


UNIVERSITÀ DEGLI STUDI DI BERGAMO

UNIVERSITÀ DEGLI STUDI DI PAVIA

Scuola di Alta formazione Dottorale

Corso di Dottorato in Scienze Linguistiche

Ciclo XXXVI

Settore scientifico disciplinare L-LIN/01 – Glottologia e linguistica

ARGUMENT STRUCTURE OF EXPERIENTIAL VERBS IN HINDI
A CONSTRUCTIONAL APPROACH

Tutor:
Prof.ssa Silvia Luraghi

Candidata:
Carnesale Lucrezia

Matricola n° 1073085

Anno accademico 2022/23

Table of Contents

List of abbreviations 6
1 INTRODUCTION 8
1.1. Motivations behind the present study 8
1.2. Aims and structure of the study 10
2. THE DOMAIN OF EXPERIENCE 14
2.1. Bodily sensations, perceptions, and cognitions: the notion of embodiment 16
2.2. The encoding of situations: event structure and argument coding 18
2.2.1. The encoding of experiential events 20
2.3. Experiential constructions in South Asian languages 27
3. THE LANGUAGE UNDER INVESTIGATION: MODERN STANDARD HINDI 35
3.1. A historical background 35
3.2. Preliminary remarks on Hindi morphosyntax 38
3.2.1. Nominal morphology 39
3.2.2. Hindi postpositions 42
3.2.2.1. Simple postpositions 44
3.2.2.2. Complex postpositions 51
3.2.3. Verbal morphology 52
3.2.3.1. Non-finite verb forms 53
3.2.3.2. Finite verb forms 57
3.2.4. Differential case marking in Hindi 60
3.2.4.1. Split ergativity 61
3.2.4.2. Differential Object Marking (DOM) 62
3.2.4.3. Differential Subject Marking 65
3.2.5. Complex predicates 71
3.2.5.1. Verb + Verb Complex predicates 71
3.2.5.2. Noun + Verb Complex predicates 74
3.2.5.2.1. Ligh verb alternations: actionality and valency 79
3.3. Complementation in Hindi 81
4. THE THEORETICAL FRAMEWORK 85
4.1. Frame semantics 85
4.2. The constructional approach 88
4.2.1. A constructional approach to argument structures 93
4.2.2. Usage-based approach 98
4.2.3. A constructional approach to Hindi 99
4.2.3.1. A constructional approach to Hindi N-V complex predicates 104
4.3. Transitivity marginality and iconicity in Hindi 110
5. METHODOLOGY 118

5.1. The corpus 121
5.1.2. The literary corpus 121
5.1.3. The HiTenTen Corpus 124
5.1.4. Corpus creation and data extraction 124
5.1.5. Sample representativeness 129
6. BODILY SENSATIONS 132
6.1. Feelings of pain and itch 134
6.2. Feelings of temperature 138
6.3. Feelings of saturation or lack (hunger and thirst) 140
6.4. Metaphorical uses 142
6.5. Discussion 146
7. PERCEPTIONS 153
7.1. The typology of the lexicalization of perceptions 153
7.2. The semantic properties of perceptions 154
7.3. Presentation of Hindi data 158
7.3.1. *Dekhnā*: the generic verbs for visual perceptions 160
7.3.2. *Sunnā*: the generic verbs for auditory perceptions 165
7.3.3. Non agentive visual and auditory perceptions 169
7.4. (Anti)causative alternation in the expression of Hindi perceptive events 178
7.4.1. The (anti)causative paradigm and verbs of perceptions 181
7.5. Complement clause constructions 185
7.6. Other sense modalities: touch, taste and smell 192
7.7. Discussion 199
8. COGNITION 206
8.1. Constructions with unexpressed roles 208
8.2. Actionality and aspect in verbs of cognition 211
8.3. Presentation of Hindi data 218
8.4. Think 219
8.4.1. Socnā 220
8.4.2. Complex predicates expressing thinking: *vicār* and *khayāl* 230
8.5. Know 239
8.5.1. Knowing as a state 239
8.6. The acquisition of knowledge: Learn/Understand 255
8.7. Remember and forget 260
8.7.1. Remember: complex predicates with *yād* 261
8.7.2. Forget 269
8.8. Discussion 273
8.8.1. The semantic contribution of complement types 274
8.8.2. The distribution of the light verbs in complex predicates 278
9. ANALYSIS AND DISCUSSIONS 282

9.1. Constructions productivity and semantic coherence 282
9.1.1. The transitive construction 285
9.1.2. The dative construction 288
9.1.3. The rise of the dative construction 291
9.1.4. Two competing constructions in the language 293
9.1.5. Constructions with limited semantic scope 295
9.1.5.1. Genitive Experiencer construction 296
9.1.5.2. The locative construction 297
9.1.5.3. The copular construction 297
9.1.5.4. The oblique Stimulus construction 298
9.2. Transitive-intransitive (anti)causative alternation 302
9.2.1. Some remarks on the origin of the extended (dative) intransitive construction 304
10. CONCLUSION 306
REFERENCES 310

List of abbreviations

1 First person

2 Second person

3 Third person

ACC Accusative

ADJ Adjective

AOR Aorist

CAUS Causative

COM Comitative

CRR Correlative

CP Conjunctive Participle

DAT Dative

DIR Direct

EMPH Emphatic

ERG Ergative

F Feminine

FUT Future

GEN Genitive

H Human

INC Inchoative

IND Indirect

INDF Indefinite

INF Infinitive

INS Instrumental

Intr Intransitive

IMP Imperative

IPRF Imperfect

LOC Locative

M Masculine

N Neuter

NOM Nominative

NN Non nominative

OBL Oblique

Obj Object

PL Plural
PRG Progressive
PRF Perfect
PRN Pronoun
PRS Present
PRTP Participle
PSP Postposition
PST Past
REFL Reflexive
REL Relative
SG Singular
SBJV Subjunctive
Subj Subject
Tr Transitive
V Verb

1 Introduction

1.1. Motivations behind the present study

The present work aims to provide an exhaustive analysis of the argument structure constructions of Hindi experiential verbs, which is currently lacking in the typological panorama. Even though the domain of experiential constructions in Indo-Aryan languages is a relatively explored area, previous studies have mostly dealt with the encoding of non-nominative subjects and thus focused mainly on the coding of the Experiencer. Previous authors (Verma and Mohanan 1990, Bhaskararao and Subbarao 2004, Kachru 1990, Montaut 1997) agree in pointing out that the dative construction (exemplified in 1) is the prototypical pattern for the encoding of experiential situations in Hindi and that the inherent semantics associated with it is the lack of control and volitionality by the Experiencer.

1. मुझे ठंड लग रही है।

mujhe	*ṭhaṁḍ*	*lag*	*rah-ī*	*hai*
1SG.DAT	cold(F.SG.NOM)	attach	PRGR-F	be.3SG.PRS

"I am cold."

Previous literature mainly focused on the dative Experiencer and showed less interest in other case-markings for this semantic role or in the encoding of other participants in experiential situations, such as the Stimulus. Therefore, a general discussion on argument structures associated with experiential verbs in Hindi is still lacking. One of the main aims of the present dissertation is to fill this gap. With this aim in mind, in the following chapters I will address different issues linked to the morphosyntactic expression of experiences in Hindi. I will investigate how different argument structure constructions are favored by some classes of verbs and rejected by other, and how they correlate with other linguistic features, such as aspect and actionality. Hindi shows a well-entrenched (anti)causative system and part of my intent will be to investigate how this system operates with verbs of experiences resulting in different construal of the same event.

One of my main claims will be that the semantic differences of experiential events are mirrored by their morphosyntactic realization in Hindi, and I will argue that the analysis of the expression of experiences will deepen our understanding of the syntactic-semantic interplay in this language. Hindi exhibits a wide range of constructions for the encoding of experiential events, both the transitive construction and intransitive constructions can be used for the encoding of an experience in this language and both the Experiencer and the Stimulus show a wide range of possible linguistic realizations. As I will discuss, Hindi exhibits a number of constructions which are generally selected

on the basis of the semantic properties of the event and of how it is conceptualized. This peculiarity of Hindi has led Montaut (2004a; 2013) to define Hindi a language with semantic alignment (following the definition given in Wichmann 2008) rather than syntactic ones. Languages displaying semantic alignments realize the various argument types according to the lexical semantic properties of the predicate (Dahl 2022: 7). In fact, as I will discuss in chapter 4.3, Hindi clearly encodes semantic roles according to semantically based parameters, and it associates each pattern of case marking to specific semantic features, mainly related to the most salient participant and to the type of the event (Montaut 2004b, Butt and King 2002, 2004, Butt 2006b). Thus, for example, the use of the transitive construction in Hindi is generally associated (but not restricted) to prototypical Agents volitionally acting and controlling the event. When the event expressed is characterized by a low degree of transitivity, another construction is chosen. Which constructions is employed depends on how the semantic properties of the event that is encoded deviate from the properties of a transitive event. Languages like Hindi have been defined iconic in previous literature (Malchukov 2005, 2015), since they use morphosyntax to reflect the semantic of the event and they associate each pattern to specific semantic properties. Given this iconic tendency, the analysis of Hindi experiential verbs is quite interesting, as experiential events display extremely variable properties both in relation to participants and in relation to how an event is construed (Luraghi 2020a: 38). The study of argument structures associated to this class of verb thus can give us interesting insights on the interplay between semantics and syntax in this language.

Given my intent to investigate the extent to which syntactic choices in this language are influenced by semantic properties, I have chosen to concentrate my analysis on the subdomains of Bodily Sensations, Perceptions, and Cognitions, as these subdomains exhibit the most significant differences from a semantic perspective. Perceptions and cognitions differ from bodily sensations in many respects, but one of the main differences lays in the semantic properties of the Experiencer, which may be endowed with a higher degree of awareness and agentivity in the case of the first two domains. The Experiencer of bodily sensations, in contrast, is never volitional and never controls the situation. From the analysis I conducted for the present study, it emerges that the fundamental parameter influencing the selection of the morphosyntactic construction used to encode experiential events is the Experiencer's degree of agentivity. In the end of this dissertation, I will try to draw some conclusions on the productivity of the constructions used in Hindi to encode experiences and I will establish the extent of their semantic coherence.

1.2. Aims and structure of the study

This dissertation is structured as follows. Following this introduction, in **chapter 2**, I deal with the state of the art on the functional cognitive domain of experience and, referring to previous typological and language-specific studies on experiential constructions, I describe the internal structure of this domain, which is typically conceived as composed by five subfields: Bodily sensations, Perceptions, Emotions, Cognitions and Volitions. I briefly review these different types of experiential situations and I then discuss the reason behind my decision to focus on the three experiential subdomains of Bodily sensation, Perception and Cognition. This gives me the occasion to deal with the notion of embodiment in cognitive linguistics. In the same chapter, I address the event structure of experiences and how they tend to be conceptualized and encoded cross-linguistically. I discuss the semantic properties and the morphosyntactic encoding of the three main components of an experiential event: i.e., the Experiencer, the Stimulus and the Expertum. I argue that both the Experiencer and the Stimulus are highly variable with respect to their semantic properties, and I discuss the consequences of this semantic variability on the morphosyntactic encoding of these two semantic roles, both in the same language and cross-linguistically. The final section of this chapter is dedicated to a brief overview of the expression of experiential events in South Asian languages, and to the discussion of one of the main features that contribute to make South Asia a linguistic area: i.e. the use of oblique Experiencers functioning as non-canonical subjects in most South Asian languages, both Indo-Aryan and not.

In **chapter 3,** after a brief introduction on the geographical and historical collocation of the language under investigation, I outline some morphosyntactic features of Hindi, in order to endow the reader who is not familiar with the language with the basic features of the grammar. In particular, I focus on nominal and verbal morphology in Hindi, and on the main features regarding alignment and case marking in the language: i.e. split ergativity, differential object marking and differential subject marking. I also discuss complex predication in Hindi: in section 3.2.5.1 I focus on verb-verb complex predicates, while in section 3.2.5.2 I focus on noun-verb complex predicates, which are particular productive in the expression of experiential events in Hindi. I do not address in this chapter features that are preliminary to the discussion of specific verbs analyzed in the following chapters, and I postpone their description to the chapter in question. For example, I deal with (anti)causative constructions in Hindi in chapter 7 dedicated to the expression of perception and I deal with the relation between aspect and actionality in the language in chapter 8, while discussing verbs of cognition.

In **chapter 4**, I expose the main tenets of the theoretical approach I rely on in this dissertation, i.e. usage-based Construction Grammar, and I discuss how this framework can be applied in the

analysis of argument structures. I deal with the notion of *Constructicon*, which is intended as the repository of the constructions of a language, conceived as a syntax-semantics continuum generated by a network of constructions. I also describe how constructions are conceived in this theoretical approach, i.e. as conventionalized pairings of meaning and form. In this framework, constructions have their own semantics and are subject to generalizations, for example they tend to correlate with specific classes of verbs according to their semantics. I also present some examples of Hindi constructions in order to show that this framework is particularly suitable for the analysis of argument structures in a language like Hindi. These examples allow me to show how the overall interpretation of a construction is arrived at by integrating the argument structure with the main verb and its various arguments. Additionally, in the same section I show how argument structures supply their own semantics to the overall construction and may contribute to construe the event in different ways. In section 4.2.3.1, I focus in particular on the constructional analysis of Hindi noun-verb complex predicates, which appear to be quite problematic with respect to their argument structure and the ambiguous status of their nominal host. In the same chapter, I also discuss the usage-based approach that leads my interpretation of the functional distributions of constructions. In particular, I expose the difference between type frequency and token frequency and how these can give us insights on the productivity and the semantic coherence of a given construction. Lastly, in section 4.3, I present the definition of Hindi transitive construction that I use in this dissertation, and I argue that this pattern is marginal in the language (at least in comparison with many SAE languages which show a much higher degree of transitivity prominence) and that the notion of iconicity accounts for this marginality in the language.

In **chapter 5,** I give some methodological remarks and I present the source of my data and the tool that I used to extract them, i.e. SketchEngine. All the examples used in this dissertation are taken from corpora. My main source of linguistic data is a corpus I collected specifically for the purpose of this analysis consisting of literary texts of the 20th century. On few occasions, I relied on data taken from other sources: in particular, I referred either to the hiTenTen corpus (consisting of text taken from the web and available on SketchEngine) or to examples taken from the previous literature. When reporting examples sourced from the literary corpus, I do not provide specific information on the source, as the literary corpus constitutes my primary and default reference. When examples are extracted from the hiTenTen corpus, I explicitly specify this beneath the sentence, while when the examples are taken from previous literature, I specify this within the text and provide detailed references. If the examples from previous literature are in Hindi I adapt them with my glosses, while if they are in other languages, I report them and the glosses exactly as they are in the source.

Chapters 6, 7 and **8** are the bulk of my dissertation as they are dedicated to the analysis of the argument structures encoding the three experiential verb classes I focus on in this work: Bodily Sensations (6), Perceptions (7) and Cognitions (8). **Chapters 6** is devoted to the analysis of verbs of Bodily Sensations in Hindi. I focus on three main sensation types: feelings of hunger/thirst, feelings of hot and cold, and feelings of sensations localized on a specific area of the body. I show that the dative construction is the prototypical construction used to encode sensations in the language, as it is the only one that can encode all these three types. I also discuss two main constructions, whose functional distribution is more limited than that of the dative construction: i.e. the locative construction expressing localized sensations and the copular construction expressing feelings of temperature. In the section dedicated to the final discussion, I deal with the verb *lagnā*, whose use is extremely pervasive in the expression of experiences in Hindi.

In **chapter 7**, I address the constructional patterns of verbs of perceptions in Hindi. Following Viberg's (1984) Modality Hierarchy, I first deal with verbs of visual perception, followed by verbs of auditory perceptions and then by the other three classes of verbs: touch, taste and smell. I argue that the main property governing the morphosyntactic encoding of perceptions in Hindi is the agentivity of the Experiencer. I show that verbs of visual perceptions differ from verbs expressing perceptions through the other sense modalities, as they allow an oblique Stimulus construction, thus aligning with typological observations according to which verbs of seeing typically show a higher degree of complexity both from a lexical and a constructional point of view. In the same chapter, I discuss the Hindi (anti)causative system and how this interacts with the expression of perceptions and experiences in general in the language. This discussion will turn to be useful also for the analysis of verbs of cognition in the following chapter. Finally, I also deal with complementation of perception verbs in Hindi, and I show that a construction that turns to be particularly relevant for the encoding of this experiential subdomain is the predicative participle construction, specifically used for the expression of direct (visual) perception of state of affairs. Beside this participial construction, a finite complement clause is also used, both with verbs of visual perception and with verbs of auditory perception.

Chapter 8 is devoted to the analysis of Hindi verbs of cognition. In this chapter, I focus mainly on verbs of knowing and thinking and on verbs of forgetting and remembering. The expression of knowing and thinking is particularly interesting because verbs expressing this cognition types are sometimes characterized by striking aspectual distributions that can give us insights on the lexical aspect of a verb and on how it construes the event. The expression of remembering is interesting as well, as this experience is expressed by a series of complex predicates realized by a nominal host and the alternation of the light verb. I will show that the alternation of the verb in complex predicates

expressing memory is particularly productive and it perfectly exemplifies how the light verb contributes its own semantics which, integrated with the semantics of the nominal host, contributes to different construals of the same event. The subfield of cognition is the most complicated with respect to the constructional network expressing it. Verbs belonging to this subdomain may occur with the transitive construction, the oblique Stimulus construction and the dative construction, which are also found with verbs expressing perceptions. Beside these constructions, a construction marking the Experiencer with the genitive is also allowed with some verbs of cognition, thus distinguishing this subdomain from that of perception. As I will argue, the reason behind this restricted functional distribution of the genitive construction lies in the semantics of the construction, which implies a stative reading and an agentive Experiencer. Also, with respect to complementation, this subdomain shows a higher degree of complexity than the domain of Perception, as besides a predicative participle clause and a finite complemental clause, it also shows predicative adjectival clauses and non-finite clauses with infinitives.

