन अब जनसंख्या में उनका प्रतिशत काफी कम है। २०१८ की जनसंख्या रिपोर्ट के मुताबिक सिंगापुर में सन् २०१८ की जनगणना के आधार पर ७४.३% चीनी, १३.४% मलय, ९.१% भारतीय और ३.२% मिश्रित जातियाँ हैं (Department of Statistics 2018)।

सन् १९६५ में सिंगापुर को जब पूर्ण स्वराज्य मिला तो कई चुनौतियाँ थीं; रोजगार, आवास आदि पर एक चुनौती थी यहाँ की बहुसंस्कृति को बनाए रखना। इसी बहुसांस्कृतिक समाज के निर्माण ने बहुभाषिता को बढ़ावा दिया। सिंगापुर में अंग्रेज़ी, चीनी, मलय और तमिल चार भाषाओं को आधिकारिक भाषा का दर्जा मिला था अतः अंग्रेज़ी को मुख्य भाषा तथा अन्य तीन को द्वितीय भाषा के रूप में पढ़ाने की शुरुआत हुई।

इस शोध में सिंगापुर की द्विभाषी नीति नहीं बल्कि दो मुख्य विश्वविद्यालयों में हिंदी भाषा को अपने पाठ्यक्रम में ऐच्छिक रूप से चुनने के कारणों, उनकी प्रेरणा और संबंध पर चर्चा की गई है। एक ओर जहाँ हिंदी सीखने वाले छात्रों के बीच भाषाई सांस्कृतिक भिन्नता है वहीं दूसरी ओर हिंदी पढ़ने की उनकी प्रेरणा में भिन्नता के बीच संबंध ‘परिवर्तनशीलता’ का रूप है।

२ सिंगापुर के विश्वविद्यालयों में हिंदी

सिंगापुर में वैसे तो स्थानीय विद्यालयों में भी हिंदी को द्वितीय भाषा के रूप में पढ़ाया जा रहा है पर इस शोध पत्र में सिर्फ विश्वविद्यालयों में हिंदी को विदेशी भाषा के रूप में सिखाने के पहलुओं पर प्रकाश डाला गया है। सिंगापुर के दो मुख्य विश्वविद्यालयों में हिंदी भाषा शिक्षण कार्यक्रम संचालित होता है; नेशनल यूनिवर्सिटी ऑफ सिंगापुर (एन यू एस) और नान्यांग टेक्नोलॉजिकल यूनिवर्सिटी

(एन टी यू)। ये दोनों विश्वविद्यालय एशिया में ही नहीं विश्व में महत्वपूर्ण स्थान रखते हैं। नेशनल यूनिवर्सिटी ऑफ़ सिंगापुर (एन यू एस) में हिंदी भाषा शिक्षण की शुरुआत सन् २००८ में हुई। नान्यांग टेक्नोलॉजिकल यूनिवर्सिटी (एन टी यू) की बात करें तो हिंदी भाषा शिक्षण सन् २०१४ से हो रहा है।

एन यू एस में हिंदी कला और सामाजिक विज्ञान संकाय की शाखा 'सेंटर फॉर लैंग्वेजेंज़ स्टडीज़' के अंतर्गत 'माईनर' और ऐच्छिक रूप में सिखाई जाती है। हिंदी के छह स्तर हैं जिन्हें प्राथमिक, मध्यम और उच्च तीन श्रेणियों में बाँटा गया है। ये छह स्तर विद्यार्थी छह सत्रों में पूरा करते हैं। यहाँ भारतीय भाषाओं में हिंदी के अलावा तमिल भाषा भी सिखाई जाती है पर वह सिर्फ़ ऐच्छिक रूप में ही उपलब्ध है क्योंकि कम रुचि के कारण सिर्फ़ प्राथमिक स्तर की कक्षाएँ ही चलाई जाती हैं।

एन टी यू में हिंदी मानविकी, कला और सामाजिक विज्ञान संकाय की शाखा 'सेंटर फॉर मोर्डन लैंग्वेजेंज़' के अंतर्गत ऐच्छिक विषय के रूप में सिखाई जाती है। एन टी यू हिंदी के दो स्तर हैं पर २०१४ से २०२० तक सिर्फ़ पहले स्तर की कक्षाएँ ही चलाई गई हैं जिसका कारण हिंदी के अगले स्तर में विद्यार्थियों की कम रुचि है। भारतीय भाषाओं में हिंदी के अलावा एन टी यू में भी तमिल भाषा की कक्षाएँ चलाई जाती हैं।

इस पत्र में कुछ ख़ास पहलुओं को ही उठाया गया है; सिंगापुर के विश्वविद्यालयों में हिंदी पढ़ने वाले छात्र आखिर हैं कौन? उनकी पृष्ठभूमि क्या है? क्या यहाँ भी अमरीका और यूरोपीय देशों की तरह 'हेरिटेज लर्नर' हैं? कक्षाओं में अलग छात्रों की पृष्ठभूमि का अनुपात क्या है? जैसा कि ज़्यादातर लोग जानते हैं कि सिंगापुर की जनसंख्या में मलय, चीनी, भारतीय और अन्य जिसमें यूरेशियन आदि मुख्य हैं, चार जातियों में विभाजित वर्ग हैं तो इस भिन्न पृष्ठभूमि का कोई अलग प्रभाव है? संक्षेप में यह भी चर्चा की गई है कि हिंदी शिक्षण के लिए किस प्रकार की सामग्री का इस्तेमाल किया जाता है और उसका भाषा सीखने में क्या प्रभाव दिखाई देता है?

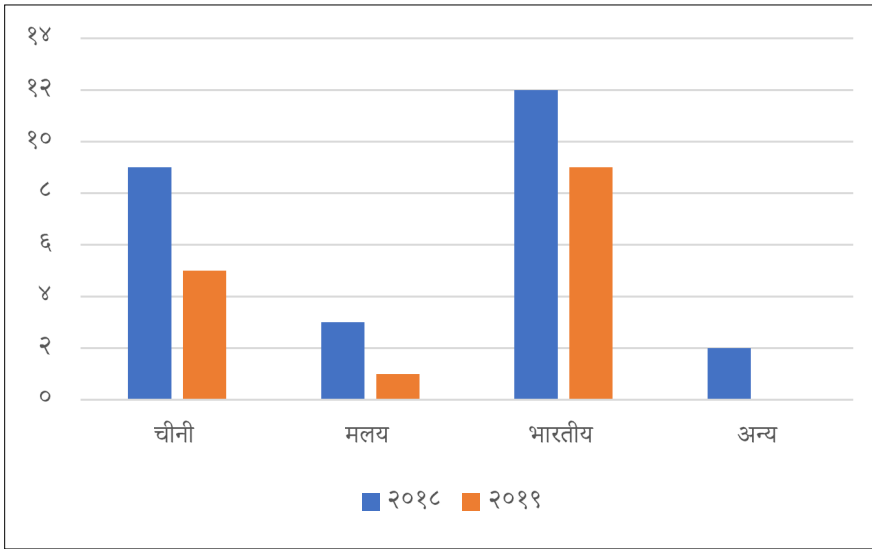
३ हिंदी सीखने वाले: जातीय परिप्रेक्ष्य

इस विश्लेषण के लिए कुछ आँकड़े दो वर्षों के भीतर इकट्ठे किये गए। इन आँकड़ों में मुख्य रूप से प्रश्नावली और छात्रों के विचार सम्मिलित हैं। छात्रों द्वारा की गई टिप्पणी भी शिक्षण सामग्री में इस्तेमाल की गई है। निष्कर्ष तक पहुँचने के लिए मुख्य रूप से एकत्रित आँकड़ों का विश्लेषण करने की आवश्यकता है इसलिए पहले कुछ आँकड़े देखते हैं ताकि यह अनुमान लगाया जा सके कि यहाँ कक्षाओं में किस पृष्ठभूमि से छात्र अधिक हैं, क्या इसका कोई ख़ास कारण है और कोई प्रभाव शिक्षण तकनीक पर है?

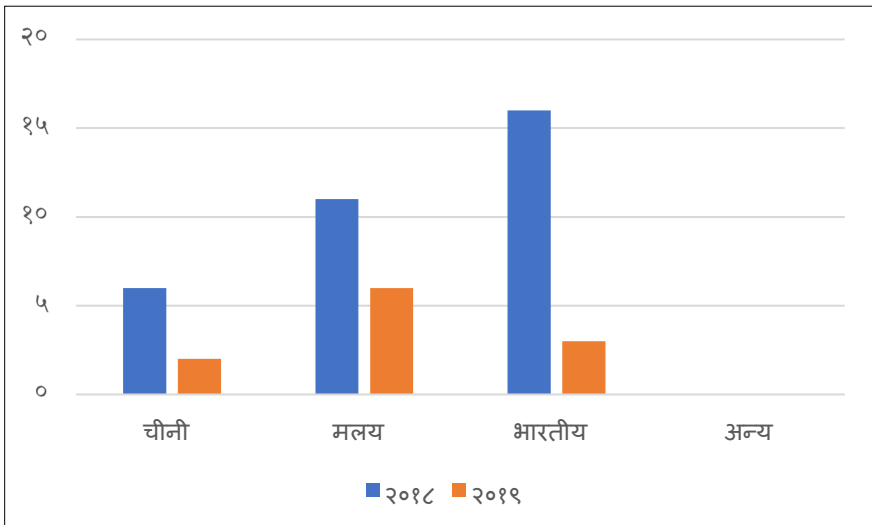
इस विश्लेषण के लिए काफी बारीकी से निरीक्षण, तुलना और सारांश, समेकन आदि किया गया है जिसकी अनुशंसा क्रेस्वाल और सिल्वरमैन ने भी की है। (Creswell 2002; Silverman 2001) हालाँकि सिर्फ़ तीन सत्रों का सहारा लिया जा रहा है पर ज़्यादातर रुझान इसी प्रकार का होता है। यद्यपि यहाँ कुछ संख्या दी गई है फिर भी उसे लगभग के रूप में ही प्रयोग किया जाएगा क्योंकि कई बार पाठ्यक्रम के बीच में कुछ छात्र जुड़ते या छोड़ देते हैं।

निम्न रेखाचित्र १ और रेखाचित्र २ के माध्यम से अलग जातीय पृष्ठभूमि वाले छात्रों की संख्या स्पष्ट की गई है।

इन दोनों आँकड़ों के अध्ययन से यह तो स्पष्ट है कि दोनों ही विश्वविद्यालयों में भारतीय छात्रों की संख्या अधिक है पर विश्वविद्यालय २ में मलय छात्रों की संख्या भी काफी है। सिंगापुर की



रेखाचित्र १ विश्वविद्यालय १- अलग जातीय पृष्ठभूमि वाले छात्र



रेखाचित्र २ विश्वविद्यालय २- अलग जातीय पृष्ठभूमि वाले छात्र

जनसंख्या में चीनी प्रतिशत अधिक है लेकिन हिंदी कक्षाओं की ओर उनका रुझान कम है, आखिर इसके क्या कारण हैं? इन कारणों पर भी चर्चा यहाँ की गई है।

सिंगापुर के दोनों ही विश्वविद्यालयों में अन्य भाषाओं की तुलना में हिंदी सीखने वालों की संख्या कम है। यूरोपीय भाषाएँ और जापानी, चीनी, कोरियाई, बहासा इंडोनेसिया जैसी भाषाएँ बहुत लोकप्रिय हैं। आखिर ऐसा क्या है जो हिंदी सीखने को प्रोत्साहित नहीं करता।

३.१ हिंदी सीखने वाले मलय छात्र

मलय छात्रों की संख्या भी कहीं कम तो कहीं अधिक है। जब मलय छात्रों से सवाल पूछे गए कि क्या वे हिंदी भाषा सीखना चाहते हैं और क्यों? उनका उत्तर स्पष्ट था कि वे आज भी बॉलीवुड के दीवाने हैं और इस फ़िल्मी प्रेम ने ही हिंदी भाषा की ओर उनको खींचा है। हालाँकि फिल्मों में 'सबटाइटल' होते हैं लेकिन वो मज़ा नहीं आता है जब तक कि कुछ बातें समझ न आएँ। अब बॉलीवुड के साथ ही अरब संस्कृति का बोलबाला इन मलय छात्रों में अधिक बढ़ा है। वे अरबी भाषा सीखने को ज़्यादा महत्त्व देने लगे हैं। चूँकि मलय भाषा की लिपि रोमन ही है तो उन्हें अरबी लिपि सीखने में अधिक उत्साह नज़र आता है।

यह भी सत्य है कि ज़्यादातर मलय छात्र धार्मिक कक्षाओं में अरबी भाषा के संपर्क में आते हैं और जब विश्वविद्यालय में उन्हें पुनः अरबी सीखने का मौक़ा मिलता है तो उसे पाना चाहते हैं। और इस फ़ेहरिस्त में हिंदी के प्रति आकर्षण ढीला पड़ जाता है। कुछ छात्रों के अनुसार –

“अब शाहरुख़ खान, सलमान खान और आमिर खान का ज़माना भी ख़त्म हो रहा है तो हिंदी सीखने का शौक भी कम पड़ रहा है। कोई उनके जैसा चाहिए जो हमें फिर से हिंदी की ओर ले चले।”

३.२ हिंदी सीखने वाले चीनी छात्र

अब बात आती है चीनी छात्रों की संख्या की। हिंदी कक्षाओं में चीनी छात्रों की संख्या काफ़ी कम रहती है। चीनी भाषा में भी लिपि सीखनी होती है और हिंदी में भी। तो कुछ छात्र कहते हैं –

“हम ऐसी भाषा सीखना चाहते हैं जिसमें लिपि न सीखनी पड़े। लिपि सीखने में काफ़ी समय चला जाता है और बोलना, बातचीत करना बहुत नहीं सीखा जा सकता। इसलिए मैं हिंदी नहीं सीखना चाहता था लेकिन फिर सोचा चलो देखता हूँ। अब लगता है यह ग़लतफ़हमी है। असल में हिंदी लिपि सीखनी आसान है।”

हिंदी के कम लोकप्रिय होने का दूसरा कारण सिंगापुर में भारत की भाषा तमिल का सरकारी कामकाज की एक भाषा होना है। यहाँ के चीनी छात्रों को और जो छात्र बाहर के विश्वविद्यालयों से आते हैं, उन्हें तमिल सीखना बेहतर विकल्प लगता है। उनका तर्क है कि अगर भारत की कोई भाषा सीखनी है तो तमिल सीखनी चाहिए क्योंकि तमिल यहाँ प्रयोग कर सकते हैं। भारत कहाँ जाना है कि हिंदी का इस्तेमाल करेंगे।

हालाँकि तमिल सीखने वालों की संख्या भी अधिक नहीं है। यह कहीं न कहीं भारत के प्रति विकसित एक सोच का नतीजा भी है। सिंगापुर में भारत से शुरुआती दौर यानी १९४०-५० में मज़दूर वर्ग ही अधिक आया और बाद के वर्षों में भी कम पढ़े-लिखे लोगों का आगमन कई छोटे कार्यों के लिए भी रहा। इस शुरुआती सामाजिक ढाँचे ने भारत और भारतीयों के प्रति बहुत सकारात्मक सोच नहीं पनपने दी जिसका कुछ असर आज तक दिखाई देता है। यह भी एक कारण है भारतीय भाषाओं के प्रति कम रुचि का।