Chapter Errore. L'origine riferimento non è stata trovata. is dedicated to the final discussion. In t his chapter I summarize the findings of the previous chapters and, relying on the type frequency and the token frequency of the constructions used in Hindi to express the three experiential subdomains I analyzed, I draw some conclusions with respect to their productivity and their semantic coherence. Focusing on the semantics and the functional distribution of the transitive and the dative construction, I argue that these two patterns seem to be in competition in modern Hindi. Relying on some data of individual verbs analyzed in the previous chapters I argue that there is reason to believe that the dative construction is overtaking some of the semantic areas that originally were under the domain of the transitive construction. I will also argue that the extension of the dative construction at the expense of the transitive one is consequence of a change in the encoding strategies that involved Hindi, and that results in privileging semantic features over grammatical relations in the marking of semantic roles. Focusing on the dative Experiencer, I briefly outline a complex diachronic picture in which the contact with Persian led to a massive insertion of complex predicates in the lexicon that, interacting with the emerging semantically constrained use of postpositions, boosted the productivity of the dative construction in the language. In the same section, I also describe the semantics of the other main constructions analyzed in the previous chapters, i.e. the locative construction, the genitive construction, the copular construction, and the oblique Stimulus construction; and I argue that their semantic coherence accounts for their distribution over the experiential verb classes.

2. The domain of Experience

The semantic-cognitive domain of experience is related to the capacity of human beings to perceive external situations and to elaborate the raw data acquired via sensations and perceptions in order to reach an understanding of the world around them. The term experience covers a large range of situations and refers to human beings' perceptive, cognitive or emotional ability. This domain is universal, as experiential situations depend upon biological and anatomical features shared by all human beings. However, the conceptualization and the linguistic expression of experiences may differ from culture to culture and consequently from language to language (Verhoeven 2007). Thus, languages around the world usually show important variations in the encoding of the same experiential situation.

Based on Verhoeven's (2007) discussion concerning types of experiential situations, I consider the cognitive domain of experience as composed by five subdomains: bodily sensations, volitions, perceptions, cognitions and emotions. Given the prominence of the experiential domain in our lives (Wierzbicka 1981), linguists have dedicated much effort to its analysis. Among the most analyzed subdomains, **emotion** boasts the largest number of specific studies (see for example Wierzbicka 1995, 1999, Harkins and Wierzbicka 2001, Athanasiadou and Tabakowska 2010, Niemeier and Dirven 1997). This subdomain is extremely complex and highly cross-linguistically heterogeneous in its internal structure, as it is strongly connected to the socio-cultural values of the speakers' community. As a consequence, the domain of emotions is also the most varied among the experiential subdomains with respect to its grammatical coding (Verhoeven 2007: 44). However, despite this complexity, emotions seem also to be subject to typological universals based on universal properties of human beings. In particular, it has been noted that languages are more prone to encode emotions that are considered basic from an ethno-psychological perspective with specific lexical units. These emotions are happiness, sadness, fear, anger, disgust, shame and surprise (see on this Izard 1977, Johnson-Laird and Oatley 1989, 1992, Ekman 1992, Ekman and Davidson 1994, Wierzbicka 1999). Other more complex emotions such as pity, jealousy or worry in contrast tend to be culture-specific and thus show a wider range of variability across languages.

Scholars have also devoted many studies to the subdomain of **cognition** (Horie 1985, Croft 1993, Filip 1996, Fortescue 2001, Goddard 2003), which refers to internal experiences linked to mental functions of human beings such as remembering, understanding, knowing, thinking, believing, and so on. The domain of cognition exhibits a high complexity, and it can be interpreted as consisting of two main groups (Verhoeven 2007, Fortescue 2001, Luraghi 2020a): mental states (such as knowing) and controlled mental activities (thinking, remembering, understanding, etc.). Horie (1985) proposes to distinguish two subdomains: cognition and conception. The first represents more stative

situations related to the presence or absence of information in the mind of the experiencer, while the second refers to more active mental processes. Similarly, the subdomain of **volition** is typically divided into two large semantic areas and includes on the one hand psychophysical concepts such as desire and want, and on the other concepts related to intentionality, such as decide or intend. This subdomain seems to be linked with other experiential subdomains. For example, a number of complex emotions, such as jealousy, seems to belong to both the subdomain of emotion and that of volition. Moreover, volitions show many similarities with cognitions as they are both typically related to a content and generally imply a self-aware experiencer, for this reason they are typically consider a specifically human prerogative.

Perceptions refer to concepts related to the acquisition of information on the external world through the five sense modalities (sight, hearing, touch, taste and smell). Perceptions can be agentive, such as look at or listen to, or non-agentive, such as see and hear and this distinction is generally encoded either lexically or constructionally (Viberg 1984, Verhoeven 2007, Croft 2012). This subdomain is linked to the domain of cognition as the acquisition of information through senses generally results in the presence of knowledge about the world in the mind of the Experiencer. As I will discuss, this link leads to typologically common metonymic extensions from perceptions to cognitions (Sweetser 1990). Adjacent to the subdomain of perceptions is the subdomain of **bodily sensation**, which concerns experiences related to the body, such as feelings of hunger or thirst, temperature (hot or cold), pain, itching and so on. Unlike the other experiential subdomains, bodily sensations are not typically associated to a source or a goal and are conceived as occurring by themselves (Verhoeven 2007: 43). They differ from other experiential subdomains also in that they are always characterized by the absence of control on the part of the experiencer (Bossong 1998).

It should be noted that the delimitation of the subdomains quickly exposed above is sometimes quite difficult and even arbitrary. Like any other cognitive-functional domain, experiential subdomains have fuzzy boundaries, which often overlap, so that a given experiential situation can be associated with more than one subdomain. For example, a situation such as desiring something indicates an emotional state, but in certain cases it may also involve intentionality (Wierzbicka 1999). Similarly, caring for someone entails thinking about that person and keeping this person in the mind, and as a consequence is also linked to remembering (Luraghi 2020a: 20). Verbs of bodily sensation or perception can often extend to other domains, for instance, as already mentioned, the domain of cognition is frequently conceptualized through a metonymic extension which has perceptions as source domain (Sweetser 1990).

2.1. Bodily sensations, perceptions, and cognitions: the notion of embodiment

In this work, I decided to focus my analysis on the Hindi expression of three experiential subdomains: Bodily sensations, Perceptions, and Cognitions. The choice was not casual, and it was driven by specific reasons. Previous scholars (Sweetser 1990, Colman 1994, Viberg 2015, Luraghi 2020a) have focused on the connection of these three experiential subdomains. Colman (1994) for example describes perception as consisting of a series of sequential stages that allows the interpretation of stimuli from the environment. This sequential process starts from the elaboration of raw sensory data acquired through the five sense modalities and arrives at mental experience reaching conscious awareness (Colman 1994: 153). These sequential stages form the experience of sensory stimuli to the elaboration and acquisition of some cognitive content is mirrored at the linguistic level. It is typologically common for example that verbs expressing sensations and perceptions are extended to encode cognitions. For instance, the two main verbs expressing knowledge in Indo-European languages are both derived from the connection of these three subdomains (Viberg 2015: 105-106). In some Germanic languages, the verb expressing generic knowledge is originally derived from the perfective form **woida* "have seen" of the Indo-European root **weid-* "see" (see for example German *wissen*, Swedish *veta* and Icelandic *vita*). While in many Romance languages, the verb expressing knowledge is derived from the Latin verb *sapēre* which originally meant "taste" (see for instance, French *savoir*, Italian *sapere* and Spanish *saber*). This verb still expresses today both the meaning "know" and the meaning "taste".

One can conceive sensations, perceptions and cognitions as located in a continuum and as being characterized by fuzzy boundaries: verbs of perception are placed in the middle of the continuum and stand between raw physical sensations on the one hand and abstract cognitive processes on the other hand. This continuum from bodily feelings to mental abilities have been explained in cognitive linguistics through the notion of embodiment. The concept of embodiment highlights the idea that our cognitive functions and abilities, including language, are deeply grounded in our bodily experiences and interactions with the physical world (Johnson 1987, Johnson and Rohrer 2007). For example, our understanding of the concept "warmth" is anchored to our sensory experiences of temperature. As a consequence, cognitive and abstract concepts tend to be metaphorically mapped on the basis of more concrete experiences. These ideas refer to the notion of conceptual metaphor elaborated by Lakoff and Johnson (1980), who emphasize that metaphors are not just literary devices but are pervasive in everyday language and point out that we use them unconsciously to structure our thoughts and shape our understanding of the world. In conceptual metaphors, a domain which is concrete and well-understood is used as the source domain for attributes and structures used to understand a less tangible target domain. For instance, the expression

"feeling down" uses a conceptual metaphor (Lakoff and Johnson 1980: 14-15) that draws on our physical experience of orientation in the real world to convey abstract concepts such as an emotional or mental state. Lakoff and Johnson argue that some conceptual metaphors are largely universal, grounded in our bodily experiences and basic cognitive functions. For example, the "UP-DOWN" metaphor occurring the English sentence *I'm feeling down today* is typologically widespread (the same metaphor for example operates also in the Italian sentence *Oggi mi sento un po' giù* "Lit. Today I am feeling a little down"). However, cultural and linguistic factors can also lead to variations in how metaphors are expressed and interpreted.

Experiential verbs constitute a very fruitful field of research on conceptual metaphors and the notion of embodiment, and they provide crosslinguistic evidence for reflexes of human experience on lexical and constructional encoding. This is true for Hindi as well. As I will discuss, perception verbs may be used to express cognitive events in Hindi (section 8.4.1) and further evidence of the link between these two domains in Hindi is also provided by the fact that visual perception verbs may also be used to mark evidentiality (section 7.5). Another interesting example of the way in which physical perception shapes our understanding of the world and determines the way we conceptualize it is represented by Hindi verbs for bodily sensations used to encode emotions that I discuss in section 6.4.

The other reason why I chose to focus on these three experiential subdomains lies in their contrasting semantic properties. Bodily Sensations, Perceptions, and Cognitions exhibit the most significant differences from a semantic perspective. As mentioned in section 1.2, one of the most relevant differences between perceptions and cognitions on the one hand and bodily sensations on the other is that the first two subdomains may be endowed with a higher degree of awareness and agentivity, while the bodily sensations never involve a volitional Experiencer that controls the situation. The main differences then are related to the semantic properties of the event and its semantic closeness with transitive prototype. Similar observations have been made by many scholars. Tsunoda (1985), for example, when discussing verbs of experiences distinguishes verbs of perceptions and cognitions on the one hand and verbs of sensations (and emotions) on the other. On a similar line, Malchukov (2005, 2015) notes that event though all experiential classes deviate from the transitive prototype (see section 4.3), bodily sensations are the most semantically distant. As he points out: "sensation predicates (such as "freeze", "be sick") deviate arguably even further from transitivity prototype than emotion predicates, since Experiencer is their only argument, while many emotion predicates ("like", "fear") take two arguments" (2005: 81). As mentioned in section 1.1, the main purpose of my dissertation is to assess the semantics of the constructions occurring with verbs of experiences in Hindi and to establish the extent to which syntactic choices are driven by semantics in

the language. I believe that focusing on the three most semantically different subdomains belonging to the experiential class is the most productive way to gather interesting insights on the semantic-syntactic interplay in the language. As I will show in the following chapters, the constructional analysis of the Hindi expression of these three experiential subdomains supports the typological observations given by Malchukov and Tsunoda. In particular, verbs of perceptions and cognitions pattern together in many respects, while bodily sensations are encoded by specific constructions that are not productive within the other two semantic subdomains. The only construction that is used to encode all three experiential subdomain is the dative construction.

2.2. The encoding of situations: event structure and argument coding

In this section, I present the theoretical approach to the study of the linguistic expression of situations that I will follow in this dissertation. This approach mainly relies on the cognitive functional perspective elaborated by many scholars since the 1970s (Fillmore 1977, Comrie 1981, Langacker 1987, 1991, Croft 2012, 2022) and starts from the assumption that the linguistic expression of a situation mirrors how the situation is conceptualized in the mind of the speakers. At a cognitive representation level, a situation consists of participants which are related to one another in a stative or dynamic relation. How a participant is related to other participants and to the situation defines its participant role. At a linguistic representation level, the situation is typically expressed by the predicate while the participants are expressed by the arguments that occur with it. At the linguistic level, participant relations are reflected in the relations between the predicate and its arguments.

Real life events and situations are very diverse and different event-types have quite different participant roles. Hence, there cannot be as many argument types as participants, given that participant roles are potentially infinite. Languages thus exploit patterns of generalization and analyze participant roles as more general roles that cut cross large classes of events. These broader categories of participant roles are called semantic roles (Croft 2012, 2022). Following the traditional convention, I will use the same label for both participants and semantic roles and I will distinguish between them by using a capital first letter for semantic roles (so for example I will use experiencer when referring to the participant, while I will use Experiencer when referring to the semantic role). Common semantic roles are Agent, Patient, Theme, Recipient, Beneficiary, Experiencer and Stimulus. These roles are typically realized as core arguments and tend to be given the grammatical status of subject, direct object and indirect object. Other peripheral semantic roles are Instrument, Comitative, Cause which are usually realized via oblique grammatical relations. For example, the event described by the sentence *I sent the letter to my sister* depicts three participants *I*, *the letter* and *my sister* which are associated to three different semantic roles: an Agent, a Theme and a Recipient respectively. Semantic

roles are abstract generalization which are invariant across languages and for this reason they can be used as comparative concepts as they are cross-linguistically applicable (Haspelmath 2010, Levin and Rappaport Hovav 2005). On the other hand, the linguistic expression of this semantic roles is language specific and may be subject to much variation. The range of variation of the linguistic encoding of semantic roles seems to be directly related to their semantic stability, meaning that the more semantically complex a role is the more variation in its morphosyntactic coding is attested both inter- and cross-linguistically. As I will discuss this is particularly true with the semantic roles of the Experiencer and the Stimulus, which shows an extremely complex and variable semantic characterization.

A participant bears certain properties independently of its roles, i.e. from its relational properties with the other participants and with the situation in its globality. Participant properties refer to features such as animate, human, abstract, specific and speech act participant (Lehmann 2002). These features have important consequences at the linguistic level as they generally determine how the argument expressing a given participant is linked to grammatical relations. Participants can be viewed as located on a scale (Figure 1 adapted from Lehmann 2002: 4) ranging from most speaker-like entities to least speaker-like entities. Participants higher in the scale are conceived as more salient and thus are usually associated to the higher grammatical relations of subject and object. Subjects encode the most salient argument and objects encode the next most salient argument.

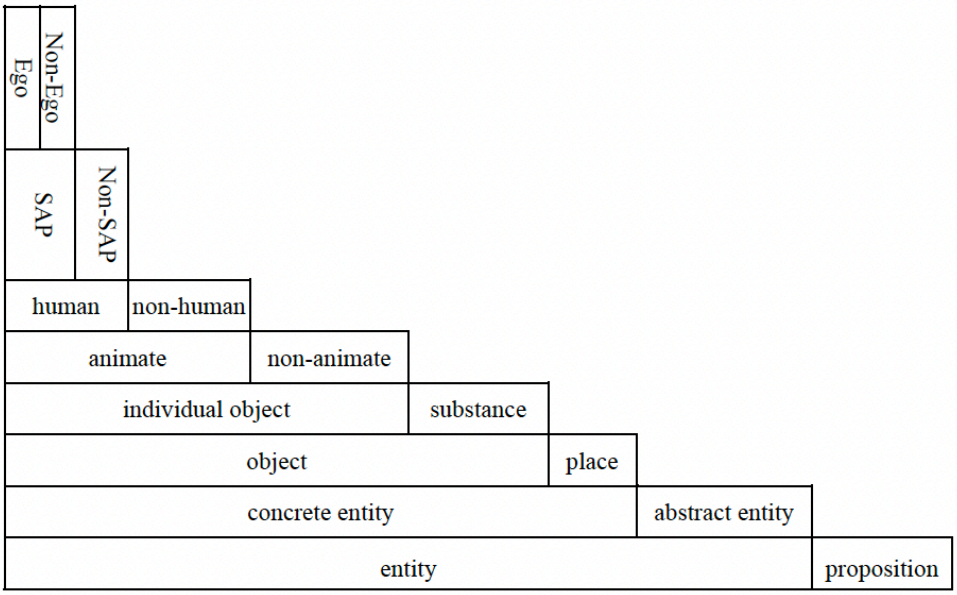

Figure 1: Participant properties (adapted from Lehmann 2002: 4).

According to their relational properties, participant roles and their generalizations (semantic roles) may show different semantic properties. Most important in this work are the semantic properties of control and affectedness. A participant which controls the situation is viewed as intentionally bringing

about the event and thus as responsible for it, while affected participants are involved in a situation which has some effects on them and that may also result in their change of state. These two semantic features are central because almost every situation can be viewed as entailing a participant either in control of the situation or affected by it and, for this reason, they allow generalizations over the different semantic roles. See for example the macrorole pair ACTOR – UNDERGOER in Role and Reference Grammar (Foley and Van Valin 1984, Van Valin and LaPolla 1997) or the Dowty's (1991) Proto-Agent and Proto-Patient roles. The prototypical semantic role in control of a situation is the Agent, while the prototypical affected role is the Patient. These two roles can be viewed as constituting the opposite edges of a two-dimensional continuum (represented in Figure 2) which ranges from the most agentive semantic roles to the most affected ones (Van Valin 1993, 2005, Dahl 2014). The Agent and the Patient role show an important characteristic crosslinguistically, that is the tendency to be consistently encoded with unitary morphosyntactic devices (see on this Dahl 2014: 183). This can be explained by their semantic stable characterization, as I will show in the following section (2.2.1) this is not the case for other more semantically unstable roles such as the Experiencer.

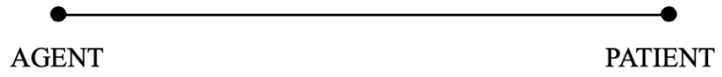

Figure 2: Prototypical Agent and Patient as poles on a scale (adapted from Dahl 2014: 183).

2.2.1. The encoding of experiential events

Following previous studies on experiential constructions (Verhoeven 2007, Luraghi 2020a, Fedriani 2012, Dahl 2014), I regard experiential situations as entailing two participants, the Experiencer and the Stimulus. Figure 3, adapted from Verhoeven (2007: 52), represents the components of an experiential situation.

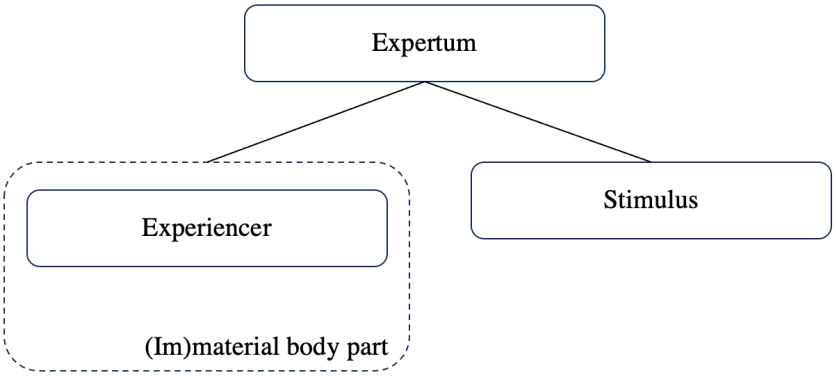

Figure 3: Components of experiential situations.