भारत बहुत सुन्दर देश है लेकिन वहाँ के दर्शनीय स्थल भी सिंगापुर के लोगों को उतना नहीं लुभा पाते जितना यूरोप और अन्य देशों के। जो छात्र हिंदी सीख रहे हैं उनमें से कई भारत जाना चाहते हैं या हिंदी भारत घूमने के लिए ही सीख रहे हैं लेकिन बड़ी संख्या को हिंदी कक्षाओं तक ले आ पाने में भारत का पर्यटन उद्योग अभी पीछे है।

३.३ हिंदी सीखने वाले अहिंदी भाषी भारतीय छात्र

अगर हम रेखाचित्र को देखें तो यह स्पष्ट हो जाएगा कि हिंदी सीखने वालों की संख्या अधिक नहीं है लेकिन उन छात्रों में भारतीय मूल के छात्रों की दोनों ही विश्वविद्यालयों में सबसे बड़ी संख्या है। आखिर भारतीय क्यों हिंदी सीखना चाहते हैं? क्या ये भी यूरोप और अमरीका की तरह हेरिटेज लर्नर हैं? सिंगापुर के दोनों ही विश्वविद्यालयों में हिंदी को सिर्फ विदेशी भाषा के रूप में ही सिखाया जाता है। यहाँ एक ही कक्षा में दो अलग तरह के छात्र न हों यही कोशिश की जाती है। अगर छात्र हिंदी ज्ञान छिपाकर कक्षा में बैठते हैं और इसका पता चल जाता है तो विश्वविद्यालय बहुत कड़ा रुख अपनाते हैं। तो आखिर ये भारतीय कौन हैं? सिंगापुर में भारतीयों में बहुत बड़ा समुदाय तमिल लोगों का है। भारत की तरह इन तमिल लोगों को हिंदी नहीं आती क्योंकि वे विद्यालय में तमिल दूसरी भाषा के रूप में पढ़ सकते हैं और हिंदी पढ़ने, सीखने का कोई कारण या ख़ास अवसर नहीं रहता। यही भारतीय हिंदी सीखने वालों में बड़ी संख्या बनाते हैं। इनमें से कई पूरी तरह से सिंगापुर में ही बसे हैं और कभी भारत जाना भी नहीं होता तो उनके लिए हिंदी ठीक वैसी ही है जैसी किसी विदेशी भाषी के लिए। वे बॉलीवुड फ़िल्में अवश्य देखते हैं पर उतना ही समझते हैं जितना कोई भी हिंदी भाषी तमिल फ़िल्म देखते समय समझेगा। भारतीय विरासत से जुड़े होने के कारण हिंदी न समझ पाने की कमी कहीं न कहीं खलती है। इसी कमी को पूरा करने के लिए अहिंदी भाषी भारतीय छात्र हिंदी सीखने आते हैं। हिंदी कक्षाओं में इनका अनुपात अधिक होता है क्योंकि दक्षिण भारत से तमिल, तेलगू, कन्नड़ कई भाषा-भाषी लम्बे समय से यहाँ हैं और हिंदी को आज भारत की प्रतिनिधि भाषा के रूप में सभी देख रहे हैं तो वे भी इस दौड़ में पीछे नहीं रहना चाहते। साथ ही सिंगापुर में हिंदी फ़िल्मों के प्रति दक्षिण भारतीय समुदाय में भी गहरी रुचि है जो उन्हें कई बार हिंदी सीखने को प्रेरित करती है। एक और कारण आजकल उत्तर भारतीयों की बढ़ती संख्या भी है।

इस भिन्न पृष्ठभूमि का शिक्षण पर किस तरह से असर होता है यहाँ यह भी जानना ठीक होगा। भारतीय पृष्ठभूमि वाले छात्रों को भारतीय संस्कृति से संबंधित काफ़ी ज्ञान होता है तो ज़ाहिर है भाषा सीखने का सांस्कृतिक पक्ष उन्हें पहले से ज्ञात होता है, तो यह आसान बन जाती है। तमिल और हिंदी के कुछ स्वर समान हैं तो पहली कुछ कक्षाएँ उनके लिए आसान हो जाती हैं। हिंदी फ़िल्में भी चीनी छात्रों के मुक़ाबले वे ज़्यादा देखते हैं तो कुछ-कुछ शब्द भी जल्दी पकड़ लेते हैं। वैसे ये अलग बात है कि परीक्षा में चीनी छात्र कई बार अधिक अच्छा करते हैं।

रोमेन ने ५ तथ्य सुझाए हैं जो भाषा के बदलाव यानी भाषा सीखने के लिए प्रेरित करने में कहीं न कहीं महत्वपूर्ण कारक हैं (Romaine 1989: 42):

१. भाषा का दर्जा
२. मनोभाव या रवैया
३. विद्यालय की भूमिका व भाषानीति
४. घर पर भाषा संरक्षण न कर पाना
५. नई पीढ़ी द्वारा अपूर्ण शिक्षा

रोमेन के इन पाँच तथ्यों पर ध्यान देने से यह तो स्पष्ट है कि भारत में हिंदी को कामकाज आदि से जोड़कर इस्तेमाल करने की स्थिति जब तक और अधिक बेहतर नहीं होगी तो विदेशी भाषा के रूप में हिंदी सीखने वालों की संख्या में भी बढ़ोतरी नहीं होगी। हिंदी के प्रति सकारात्मक सोच विकसित करने की ज़रूरत है। भाषा को रोज़गार या देश की मुख्य धारा से जोड़े बिना उसके प्रति रवैये को बदला नहीं जा सकता। 'रोमेन' के चौथे व पाँचवें बिंदु पर अगर ध्यान दें तो अंग्रेज़ी के बढ़ते चलन व वर्चस्व ने भारत में हिंदी को प्रभावित किया है। आज भारत में ख़ासकर शहरों में लोग कई वाक्यों में अंग्रेज़ी शब्दों का समावेश कर ही देते हैं। भले ही हिंदी पढ़ रहे हैं पर भविष्य में उसकी उपयोगिता बातचीत से ज़्यादा कहीं नज़र नहीं आती। आज घरेलू भाषा का छात्रों के जीवन में कोई व्यावहारिक पहलू नहीं है। न तो स्थानीय न राष्ट्रीय स्तर पर इसका कोई मूल्य है। भाषा में विद्वत्ता हासिल करने पर भी कोई सामाजिक या नौकरी संबंधी लाभ दिखाई नहीं देता इसलिए भारत में हिंदी के प्रति रवैया जब ठीक नहीं है तो कोई विदेशी क्यों सीखना चाहेगा? छात्रों को तो यही लगता है कि अगर कभी भारत घूमने गए तो अंग्रेज़ी से काम चल जाएगा।

भाषा समय के साथ बदलती रहती है क्योंकि अपने साथ जोड़ने व त्यागने का भाव भाषा बड़ी आसानी से अपना लेती है। भाषा को बरकरार रखने के कई कारकों में से एक रिश्तेदारों से मिलने अपने देश जाना है। सिंगापुर में भी यह बात स्पष्ट होती है क्योंकि जिन भारतीय छात्रों से विचार लिए गए उनमें से ३० प्रतिशत पिछले पाँच वर्षों में कम से कम एक बार भारत अवश्य गए हैं। भारतीय होने के सूत्र की मजबूती भारत में अपने रिश्तेदारों से मिलकर ज़्यादा दृढ़ हो जाती है। उस समय हिंदी भाषा अपनी महत्वपूर्ण भूमिका निभाती है क्योंकि भले ही आज भारत विदेश हो गया हो, जहाँ अंग्रेज़ी का बोलबाला बढ़ गया है, पर दक्षिण में भी कई बार उन्हें हिंदी की थोड़ी बहुत ज़रूरत पड़ जाती है। दोरियन के अनुसार भाषा को बरकरार रखने के दो महत्वपूर्ण कारक व्यावहारिकता एवं सामाजिक लाभ भी हैं (Dorian 2014: 207)।

इसका अर्थ स्पष्ट है कि अगर भाषा के द्वारा रोज़गार नहीं है या कोई सामाजिक लाभ नहीं है तो उस भाषा को सीखने का चाव कम हो जाता है। सिंगापुर में हिंदी सीखकर रोज़गार की कल्पना अभी रूप नहीं ले सकती और विश्व के अन्य देशों में भी हिंदी के बल पर काम नहीं मिलता। हालाँकि यूरोपीय और अन्य एशियाई भाषाओं में यह संभव है जिसकी वजह से उनके प्रति अलग दृष्टिकोण है। हालाँकि सिंगापुर में कई बार प्रेम प्रसंग के कारण यह भाषा सीखी जा रही है। आज अंतर्राज्यीय विवाह की प्रथा अधिक है और प्रेम जिससे हो जाए उसके लिए आज का युवा बहुत कुछ करने को तैयार रहता है। कभी-कभी भाषाई भिन्नता वाले माता-पिता हों तो ऐसी 'कोड' शब्दावली का प्रयोग करना चाहते हैं

जिसे सिर्फ वही समझे जिसको समझाना उनका उद्देश्य होता है और वहाँ भी हिंदी एक छोटी भूमिका निभाती है। आजकल बहुभाषावादी होना कई लोगों के लिए इसी प्रकार की सहूलियत दे रहा है।

४ हिंदी सीखने वाले छात्रों के विचार और हिंदी रुझान

जो छात्र हिंदी सीख रहे हैं जब उनसे विचार लिए गए कि आखिर उन्होंने हिंदी क्यों चुनी और यह भाषा भविष्य में कैसे काम आएगी? आँकड़ों को बार-बार विश्लेषित किया गया और कई तरीके से तुलना भी की गई (Silverman 2001)। उनके उत्तरों में कुछ भिन्नता देखने को मिली। ये विचार पहले स्तर पर हिंदी सीख रहे छात्रों से लिए गए हैं। छात्रों ने विचार अंग्रेजी में दिए जिसका हिंदी अनुवाद किया गया है। छात्रों के विचार प्रश्नावली के साथ ही, बड़े वाक्यों और बातचीत के अलग रूप में थे। अतः उसका बारीकी से विश्लेषण किया गया (Creswell 2002)। यह नमूना चालीस छात्रों के विचारों पर आधारित है। नमूने का आकार बहुत बड़ा नहीं है पर जो निष्कर्ष निकले हैं, उनका काफी हद तक व्यापक स्तर पर इस्तेमाल किए जा सकते हैं। उनके उत्तरों में भिन्नता है और काफी विस्तृत क्षेत्र को दिखाता है।

५ हिंदी सीखने के कारण

छात्रों से पूछे गए प्रश्न—“आप हिंदी क्यों सीख रहे हैं?” और “यह भाषा भविष्य में कैसे काम आएगी?”—के संबंध में जो विचार आए, उनका विश्लेषण किया गया तो कुछ पहलू बार-बार दिखे पर कुछ नई बातें भी दिखीं। कुछ आम चीजें हर भाषा के सीखने के पीछे कारण होती हैं। कुछ बातें हर भाषा के सम्बन्ध में बदल जाती हैं। ज़्यादातर छात्रों ने एक से अधिक कारण दिए, जिनकी वजह से वे हिंदी के प्रति खिंचे हुए हैं।

५.१ बॉलीवुड

हिंदी भाषा सीखने के आम कारणों में सबसे अधिक छात्रों ने फ़िल्मी प्रेम ज़ाहिर किया। लगभग ६० प्रतिशत छात्र भारत की फ़िल्मी हस्तियों और फ़िल्मी दुनिया से इस क़दर प्रभावित हैं कि भाषा सीखना चाहते हैं। हिंदी फ़िल्में और गीत भी बहुत आकर्षित करते हैं। बॉलीवुड ने हिंदी भाषा के प्रति एक सकारात्मक रवैये का विस्तार किया है। फ़िल्मी संवाद समझना, गीतों के बोल समझना फ़िल्म देखने को अधिक दिलचस्प बनाता है। ‘सबटाइटल’ कई बार अर्थ का भाव और उसकी गहराई नहीं दे पाते इसलिए छात्र थोड़ी सी हिंदी ही सही पर सीखना चाहते हैं। यहाँ सिर्फ़ दो-तीन उदाहरण दिए जा रहे हैं पर लगभग सभी ने इसी तरह की बात कही है।

“हिंदी फ़िल्में देखने की शुरुआत बहुत पहले हुई और मैं तब से ही हिंदी सीखना चाहता था क्योंकि मुझे गीतों में बहुत अधिक दिलचस्पी है। और एक दिन उत्तर भारत ज़रूर जाना होगा तो यह भाषा मेरी मदद करेगी।”

“मुझे हिंदी फ़िल्में देखना और गाने सुनना बहुत अच्छा लगता है।”

“हिंदी फ़िल्में बचपन से देख रहा हूँ तो जब मौक़ा मिला तो लगा हिंदी सीखनी चाहिए। हिंदी सुनने में बहुत अच्छी लगती है जैसे ‘फ़राज़’ फिल्म में जो कविता है, अब मैं कुछ समझ सकता हूँ तो मुझे बहुत अच्छा लगता है।”

बॉलीवुड के प्रति आकर्षण भारतीय और मलय पृष्ठभूमि वाले छात्रों में ही दिखी। हाँ दो-तीन ऐसे छात्र भी मिले जो चीनी हैं लेकिन बॉलीवुड से प्रेरित और आकर्षित हैं। उनमें से एक छात्र कुछ वर्षों तक पिताजी की नौकरी के कारण कोलकाता में रह चुका है। भारतीय परिवेश में रहने के कारण उसे हिंदी, बांग्ला और उड़िया के प्रति भी लगाव है। उसने हिंदी कभी सीखी नहीं इसलिए अब सीख रहा है पर हिंदी फ़िल्में देखी हैं और गाने सुनता है। भारतीय संस्कृति से भी वह बहुत अधिक प्रभावित है जिसका प्रमाण उसके नाम में दिखता है। वह चीन से है पर उसने अपना उपनाम अर्जुन रखा है और हिंदी कक्षा में सब उसे अर्जुन ही पुकारते हैं। दूसरी चीनी छात्रा को अलग-अलग संगीत पसंद है। वह कई भाषाओं के गीत सुनती है और इस तरह उसे बॉलीवुड से प्रेम है।

५.२ नई भाषा सीखना

आजकल नई भाषा सीखने का चलन अधिक है। ऐसा नहीं है कि पहले नहीं था पर तब अवसर कम थे तो जो लोग वाकई भाषा-संस्कृति से प्रेरित होते थे, ज़्यादातर वही सीखते थे। कई बार उस भाषा में वे महारत भी हासिल करते थे। आज अवसर अधिक हैं। छात्र आसानी से विश्वविद्यालयों में नई भाषा विशेषज्ञों से सीख सकते हैं। लिए गए विचारों में से कई लोग सिर्फ़ एक नई भाषा सीखने की वजह से हिंदी सीख रहे हैं। हाँ उन्हें देवनागरी लिपि ने अवश्य सम्मोहित किया है।

“नई भाषा सीखने से ‘कौग्नितिव स्किल्स’ बेहतर होता है तो मैं नई लिपि वाली भाषा सीखना चाहता था ताकि हिंदी लिपि भी पढ़ सकूँ और कौशल भी बढ़ा सकूँ।”

“चीनी भाषा से यह बहुत अलग है इसलिए मेरे लिए सीखना बहुत दिलचस्प है।”