The Experiencer is the main participant of the event, a prerequisite for an experience to happen, and for this reason it is always present on a semantic-cognitive level. The Stimulus is the second participant (Blansitt 1978), less salient than the Experiencer, it is absent in many experiential situation types (for example in bodily sensations like "to be cold") and it can be omitted in the linguistic encoding, even if present on the semantic-cognitive level, like in "Let me think". The semantic core of an experiential situation is called Expertum and it is typically lexicalized in the predicate. The Expertum is the perception, cognition, bodily sensation or feeling that is experienced by the Experiencer. In Table 1, the components of an experiential situation are exemplified in English sentences for each experiential subdomain.

Table 1: Components of the five experiential types.

	EXPERIENCER	EXPERTUM	STIMULUS
Bodily sensations	The child	is cold	
Perceptions	I	heard	a sound
Cognitions	John	understood	the lesson
Emotions	Marco	loves	Sara
Volitions	Sara	wants	an ice-cream

The Experiencer's participant must be sentient, as it must be able to undergo and experience an internal state or change of state. For this reason, the prototypical Experiencer is animate and human. As a consequence of its participant's properties, the Experiencer is always high in the salience scale (Figure 1) and is typically a better candidate than the Stimulus to be linked to the subject relation

(Butt, Grimm and Ahmed 2006: 14). As regard the semantic properties, the Experiencer shares an important feature with Agents, that is animateness, yet it is also semantically similar to the Patient in many respects: typically, it is not volitional and does not control the event and it is affected by it. The situation is even more complex as even in experiential events in which the Experiencer shows some degree of agentivity and brings about the event (for example in agentive perceptions or cognitions), s/he is not a prototypical Agent as s/he also receives the consequences of the event. As Dahl (2014) points out, given this unstable semantic characterization the Experiencer can be conceived as located in between the Agent and the Patient semantic roles in the Agent-Patient scale in Figure 4.

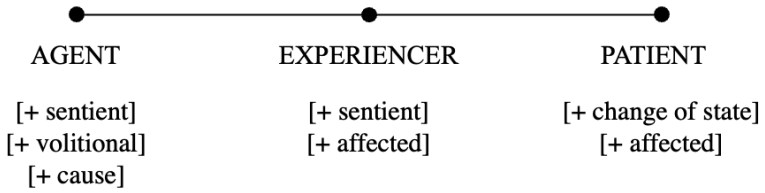

Figure 4: Experiencer location on the Agent-Patient scale (adapted from Dahl 2014).

The semantic properties of the Experiencer are subject to much variation also because they partly depend on the type of situation, and partly on the way in which a specific situation is construed (Luraghi 2020a). According to the type of event and to the way in which we construe it, Experiencers can exert some degree of control on the situation. For example, in an agentive perception (as in *I am listening to music*) the Experiencer is volitional and controls the experience, while in a bodily sensation (as in *I am hungry*) it is non-volitional, it does not control the situation, and it is physically affected by it. Table 2, taken from Luraghi (2021: 38), shows the variable semantic properties of the Experiencer comparing them with the properties of other more stable semantic roles: Agent, Recipient/Beneficiary and Patient. The ambiguous status of the Experiencer is reflected by its vague characterization in previous works on experiential verbs. For example, Næss (2007) defines Experiencers as volitional undergoers, Smith (1993) refer to them as showing a bilateral involvement, while Fedriani (2012: 2) speaks of a "semantic paradox realized by the Experiencer, who is at one and the same time animate, usually human, but non-instigating, and who undergoes something that happens to him or her".

Table 2: The semantic properties of the Experiencer and other sematic roles (from Luraghi 2014: 111).

	Experiencer	Agent	Recipient/Beneficiary	Patient
Humanness	+	+	+	+/-
Control	+/-	+	+	-
Volition	+/-	+	-	-
Affectedness	+	-	+	+
Change of State	+/-	-	-	+

The variability of Experiencer's properties correlates with the fact that many languages across the world do not have a specific case-marking for this semantic role, as opposed to other roles such as Agent, Patient or Recipient. As Verhoeven (2007) points out, there is a clear cross-linguistic tendency to encode experiential situations through the extension of constructions that are prototypically associated to other states or events, like transitive constructions or possessive constructions. Additionally, at the same time, she highlights that most languages lack a specific construction for experiential events. This led Haig (2009: 6) to define Experiencers as parasitic on other semantic roles, in the sense that they tend to co-opt the morphosyntactic coding of other roles.

In examples 1-10, I exemplify some of the different construction types which are used frequently by the languages of the world to encode experiences. Sentence 1 and 2 exemplify a very common construction which shows a single argument (the Experiencer) encoded in the nominative and lexicalizes the experience in an adjective or an intransitive verb. Sentence 3 represents a construction in which the body part of the Experiencer is encoded as the subject of a copular construction and the Experiencer is encoded as an attributive possessor. These three construction types are typical of bodily sensations as these experiential situations generally involve one participant only. Sentence 4 is an example of a transitive construction in which the Experiencer is encoded as the subject, while the Stimulus as the direct object. Inverse transitive constructions such as sentence 5 are used as well, even if far more rarely: in this construction the Experiencer is encoded as a direct object, while the Stimulus or the Expertum function as the subject. Experiencers are also very frequently encoded as possessors, either in predicative possessive constructions which are metaphorically extended to express abstract experiences, as in 6, or in attributive possessive constructions as in 7, in which the experience is encoded as the subject and the Experiencer is in the genitive and encoded as the possessor of the experience. A variant of this construction is in 8, in which the Experiencer is still encoded in the genitive but in this case it is linked to a bodily part marked with the locative and

conceptualized as the place where the experience is located. Similar to this latter construction is the external possessor construction (Luraghi 2020b, Haspelmath 1999) given in 9, which construes the bodily part as an Agent and the Experiencer, which is the possessor of the bodily part, is in the oblique case. Lastly, another typologically very common construction is the dative construction given in 10 in which the Experiencer is marked with the dative case and the Expertum (or the Stimulus) is in the nominative.

1 **Copular construction (Experience as subject)**

I am cold.

2 **Single nominative/absolutive argument construction**

Nae-ka *chup-ta.*

1.SG-NOM cold-DECL

"I am cold." (Korean, from Lehmann et al. 2000:72, cited in Verhoeven 2007: 74)

3 **Copular construction (Body part as subject and Experiencer as attributive possessor)**

Ua *faanoanoa* *o='u* *lagona.*

PRF sad POSS=1.SG feeling

" I am sad. Lit. My feelings are sad." (Samoan, Mosel and Hovdhaugen 1992:771)

4 **Transitive construction (Experience as subject)**

Sarah loves dogs.

5 **Inverse transitive construction (Expertum as subject, Experiencer as object)**

mete *i-kam* *yo.*

disease 3.SG-get 1.SG.ACC

"I am sick" (Mbula, Austronesian Language; from Bugenhagen 2001:73, cited in Verhoeven 2007: 81).

6 **Possessive Constructions**

Ho *fame.*

have.1SG.PRS hunger

"I am hungry." (Italian)

7 **Experiencer as possessor of the Expertum, Expertum as subject**

uyku-m *gel-di*

sleep.POSS.1SG has.come

"I am sleepy." Lit. "My sleep has come." (Turkish, from Fedriani 2012: 5)

8 Experiencer as possessor of a body part, Expertum as subject

mere	*sir=meṁ*	*dard*	*hai*
1SG.GEN-M.OBL	head.M.SG.OBL=in	pain(M)SG.NOM	be.3SG.PRS

"I have an headache (Lit. In my head there is pain.)"

9 Experiencer as External possessor (Body part as subject)

Mi fa male la testa.

1SG.DAT do.3SG.PRS pain the.F head(F)

"I have a headache." (Italian)

10 Experiencer in an oblique case and Stimulus in the nominative/absolutive

Mér	*smakkaðist*	*hákarlinn*	*vel.*
me.DAT	tasted	shark-the.NOM	well

"I liked the taste of fermented shark." (Icelandic, taken from Barðdal 2011: 65)

The examples given above are just a handful of the constructions that are found in languages around the world (for a detailed list of experiential constructions see Verhoeven 2007: 69-86). This lack of homogeneity exists both crosslinguistically and within the same language and has led some scholars to even doubt the usefulness of postulating the existence of such a semantic role. Dik (1981), for example, claims that he "doubt[s] whether a special semantic function of Experiencer is needed […], it turns out that both within and across languages, these states of affairs, which we shall call "experiences", are expressed in different ways. […] My impression is that it would be an oversimplification to assume that there is just one underlying structure for experiences, of the form "Experiencer–experiences–Experienced" (1981: 42–43).

In sum, the semantic role of the Experiencer displays much semantic complexity, as it shows ambiguity with respect to many semantic properties. The lack of a definite semantic characterization is mirrored by the lack of specific morphosyntactic expressions both across languages and within the same language. The Stimulus, on the other hand, covers a larger range of participant types than the Experiencer. It can be the inanimate element that triggers a perception (as in 11), it can be the person toward which an emotion is directed (as in 12), or it can be the content of a cognitive state or process (as in 13). Even states of affairs can function as Stimuli.

11 *She heard [a thunder].*

12 *John loves [his siblings] very much.*

13 *I was thinking [about you].*

From the point of view of semantic properties, much like the Experiencer, the Stimulus is highly variable. Croft (1993, 2003, 2022) points out that the Stimulus role can be conceptualized either as the trigger of a situation which alters the (mental) state of the Experience, or as the target toward which the Experiencer's attention is directed (see on this also Osmond 1997). Notably, this different conceptualization of the Stimulus correlates with specific differences in the semantic properties of the Experiencer. When the Stimulus is conceptualized as a cause and thus triggers the experience, the Experiencer is conceived as lacking control and volitionality and as affected by the event initiated by the Stimulus. In contrast, when the Stimulus is conceptualized as the target content of the experience, the Experiencer shows some agentive properties as it volitionally directs its attention toward the Stimulus. Croft (1993) represents these two different cause-target construals of the Stimulus and the Experiencer through the schema in Figure 5.

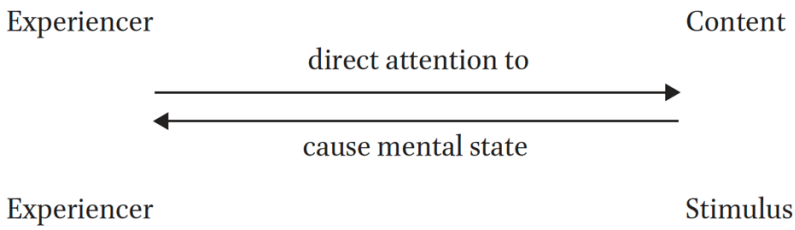

Figure 5: The different cause-target construals of the Stimulus and the Experiencer (adapted from Croft 1993: 64).

Additionally, the Stimulus is semantically different from a Patient. When it is conceptualized as a cause it cannot be assimilated to a Patient as it cannot be affected by the experience it is responsible of. While when it is conceptualized as a target concept it differs from the Patient role as it is not physically affected by the event. This indeterminateness with respect to the semantic properties of volitionality and affectedness has led Næss (2007) to conceive the Stimulus role as neutral, as it is not necessarily associated to specific semantic (and participant) parameters.

Fedriani (2012: 34-35) divides semantic roles into three groups, on the basis of their participant properties (inherent properties in her terminology) and semantic properties (relational properties in her terminology). The three groups are: 1. Relation-based roles, that are mainly determined on the basis of relational/semantic properties, 2. Participant-based roles, mainly defined by inherent properties and 3. Meta-roles which are not defined either by their inherent or their relational properties. According to Fedriani, the semantic role of the Patient is a perfect example of relation-based roles, as it is underspecified in relation to participant properties, and it is mostly

characterized by its relation to the other participants and the situation in its globality. Relation-based roles such as the Patient tend to show a high level of consistency in their morphosyntactic encoding. As Fedriani points out, the Experiencer and the Stimulus are not relation-based roles, as they are not univocally defined by relational/semantic properties. The Experiencer is not specified with respect to its control, agentivity, and volitionaliy, but it is inherently characterized by the fact of being animate and preferably human, thus it is a participant-based role. While the Stimulus is a meta-role, as it is underspecified in relation to its participant's properties, and it is also underspecified with respect to its relations within the event. On the line of other scholars (Verhoeven 2007, Dahl 2004, Luraghi 2020a, Lehmann 1991 and many others), Fedriani thus concludes that the lack of relational/semantic properties' specification results in the lack of a well-defined linguistic coding.

This ambivalence in the semantic properties of experiences, which can be interpreted either as entailing a non-agentive Experiencer receiving the consequences of an event triggered by an external source, that is the Stimulus, or as entailing an agentive Experiencer volitionally bringing about some mental or perceptual action, are mirrored in two main tendencies in the morphosyntactic encoding of experiences. As Bossong (1998) points out, languages of the world may assimilate the Experiencer to an Agent and encode it as a nominative subject or they may assimilate it to a Patient/Recipient and encode it with an oblique marking. In Bossong's typology, the first type of Experiencer encoding results from a generalizing strategy, which uses the same case-marking to express the main argument in a sentence regardless of its semantic role, while the second type of Experiencer encoding derives from an inverting strategy, in which not the Experiencer, which is the most salient argument, but the Stimulus is encoded as a nominative subject. Bossong did a typological study comparing the expression of experiential events in 40 languages and concluded that in SAE (Standard Average European) languages the generalizing strategy is predominant (see on this also Dahl 1990 and Haspelmath 1998). South Asian languages on the other hand show a quite different characterization, as the most productive strategy in these languages is the inverting one. South Asian languages are well known for showing a heavy preference for oblique subjects and the next section is devoted to a brief exposition of experiential constructions in South Asian languages, with a focus on Indo-Aryan languages.

2.3. Experiential constructions in South Asian languages

A common feature of many South Asian Languages is the tendency to encode the Experiencer in oblique cases. This tendency is so pervasive in these languages that it has been identified as one of the shared properties that contribute to define South Asia as a linguistic area (Masica 1976, 2001,

Hock 2016a, Subbarao 2012, Verma and Mohanan 1990, Verma 1976). The South Asian linguistic area consists of many unrelated languages belonging to different linguistic families which developed several shared features as a consequence of a long-lasting contact. South Asian linguistic families comprehend Indo-Aryan, Dravidian, Tibeto-Burman, Munda and Dardic. Languages belonging to these families are spoken in a vast region that spreads from Afghanistan in the West to Bhutan and Bangladesh in the East covering all the territories in the middle: Pakistan, India, Nepal, Tibet, Bhutan and Sri Lanka (see on this Masica 1976, 1991).

Oblique Experiencers are particularly frequent in Indo-Aryan and Dravidian languages, while they are quite rarely attested in Tibeto-Burman languages (Masica 1976, Bickel 2004) although Hook (2014) argues that some Tibeto-Burman languages which are in proximity with the Indo-Aryan area show oblique Experiencers as well. Sentences 14-23 exemplify this construction in some Indian languages. As the examples show, there are two main types of constructions occurring with an oblique Experiencer. In the first, the Stimulus is in the nominative and the Expertum is lexicalized in a simple verb (as in 14, 17, 18, 19 and 21). In the second construction, the nominative element is the Expertum, not the Stimulus, and the verb carries a vague semantics such as "happen" or "feel" and functions as a light verb in a noun - verb sequence. This latter construction typically encodes bodily sensations or experiential types that are not generally associated to a cause or a specific content, and for this reason do not encode a Stimulus (see for example 15, 16 and 20). The oblique Experiencer is most commonly marked with the dative, but also other case markings are possible in South Asian languages. In Bangla for example the prototypical case for the Experiencer is the genitive (Hock 2016a, Klaiman 1980). Example 22 shows a genitive Experiencer and it is interesting because it is also a case of an impersonal construction (a construction type that is quite rare typologically, see Fedriani 2012: 48) in which the Stimulus is marked with the accusative and there is no nominative element. A genitive marking of the Experiencer is also found in Assamese (23) and Oriya (Masica 1991: 346, Bickel 2004: 88).

14 *mini-laa* *ravi* *aavD-t-o*
Mini.3SG.F-DAT Ravi.3SG.M like-PRS-3SG.M
"Mini likes Ravi." (Marathi, Indo-Aryan, adapted from Wali 2004: 245)

15 *hun-kaa* *bhukh* *lag-l-ain(h)*
3HON-DAT hunger feel-PT-3HON.NN
"He felt hungry." (Maithili, Indo-Aryan, adapted from Yadava 2004: 254)

16 *rinar* *Dim* *bhalo* *lage*

Rina.IND egg/s good feels

"Rina likes eggs." (Bangla, Indo-Aryan, adapted from Dasgupta 2004: 131)

17 *paaDoši-ne* *maraaThi* *aawD-š-e*

neighbor-DAT Marathi know-FUT-3

"The neighbor will know Marathi."(Gujarati, Indo-Aryan, adapted from Mistry 2004: 5)

18 *maalati-ki* *bazaar-loo* *endaroo$_i$* *kanipincee-ru$_i$*

Malati.3SG.NOM-DAT market-in many were.visible.3PL.H

"Malati saw many people in the market." (Telugu, Dravidian, adapted from Subbarao and Bhaskararao 2004: 162)

19 *kuTTi-l'l'ə* *taNukk-unnu*

child-DAT feel.cold-PRES

"The child feels cold." (Malayalam, Dravidian, adapted from Jayaseelan 2004: 230)

20 *makkaL-ige* *tama-g-ee* *naacike* *aayi-t-u*

children-DAT selves-DAT-EMPH shame happen-PST-3NSG

"The children themselves felt ashamed." (Kannada, Dravidian, adapted from Amritavalli 2004: 6)

21 *avan-ukku* *pasi-kkar-tu*

he-DAT hunger-PRS-3N.SG

"He is feeling hungy." (Tamil, Dravidian, adapted from Lakshmi Bai 2004: 246)

22 *amar* *toma=ke* *cai*

1.GEN you=ACC wants

"I need you." (Bangla, Indo-Aryan, taken from Klaiman 1980: 275)

23 *mor* *piyāha* *lāg-is-e*

1SG.GEN thirst.NOM perceptible-IPRF-3

"I am getting thirsty." (Assamese, Indo-Aryan, taken from Sharma 1963: 122, cited by Bickel 2004: 87).