५.३ भारत भ्रमण और भारत को जानना

भारत घूमना और लोगों से बातचीत करना भी हिंदी भाषा सीखने वाले छात्रों के बीच काफी लोकप्रिय जवाब रहा। भ्रमण और भाषा का काफी पुराना संबंध रहा है। भ्रमण के दौरान भाषा की अनिवार्यता से सभी सहमत हैं और लोगों ने किसी भी नए देश की सैर से पहले उस देश की भाषा में कुछ वाक्य सीखने की सलाह दी है। आज दुनिया अधिक छोटी हो गई है और ज़्यादा से ज़्यादा लोग दूसरे देशों की सैर करना चाहते हैं। हालाँकि भारत जितना सुन्दर और विविधता भरा है, वह अपना उतना प्रचार नहीं कर पाया है। आज भी भ्रमण के लिए लोगों की पसंद में भारत पहले स्थान पर नहीं आता। भारत भ्रमण के आकर्षण के कुछ खास कारण ही ज़्यादातर दिखाई देते हैं जिनमें संस्कृति मुख्य है। धीरे-धीरे अब यह क्षितिज भी व्यापक हो रहा है और भारत भ्रमण के कारणों में

कुछ भिन्नता भी दिखने लगी है। यहाँ चीनी छात्रों में भी भारत भ्रमण का रुझान दिखा। भारतीय मूल के छात्रों में भी उत्तर भारत देखने की चाह स्पष्ट दिखी और इन्हीं कारणों से वे हिंदी कक्षा में उपस्थित हैं।

“मैं एक दिन भारत घूमना चाहता हूँ। हिंदी सीखने से वहाँ के लोगों की बात शायद कुछ-कुछ समझ पाऊँ और सबसे ज़रूरी है इससे मुझे भारतीय संस्कृति का भी ज्ञान मिल रहा है।”

“हिंदी सीखने से मैं लोगों से बातचीत कर पाऊँगी। हिंदी भारत में सबसे ज़्यादा बोली जाने वाली भाषा है और सिंगापुर में भी काफ़ी लोग बोलते हैं। मेरी हिंदी गीतों में बहुत रुचि है”।

५.४ भारतीय संस्कृति और देवनागरी लिपि

भारतीय संस्कृति के प्रति आकर्षण प्राचीन समय से ही पश्चिम के लोगों में काफी रहा है। वाराणसी के घाटों पर विदेशियों की भीड़, कहीं किसी गली में रहकर संस्कृत और हिंदी सीखते लोगों के बारे में काफी कुछ लिखा जा चुका है। भारतीय संस्कृति में जितनी भिन्नता है उतनी शायद ही किसी और संस्कृति में हो। भारतीय संस्कृति के नज़दीक आने के लिए छात्र अन्य भारतीय भाषाओं के बजाय हिंदी चुनते हैं। उन्हें यह अहसास है कि हिंदी भाषा ही भारत की संस्कृति को समझने में काम आएगी। धर्म सिंगापुर में छात्रों को अधिक प्रभावित नहीं करता पर संस्कृति के कई पक्ष बहुत अधिक प्रभावित करते हैं। संस्कृति जानने की अधिक चाह चीनी और कुछ अहिंदी भाषी भारतीय छात्रों में दिखी। मलय छात्र संस्कृति से कुछ कम प्रभावित दिखे। संस्कृति के साथ ही देवनागरी लिपि भी आकर्षण का एक पहलू है। हालाँकि कई लोग लिपि के चलते हिंदी कक्षा की ओर देखते भी नहीं पर अगर वे आ जाते हैं तो उनका दृष्टिकोण भी बदलता दिखा है। देवनागरी लिपि आसानी से सीखी जा सकती है और जब छात्र लिपि सीख लेते हैं तो एक अलग उपलब्धि का अहसास होता है। पहले स्तर की कक्षाओं में शुरू के कुछ हफ़्तों में दोनों ही विश्वविद्यालयों में लिपि सिखा दी जाती है। छात्र ‘ट्रान्सलिट्रेशन’ यानी रोमन लिपि के द्वारा सही उच्चारण करने की विधि का इस्तेमाल कर सकते हैं पर देखा गया है जब उन्हें देवनागरी लिपि समझ आ जाती है तो वे भाषा के प्रति अधिक उत्साहित नज़र आते हैं। वे लिपि का अधिक से अधिक इस्तेमाल करना भी चाहते हैं। हाँ जब कक्षा में ‘एक्टिविटी’ के लिए कुछ जल्दी में लिखना होता है तो वे कई बार ‘ट्रान्सलिट्रेशन’ का इस्तेमाल करते हैं।

“भारतीय संस्कृति में ‘रिचनेस’ है मैं उसके बारे जानना चाहती हूँ।”

“भारत की संस्कृति और संगीत मुझे बहुत आकर्षित करते हैं।”

“भारतीय उपमहाद्वीप की कविताएँ मुझे बहुत आकर्षित करती हैं। अंग्रेज़ी में पढ़कर मज़ा थोड़ा कम होता है इसलिए हिंदी सीख रहा हूँ। मेरी भारतीय इतिहास और राजनीति में भी बहुत रुचि है।”

“मैं हिंदी सीख रहा हूँ क्योंकि इसकी लिपि बहुत सुंदर है और मैं संस्कृति को समझना चाहता था।”

जहाँ एक ओर छात्रों ने भाषा सीखने में लिपि को रुकावट की तरह सोचा वहीं लिपि के कारण भाषा सीखने की बात लगभग २० प्रतिशत छात्रों ने की जो यह दिखाता है अलग लिपियाँ कई बार अधिक आकर्षित करती हैं। संस्कृति के प्रति खिंचाव उसकी भिन्नता और विविधता के कारण है। साथ ही सिंगापुर में हाल के कुछ वर्षों में उत्तर भारतीय संस्कृति काफी ज़ोर पकड़ रही है और लोगों को आकर्षित कर रही है। होली का त्योहार एक अच्छा उदाहरण है। आज सिंगापुर में कई जगहों पर होली का बड़ा आयोजन किया जाता है जिसमें चीनी, अमेरिकी, यूरोपीय, अन्य कई देशों के लोग, अहिंदी भाषी भारतीयों के साथ ही कुछ मलय छात्र भी नज़र आते हैं। इस तरह से संस्कृति का एक पहलू भाषा से जुड़ जाता है और हिंदी और संस्कृति के बारे में कुछ बातें तो अवश्य ही पता चलती हैं।

५.५ दोस्तों एवं रिश्तेदारों से बातचीत

दोस्तों और रिश्तेदारों से बात करना भी हिंदी सीखने के कारणों में से एक खास कारण है। जैसा कि शुरू में ही हम देख चुके हैं कि सिंगापुर में हिंदी सीखने वाले छात्रों में एक बड़ा वर्ग अहिंदी भाषी भारतीयों का है। उनमें से कई छात्रों के रिश्तेदार भारत में रहते हैं और कई बैंगलुरु, मुंबई, पुणे आदि स्थानों में रहते हैं। वहाँ हिंदी बोलने का चलन है और छात्रों को लगता है कि जब वे उनसे मिलते हैं तो हिंदी न जानने के कारण कई बातें, मज़ाक, कई चीज़ों में पीछे रह जाते हैं। वे सीखकर इस बाधा को दूर करना चाहते हैं।

“मेरे दोस्त हिंदी भाषी है और हर समय हिंदी में बात करते हैं जब उन्हें लगता है कि यह दूसरों को न समझ आए। मैं चाहती हूँ कि हिंदी सीख लूँ तो उनकी बातें समझ पाऊँगी और उनकी संस्कृति भी अधिक समझना चाहती हूँ।”

“मेरे काफ़ी रिश्तेदार भारत में रहते हैं और हिंदी बोलते हैं। मैं उनसे हिंदी में बात करना चाहती हूँ।”