These constructions align with a general tendency in South Asian languages to not use the transitive construction to encode situations which involve a non-volitional participant, since the transitive

pattern is typically associated with an agentive participant (Masica 1991: 350). The oblique first participant construction is associated to a wide range of situations, all depicting events which are not conceived as brought about by the participant, but as happening to them without their control and volitionality. The semantic-functional domains of oblique subjects in Indian languages are summarized by Subbarao and Bhaskararao (2004) as follows:

a. Psychological states and emotions
b. Physiological and mental ailments
c. Natural phenomena pertaining to the body
d. Perceiver of visual and auditory actions
e. Expression of possession and kinship
f. Expression of obligation and necessity
g. Acquisition of knowledge and skill
h. Part-whole relationship

Verhoeven (2007: 71) proposes a classification of the constructions used by the languages of the world for the expression of experiences. She singles out four types based on the different orientation of the predicate toward one of the components of the situation and she identifies the goal of the orientation in the element that is given the subject function in the sentence. In this way, she is able to distinguish between Experiencer-oriented predicates, body part-oriented predicates, Stimulus-oriented predicates and Expertum-oriented predicates. It would be tempting to associate the constructions presented in the examples 14-23 above with Stimulus-oriented predicates or Expertum-oriented predicates. However, this is not the case, as scholars of Indian languages (Kachru and Bhatia 1976, Bhaskararao and Subbarao 2004, Verma and Mohanan 1990, Masica 1991) generally agree in considering these oblique Experiencers as non-nominative subjects. Among the tests that have been proposed for the diagnostics of subjecthood (Subbararo 2012) the following are generally agreed upon:

i. The oblique Experiencer is coreferential of possessive reflexive.
ii. The oblique Experiencer is coreferential with the unexpressed subject of conjunctive participles.
iii. The oblique Experiencer is the leftmost argument in pragmatically unmarked sentences.
iv. The oblique Experiencer is marked with the genitive case in nominalizations.
v. The oblique Experiencer is coreferential with the unexpressed nominative subject in coordinated sentences.

South Asian languages show a tendency to mark the Experiencer with an oblique case, but they still recognize the Experiencer as the more salient argument and thus assign it the syntactic status of a subject. I conceive these South Asian oblique Experiencer constructions as Experiencer-oriented predicates, as the tests listed above show that the syntactic pivot (Dixon 1994) in these sentences is still the Experiencer even if it is marked with the oblique case.

Oblique Experiencer constructions are not the only strategy used in South Asian languages for the expression of experiential events. Nominative single argument constructions, transitive constructions, oblique second argument constructions and other construction types can be used as well. For example, as I will discuss more in detail in the next chapters, Hindi experiential verbs show a wide range of case marking patterns. The dative Experiencer construction is the prototypical in Hindi, and it is instantiated by two different construction types: encoding either the Stimulus or the Expertum as a nominative NP. Notably, the dative construction mostly occurs in Hindi with intransitive noun verb complex predicates, in which the noun lexicalizes the semantic core of the experience (24). Other markings on the Experiencer can be used as well (such as the genitive, see on this also Montaut 2016, Kachru 1990), and these different case markings supply a semantic change in the construal of the event. Besides oblique Experiencer constructions, the transitive construction (25) is frequently used as well, especially for perceptions and cognitions. And other less frequent constructions exist in Hindi, such as the nominative single argument construction (26), and the oblique Stimulus constructions (27). Additionally, different types of more complex constructions are found such as the locative construction (28) encoding the Experiencer as an attributive possessor of a bodily part which is in the locative case and conceptualized as the place where the experience is located.

24 रग्घू को इस समय मर्मान्तक पीड़ा हो रही थी।

ragghū=ko	*is*	*samay*	*marmāntak*	*pīṛā*	*ho*	*rah-ī*
ragghu=DAT	this.OBL	time	piercing	pain(F.SG.NOM)	be	PRGR-F
th-ī.						
be.PST-F.SG						

“Ragghu was in piercing pain at that moment.”

25 गॉंव-वालो की फरियाद कौन सुनता!

gāṁv-vāl-oṁ=kī	*phariyād*	*kaun*	*sun-t-ā!*
villager(M)-PL.OBL=GEN	complaint(F.SG.NOM)	who.NOM	listen-IPRF-M.SG

“Who listens to the complaints of the villagers?”

26 वह बच्चा प्यासा है।

vah	*bacc-ā*	*pyās-ā*	*hai*
that.NOM	child(M)-SG.NOM	thirst-M.SG	be.3SG.PRS

"That child is thirsty."

27 ओंकारनाथ ने दु:खी आंखों से पत्नी की ओर देखा।

oṁkārnāth=ne	*du:kh-ī*	*ānkh-oṁ=se*	*patnī=kī_or*	*dekh-ā.*
onkarnath=ERG	sad-F	eyes(F)-PL.OBL=INS	wife=towards	look_at-PRF.M.SG

"Onkarnath looked at his wife with sad eyes."

28 शाम को उसके पेट में दर्द होने लगा।

śām=ko	*us=ke*	*peṭ=meṁ*	*dard*
evening=at	3SG.OBL=GEN	stomach(M.SG.OBL)=in	pain(M.SG.NOM)
ho-n-e	*lag-ā*		
be-INF-OBL	start-PRF.M.SG		

"In the evening he started having a stomachache."

The pervasive use of oblique subjects in Hindi and in Indo-Aryan in general seems to be an innovation, as many scholars pointed out that Sanskrit did not show any trace of oblique subject constructions (Hook 1990, 1991, Dahl 2014, Butt and Deo 2013). Dahl (2014) investigated the expression of experiences in Vedic Sanskrit, across the five subdomains of sensations, perception, cognitions, volitions and emotions. He shows that Vedic displayed three different construction-types: the nominative-Experiencer construction, the dative-Experiencer construction and the accusative-Experiencer construction. However, even if the language allows oblique markings on the Experiencer, these case markings are heavy marginal and limited to specific subdomains, and the language displays a very strong preference to mark the Experiencer with the nominative. In particular, Dahl argues that the dative marking (as in 29) is allowed in Vedic only for the expression of positive emotions such as taste (*svad-*) or seem/please (*chand-*), while the accusative-Experiencer construction (as in 30) is typically used for the expression of mental or bodily states and emotions. However, it should be noted that this latter construction mainly occurs with verbs displaying a causative morphology and as Dahl points out the accusative marking on the Experiencer is most probably triggered by the causative form of the verb. While the nominative Experiencer can occur either in a single argument construction, or with a Stimulus. The Stimulus can be marked either with the accusative (as in 31) or with the dative

(as in 32). Dahl notes that the nominative-accusative construction shows much broader lexical variation than the nominative-dative construction and it seems to be the preferred construction for the expression of experiential situations in Vedic, while the other markings on the Experiencer show a limited scope. More importantly, there is no evidence to consider the oblique experiencers in these constructions as non-nominative subjects (Hock 1990, Butt and Deo 2013).

29 Dative Experiencer

cakraṃ	*yad*	*asya*	~	*apsv*	*ā*	*niṣattam*	*uto*
wheel.NOM	when	he.GEN		waters.LOC	in	rest.in.PPP	and

tad	*asmai*	*madhv ic*	*cacchadyāt /*
that.NOM	he.DAT	honey.NOM	even seem.OPT.SG

"When his wheel is set fast in the waters, even that would seem like honey to him." (Rigveda X 73.9b after Klein 1985: 453, adapted from Dahl 2014: 191)

30 Accusative Experiencer

pra	*bodhaya*	*jaritar*	*jāram*	*indram /*
forth	arouse.IMP.2SG	singer.VOC	friend. ACC	Indra.ACC

"O singer, arouse (our) friend Indra!" (Rigveda X 42.2, adapted from Dahl 2014: 196)

31 Nominative Experiencer - Accusative Stimulus

yaṃ	*jīvam*	*aśnavāmahai*	*na*
who.ACC	alive.ACC	reach.SBJV.1PL	not

sa	*riṣyāti*	*pūruṣaḥ //*
he.NOM	become.hurt.SBJV.3SG	man.NOM

'He whom we shall encounter alive that man shall not become hurt." (Rigveda X 97.17cd, adapted from Dahl 2014: 190)

32 Nominative Experiencer - Dative Stimulus

ayaṃ	*ha*	*tubhyaṃ*	*varuṇo*	*hṛṇīte*
this.NOM	indeed	you.DAT	Varuna.NOM	be.angry.PRS

"Indeed this Varuna a is angry at you." (Rigveda VII 86.3, adapted from Dahl 2014: 190)

Given the pervasive use of oblique subjects in modern Indo-Aryan languages and its absence in Old Indo-Aryan (OIA), many scholars have focused on investigating of the emerging of this phenomenon

in South Asian languages. Barðdal and Eythòrsson (2009), Barðdal and Smitherman (2012) and Barðdal (2013), for example, propose a diachronic application of Construction Grammar to oblique subject construction. They argue that OIA had in fact oblique subjects and that New Indo-Aryan dative Experiencer constructions are not an innovation and are inherited from Proto-Indo-European, like Icelandic dative subjects.

However, Butt and Deo (2013) note that this view is problematic in many respects. First, it is assumed that there is a uniform case-marking system that New Indo-Aryan inherited from Old Indo-Aryan. However, this is not the case, as the complex inflectional system existing in Vedic and Sanskrit was lost in Middle Indo-Aryan and is reduced to a binary inflection opposition between a direct and an oblique case in NIA languages (Butt 2006b, Butt and Ahmed 2010, Montaut 2013, Reinöhl 2016). The old inflectional case-marking system is replaced by postpositions in modern Indo-Aryan languages. Hence, one cannot trace a contiguous line connecting OIA dative arguments to NIA dative arguments. Moreover, as mentioned above, oblique Experiencers in Sanskrit were not oblique subjects (Hock 1990, 1991), so there is no reason to believe that an oblique subject construction was inherited from Old Indo-Aryan by New Indo-Aryan languages. Butt and Deo (2013) suggest that the oblique case marking on Experiencers is an innovation that appeared after the rise of ergativity in Indo-Aryan as a "part of a larger semantic-based Differential Case Marking system" which led to semantically motivated case alternations in all South Asian languages (see on this also Butt 2006b, Butt and Ahmed 2010).

In particular, focusing on Marathi, they argue that dative subjects developed in three ways: 1. From verbs expressing change of state in Sanskrit which developed into experiencer verbs; 2. From Sanskrit verbs that were originally intransitive which acquired an Experiential reading, 3. And from nominative Experiencers transitive predicates which were reanalyzed as dative Experiencers predicates. I will come back to these issues in chapter **Errore. L'origine riferimento non è stata trovata.**, where I try to make some observations of the rise of the dative construction in Hindi and its spread in the grammar.

3. The language under investigation: Modern Standard Hindi

3.1. A historical background

Hindi is a New Indo-Aryan language predominantly spoken in the northern regions of India. It is one of the most widely spoken languages worldwide, with over 300 million speakers who use it as their mother tongue. The number of speakers increases significantly when non-native speakers are included. The highest concentration of native speakers resides in the area referred to as *Hindi belt*, which comprises the Indian states of Uttar Pradesh, Uttaranchal, Madhya Pradesh, Chhattisgarh, Bihar, Jharkhand, Haryana, Rajasthan, Himachal Pradesh, and the Delhi territory. In urban areas outside these states, several non-standard Hindi varieties are also spoken, such as those found in Mumbai, Kolkata, and Hyderabad. Recent diasporas have also led to the emergence of Hindi-speaking communities scattered across Europe, the US and Canada (Shapiro 2003). As is the case for all languages with extensive reach, the term 'Hindi' denotes a linguistic continuum situated in the northern regions of the Indian subcontinent, rather than a language with well-defined boundaries. As early as the beginning of the 20th century, Grierson (1904: 3) observed that "the name Hindi is popularly applied to all the various Aryan languages spoken between the Panjab on the west and the river Mahānandā on the east; and between the Himalayas on the north and the river Narbada on the south."

A large number of regional varieties belong to this continuum. The standard Hindi variety, recognized as the official language by the Indian government and taught in schools, is based on the Western variety called *khaṛī bolī*, spoken mainly in Delhi and the areas surrounding the capital (McGregor 1977, Chatterji 1960, Schmidt 2003, Drocco 2019). This variety has been referred to with many names: Urdu, Hindi/Hindavi, Hindustani. *Urdu* is a Turkish word from the locution *zabān-e-urdū-e-mu'alla* meaning "the language of the high camp" that was used to identify the language spoken by the Moghul soldiers allocated in the area around the Red Fort in Delhi around the 17th century. *Hindi*/*Hindavi* is a foreign term that was used by the Turkish and Afghan invaders to refer to the language spoken by the populations inhabiting the northern regions of India (Montaut 2004b: 2). While *Hindustani* was the term used by the British colonizers to identify the lingua franca spoken in the territory called Hindustan, which at that time also comprised the territories that today are in Pakistan.

The relationship between the standard language and regional dialects in Hindi is intricate, with various dialects attaining high literary status. During the 16th and 17th centuries, literature in the northern regions was primarily written in *Braj bhāṣā* and *Avadhī*, rather than in the dialect of the areas surrounding the capital. These two languages served as vehicles for different types of literature.

Braj bhāṣā was the language of the city of Mathura and was primarily used for composing Kṛṣṇa's devotional lyrics. According to the tradition, Kṛṣṇa was born in Mathura, and thus Braj bhāṣā was chosen as the language used for the literary production consecrated to him, whereas literature in *Avadhī* found its roots in the city of Ayodhya and was used for composing the *mahākāvya*, the great Indian heroic poem.

In the 18th century, the *khaṛī bolī* dialect became the most widely understood variety in the northern regions of India due to medieval religious practices. As some scholars pointed out (Narula 1976, Montaut 2013), religion had a profound effect on the linguistic evolution of India: many *sadhus*, *yogis*, and pilgrim saints *nirguṇ* who wandered around the northern region of the Indian sub-continent spread the *khaṛī bolī* dialect. In order to make preaches and devotional songs accessible to the greatest number of people, these sadhus abandoned the elements of their language and developed a common vocabulary influenced by the more familiar *khaṛī bolī* variety. Considering that *sadhus* constituted one fifth of the population in northern India during feudal times (Narula 1976), it is not surprising that religion contributed so much to the diffusion of the *khaṛī bolī* variety.

The official language of Pakistan, called Urdu, is also related to the *khaṛī bolī* variety. Hindi and Urdu are the result of a political process of linguistic differentiation that occurred between the 19th and 20th centuries. The standardization process of Hindi began in the early 19th century and was primarily initiated by the linguistic policies of the British Empire. After establishing the Fort William College in Kolkata at the beginning of the 19th century, the colonizers felt the necessity of mastering the language spoken by Indians and introduced courses to educate their administrators. The British thus started a course for the linguistic variety they called 'Hindustani', which was the *khaṛī bolī* dialect written in the Persian script and heavily influenced by Arabic and Persian terms. This was the language that was used in the Moghul courts and that served as the empire's second official language alongside Persian. In 1802, the Fort William College began offering courses in Hindi, which was then understood as a distinct language from Hindustani and was actually a variant of the *khaṛī bolī* dialect, with lesser Persian influence and written in the Devanagari script. As McGregor (1967: 116) points out:

> From experience, it was also realized that some place should be made for the study of a form of language which should not rely on the Perso-Arabic vocabulary of the lingua franca but come nearer in its vocabulary to the various local dialects of the Hindi language area. [...] Since Khari Boli was now widely understood in northern India, it was perhaps inevitable that Khari Boli rather than Braj Bhasha should become the basis of this new style.

The establishment of two different teachings at the Fort William College for what was essentially the same language reinforced the belief of the existence of two distinct languages: Hindustani/Urdu, the language spoken by Indian Muslims, and Hindi, the language of the Hindu community. This idea, however, had no foundation, as *khaṛī bolī* was the union of two idioms: the variety spoken in the capital and the variety of rural areas, with the only difference that in Delhi, where the economic, cultural, and administrative contacts with the court were stronger, the tendency to Persianize the language prevailed, while in the countryside there was a strong influence of regional idioms, with the extensive use of Indo-Aryan terms (Milanetti 2002). British linguistic policies institutionalized a distinction that had never existed, convincing the Indians themselves that the origin of this distinction was based on identity: thus, Urdu became the language of the Muslim community, while Hindi became the language of the Hindu community. Note, however, that some scholars pointed out that there was no deliberate intention from the British Empire to divide Hindi and Urdu. For example, Rai (1984) emphasizes that the British simply aligned to a distinction that was already acknowledged as back as the 1755, when the Urdu poet Shah Hatim in his edicts codified the transformation of the colloquial speech spoken in the 14th that emerged from the mix of Panjabi, Harianvi, Braj and Rajasthani into the Persianized speech spoken by the elites in the 17th -18th (see on this also Montaut 2004a: 3). This means that the British colonizers simply recognized and formalized a pre-existing distinction and institutionalized it.

In the late 19th century, the differentiation between Hindi and Urdu was accentuated by Hindu nationalists who began a process of Sanskritization of the Hindi language. The aim was to distance Hindi from Urdu, which was associated to Muslims, and to create what they called *śuddh hindī*, a "pure" Hindi not "contaminated" by Arabic-Persian terms. This process of Sanskritization was driven by Hindu nationalist movements that emerged in the last quarter of the 19th century and developed into pro-Hindi movements. The proponents of the movement claimed that Hindi and Urdu were two distinct languages, and that the overcoming of Hindi over Urdu would represent the redemption of the Hindu community from their ancient Muslim oppressors. These nationalists sought to bring India back to the form it had during the Gupta Empire, considered the golden age of ancient India, before the arrival of Islam. Since the language of the great ancient literature of the Gupta empire and previous literature was Sanskrit, the process of Sanskritization of the language was begun (King 1994, Orsini 2010, Consolaro 2003).

The extreme outcome of this linguistic policy was the so-called process of *tatsemization*, which entailed replacing *tadbhava* terms and Arabic-Persian terms with corresponding Sanskrit forms (called *tatsama*). *Tadbhavas* are terms resulting from natural phonological changes occurred in the Prakrit dialects, phonetically characterized by the elimination of consonantal clusters and diphthongs

and are therefore New Indo-Aryan words inherited from Middle Indo-Aryan languages, whereas *tatsamas* are Sanskrit terms transplanted into Hindi in their original form. The result of this *tatsemization* was an artificial non-colloquial language, politically constructed at the table by nationalists. After the partition between India and Pakistan in 1947, Muslim nationalists initiated a mirroring purification process. This process of Persianization involved cutting off all *tadbhava* terms and replacing them with Arabic-Persian terms. The outcome of this process was an Urdu language that was distinct from its original *kharī bolī* matrix, as much as Sanskritized Hindi was. As Narula (1976: 83) lucidly described:

> The controversy about Urdu and Hindi being the languages of two different groups in the same areas gets easily solved when it is realised that both Urdu and Hindi are non–colloquial languages created out of the same speech – Khaṛi Boli – of Delhi and the neighbourhood.