“अब सिंगापुर में बहुत लोग हिंदी बोलते हैं इसलिए हिंदी सीखना उपयोगी होगा।”

दोस्तों की बातें अगर न समझ आए तो बड़ी खीझ मचती है और ऐसा यहाँ भी होता है। सिंगापुर को ‘एजुकेशन हब’ की संज्ञा दी जाती है। यह भारत से नज़दीक है और यहाँ के विश्वविद्यालयों की गणना बहुत अच्छे विश्वविद्यालयों में की जाती है। शायद यही कारण है कि सिंगापुर में भारत से आकर पढ़ने वालों की संख्या काफी है। वे ज़्यादातर हिंदी बोलते हैं। कई बार अपने दोस्तों की बातें, चुगली आदि समझने के लिए हिंदी सीखने आते हुए देखा गया है। दोस्ती कभी-कभी प्यार में बदल जाती है और कुछ छात्र अपने भावी जीवनसाथी, उसकी संस्कृति को और बेहतर तरीके से समझने के लिए भी हिंदी सीखते हैं।

५.६ भारत की अर्थव्यवस्था और रोज़गार के अवसर

भारत की अर्थव्यवस्था और रोज़गार के अवसर यह रुझान हाल के वर्षों में देखने को मिला है। चीनी छात्र भी भारत को उभरती हुई सबसे बड़ी अर्थ व्यवस्था के रूप में देखने लगे हैं। भारत में काम करने को लेकर आज भी बहुत उत्साह नहीं है लेकिन इस विकल्प को कोई छोड़ना भी नहीं चाहता। कंप्यूटर जैसे कोर्स पढ़ने वाले छात्र ये मानते हैं कि भविष्य में उन्हें भारतीयों के साथ भी काम करना हो, तो ऐसे में हिंदी काम आएगी। भारत का बाज़ार भी बहुत बड़ा है। रोज़गार के अवसर मिलने पर भाषा को रुकावट नहीं बनने देना चाहते। भारत पर सबकी नज़र है और सभी हिंदी को ही भारत की भाषा के रूप में देखते हैं। यह भी यथार्थ है कि भारत में अंग्रेज़ी के बल पर काम किया जा सकता है इसलिए हिंदी सीखने वालों की संख्या अधिक नहीं है। सबको यह मालूम है कि हिंदी की ज़रूरत काम के लिए भले न हो पर बाज़ार और वहाँ लोगों एवं संस्कृति को समझना भी आवश्यक है इसलिए हिंदी सीखने की कोशिश की जा रही है।

“भारत में काम करना भी एक विकल्प है तो हिंदी भाषा उसके लिए ज़रूरी है।”

“आज अगर देखा जाए तो भारत को विश्व की सबसे शक्तिशाली और तेज़ी से आगे बढ़ती हुई अर्थव्यवस्था के रूप में देखा जा रहा है। और भारत में काम के बहुत मौक़े बन रहे हैं इसलिए संवाद के लिए कुछ आधार होना चाहिए। भाषा पूरी न भी आए तो कम से कम संस्कृति की बेहतर समझ हो ताकि भारतीयों के साथ अच्छे कामकाजी संबंध बन सकें।”

“चीन के बाद भारत की अर्थव्यवस्था काफ़ी तेज़ी से बढ़ रही है इसलिए हिंदी सीखना बहुत उपयोगी होगा।”

“मैं कंप्यूटर साइंस का छात्र हूँ और भारत के लोग भी इसमें बहुत अच्छे होते हैं। अब भारत में काम के अवसर भी बढ़ रहे हैं तो शायद एक दिन मुझे भी किसी भारतीय के साथ काम करना हो। और ऐसे में हिंदी काम आएगी।”

६ हिंदी अध्यापन सामग्री


हिंदी सीखने के क्या-क्या मुख्य आकर्षण हैं इन पर चर्चा करने के बाद एक और बात सामने आती है और वह है हिंदी अध्यापन में प्रयुक्त सामग्री और उनका प्रभाव। हालाँकि यह विषय काफ़ी विस्तृत है और इस पर पूरे एक शोध पत्र की आवश्यकता है इसलिए यहाँ बहुत ही संक्षेप में अध्यापन की बात की जायेगी। क्या अध्यापन सामग्री का प्रभाव भी छात्रों की संख्या और रुझान पर पड़ता है? अध्यापन के लिए सिर्फ एक ही विश्वविद्यालय के आँकड़ों का अभी इस्तेमाल किया जा रहा है। इसके लिए छात्रों की टिप्पणी का प्रयोग किया गया है। अगर संक्षेप में कहा जाए तो छात्र भाषा के साथ ही संस्कृति के प्रति अधिक उत्साहित दिखते हैं। उनके लिए सिर्फ़ किताबों का इस्तेमाल पर्याप्त नहीं होता। वे उस भाषा के समाज को असली रूप में देखना चाहते हैं। जैसे अन्य सभी देशों में होता है वैसे ही सिंगापुर में भी कुछ निर्धारित पुस्तकें हैं और उनके अलावा मीडिया और तकनीकी

का भरपूर प्रयोग होता है। ऐसा देखा गया है कि जब छात्र 'मॉडूल' के बारे में लिखते हैं तो वे उन्हीं बातों को याद रखते हैं जो उन्हें भाषा के असली रूप से परिचित करवाती हैं। अगर कोई ख़ास वीडियो या गीत देखा हो और उससे भाषा-संस्कृति के बारे में कुछ सीखा हो तो वह ज़्यादा प्रभावी होता है। यूरोपीय या अन्य कई भाषाओं में शिक्षण सामग्री बेहतर रूप में विकसित है और हिंदी में अवश्य इसका अभाव दिखाई देता है पर सामग्री के कारण हिंदी भाषा सीखने आना या न आना अभी बहुत स्पष्ट नहीं है और इस पर शोध जारी है। इसका कारण यह भी है कि सिंगापुर में भिन्न तरह की सामग्री हिंदी सिखाने के लिए प्रयोग की जाती है। भाषा शिक्षण सामग्री में ऑडियो-वीडियो रिकार्डिंग ख़ास तैयार किये जाते हैं। पॉडकास्ट, वेबकास्ट आदि का इस्तेमाल सीखने की प्रक्रिया को रोचक बनाता है साथ ही कई तरह के 'एक्टिविटी' कार्ड्स ख़ास ज़रूरत के हिसाब से तैयार किये गए हैं जिन्हें समय-समय पर बदला जाता है। विश्वविद्यालयों के 'ऑनलाइन पोर्टल' का इस्तेमाल करके आजकल की पीढ़ी को 'मोबाइल लर्निंग' की सहूलियत भी दी जा रही है। ऐसी कई सामग्रियों के प्रयोग से अवश्य ही सीखने-सिखाने की प्रक्रिया अधिक दिलचस्प बनती है जिसका ज़िक्र छात्र कई बार अपने 'फ़ीडबैक' में करते हैं।