Today, Hindi and Urdu are officially recognized as two distinct languages, and they express different literary traditions and cultural contexts. However, they share the same morphological and syntactic bases, differing only in vocabulary and script. Hindi employs the Devanagari script, which was once used for Sanskrit but has been adapted to the phonology of the modern language. Moreover, Hindi has a rich vocabulary derived from Sanskrit. On the other hand, Urdu uses the Persian alphabet, modified to accommodate phonemes absent in Persian, such as retroflex consonants, and has a vocabulary rich in Persian and Arabic terms. Besides these differences, on a colloquial level these two languages are the same: speakers of Urdu and Hindi can understand each other easily, and the significant differences between the two languages mainly relate to the literary and religious domains.

3.2. Preliminary remarks on Hindi morphosyntax

This chapter provides a general introduction to Hindi morphosyntax. In particular, I will deal with nominal morphology and the use of postposition in section 3.2.1 and 3.2.2; while in section 3.2.3, I introduce Hindi verbal morphology. I then move on to discussing the extensive use of differential case markings in the language, focusing on split ergativity and differential object marking (section 3.2.4). In the same section, I also address non-canonical alignments and the use of non-nominative subjects. In section 3.2.5 I introduce complex predication in Hindi: in section 3.2.5.1 I focus on verb-verb complex predicates, while in section 3.2.5.2 I focus on noun-verb complex predicates, which are particular productive in the expression of experiential events in Hindi.

Hindi a is head-final language: it has postpositions and its modifiers (adjectives and genitives) precede the modified elements. The verb is always in final position when the sentence is unmarked.

Adjuncts and adverbial elements can be placed anywhere in the sentence, but they rarely occur after the verb. An example of a canonical unmarked sentence in Hindi is 33.

33. कल रतन को रुपये दे दूँगा।

kal	*ratan=ko*	*rupaye*	*de*	*dūṁ-g-ā*
tomorrow	ratan=DAT	rupee(M.PL.NOM)	give	give.1SG-FUT-M.SG

"I will give the rupees to Ratan tomorrow."

Hindi is a SOV language, in the sense that in pragmatically unmarked clauses the order of the constituents is subject-object-verb. However, word order in Hindi is free: constituents can occur in any position in the sentence, depending on information structure.

Hindi is a split-ergative language, in which the split is based on verbal aspect: the language exhibits a nominative-accusative alignment in the imperfective aspect and an ergative-absolutive alignment in the perfective aspect. However, syntactically Hindi remains an accusative language (Butt and Ahmed 2008, Bitter and Hale 1996). This means that the ergative argument in the ergative sentence behaves syntactically as a subject with respect to control, coordination, anaphora, and so on. Hindi also shows Differential Object Marking (DOM): the direct object is marked when it is human and/or specific (Butt 1993). Moreover, the language exhibits a tendency to use non-nominative subjects (Kachru 1990, Mohanan 1994, Butt and King 2002), mostly dative subjects (for Experiencers). However, locative subjects (for Possessors in alienable possession), genitive subjects (for Possessors in inalienable possession) and instrumental subjects (for ineffective Agents) are also frequently found. For a more detailed analysis of ergativity, DOM and non-nominative subjects see section 3.2.4 below.

3.2.1. Nominal morphology

Nouns in Hindi are inflected for number and case. These grammatical categories are not always overtly marked, however, even if many nouns have zero markers, their grammatical categories have consequences for agreement patterns. All nouns in Hindi are assigned to either masculine or feminine gender: with animate nouns, grammatical gender corresponds to the referent's gender, while with inanimate nouns gender is arbitrary. Both masculine and feminine nouns are divided into two classes according to the endings they take in the singular, direct form: masculine nouns may end in *-ā* or in various other endings; while feminine nouns may end in *-ī* or in various other endings. Hindi nouns have two numbers: singular and plural.

Inflection in Hindi is reduced to a binary opposition between a direct case used for subjects and direct objects and an oblique case used when nouns are followed by postpositions. Table 3 shows the inflection of each class of nouns according to the categories of gender, number and case.

Table 3: Inflecting paradigm nouns in Hindi.

	Masculine in *-ā*	**Masculine non in *- ā***	**Feminine in *-ī***	**Feminine non in *-ī***
SG. DIR.	*bacc-ā* (child.M)	*ghar* (house)	*bacc-ī* (child.F)	*kitāb* (book)
SG. OBL.	*bacc-e* (child.M)	*ghar* (house)	*bacc-ī* (child.F)	*kitāb* (book)
PL. DIR.	*bacc-e* (child.M)	*ghar* (house)	*bacc-iyāṁ* (child.F)	*kitāb-eṁ* (book)
PL. OBL.	*bacc-oṁ* (child.M)	*ghar-oṁ* (house)	*bacc-iyoṁ* (child.F)	*kitāb-oṁ* (book)

Adjectives are divided into two inflectional classes: adjectives ending in *-ā* in the masculine singular direct form and adjectives ending with any other form. Adjectives inflect according to a simplified version of the nominal inflecting system, as shown in Table *4*. Adjectives in *-ā* form the corresponding feminine alternating *-ā* with the feminine ending *-ī*; while adjectives that do not end in *-ā* do not distinguish between masculine and feminine. Adjectives inflect according to the same categories of noun inflection, but the adjectival paradigm is highly reduced: adjectives in *- ā* show the same form (ending in *-e*) for the singular oblique and the plural (both direct and oblique), while the feminine correspondent shows the ending *-ī* for all cases and numbers. Adjectives not ending in *-ā* do not inflect.

Table 4: Inflecting paradigm of adjectives in Hindi.

	Adjectives ending in *-ā*		**Adjectives not ending in *-ā***	
	Masculine	**Feminine**	**Masculine**	**Feminine**
SG DIR	*baṛ-ā thail-ā*	*baṛ-ī mez*	*lāl thail-ā*	*lāl mez*
	(big bag)	(big table)	(red bag)	(red table)
SG OB	*baṛ-e thail-e*	*baṛ-ī mez*	*lāl thail-e*	*lāl mez*
	(big bag)	(big table)	(red bag)	(red table)
PL DIR	*baṛ-e thail-e*	*baṛ-ī mez-eṁ*	*lāl thail-e*	*lāl mez-eṁ*
	(big bag)	(big table)	(red bag)	(red table)
PL OB	*baṛ-e thail-oṁ*	*baṛ-ī mez-oṁ*	*lāl thail-oṁ*	*lāl mez-oṁ*
	(big bag)	(big table)	(red bag)	(red table)

Since inflection only distinguishes between a direct case and an oblique case, Hindi uses a large set of postpositions to distinguish syntactic and semantic functions. Hindi cases and their postpositions

are shown in Table 5 (for an account of Hindi postposition see section 3.2.2 below). The nominative is the only case that can trigger agreement with the verb.

Table 5: Cases in Hindi.

Case	Noun form	Postposition
Nominative	Direct	*Ø*
Ergative	Oblique	*ne*
Accusative (DOM)	Oblique	*ko*
Dative	Oblique	*ko*
Genitive	Oblique	*kā* / *ke* / *kī*
Locative	Oblique	*meṁ* "in", *ke pās* "near"
Instrumental	Oblique	*se*

The pronominal system does not involve gender distinctions, but Hindi pronouns inflect for number and case (direct and oblique). The direct form is used when the pronoun stands in the nominative and is the subject of the sentence. Note that the direct case is also used to encode direct objects; however Hindi exhibits Differential Object Marking, so the direct object is marked when it is animate and/or specific/definite. The consequence of this is that pronouns, which are at the leftmost side of the animacy and individuation hierarchy[1] (Croft 2003: 130), rarely stand in the direct form when they function as the direct object of a sentence. The oblique form of the pronoun is used when it is followed by any of the postpositions shown in Table 5. When it stands in the dative, the pronoun may be in the oblique form and followed by the postposition *ko*, or it may show a reduced form (see the second column of Table 6). Moreover, when used in the genitive, the possessive form of the pronoun is required, and the oblique form followed by the genitive postpositions *-kā/-ke/-kī* is not allowed. The possessive form is inflected according to the categories of gender, number and case, and it follows the paradigm of the adjectives ending in *-ā*: for example: *merā* (my.M.SG.DIR), *mere* (my.M.SG.OBL, M.PL.DIR/OBL), *merī* (my.F.SG.DIR/OBL, F.PL.DIR/OBL). The ergative form of the pronouns follows an exceptional paradigm and does not always require an oblique form, as shown in the fourth column of Table 6. The 3rd person pronouns are identical to the demonstratives: *vah* "he/she" and also "that"; *ve* "they" and also "those".

[1] Following Croft (2003: 130) the animacy and individuation hierarchy is as follows: 1st/2nd person pronouns, 3rd person pronouns > proper name > human common nouns > non-human animate common nouns > inanimate common nouns.

Table 6: Hindi pronominal paradigm.

	Direct	Oblique	Dative (reduced form)	Ergative	Possessive (M.SG)
1SG	*maiṁ*	*mujh*	*mujh=[2]ko (mujhe)*	*maiṁ=ne*	*merā*
2SG	*tu*	*tujh*	*tujh=ko (tujhe)*	*tu=ne*	*terā*
3SG	*vah*	*us*	*us=ko (use)*	*us=ne*	*uskā*
1PL	*ham*	*ham*	*ham=ko (hameṁ)*	*ham=ne*	*hamārā*
2PL	*tum*	*tum*	*tum=ko (tumheṁ)*	*tum=ne*	*tumhārā*
3PL	*ve*	*un*	*un=ko (unheṁ)*	*unhoṁ=ne*	*unkā*
HON	*āp*	*āp*	*āp=ko (āpko)*	*āp =ne*	*āpkā*

Lastly, note that the plural form of the pronouns in Hindi may also be used to refer to singular referents, especially when the speaker wants to show politeness and convey respect towards the person referred to by the pronominal form. The plural form of the pronoun also implies that the verb shows plural agreement. Plural agreement may also occur when the person you want to refer to with respect is expressed by a singular noun or a proper name.

3.2.2. Hindi postpositions

As mentioned in section 3.2.1, nominal inflection in Hindi only shows a binary opposition between a direct form and an oblique form: the direct form is used when the noun has the syntactic function of subjects or direct objects, while the oblique form is used when the noun has any other syntactic function. In order to be further specified, oblique syntactic functions and semantic functions require the noun to be followed by a postposition. The only two oblique functions that do not require a postposition are the allative function and the adverbial temporal function, which are expressed by the mere oblique form of the noun: see the allative function of *mere ghar* "to my house" in sentence 34 and the temporal adverbial function of *us din* "that day" in sentence 35.

34. आप कल मेरे घर आइएगा।

āp	*kal*	*mer-e*	*ghar*
2SG.HON	tomorrow	1SG.GEN-M.SG.OBL	house(M.SG.OBL)

[2] I follow Butt and King (2004) and consider Hindi simple postpositions as clitics, therefor I will gloss them with the symbol = according to the Leipzig Glossing Rules (see on this also section 3.2.2.1).

ā-ie-g-ā

come-2SG.HON-FUT-M.SG

"You will come to my house tomorrow."

35. उस दिन भी मैं तुमसे बहुत-बहुत बातें करना चाहती थी।

us	*din*	*bhī*	*maiṁ*	*tum=se*	*bahut-bahut*
that.OBL	day(M.SG.OBL)	also	1SG.NOM	2PL=COM	many-many
bāt-eṁ	*kar-nā*	*cāh-t-ī*			*th-ī*
speech(F)-PL.NOM	do-INF	want-IPRF-F.SG			be.PST-F.SG

"That day as well, I wanted to talk to you a lot."

Hindi postpositions are divided in two groups: simple postpositions and compound postpositions. Simple postpositions are polysemic and can express both argumental functions and non-argumental functions, while compound postpositions typically express adjunct functions and have more specific meaning.[3] Case marking in Hindi is thus realized through an inflecting system operating on the nouns which distinguish only direct syntactic functions from oblique ones, and the addition of postpositions is used to further specify the oblique functions.

This Hindi case system has been accounted for in different ways, but all scholars agree in identifying a threefold system. Masica (1991), for example, distinguishes three Layers of case-marking: Layer I is the level of the binary inflecting system opposing a direct and an oblique case; Layer II is what I called simple postpositions, while Layer III covers compound postpositions consisting of the genitive *ke* (or *kī*) and a nominal part. Mohanan (1994) proposes a similar classification distinguishing three levels in Hindi case-marking system: 1. Stem forms: oblique and direct form of the noun; 2. clitics: simple postpositions and 3. Postpositions: complex postpositions with a nominal part. A similar distinction is assumed by Butt and King (2004) who, elaborating on what had previously been pointed out by Mohanan (1994), make a case for considering the case-markers of intermediate level as clitics and not as affixes. They note that evidence from coordination, stress and an intervening focus particle (*hī* 'only') support the hypothesis of the clitic status of the intermediate level.

In this work I will follow a similar classification of Hindi case-marking system and I will distinguish between inflection, postpositions and complex postpositions. The next sections give an

[3] With the exception of *-ke pās* which is used to encode the Possessor argument in predicative possessive constructions.

exposition of the use of simple postpositions (section 3.2.2.1) and complex postpositions (section 3.2.2.2) in the language.

3.2.2.1. Simple postpositions

Table 7 summarizes Hindi simple postpositions and their use. Only the ergative postposition *ne* is not polysemic in Hindi, while all other postpositions can express more than one semantic function (both argumental and non-argumental).

Table 7: Hindi cases and their functions.

Postposition	Case	Function
Ø	Nominative	Subject in imperfective transitives, direct object
ne	Ergative	Subject in perfective transitives
ko	Accusative	Direct object (with animate/definite nouns DOM)
ko	Dative	Recipient, Beneficiary, Experiencer
se	Ablative	Comitative, Instrument, Delative
kā/ke/kī	Genitive	Possessor, Relatum
meṁ	Inessive	Spatial location, temporal location
par	Superessive	Spatial location, temporal location

The postposition ***ne*** is used to mark agent subjects in perfective transitive sentences, an example is given in 36. For a discussion on Hindi ergative alignment, see section 3.2.4.

36. तुम ने मेरा अभिमान तोड़ दिया।

tum=ne	*mer- ā*	*abhimān*	*toṛ*	*diy-ā*
2PL=ERG	my-M.SG.NOM	pride(M.SG.NOM)	break	give.PRF-M.SG

"You hurt my pride."

There is no agreement regarding the origin of the postposition *ne*. Scholars trace the origin of Hindi ergativity to the past passive participle construction in *-tà* of Classical Sanskrit. For this reason, the first suggestion regarding the source of the postposition *ne* was the instrumental ending of the *-a* stems in Sanskrit: i.e. *-ena,* occurring in the *tà* participle construction that later developed into the ergative construction of New Indo-Aryan languages (Trumpp 1872). However, scholars soon realized that, although this explanation was supported by syntactic material, it was not likely from a phonological point of view. Moreover, it is typologically quite uncommon that an inflectional ending

develops into a postposition (Butt 2006b, Reinöhl 2016). Other proposals trace the origin of *-ne* back to lexical sources or to borrowed forms. Reinöhl (2016), for example, makes a case for a borrowed source: she is not sure about what neighboring language or dialect the postposition might come from, but she points out that both old Rajasthani and Haryani shows *-ne* (or similar forms) for the encoding of both dative and ergative. Moreover, Drocco (personal communication) notes that the same use of the postposition *ne* is also present in some contemporary Eastern Rajasthani dialects. Butt (2006b: 80-81), instead, proposes that it could be derived from the locative (*janiyee*) of the Skt. *janiyā* meaning "for the sake of, because of". Butt also tries to give a different explanation for the origin of the postposition, connecting it to an original oblique/dative case rather than to an instrumental: she points out that many modern Indo-Aryan languages closely related to Hindi exhibit strong correlations between dative and ergative case markings; see for example: Panjabi dat. *nũ* vs erg. *ne*, Gujarati dat. *ne*/*nẽ* vs erg. *-e* (< old *nē*), Nepali dat. *lāī* vs erg. *le*. In particular, Butt assumes that the ergative alignment at first required an oblique subject, that was formally identical to the dative, and she proposes that the modern ergative *ne* was introduced to encode the non-nominative subject when occurring in an ergative construction, in order to distinguish it from the dative (a suggestion also made by Verbeke and De Cuypere 2009 and firstly proposed by Beames 1970 [1875]).

Many scholars pointed out that the use of the ergative postposition in Hindi does not only occur with transitive perfective predicates and that the ergative marking is sensitive to both syntactic and semantic features (Butt and King 2002, De Hoop and Narashiman 2005). Indeed, even if ergativity in Hindi is triggered by aspectual properties, the postposition *ne* may also be used with non-transitive verbs. Intransitive sentences using an ergative marking of the most salient participant may occur in Hindi and in these cases the ergative clearly expresses a high level of agentivity. For instance, the single argument of many "body emission" verbs can optionally be marked with the ergative case. When this happens, the ergative case-marking encodes a more Agent-like argument: volitional and in control of the event, as the contrasting examples 37 and 38 show (see on this Butt and King 1991, Mohanan 1994, de Hoop and Narasimhan 2005, Butt 2006b).[4]

37. *laṛkī* *chīllā-ī*
girl(F.SG.NOM) scream-PRF.F.SG
"The girl screamed."

[4] This is a classical instance of the unergative verb class first identified by Perlmutter in his seminal paper on the Unaccusative Hypothesis (Perlmutter 1978).

38. *laṛkī=ne* *chīllā-yā*
girl(F.SG.OBL)=ERG scream-PRF.M.SG
"The girl screamed (purposefully)."

The postposition ***ko*** is used to mark both the accusative and the dative, and to express a few other spatial and temporal functions (on the functional distinction between accusative and dative see section 3.2.4). Scholars (Butt and Ahmed 2010, Reinöhl 2016) agree to trace the origin of this postposition back to the locative form (*kakṣe*) of the Sanskrit noun *kakṣa* "side of the body, armpit", a hypothesis first proposed at the end of the 19th century (Beames 1970 [1875], Kellogg 1972 [1875]). This means that the modern dative postposition has its origin in a spatial marker derived from a body-part noun. Body-part nouns are quite common sources for the origin of spatial adpositions: as we will see in the following part of this section, other postpositions in Hindi developed through similar paths. As Ahmed (2006, 2009) and Butt and Ahmed (2011) notice, non-spatial uses originated from the extended meaning based on metaphorical or metonymic processes: a Recipient is metaphorically reinterpreted as the place towards which the event is directed. Typologically, this is a quite common metaphor (Haspelmath 2003, Luraghi 2014). Examples of the usages of the postposition *ko* are given below in sentences 39-42. For a thorough repertory of the usages of the dative postposition see among others Ahmed (2006).

39. Temporal adverbial
हम लोग रात को ज़मीन पर सोते हैं।

ham	*log*	*rāt=ko*	*zamīn=par*
1PL.NOM	people(M.PL.NOM)	night(F.SG.OBL)=at	floor(F.SG.OBL)=on
so-t-e	*haiṁ*		
sleep-IPFR-M.PL	be.3PL.PRS		

"At night we sleep on the floor."

40. Dative
उसने सिपाही को पैसे दे दिए।

us=ne	*sipāhī=ko*	*pais-e*	*de*	*diy-e*
3SG=ERG	soldier(M.SG.OBL)=DAT	money(M)-SG.NOM	give	give.PRF-M.PL

"He gave the money to the soldier."

41. Allative, Spatial goal

sāmān	*ghar=ko*	*pahūṁc*	*gay-ā*
luggage(M.SG.NOM)	home(M.SG.OBL)=at	reach	go.PRF-M.SG

"The luggage reached home." (taken from Ahmed 2006: 3, my glosses)

42. Accusative (DOM)

मैं इत्मीनान से बैठकर उस किताब को पढ़ने लगा।

maiṁ	*itmīmān=se*	*baiṭh=kar*	*us*	*kitāb=ko*
1SG.NOM	leisure(M.SG.OBL)=with	sit=CP	that.OBL	book(F.SG.OBL)=ACC
paṛh-n-e	*lag-ā*			
read-INF-OBL	attach-PRF.M.SG			

"I sat down leisurely and started reading that book."

The postposition ***kā* (/*ke*/*kī*)** is the genitive postposition and functions mainly as a N-N relational marker expressing several meanings: whole-part relationship, kinship, possessive relationship, and in general some kind of relationship between two entities. The genitive postposition is attached to the oblique form of a noun to form a PP and it agrees in gender, number and case with the head-noun. The noun of the PP and the postposition thus form an adjectival unit inflecting as an adjective in *-ā*. The form *kā* is the masculine singular form, the masculine plural is *ke*; the feminine form is *kī* and it is the same for both the singular and the plural (see examples 43a-c). Previous literature acknowledge that the inflection of the genitive postposition is due to its participial origin, however scholars do not agree on which participial form it comes from. Some linguists (Montaut 2004, Oberlies 1998) suggested that it derives from the past participle of the Sanskrit root *kṛ-* "do": i.e. *kṛta* (*kṛta* > *kṛitya* > *kiya* > *kā*); while others (Bubenik 1998) propose that it comes from the Sanskrit future participle in *-ya-*: *kārya*. Reinöhl (2016: 63) notices, following previous proposals (Oberlies 1998, Kellogg 1972 [1875]), that the attested variants *kṛtaka-* or *kāryaka-* are more likely from the point of view of the semantic evolutions, since *-ka-* is a suffix deriving adjectives.

As mentioned in section 3.2.1 above, when the genitive occurs with a personal pronoun, the possessive form of the pronoun is required, and not its oblique form followed by the postposition, examples are given in (43d) and (43e).

43.

a.

bacc-e	*k-ā*	*ḍibb-ā*
child(M)-SG.OBL	GEN-M.SG.NOM	box(M)-SG.NOM

"The child's box."

b. *bacc-e* *k-e* *ḍibb-e*
child(M)-SG.OBL GEN-M.PL.NOM box(M)-PL.NOM
"The child's boxes."

c. *bacc-e* *k-ī* *kitāb*
child(M)-SG.OBL GEN-F book(F.SG.NOM/OBL)
"The child's book."

d. *mer-ā* *ḍibb-ā*
1SG.GEN-M.SG.DIR box(M)-SG.NOM
"My box."

e. *mer-ī* *kitāb*
1SG.GEN-F book(F.SG.NOM/OBL)
"My book."

The postposition ***meṁ*** is used to mark the inessive locative, as shown in example 44. It may also be used with some temporal meaning, in particular in relation to months (*māgh meṁ* "in the month of Māgh"), seasons (*garmiyoṁ meṁ* "in summer") and years (*1947 meṁ* "in 1947").

44. मगर वह अभी बम्बई में है।

magar	*vah*	*ab=hī*	*bambaī=meṁ*	*hai*
but	3SG.NOM	now=EMPH	Bombay=in	be.PRS.3SG

"But he is in Bombay now."

According to previous works (Reinöhl 2016: 54), this postposition derives from the Sanskrit form *madhye*, the locative singular of *madhya-* "placed in the middle, central". The noun *madhya* was also used with the meaning "waist": the development of the postposition *meṁ* thus resembles that of other Hindi postpositions from a noun indicating a part of the body to spatial adposition (as mentioned above the same happened for the postposition *ko* < Skt. *kakṣa* "side of the body, armpit"[5]). There are some doubts on the origin of the nasalization of the vowel in *meṁ*, unlikely derived from the form *madhye*. For this reason, some scholars (Kellogg 1972 [1875]: 132) proposed to trace the origin of the postposition *meṁ* back to the accusative form of the noun *madhyam*, however Reinöhl (2016) excludes this possibility and argues for the form *madhye*, which is semantically more plausible: she points out that the spatial meaning of Hindi *meṁ* resembles that of a Sanskrit locative (inessive) rather

[5] The same grammaticalization path gave origin to the compound postposition *ke pās*, derived from the Skt. noun *parśva.*

than a Sanskrit accusative (allative), and she also notices that forms of spontaneous nasalizations are particularly frequent in Hindi[6].

The postposition ***se*** expresses a wide range of functions, and it is variably glossed. Its main usages are the comitative (sentence 45), the instrumental (sentence 46) and the delative (sentence 47). It is also used to express concrete or abstract sources, and with emotion verbs it may express the Stimulus that triggers an emotion (see example 48). It also forms manner adverbials when attached to nouns (for example, *jaldī se* "quickly" *dhīre se* "slowly").

45. Comitative use

उसने पति से यह सारी कथा सुनायी।

us=ne	*pati=se*	*yah*	*sār-ī*	*kathā*	*sunā-yī*
3SG=ERG	husband(M.SG.OBL)=COM	this	whole-F	story(F.SG.NOM)	tell-PRF.F.SG

"She told the whole story to her husband."

46. Instrumental use

उसने विशवास-भरी आँखों से देखा।

us=ne	*viśvās-bhar-ī*	*āṁkh-oṁ=se*	*dekh-ā*
3SG.OBL=ERG	faith-full-F	eyes(F)-PL.OBL=INS	see-PRF.M.SG

"He looked at her with eyes full of faith."

47. Delative use

मैं उसके साथ घर से गया।

maiṁ	*us=ke sāth*	*ghar=se*	*gay-ā*
1SG.NOM	3SG.OBL=with	house(M.SG.OBL)=from	go.PRF-M.SG

"I left the house with him/her."

48. Source

मैं दुश्मनों से नहीं डरती।

maiṁ	*duśman-oṁ=se*	*nahīṁ*	*ḍar-t-ī*
1SG.NOM	enemy(M)-PL.OBL=from	not	fear-IPRF-F.SG

"I am not afraid of enemies."

[6] An example is the noun *sāṁp* "serpent", derived from the Sanskrit noun *sarpa* (Reinöhl 2016: 55).

The polysemy of the ablative marker is typologically a common feature and the closeness of concepts such as Source, Instrument, Companion and Manner has been thoroughly investigated in previous studies (Heine et al 1991, Luraghi 2001, Stolz 2001, Haspelmath 2003). The syncretism of instrumental and comitative for instance can be explained by the "companion metaphor" firstly identified by Lakoff and Johnson (1980) according to which an INSTRUMENT is a COMPANION (on the syncretism of Instrument, Comitative and other semantic roles see Luraghi 2001).

There is no agreement on the origin of the postposition *se*. Reinöhl (2016: 57) discusses a list of previous suggestions on possible Sanskrit forms: Kellogg (1972[1875]) suggests to trace the origin of *-se* back to the locative form *saṅge* of *saṅga* "contact, relation", while Oberlies (1998) proposes the adverb *samaṁ* "together" as a possible source. Turner (1971), in contrast, makes a case for the Old Indo-Aryan *sahita-* "accompanied". All suggestions fit the semantic range of *-se* in Hindi and they are plausible semantic sources for this postposition.

The postposition *se* may also be used to encode non-prototypical agents that act in an event but are involuntary or inefficient (Montaut 2004a, 2004b, see also Masica 1996, who speaks of incapacity of the Agent). This use probably derives from a metaphoric extension according to which an involuntary agent is conceptualized as the Source of an event. Consider example 49, taken from Montaut (2004a: 211): in the first sentence (a) the Agent is marked with the ergative, the speaker here believes that the interlocutor purposefully killed a person; in the second sentence (b), the Agent is marked with the instrumental: the speaker is saying that the event was an accident and that he killed the person unconsciously.

49. Instrumental Agent vs Ergative Agent:

a: *tum-hīṁ=ne* *us-kā* *khūn* *kiy-ā*
2SG-EMPH=ERG 3SG.OBL-GEN blood(M.SG.NOM) do.PRF-M.SG
a: "It's you who murdered him."

b: *sāhab* *maiṁ=ne* *us-kā* *khūn* *nahīṁ* *kiy-ā,* *mujh=se*
sir 1SG=ERG 3SG-GEN blood(M.SG.NOM) not do.PRF-M.SG 1SG=INS
ho *gay-ā*
be go.PRF-M.SG
b: "Sir, I did not kill him, it happened by myself (I did it unconsciously)."

Another use of the postposition *se* is to mark inefficient agents: this construction has been labelled passive of disability and it is exemplified in sentence 50 (Glassman 1976, Van Olphen 1980, Butt and King 2004).

50. *mujh=se* *darvāz-ā* *nahīṁ* *khul* *gay-ā*
1SG.OBL=INS door(M)-SG.NOM not be_open go.PRF-M.SG
"I couldn't open the door."

In addition to the postpositions listed in Table 7 and exposed above, Hindi exhibits two other simple postpositions: *par* "on" and *tak* "till". The postposition *tak* "till" expresses spatial or temporal limits, such as in *kal tak* "until tomorrow". The postposition *par* "on" can encode spatial relations such as *mez par* "on the table" or *ghar par* "at home". Moreover, *par* may also be used to express the Stimulus towards which an emotion is directed as exemplified in example 51.

51. उसे गोबर पर गुस्सा आता।
use *gobar=par* *guss-ā* *ā-t-ā*
3SG.DAT gobar=on anger(M)-SG.NOM come-IPRF-M.SG
"He gets angry at Gobar."

3.2.2.2. Complex postpositions

Complex postpositions are composed by the genitive postposition followed by an adverb or by a noun in the oblique form, see the examples in the Table 8. The genitive postposition may be in the masculine oblique form or in the feminine form depending on the element that follows it (recall from section 3.2.2.1 that the genitive postposition agrees with the element on its right). If the element following the genitive is a feminine noun the oblique feminine form of the genitive postposition – *kī* is employed, while if the element is a masculine noun or an adverb, then we have the oblique masculine form -*ke*. Table 8 shows a list of some of the most common complex postpositions in Hindi.

Table 8: List of some of the most common Hindi complex postpositions.

Postposition	Meaning	Postposition	Meaning
ke pās	near	*kī taraph /kī or*	towards
ke lie	for	*ke bād*	after
ke sāth	with	*ke daurān*	during
ke binā	without	*ke upār*	above
ke andar	inside	*ke nīce*	under
ke bāhar	outside	*ke āge*	ahead
ke sāmne	in front of	*ke pīche*	behind

When complex postpositions follow a noun, they are attached to its oblique form, as in 52, but when they follow a personal pronoun, the possessive form of the pronoun is required, rather than its oblique form followed by the genitival postposition, as in sentence 53. Complex postpositions generally express non-argumental functions in Hindi. This is a quite common typological division of labor among languages of the world for the expression of semantic roles (Kittila et al 2011: 9). Argumental semantic roles, such as Agents or Experiencers are expressed by formally less heavy linguistic units (simple postpositions), while peripheral semantic roles are expressed by the heavier ones (complex postpositions).

52. मेहता के पास सामान तो ज्यादा न था।

mehtā=ke	*pās*	*sāmān*	*to*	*jyādā*	*na*	*th-ā*
mehta=GEN	side	belonging(M.SG.NOM)	the	many	not	be.PST-M.SG

"Mehta didn't have many belongings."

53. लेकिन मेरे पास नगद नहीं है।

lekin	*mere*	*pās*	*nagad*	*nahīṁ*	*hai*
but	1SG.GEN	side	cash(M.SG.NOM)	not	be.3SG.PRS

"But I have no cash."

3.2.3. Verbal morphology

The citation form of the Hindi verb is the infinitive[7], formed by the root followed by the suffix *-nā*: for example: *ho-nā* "be", *khā-nā* "eat", *cal-nā* "walk", etc. The root of the verb is the base from which all other verb forms are derived. Most verb forms show analytic constructions consisting of a participle (imperfective or perfective) and the auxiliary verb *honā* "be. The participle agrees with the subject in gender and number, whereas the auxiliary agrees in number and person. The participle contributes aspectual information, while the auxiliary adds information on tense and mood. Verb forms that are not analytic (and do not present a participle) may not show gender agreement, but always agree in person and number with the verb (see section below for an exposition of finite verb forms in Hindi).

As Butt and King (2004) point out, agreement does not uniquely identify a grammatical function in Urdu/Hindi (see also Mahajan 1991, Mohanan 1994, Butt 1995). Agreement is according both to grammatical functions and case marking: the verb agrees with the nominative argument,

[7] Note that many scholars do not use the infinitive form but the root as citation form.

irrespective of its role. If there are two nominative arguments the verb agrees with the subject, i.e. the argument which in the perfective assumes the ergative case (as in example 54). If the subject is not in the nominative, and the second argument is in the nominative, the verb agrees with second argument (as in 55). If there are no nominative arguments, the verb shows the "default" masculine singular form (56).

54. बाबू के सिवा वह और किसी को न जानती थी।

bābū=ke sivā	*vah*	*aur*	*kisī=ko*	*na*
babu=except	3SG.NOM	other	someone.OBL=ACC	not
jān-t-ī	*th-ī*			
know-IPRF.F	be.PST-F			

"She did not know anyone except Babu."

55. उन्होंने चुपचाप बैठकर खाना खाया।

unhoṁ=ne	*cupcāp*	*baiṭh=kar*	*khan-ā*	*khā-yā*
3PL.OBL=ERG	silently	sit=CP	food(M)-SG.NOM	eat-PRF.M.SG

"They ate sitting in silence."

56. मैंने सोचा कि तुम बैठी-बैठी ऊब गयी।

maiṁ=ne	*soc-ā*	*ki*	*tum*	*baiṭh-ī-baiṭh-ī*
1SG=ERG	think-PRF.M.SG	that	2SG.NOM	sit-PRF.F.PL- sit-PRF.F.PL
ūb	*gay-ī*			
bored	go.PRF-F			

"I thought you got tired sitting there."

3.2.3.1. Non-finite verb forms

Non-finite verb forms are very frequent in South Asian languages: as many scholars pointed out (Masica 1976, Subbarao and Arora 2009, Hock 2016b) their extensive use is one of the distinguishing pan-Indian features. Hindi has a strong tendency to use non-finite subordination devices, and in this respect, it seems to conform to the idealized picture of "strict OV" (Hock 2016b: 576). Hindi typically uses a special non-finite verb form, the conjunctive participle, which can be realized in two ways:

a. with the root followed by the suffix *-kar* (this suffix also has an allomorph *-ke*, which is obligatory only with the root *kar-* "do", for euphony reasons).

Ex. *dekhnā* “see” > *dekhkar*, *socnā* “think”> *sockar*; but *karnā* “do” > *karke*, not **karkar*.

b. with the bare root (this use depends on stylistic choices, and it is less frequent in spoken Hindi).
Ex. *dekhnā* “see” > *dekh*, *socnā* “think” > *soc*; *karnā* “do” > *kar*.

The conjunctive participle in *-kar* can be used with an adverbial modifying function expressing contemporaneity – thus resembling gerunds of SAE languages (see Haspelmath and König 1995). However, the main function of this construction is to express sequential actions (Subbarao and Arora 2009, Masica 1991, Hock 2016a, Montaut 2004). The conjunctive participle has traditionally been interpreted as expressing anteriority in relation to the verb in the main clause and thus fulfilling a narrative-chaining function. For this reason, the term used in previous Indian grammatical tradition is *pūrvkālik kr̥dant* “anterior participle” (Montaut 2004: 93). Some examples of the double interpretation of this verb form are given below. Sentence 57 can be easily interpreted as expressing both anteriority and simultaneity. Sentence 58 is a clear example of anteriority interpretation, while sentence 59 clearly expresses simultaneity.

57. यह सोचकर उसने बोला ठीक है।

yah	*soc=kar*	*us=ne*	*bol-ā*	*thīk*	*hai.*
this	think=CP	3SG=ERG	say-PRF.M.SG	fine	be.PRS.3SG

“Thinking that he said: “Ok”.
“Having thought about that he said: ‘Ok’”

58. गांव में पहुंचकर वह पहले एक किसान के घर गया।

gāṁv=meṁ	*pahuṁc=kar*	*vah*	*pahle*	*ek*	*kisān=ke*
village(M.SG.OBL)=in	arrive=CP	3SG.NOM	firstly	one	farmer(M.SG.OBL)=GEN
ghar	*ga-yā*				
house(M.SG.OBL)	go.PRF-M.SG				

“After reaching the village he first went to a farmer's house.”

59. वह दौड़कर घर आया।

vah	*dauṛ=kar*	*ghar*	*ā-yā*
3SG.NOM	run=CP	home(M.SG.OBL)	come-PRF.M.SG

“He came home running.”

According to Davison (1979, 1981), Hindi conjunctive participles have three distinct main functions: e.g., coordinating, subordinating and adverbial function. Conjunctive participles with coordinating function encode a coordinative relation between the clauses and can often be paraphrased with coordinative conjunctions such as "and then" or with a clause introduced by "when" (as in sentence 58). Conjunctive participles with subordinative function express a subordinate relation with the matrix clause, mainly encoding a causal meaning, and still implying anteriority of the action expressed by the conjunctive participle. See the example below (60). Conjunctive participles with adverbial subordination typically express simultaneity and manner (as in 59 above).

60. पत्रों में लेख लिखकर मेरा निर्वाह नहीं हो सकता।

patr-oṁ=meṁ	*lekh*	*likh=kar*
Newspaper(M)-PL.OBL=in	article(M.PL.NOM)	write=CP

merā nirvāh nahīṁ ho saktā.