७ निष्कर्ष

निष्कर्ष रूप में कहा जा सकता है कि यह पूरा शोध छात्रों के अनुपात और रुझान स्पष्ट कर देता है। हिंदी कक्षा में आखिर अहिंदी भाषी भारतीय छात्र अधिक क्यों हैं और चीनी छात्रों की संख्या कम क्यों है? हिंदी सीखने वाले छात्रों में अधिक संख्या भारतीय मूल के छात्रों की ही दिखाई देती है जो ज़्यादातर द्रविड़ भाषा परिवार की पृष्ठभूमि से आते हैं। अहिंदी भाषी भारतीय छात्रों का भारतीय सांस्कृतिक परिचय अधिक होता है तो कई बार इससे उन्हें भाषा सीखने में आसानी होती है। उनके बाद चीनी मूल के छात्रों की संख्या रहती है साथ ही मलय मूल के छात्र भी हिंदी सीखने के प्रति रुचि रखते हैं। चीनी और मलय मूल के छात्र चूँकि सांस्कृतिक पृष्ठभूमि में एकदम भिन्न होते हैं तो हिंदी भाषा के प्रति उनके रुझान का एक कारण यह बन जाता है। वे छोटी-छोटी बातों में भी बड़ी रुचि लेते हैं। सामान्य रूप से हिंदी सीखने के लिए कुछ बातें अधिक दिखी हैं; बॉलीवुड, यात्रा-भ्रमण, संस्कृति, देवनागरी लिपि, दूसरों को समझना, प्रेम-दोस्ती, रिश्तेदारों से बातचीत, भारत की अर्थ-व्यवस्था आदि। ये मुख्य पहलू हैं जो सामने उभरकर आए हैं जिनके कारण हिंदी सीखी जा रही है। आज सीखनेवालों की संख्या भले ही कम हो पर भारत को उभरती शक्ति के रूप में देखने वाले लोगों को विश्वास है कि हिंदी सीखने वालों की संख्या तो बढ़ेगी ही साथ ही यह रुझान संस्कृति को और बेहतर तरीके से जानने में सहायक होगा।

ORCID®

Sandhya Singh  <https://orcid.org/0000-0002-1723-0924>

संदर्भ-सूची

- Creswell, John W. 2012 [2001]. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. 4th ed. Boston, MA: Pearson.
- Dorian, Nancy C. 2014 [1982]. “Language Loss and Maintenance in Language Contact Situations”, in: Dorian, Nancy. *Small Language Fates and Prospects: Lessons of Persistence and Change from Endangered Languages. Collected Essays*. (Brill’s Studies in Language, Cognition and Culture 6.) Leiden; Boston 2014, 203–222.
- Chew, Ernest C. T. & Edwin Lee (eds) 1991. *A History of Singapore*. Singapore: Oxford University Press.
- Romaine, Suzanne 1989. *Bilingualism*. (Language in Society 13.) Oxford (et al.): Basil Blackwell.
- Silverman, David 2001. *Interpreting Qualitative Data: Methods for Analysing Talk, Text, and Interaction*. 2nd edn. London; Thousand Oaks, California: Sage Publications.
- Stathopoulou, Maria 2015. *Cross-language Mediation in Foreign Language Teaching and Testing*. Bristol: Multilingual Matters.
- Department of Statistics, Singapore 2018. *Population Trends 2018*. Singapore: Department of Statistics. (<<https://www.singstat.gov.sg/>>, accessed: October 18, 2019).

On the Contributors

Anvita Abbi, Ph. D. (Cornell University, USA). An advisor to UNESCO on language issues for long, Prof. Abbi is an eminent linguist and social scientist belonging to the family of Hindi writers. She has been visiting professor in Universities across Europe, Australia, Canada, and America. She taught Linguistics at the Jawaharlal Nehru University, India, for 38 years. She has carried out first-hand field research on all the six language families of India, extending from the Himalayas to the Andaman and Nicobar Islands, to identify Indian linguistic typology and shared features across diverse Indian languages. An editor and author of 25 books published nationally and internationally, including *Voices from the Lost Horizon* (2021 Niyogi Books), *Unwritten Languages of India* (2017 Sahitya Akademi), *A Grammar of the Great Andamanese Language. An Ethnolinguistic Study* (2013 Brill, Netherlands), *Endangered Languages of the Andaman Islands* (2006 Lincom Europa, Germany), *A Manual of Linguistic Fieldwork and Structures of Indian Languages* (2001 Lincom Europa, Germany) and a collection of short stories *मुट्टी भर पहचान* (1969 Vani Prakashan). She has been an acclaimed short story writer in Hindi. Her work on tribal and other minority languages of South Asia has been exemplary and has bagged several national and international awards, including the *Padma Shri* in 2013 by the President of India and the *Kenneth Hale Award* in 2015 by the Linguistic Society of America for “outstanding lifetime contributions to the documentation and description of languages of India”.

Email: anvitaabbi@gmail.com

प्रोफेसर रमा कान्त अग्निहोत्री कुछ साल पहले दिल्ली विश्वविद्यालय से सेवानिवृत्त हुए। आजकल विद्या भवन सोसायटी, उदयपुर के साथ कार्यरत हैं।

Rama Kant Agnihotri, D.Phil. (York, UK) retired as Professor and Head, Department of Linguistics, University of Delhi, Delhi. He is interested in and has taught and written extensively about Applied Linguistics, Morphology, Sociolinguistics and Research Methods for several years. He was the Chairperson of the NCERT National Focus Group on the *Teaching of Indian Languages*.

He co-edits the Sage series on Applied Linguistics; is on the Editorial Board of *Contemporary Education Dialogue* (Sage), and is one of the Chief Editors of *Language and Language Teaching*. At present, he is Professor Emeritus at Vidya Bhawan Society, Udaipur. His books, among others, include *Second Language Acquisition: Socio-cultural and Linguistic Aspects of English in India* (1994 Sage, Delhi, ed. with A. L. Khanna), *Hindi Morphology: A Word-based Description* (1997 Motilal Banarsidass, Delhi, with Rajendra Singh), *Problematizing English in India* (1997 Sage, Delhi, with A.L. Khanna), *Adult ESOL Learners in Great Britain* (1998 Multilingual Matters, UK, with A. L. Khanna, M. K. Verma and S. K. Sinha), *Noam Chomsky: The Architecture of Language* (2001 Oxford University Press, Delhi, ed. with N. Mukherjee and B. N. Patnaik) and *Hindi: An Essential Grammar* (2007 Routledge, London).

Email: agniirk@yahoo.com

Liudmila Khokhlova is an associate professor at the Department of Indian Philology, Institute of Asian and African Studies, Moscow State University, Russia. She teaches from undergraduate to PhD level. Among the courses are Hindi, Urdu and Punjabi, General linguistics, lectures and seminars on history, typology, grammar and semantics of Indo-Aryan languages (e. g. Syntactic typology of New Indo-Aryan languages, Historical development of Western New Indo-Aryan languages; Semantic structure of Hindi/Urdu, The Language of *Adi Granth* and *Janam Sakhies*). She has extensively published (in Russian, English and Hindi) on Indo-Aryan linguistics, particularly on the morpho-syntactic history of New Indo-Aryan languages (Hindi-Urdu, Rajasthani, Punjabi and Gujarati) and on phonetics, morphology, syntax, semantics and sociolinguistic problems of Modern Western Indo-Aryan languages. She has also contributed to Sikh studies and translated some parts of sacred Sikh texts into Russian. Her textbooks, such as Hindi Reader/हिंदी पाठ्यपुस्तक (Khokhlova et al., Moscow 2011) and *A Textbook of the Punjabi Language*, 3 vols (with N. Dulai and O. N. Kohl, Moscow 1990) have significantly enhanced the teaching of these languages in Russia. Awards: Punjabi Saath, London, for contribution to Punjabi studies in Russia; Indian Embassy in Moscow for contribution to Hindi studies in Russia.

Email: khokhl@iaas.msu.ru

Ekaterina Kostina is a senior teacher at the Department of Indian Philology, Faculty of Oriental Studies, St. Petersburg State University. She conducts classes in Hindi, Bengali and Sanskrit. She is the editor and one of the authors of the *Primary Hindi Course* (2016) and the author of the *Theoretical Hindi Grammar* (2018) (both in Russian). She has penned a number of papers on languages and literatures of South Asia and teaching methods. Her main field of scholarly interest is the grammar of Indo-Aryan languages. She has also published on the history of Russian Indology, Indian “parallel cinema”, and information technologies in language teaching.

Email: e.kostina@spbu.ru

Annie Montaut, a former fellow of Ecole Normale Supérieure, is Emeritus Professor at Institut National des Langues et Civilisations Orientales, Paris, and member of the laboratory Structure and Dynamics of Languages (CNRS/INALCO/IRD). She has extensively published in Hindi linguistics, particularly on tense and aspect, syntax and semantics, discourse particles and grammaticalization processes, both in English and French (*Hindi*, in the Series Langues du Monde, Société de Linguistique de Paris. Louvain: Peeters 2012; *La Saillance/Saliency* (co-ed with K. Haude), Special Issue of *Faits de langue*. Paris: Ophrys 2012). She is also working on translation and literary traductology and has translated about twenty Hindi novels, drama and poetry collections from Hindi to French (including Nirmal Verma, Krishna Baldev Vaid, Geetanjali Shree, Anupam Mishra, Jacinta Kerketta, Kedarnath Singh). Awards: Gargi Gupta Award, Delhi, Bharatiya Anuvad Parishad 2012; Kamil Bulke Samman, Bhopal 2019.

Email: annie.montaut@inalco.fr; annie.montaut@ehess.fr

Tatiana Oranskaia, PhD (Leningrad/Saint Petersburg State University) is prof. (em.) at the Department of Culture and History of India and Tibet, University of Hamburg. Before joining the German university in 1998, she had conducted research and taught in various positions at the Department of Indian Philology at her alma mater, St. Petersburg (Leningrad) State University. Her numerous publications (in Russian, English, German and Hindi), primarily on Indo-Aryan linguistics and local Hindu cults, include nine monographs and edited volumes. Among them are *Pronominal Clitics in Indo-Iranian Languages* (1991); *New Indo-Aryan Languages* (Oranskaia et al. eds, Moscow 2011) and *Divinizing in South*

Asian Traditions (with D. Dimitrova, eds, Routledge 2018). She taught a range of courses in Indo-Aryan linguistics and languages, literature and cultural history of South Asia. She is a member of several editorial boards and academic associations. Among her awards and honours are हिंदी सेवी/Hindī sevī (MGAHV 2018), Rabindranath Tagore Award (German-Indian Society 2019) and ICCR Senior Fellowship (2016).

Email: tatiana.oranskaia@uni-hamburg.de

अनिल कुमार पाण्डेय भाषाविज्ञान एवं भाषाप्रौद्योगिकी विभाग, भाषाविद्यापीठ, महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय, वर्धा, भारत में प्रोफेसर के पद पर कार्यरत हैं। वे हिंदी माध्यम से पढ़ाते हैं। इनकी एक पुस्तक *हिंदी संरचना के विविध पक्ष* 2010 में प्रकाशित है। *वर्धा हिंदी शब्दकोश* के द्वितीय संस्करण का 2019 में इन्होंने संपादन किया है। इनका विषय क्षेत्र हिंदी भाषाविज्ञान (ध्वनिविज्ञान, रूपविज्ञान एवं वाक्यविज्ञान) है। इनके बहुत सारे भाषा विज्ञान से संबंधित लेख विभिन्न पत्रिकाओं में प्रकाशित हैं।

Anil Kumar Pandey, Professor at the Department of Linguistics and Language Technology, School of Language, Mahatma Gandhi International Hindi University, Wardha, conducts research and teaches courses in Hindi phonology, morphology and syntax. He uses Hindi as the language of instruction. Among his numerous publications is the monograph *Aspects of the Formal Grammar of Hindi* (2010, in Hindi). Prof. Pandey is the editor of the *Wardha Hindi Dictionary* (2nd edition, 2019).

Email: pandey_anilkumar61@rediffmail.com

Aaricia Ponnet is a PhD researcher at Ghent University's Department of Linguistics and an affiliated researcher at the Department of Indian Languages and Cultures. Her research, funded by the Research Foundation – Flanders, focuses on morpho-syntactic development in Hindi as a foreign language. She holds an MA in Indian studies (2015), graduated as an ICCR scholar for Kathak dance (2017), speaks fluent Hindi and regularly stays in India for her academic and artistic projects.

Email: Aaricia.Ponnet@ugent.be

† **Vashini Sharma** (1944–January 2, 2021), PhD, retired in 2006 as a professor at the Central Hindi Institute, Agra. She was born in Hyderabad and received there her education. After marriage, she moved to Agra to join Central Hindi Institute in 1973. After her retirement, she continued to serve the cause of Hindi and always believed that Hindi has a bright future as it is definitely the link and cultural language of the nation. Academically she was active till her last breath and was a regular participant in various national and international platforms. She was very active on social media, also with a couple of blogs dedicated to language, literature and culture. A bit introvert, many avenues are yet to be unearthed by her friends and well-wishers. She dedicated her entire life to the cause of the development of Hindi and left behind her a vacuum that is difficult to fill.

Dr Vikrant Shastri, son of Prof. Vashini Sharma, March 28, 2021

Sandhya Singh, PhD (Banaras Hindu University), is the convenor for the Hindi and Tamil language program at the Center for Language Studies, National University of Singapore. Her research interests lie in sociolinguistics, in particular, the Indian Diaspora and the use of Hindi and Bhojpuri in South-East Asia. She is the founder and president of Sangam Singapore, a non-profit organization that aims to spread knowledge of languages, literatures, and cultures in Singapore. As part of this, she is the editor of *Singapore Sangam*, the pioneer Hindi magazine from Singapore. Regarding tertiary education, she is part of the management committee in the Hindi Society (Singapore). For her work, she has been awarded several *Hindi sammaan*.

Email: sandhyasingh077@gmail.com

Saartje Verbeke worked as a visiting professor at Ghent University, in the department of Indian Languages and Cultures, teaching courses on the modern languages and culture of South Asia. She completed her PhD on ergativity in Indo-Aryan in 2011 and has been working since on several linguistic research projects about Indo-Aryan languages from a cross-linguistic, historical and language acquisition perspective. Her area of specialization is grammatical phenomena on the syntax-semantics interface, such as word order and referential hierarchies in (Old) Kashmiri, and the evolution of case marking from Middle Indo-Aryan to Early New Indo-Aryan.

Email: saar.verbeke@gmail.com

Heinz Werner Wessler studied Indology, Religious Studies and Musicology in Bonn, Delhi and Zurich. Habilitation at the University of Bonn in 2009 with a study on Dalit literature in Hindi. Since 2010 initially as a visiting professor, since 2014, as a full professor at Uppsala University (Sweden). Areas of interest: Hindi and Urdu literature, lexicography, history of ideas, religious history of South Asia, modern South Asian culture, music.

Email: heinzwerner.wessler@lingfil.uu.se

The present volume, which comprises seven English contributions and four in Hindi, addresses issues of linguistic variation – a phenomenon central to the study of language use – in regard to the major official language of India. It combines multiple theoretical and pragmatic approaches to a variety of linguistic phenomena conceptualised under the designation ‘Hindi’ and attempts to obtain a more accurate portrayal of the changing reality behind this versatile taxonomic term. In doing so, this volume provides insight into how the forms and functions of Hindi are changing across borders within and outside of India; the concept of language contact is thus present in several of the studies. The analyses are based upon data observed in written texts, including manuscripts, and elicited from instances of oral speech.

The contributions, prepared by established and emerging scholars from several Asian and European countries, investigate functional aspects of the regional, social, and cultural forms of Hindi and how they interact in differing contexts, time periods, and types of communication. A similar vantage point is being adopted in the investigation of possibilities and constraints of formal variation in the grammatical structures of Standard Hindi. In addition to providing analyses from the perspectives of both general linguistics and sociolinguistics, the book discusses issues associated with teaching Hindi from the perspective of language variation.

The volume is distinguished by its innovative character in terms of both the data utilized in it and the width of its scope, and aims to contribute to a better understanding of ‘Hindi’ as a concept as well as of the general principles of linguistic variation.

ISBN 978-3-948791-26-1



9 783948 791261