I cannot make a living

"I cannot make a living by writing articles for newspapers."

This multifunctionality of the Hindi narrative conjunctive participle in *-kar* aligns with typological observations. Nedjalkov (1995), for example, notices that narrative converbs[8] (as defined in Haspelmath and König 1995) may also be used contextual functions. Additionally, Bickel (1998) points out that languages spoken in Asia show a tendency to use the same non-finite form to express both narrative and adverbial modifying functions.

Hindi displays also two other non-finite constructions with a subordinating function that are formed with the imperfective and perfective participle respectively. The imperfective participle consists of the root of the verb, followed by the aspectual marker *-t-*. The perfective participle is formed by merely attaching the endings for gender and number agreement to the root of the verb,

[8] As many scholars have pointed out (Bickel 1998; Ylikoski 2003, Coupe 2009), it is not easy to propose a universal definition of the term converb. The boundaries of the category are quite fuzzy and they blur into other categories such as adverbial participles, medial verbs, absolute constructions, and infinitival constructions. Given the typological and language-specific complexities that characterize the category of converbs, Bickel (1998: 394-395) wonders if it is useful to consider this as a cross-linguistically valid notion. He points out that typological studies suggest that there are at least two types of converbs, and he declares himself skeptical whether to label them under the same term. One type is the European converb, consisting of non-finite verb forms specialized in adverbial modifying subordination. The other type of converb is the Asian converb, which systematically conflates adverbial modification and narrative function.

without any derivational aspectual suffix. Depending on the phonological structure of the root, a *-y-* may be added between the ending and the root for euphonic reasons (see for example *likh-ā* "wrote, M.SG" vs. *khā-y-ā* "eat, M.SG"). When used with adverbial function, Hindi participles can either be in the oblique form (thus ending in *-e*) or show agreement with the subject by the adding of gender and number endings (*-ā* for singular masculine; *-e* for masculine plural and oblique; *-ī* for feminine). These two forms derive from Sanskrit present participles in *anta* (> *-t-*) and the past participle in *-ita* (> *-Ø-*) respectively (Montaut 2004: 95). See the examples for the verbs *dekhnā* "see" and *socnā* "think" below.

a. Imperfective participle:

Ex. *dekhnā* "see" > M.SG: *dekhtā* (*huā*), M.OB/M.PL: *dekhte* (*hue*); F: *dekhtī* (*huī*)
socnā "think" > M.SG: *soctā* (*huā*), M.OB/M.PL: *socte* (*hue*), F: *soctī* (*huī*).

b. Perfective participle

Ex. *dekhnā* "see" > M.SG: *dekhā* (*huā*), M.OB/M.PL: *dekhe* (*hue*), F: *dekhī* (*huī*)
socnā "think" > M.SG: *socā* (*huā*), M.OB/M.PL: *soce (hue*), F: *socī* (*huī*).

Both participles may be followed by the past participle of the verb *honā* "be" (*huā* / *hue* / *huī*), without major changes in the meaning of the participial construction. Most of the time this addition intensifies the continuity (imperfective aspect) of the event expressed by the imperfect participle and the perfectivity of the event expressed by the perfect participle (Montaut 2004). Inflected participles used with an adverbial function may have various circumstantial functions. However, the imperfective adverbial participle generally indicates that the action expressed by the participle and that expressed by the main verb are simultaneous (as in 61), while the perfective adverbial participle generally indicates anteriority of the action (as in 62 and 63).

61. उन्होंने हंसते हुए कहा दो बार में ग्राम प्रधान का चुनाव हार चुका था।

unhoṁ=ne	*haṁs-t-e*	*hu-e*	*kah-ā:*
3PL.OBL=ERG	laugh-IPRF-M.SG.OBL	be.PRF-M.SG.OBL	say-PRF.M.SG

'*do bār maiṁ grām pradhān kā cunāv hār cukā thā'.*
"While laughing, he said, 'Twice I had lost the election for village head'."

62. वह लाठी लिए मेरे पास आ गया।

vah	*lāṭhī*	*li-e*	*mere*	*pās*	*ā*
3SG.NOM	stick(F.SG.NOM)	take.PRF-M.SG.OBL	1SG.GEN	side	come
ga-yā.					
go.PRF-M.SG					

"After taking a stick he came next to me."

63. वह मुँह नीचे किए हुए लेटा है।

vah	*muṁh*	*nīce*	*ki-ye*	*hu-e*
3SG.NOM	face(M.SG.NOM)	down	do.PRF-M.OBL	be.PRF-M.OBL
leṭ-ā	*hai*			
lie-PRF.M.SG	be.3SG.PRS			

"He is lying with his face bent down." Lit. "He is lying after having bent his face down."

3.2.3.2. Finite verb forms

Hindi verb forms can be grouped into two main classes: verb forms that are realized immediately from the root of the verb by adding suffixes (see for example the simple future in sentence 66) and verb forms that are realized periphrastically by means of auxiliaries combined with the imperfective (example 64) or perfective (example 65) participle of the main verb (Kellogg 1972 [1875]: 224; Van Olphen 1975). Non-periphrastic verb forms do not show overt markers of the aspectual characterization: these forms in Hindi are the future, the subjunctive and the imperative.

The tense category is characterized by a three-way distinction: present (denoting habitual actions or processes, and natural processes), past, and future (Kachru 2006, Butt and Rizvi 2008). Past and present are periphrastic and are formed by the participle followed by the past and present form respectively of the auxiliary *honā* "be". The future is not periphrastic, and it is formed via suffixation. The future suffix is – *gā* (/-*ge* /-*gī* /-*gīṁ*) and changes according to gender and number agreement: -*gā* is for the masculine singular, -*ge* for the masculine plural, -*gī* for the feminine singular and -*gīṁ* for the feminine plural. This suffix is attached to the subjunctive form of the verb, which is marked according to number and person. Hence, the future marks the number twice. This situation, which is typologically uncommon, is easily explained when considering that the suffix developed from a participle of the Sanskrit verb *gā* "go" (Butt and Rizvi 2008) through a quite common reanalysis of verbs expressing "go" as future auxiliaries (Bybee, Perkins and Pagliuca 1994).

64. मेरी मां की तो मेरे पास हर आठवें दिन चिट्ठी आती है।

merī	*māṁ=kī*	*to*	*mere*	*pās*	*har*	*āṭhaveṁ*
1SG.GEN-F	mother(M.SG.OBL)=GEN	EMPH	1SG.GEN	side	every	eight

din	*ciṭṭh-ī*	*ā-t-ī*	*hai*
day	letter(F)-SG.NOM	come-IPRF-F	be.3SG.PRS

"I get a letter from my mother every eighth day."

65. वह सन्त पालम के पास आया और अपना मस्तक उनकी छाती पर रख दिया।

vah	*sant*	*pālam=ke*	*pās*	*ā-yā*
3SG.NOM	Saint	Palam=GEN	side	come-PRF.M.SG

aur apnā mastak un=kī chātī par rakh diyā.

and placed his head on his chest.

"He came to Saint Palam and placed his head on his chest."

66. मैं कल सबेरे रूपये लेकर आऊंगी।

maiṁ	*kal*	*sabere*	*rūpy-e*	*le=kar*
1SG.NOM	tomorrow	morning	rupee(M)-PL.NOM	take=CP

ā-ūṁ-g-ī .

come-1SG-FUT-F.SG

"I will bring the money tomorrow morning."

Hindi clearly distinguishes three main aspects: imperfect, perfect and progressive. These three aspectual forms have a broad distribution over the Hindi verbal system and combine with any tense and mood form. The imperfect forms consist of the imperfect participle (section 3.2.3.1) combined with the tense auxiliary to form present imperfect and past imperfect. Perfective forms are realized by combining the perfect participle with the tense auxiliary. When the verb is in the perfective past, it may indicate either that the action occurred before some past time or in a past time which does not lead directly to the present.

The progressive aspect is indicated by the auxiliary *rahā* (/ *rahe* / *rahī*) that follows the bare root of the verb and inflects according to number and gender (see the example in 67). The progressive auxiliary *rahā* is the grammaticalized form of the perfective participle of the verb *rahnā* "stay". The evolution from verbs of meaning "standing" or "remaining" to auxiliaries for progressive aspect is typologically well-established (Bybee and Perkins and Pagliuca 1994). Deo (2006) argues that the imperfective aspect in Indo-Aryan originally could express both progressive and non-progressive (habitual) events and that in the 19th century ca. its use was restrained to expressing only non-

progressive meanings, indicating that the recruiting of *rahnā* as a progressive marker is thus a recent phenomenon.

67. खिडक़ी से बहुत ठंडी हवा आ रही थी।

khiṛkī=se	*bahut*	*ṭhaṁḍ-ī*	*havā*	*ā*	*rah-ī*
window(F.SG.OBL)=from	very	cold-F.SG	wind(F.SG.NOM)	come	PRG-F
th-ī.					
be.PST-F.SG					

"A very cold breeze was coming from the window."

The perfective aspect refers to completed events. Both the habitual and progressive aspects refer to uncompleted events. Following Comrie (1976) and Van Olphen (1975) Hindi aspectual forms may therefore be listed in the following manner:

1. imperfective
    a. habitual
    b. progressive
2. perfective

Besides these three main aspects, other forms with a more limited distribution are present in Hindi, such as inceptive, continuative, durative, frequentative, and completive. These forms are realized by different auxiliaries. For example, the inceptive is formed by the oblique infinitive form of the verb followed by the participial forms of the auxiliary *lagnā* "be attached, adhere" and tense auxiliaries (example 68), while the completive is realized by the auxiliary *cuknā* "pick up" combined with the bare root of the main verb and followed by tense auxiliaries. The frequentative is realized by adding to the perfect participle of the main verb the imperfect participle of the auxiliary *karnā* followed by the tense auxiliary. The iterative is realized with the same auxiliary used for the progressive (*rahnā* "stay") combined with the imperfective participle of the main verb.

68. मैं बहुत जल्दी-जल्दी सोचने लगा।

maiṁ	*bahut*	*jaldī-jaldī*	*soc-n-e*	*lag-ā.*
1SG.NOM	very	quicly-quicly	think-INF-OBL	start-PRF.M.SG

"I started thinking very quickly."

Hindi shows six mood distinctions: imperative, subjunctive, indicative, presumptive, contingent, and past contingent. I will not deal here with all the modal forms found in Hindi, but I will focus on the

imperative and the subjunctive, which are the most frequent. In Hindi there are many imperative forms, none of these forms participates to the aspectual distinctions. The imperative forms found in Hindi are given below.

- Second person singular imperative: formed by the bare root (*dekh* "look")
- Second person plural imperative: formed by the root followed by the ending -o (*dekho* "look")
- Honorific imperative: formed by the root followed by the ending -*iye* (*dekhiye* "look")
- Future imperative: formally identical to the infinitive form of the verb (*dekhnā* "look")
- Honorific future imperative: formed by the root followed by the ending -*iye* -*iyegā* (*dekhiyegā*).

In Hindi there are two types of subjunctives: aspectual subjunctive and non-aspectual subjunctive. Non-aspectual subjunctive is realized via inflection of the verb according to person and number (see example 69). Aspectual subjunctive is realized by imperfective or perfective participles followed by the auxiliary *honā* "be" in the subjunctive form.

69. बस भले आदमी का संग चाहती हुँ, जो मुझे अपना समझे और जिसे मैं भी अपना समझूं।

bas bhale ādmī=kā saṁg cāhtī hūṁ,
I just want the company of a good man,

jo	*mujhe*	*apnā*	*samjhe*	*aur*	*jise*
REL.NOM	1SG.ACC	REFL	consider-SBJV.3SG	and	REL.ACC
maiṁ	*bhī*	*apnā*	*samjhūṁ.*		
1SG.NOM	too	REFL	consider-SBJV.1SG		

"I just want the company of a good man, who considers me his own and whom I also consider mine".

3.2.4. Differential case marking in Hindi

The case marker of a noun in Hindi may depend both on syntactic properties and semantic properties. As mentioned in section 3.2, Hindi is a split-ergative language and also displays Differential Object Marking. Hence, case marking of an argument may be determined by the verb, by the semantic role of the argument (Differential Subject Marking), by the participant properties of the argument (Differential Object Marking) and by aspectual properties (ergativity).

Butt and King (1991, 2002, 2004) propose a thorough account of Differential Case Marking (DCM) in Hindi. They argue that the traditional opposition between structural and lexical cases is insufficient to account for the intricate relationship between morphology, semantics and syntax in

Hindi. They propose a Differential Case Marking Theory (DCT) that examines the entire case system of a language, not defining it based solely on structural cases (i.e. nominative, accusative, ergative), but by also taking into account the role of datives, genitives, instrumentals, etc. (Butt 2006b: 71). It is well known that Differential Case Marking is a pan-Indian feature that characterizes many South Asian languages, not only belonging to the Indo-Aryan branch (see Verma 1976, Bhaskararao and Subbarao 2004). Butt and Ahmed (2008: 10), for example, note that "semantically motivated case alternations can be identified in modern South Asian languages, irrespective of the genetic type of the language, the particular morpho-syntactic realization of the case system and whether the language contains an ergative case or not." They note that in all South Asian languages these case alternations depend on the interaction of many factors, including animacy, specificity/definiteness, tense/aspect, but also refer to phenomena associated with modality, etc.

In the following sections, I will address Hindi split ergativity (section 3.2.4.1), DOM (section 3.2.4.2) and DSM (section 3.2.4.3).

3.2.4.1. Split ergativity

Hindi is a split ergative language (Dixon 1979, 1994): transitive predicates show a nominative-accusative alignment under certain circumstances and an ergative-absolutive alignment under other circumstances. The split is based on the aspect of the verb: perfective transitive predicates require the ergative alignment, in which the subject is in the oblique case and followed by the ergative postposition *ne*. Any other predicate type requires the nominative alignment (Davison, 1999, Drocco 2008, Verbeke 2013). Examples are given in sentence 70, 71 and 72 below:

70. Perfect intransitive predicate

laṛk-ī	*bāzār*	*gay-ī*
girl(F)-SG.NOM	market(M.SG.OBL)	go.PRF-F.SG

"The girl went to the market."

71. Imperfect transitive predicate

laṛk-ī	*bāzār=meṁ*	*cāval*	*karīdh*	*rah-ī*
girl(F)-SG.NOM	market(M.SG.OBL)=in	rice(M.SG.NOM)	buy	PRG-F

hai

be.PRS.3SG

"The girl is buying rice at the market."

72. Perfect transitive predicate

laṛk-ī=ne	*bāzār=meṁ*	*cāval*	*karīdh-ā*
girl(F.SG.OBL)=ERG	market(M.SG.OBL)=in	rice(M.SG.NOM)	buy-PRF.M.SG

"The girl bought rice at the market."

In Hindi, the verb always agrees with the highest unmarked argument in the sentence (Mohanan 1994, Bhatt 2005, Butt 2014, Kulkarni 2011). Therefore, in a nominative-accusative sentence, the verb agrees with the subject and not with the object; while in an ergative-absolutive sentence, it agrees with the unmarked object, since the subject is marked with the ergative. Note that the object may be marked because of differential object marking (see section below 3.2.4.2): in such cases there is no unmarked element in the sentence and the verb stands in the masculine singular default form.

In literature on alignments, the term nominative is generally used to indicate the unmarked element that triggers agreement in a nominative-accusative alignment (the subject), while the term absolutive is used to refer to the unmarked element that triggers agreement in an ergative-absolutive alignment (the direct object). However, as some scholars have pointed out (Butt and King 2002, Mohanan 1994), this distinction is not insightful in Hindi as there is no morphological difference between the subject of the imperfective transitive verb (when the nominative-accusative alignment is required) and the object of the perfective transitive (when the ergative-absolutive alignment is required). They are both unmarked and trigger agreement with the verb (see Mohanan 1994 and Woolford 1997 for a discussion). For this reason, I will follow previous scholars (Mohanan 1994, Butt 1995, Butt and King 2002) and refer both to the subjects of nominative-accusative alignments and to the objects of ergative-absolutive alignments as *nominative*.

3.2.4.2. Differential Object Marking (DOM)

Besides split ergativity, Hindi shows Differential Object Marking (DOM; McGregor 1972, Butt 1993, Masica 1991, Mohanan 1994). Since the 1980s, the term DOM has been used to identify the phenomenon according to which some languages mark their (direct) objects in more than one way, depending on their properties (Bossong 1985, 1991, Comrie 1989, Aissen 2003). In Hindi, when the object is human and/or specific, it is in the oblique case and followed by the postposition *ko* (see on this Butt 1993). Compare sentence 73 and 74, in which the direct object is marked, with examples 71 and 72 above, in which the direct object is in the nominative. Example 74 also shows the default agreement pattern on the verb: the verb is in the perfective aspect, hence the subject is marked with the ergative postposition *ne*; while the object is followed by the postposition *ko*. As no unmarked argument is available to trigger agreement, the verb stands in the default masculine, singular form (*mārā*) .

73. *vah* *aurat* *tumhāre* *cāval=ko* *khā* *rah-ī*
hat woman(F.SG.NOM) 2PL.GEN rice(M.SG.OBL)=ACC eat PRG-F
hai
be.PRS.3SG
"That woman is eating your rice."

74. *us* *aurat=ne* *bacc-ī=ko* *mār-ā*
that.OBL woman(F.SG.OBL)=ERG child(F)-SG.OBL=ACC hit-PRF.M.SG
"That woman hit the girl."

Recall from section 3.2.2.1. that the accusative case and the dative case in Hindi are both marked with the postposition *ko* (as shown in 75 and 76).

75. *ham=**ko*** *pāni* *de* *d-ījie.*
1PL=**DAT** water(M.SG.NOM) give give.IMP.HON
"Please give us some water."

76. *kutt-e=ne* *aurat=**ko*** *kāṭ-ā.*
dog(M)-SG.OBL=ERG woman(F.SG.OBL)=**ACC** bite-PRF.M.SG
"The dog bit the woman."

Despite their formal identity, many scholars insisted on the distinction between dative and accusative from a functional point of view (Mohanan 1994, Butt 1995, Narasimhan 1998). For example, in order to prove their distinction, Mohanan (1994: 91-96) lists the behavioral and functional differences given below.

a) The dative does not alternate with the unmarked case depending on the properties of the participant; the accusative, instead, is allowed only when the participant is [+SPECIFIC] or/and [+ANIMATE].
b) In passivization processes, the dative is preserved, while the accusative is replaced by the nominative case. This shows that the dative is an indirect case, while the accusative is a direct one.[9]

[9] Mohanan (1994) notes that in some dialects of Modern Standard Hindi the accusative *ko* is preserved when passivization occurs. However, she also points out that this fact cannot be held as a proof of the identity between dative and accusative in such dialects, because the accusative still depends on argumental properties, while the dative depends on participant's properties.

c) Unlike the accusative, the dative shows some properties of grammatical subjecthood: for example, it can function as the antecedent of the reflexive pronoun *apnā*, and it can be coreferential with the implicit subject of conjunctive participles in *-kar* (see examples 85-89 below).

Not all scholars agree in considering the two functions of the postposition *ko* as two distinct cases, for example Mahajan (1990) treats all elements marked with the *ko* as datives. Moreover, it is cross-linguistically common that languages with differential object marking use the marker of the dative case to mark direct objects and this indicates a diachronic connection in these languages between the dative and the accusative. Even though this is probably valid also for the Hindi postposition *ko*, this issue is beyond the purpose of this dissertation. As synchronically in modern Hindi these two functions are well distinguished (Butt 1995: 18), in this dissertation I align with the well-established previous literature on Hindi/Urdu that treats the two uses of the postposition *ko* as two different cases.

Since Hindi exhibits both split ergativity and DOM, a transitive construction may show four instantiations in the language, according to the combination of the two parameters of perfectivity and specificity/animacy of the direct object. These four instantiations are shown in examples 77a-d and are summarized in Table 9.

77.

a) *vah* *laṛk-ā* *har roz* *roṭ-iyāṁ* *khā-t-ā* *hai*
that.NOM boy(M)-SG.NOM every day roti(F)-PL.NOM eat- IPRF-M.SG be.3SG.PRS
"That boy eats rotis every day."

b) *us* *laṛk-e=ne* *kal* *roṭ-iyāṁ* *khā-yī*
that.OBL boy(M)-SG.OBL=ERG yesterday roti(F)-PL.NOM eat-PRF.F
"That boy yesterday ate rotis."

c) *vah* *laṛk-ā* *har roz* *mer-ī* *roṭi-yoṁ=ko*
that.NOM boy-M.SG.NOM every day 1SG.GEN-F roti(F)-PL.OBL=ACC
khā-t-ā *hai*
eat-IPRF-M.SG be.3SG.PRS
"That boy eats my rotis every day."

d) *us* *laṛk-e=ne* *kal* *mer-ī* *roṭi-yoṁ=ko*
that.OBL boy(M)-SG.OBL=ERG yesterday 1SG.GEN-F roti(F)-PL.OBL=ACC
khā-yā
eat-PRF.M.SG
"That boy yesterday ate my rotis."

Table 9: Instantiations of the transitive frame in Hindi.

Coding Frame	Ergativity	Marked object
1-nom 2-nom V.subj[1]	×	×
1-erg 2-nom V.subj[2]	✓	×
1-nom 2-acc V.subj[1]	×	✓
1-erg 2-acc V[3SG.M]	✓	✓

Note that DOM and ergativity occur with any transitive verb, i.e. any bivalent verb that can have two nominative arguments. This means that they occur regardless of the semantic roles behind the grammatical relations, hence they do not depend on the semantic properties of the event, but on the aspectual properties of the predicate and/or on the properties of the participants. For this reason, in the present study such alternations are of secondary importance, and I will disregard this type of variation in my analysis.

3.2.4.3. Differential Subject Marking

In Hindi many predicates require their subject to be encoded with an oblique case. In particular, the dative case is prototypically associated with the Experiencer subject and the genitive case is prototypically used to mark the Possessor subject. However, what is remarkable in Hindi, is that, as scholars know well, all the cases listed in Table 5 can mark the subject in this language. Which case marks the subject depends on the semantic properties of the subject argument and of the event. This phenomenon is known as Differential Subject Marking (DSM). A clear example of DSM is the dative/ergative alternation expressing changes in the modality of the event, discussed by Bashir (1999) and Butt and King (2002). In this case the alternation between the ergative and the dative encodes an alternation between volition and obligation. Consider the examples below (example 78 and 79) taken from Butt (2006b: 71):

78. *nādyā=ne* *zu* *jā-nā* *hai*
nadya=ERG zoo go-INF be.PRS.3SG
"Nadya wants to go to the zoo."

79. *nādyā=ko* *zu* *jā-nā* *hai*
nadya=DAT zoo go-INF be.PRS.3SG
"Nadya has/wants to go to the zoo."

Another example is the alternation between genitive and locative Possessors. In Hindi, possessive relations are encoded through different case marking on the Possessor depending on the semantic properties of the possessive relationship. In the case of inalienable possession the Possessor is marked with the genitive (as in 80), while in the case of alienable and temporary possession the Possessor is expressed by the locative postposition *ke pās* (as in 81; Carnesale 2022, see also Sulger 2015).

80. हमारे पास बेचने को भूसा नहीं है।

hamāre	*pās*	*bec-n-e=ko*	*bhūs-ā*	*nahīṁ*	*hai*
1PL.GEN	side	sell-INF.OBL=to	straw(M)-SG.NOM	not	be.3SG.PRS

"We have no straw to sell."

81. उनकी तीन लड़कियाँ हैं।

un=kī	*tīn*	*laṛk-iyāṁ*	*th-īṁ*
3PL.OBL=GEN	three	daughter(F)-PL.NOM	be.PST-F.PL

"He had three daughters."

Since Keenan's pivotal paper (1976), subjecthood has been conceived as deriving from a cluster of properties, distributed among the semantic, the morphological and the syntactic level, that can either conflate all in a single argument or be split between two arguments. This led many scholars to question the existence of the subject as a universal category, since this notion becomes problematic when the properties are not attached to a single element (Li and Thomson 1976). Two types of languages can thus be singled out: cumulative and non-cumulative languages. In cumulative languages all properties of subjecthood conflate on a single argument of the verb. In these languages, the subject it is not restricted to any thematic role: Agents, Experiencers, Patients, Recipients, Themes, etc. may all be encoded as full subjects. In non-cumulative languages, in contrast, subjecthood properties are split between two arguments. Hindi primarily employs non-comulative strategies: it tends to distribute the properties of subjecthood to more than one argument according to the semantic properties of the most salient participant in the event (see Montaut 2004 for a detailed discussion).

The concept of *salience* is not self-explanatory, and the term has been used with many different meanings. In this work, I use the adjective *salient* with the interpretation given by Lehmann (2002), who considers it as strictly connected to the properties of animacy and empathy[10]. As he

[10] Many other definitions have been proposed, for example according to Croft (2022: 172) "It appears that the primary factor determining argument status and expression as argument phrases is their degree of salience or topicality to the

points out (2002: 3) “if stripped of their relations, the participants of a situation are classified by their properties. The classification imposed by the grammars of natural languages is oriented towards the ego: participants are rated as to their similarity with the speaker. The result is not really a classification, but a hierarchy. The speaker is most empathic with himself and least empathic with entities which are not even thing-like.” Figure 6, already given in section 2.2.1 and presented again below, shows the empathy hierarchy proposed by Lehmann.

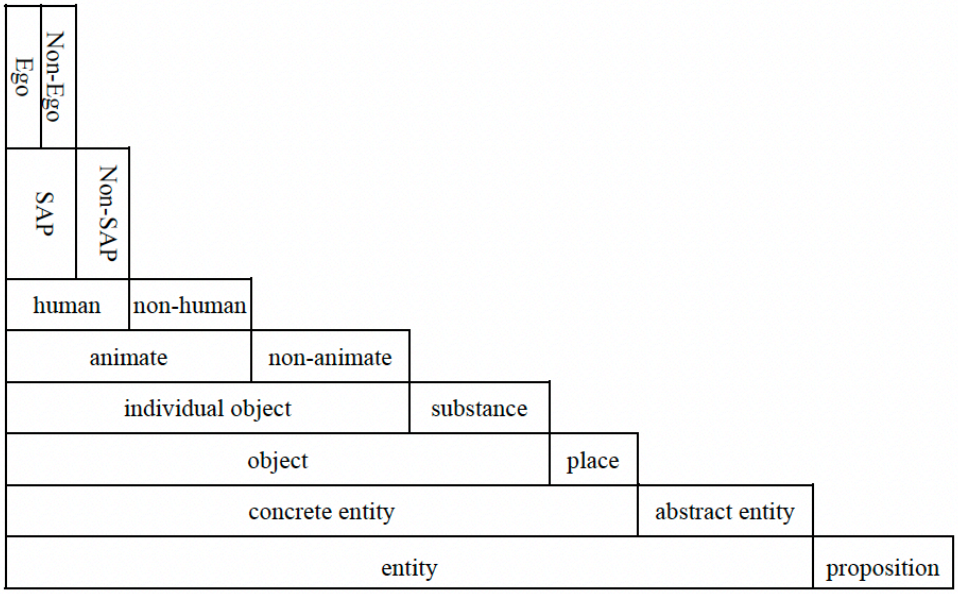

Figure 6: Empathy Hierarchy (Lehmann 2002: 4).

The central point of the hierarchy is the opposition between animate participants and inanimate ones. Animate participants serve as the initial step of an ascending hierarchy that leads to maximally empathic participants: non-human > human (> non-SAP > SAP; where *SAP* stands for *Speech Act Participants*). Conversely, inanimate participants serve as the starting point of a downward scale leading to minimally empathic participants. The higher the participant is on the empathy scale, the more salient it is in the speaker’s mind.

As discussed by Montaut (2004), Hindi encodes the most salient argument in the sentence with the most semantically fitting case-marking, thus privileging semantic roles over syntactic relations and it assigns the nominative case to the less salient entity by default. Consider for example the case of the expression of possessive relations in Hindi already mentioned above. In a predicative possessive construction, two arguments are encoded: the Possessor and the Possessee. The Possessor,

interlocutors in the discourse. The salience of a referent pertains to the degree of attention directed to that referent by the interlocutors at a given point in the discourse.”

which is prototypically human, is always in a higher position on the empathy scale than the Possessee (Heine 1997, Stassen 2009, Keidan 2008). Thus, the Possessor is the most salient element in the possessive relation, hence it is encoded with the most semantically fitting case-marking. Since possessive relations are much varied from a semantic perspective, the Possessor may receive different case-markings according to its semantic properties. When the possessive relation is alienable, the Possessor is marked with the adessive-locative postposition *ke pās*, in a construction typologically quite common (Heine 1997). When the relation is inalienable, the Possessor is marked with the genitive case. The Possessee always stands in the nominative.

Montaut (2004: 213) points out that full subjecthood in Hindi is restrained to transitive constructions (in the imperfective aspect) and to predicates with a single argument. In these cases, morphological and syntactic features conflate on the same element in the sentence, and we find arguments in the nominative that show subject's behavioral properties. In all the other cases, the properties of subjecthood are split: morphological coding properties are assigned to the less salient argument in the predicate (that is marked with the nominative and agrees with the verb), while syntactic behavioral properties are assigned to the more salient argument marked with an oblique case. Properties of the Hindi canonical subject are (see Mohanan 1994, Montaut 2013):

1. It stands in the nominative.
2. It triggers verb agreement.
3. It is in the first position in unmarked sentences.
4. It can be passivized.
5. It controls coreference with the reflexive *apnā*.
6. It controls coreference with the subject of the conjunctive participle in *-kar* and of adverbial participles.
7. It can undergo conjunctive participle reduction.
8. It controls Equi-deletion.
9. It can be subject to Equi-deletion.
10. It is preferably omitted in coordination.

Referring to these behavioral properties, various scholars (Kachru 1990, Mohanan 1994, Montaut 2004) proposed tests to identify non-canonical subjects in Hindi. Some of these tests are: 1. the non-nominative subject is the coreferential antecedent of the reflexive adjective *apnā*; 2. the non-nominative subject governs coreference with the implicit subjects of adverbial participles or conjunctive participles in *-kar;* 3. the non-nominative subject is coreferential with an implicit nominative subject. Moreover, the unmarked position of a non-nominative subject is at the beginning

of the sentence, a position generally filled by the subject in non-pragmatically marked sentences in Hindi. Note that non-nominative subjects do not all show the same properties; the behavioral properties of non-nominative subjects in Hindi are summarized in Table 10.

Table 10: Non-nominative subjects in Hindi and their behavioral properties.

Behavioral property	**NOM**	**ERG**	**DAT**	**LOC**	**GEN**	**INS**
Agreement	✓					
Initial position	✓	✓	✓	✓	✓	✓
Passivization	✓	✓				
Coreference with reflexives	✓	✓	✓	✓	✓	(✓)[11]
Coreference with conjunctive participles	✓	✓	✓			(✓)
Coreference with adv participles	✓	✓	✓	✓	✓	(✓)
Conjunctive participle reduction	✓	✓	(✓)			
Undergo equi-NP deletion	✓	✓	✓			
Control equi-NP deletion	✓	✓	✓	✓	✓	
Omission	✓	✓	✓			

As shown in Table 10,, the ergative subject only lacks verb agreement, but displays all other subjecthood properties. Dative subjects do not allow passivization, but they show all other properties. They may be omitted in coordination with nominative subjects, and they can control and can be subject to equi-deletion. However, conjunctive participle reduction with dative subjects is only allowed when the main verb as well requires the experiencer in the dative case, as in example 82 adapted from Montaut (2013: 92).

82. *bhāī=ko* *ciṭṭh-ī* *mil=kar* *khush-ī*
3SG.DAT letter(F)-SG.NOM receive=CP happiness(F)-SG.NOM.F
hu-ī
be.PRF-F.SG
"Brother was happy getting the letter."

[11] The brackets are used to indicate that instrumental subjects are not a homogeneous class and behave differently according to the semantic properties of the argument they encode (as exemplified in 83 and 84).

Genitive and locative subjects, which mark Possessors, exhibit fewer properties than dative subjects. For example, they can never undergo conjunctive participle reduction and they can never be omitted in coordinated sentences with subjects marked differently. However, they still control coreference, both with reflexive pronouns and with conjunctive participles, adverbial participles and indefinite subordinates. Instrumental subjects, which encode Agents with low agentivity and low volitionality, are not a homogeneous class (Montaut 2013: 93). For example, involuntary Agents do not control coreference (example 83), while inefficient Agents do (example 84).

83. *mujh=se* *apne* *kamr-e=kā* *darvāz-ā* *nahīṁ* *khul* *gay-ā*
1SG=INS REFL room(M)-SG.OBL= GEN door(M)-SG.NOM not open go.PRF-M.SG
"I couldn't open the door of my (REFL) room."

84. *mujh=se* **apne* *kamr-e=kā* *darvāz-ā* *ṭūṭ* *gay-ā*
1SG=INS REFL room(M)-SG.OBL=GEN door(M)-SG.NOM break go.PRF-M.SG
"I broke the door of my (REFL) room."

Since in the following chapters I will frequently deal with experiential dative subjects, some examples of subjecthood tests for dative Experiencers are given below (85-89; examples 88 and 89 are adapted from Kachru 1990: 64):

85. Coreference with the reflexive pronoun
मुझे अपनी कायरता पर लज्जा आ रही है।
mujhe *apn-ī* *kāyartā=par* *lajjā* *ā* *rahī*
1SG.DAT REFL-F cowardice(F.SG.OBL)=on shame(F.SG.NOM) come PRG.F
hai
be.3SG.PRS
"I am ashamed of my cowardice."

86. Coreference with conjunctive participles in *-kar* or imperfective/perfective participles
यह सोचकर मिझे लाजा हुई।
yah ***soc=kar*** *mujhe* *lajjā* *hu-ī*
this.NOM **think=CP** 1SG.DAT shame(F.SG.NOM) be-PRF.F.SG
"After thinking this I felt ashamed."

87. Coreference with the non-overt nominative subjects of coordinated clauses

us= ko	*ṭhaṇḍ*	*lag-ī*	*aur*	*(vah)*
3SG.OBL=DAT	cold(F.SG.NOM)	adhere-PRF.F.SG	and	(3SG.NOM)

andar gay-ā

inside go.PRF-M.SG

"He felt cold and went inside."

88. It controls Equi-deletion

lark-e= ko	*film*	*dekh-nā*	*pasand*	*hai*
boy(M)-SG.OBL=DAT	film(F.PL.NOM)	watch-INF	liked	be.PRS.3SG

"The boy likes to view films."

89. It can be subject to Equi-deletion

larke=ne	*film*	*pasand*	*ā-n-e=kī*
boy(M)-SG.OBL=ERG	film(F.PL.NOM)	liked	come=INF-OBL=GEN

carcā *nahīṁ kī*

mention(F.SG.NOM) not do.PRF.F

"The boy did not mention (his) liking the film."

3.2.5. Complex predicates

One of the pan-Indian features is the massive presence of complex predicates. This phenomenon is common to many if not all South Asian languages and according to Masica (1991) is one of the features that contributes to making South Asian a linguistic area. In South Asian linguistics, the term *complex predicate* covers a large range of predicate types, but two main categories can be singled out: 1. complex predicates composed by a sequence of two verbs (Butt 1995, Butt and Geuder 2001, Butt and Ramchand 2003, Drocco 2018, Drocco and Tiwari 2020a, 2020b), and 2. complex predicates composed by a sequence of a non-verbal element and a verb (Mohanan 1994, Butt and Ahmed 2010, Kulkarni 2011, Montaut 2016). In the next paragraph, I provide a brief description of complex predicates in Hindi: in section 3.2.5.1, I discuss V-V complex predicates, while in section 3.2.5.2, I address N-V complex predicates.

3.2.5.1. Verb + Verb Complex predicates

Verb -Verb complex predicates consist of a sequence of a main verb and a light verb. The light verb, in the previous literature, has also been called vector or compound verb, while the main verb has been

frequently referred to as polar verb (see for example Hook 1974, 1991, Drocco and Tiwari 2020a, 2020b). In this dissertation, I will refer to light the verb in V-V complex predicates with the term vector verb in order to distinguish it from the light verb of N-V complex predicates. While I will call the main verb simply as main verb. The formal structure of such a construction is as follows: the main verb comes first and stands in its root form, while the vector verb follows it in its finite form, marked with the information regarding TAM properties. All verbs can function as main verbs in this construction, but the class of vector verbs is closed. The most common vector verbs are: *jānā* "go", *ānā* "come", *denā* "give" and *lenā* "take". The vector verb modifies the main verb in various ways: it may add nuances of perfectivity, but may also add semantic features such as benefactivity, volitionality, directionality and other properties such as attitudinal (Masica 1976, 1991; Abbi and Gopalakrishnan 1991; Butt 1995; Butt and Geuder 2001; Drocco and Tiwari 2020a, 2020b). Since one of the main functions of this formation is to convey completeness of the action expressed by the main verb, V-V complex predicates tend to correlate with the perfective or the imperative forms (Masica 1991, Chakraborty 1992), while they are less frequent with the future form and quite rare with the progressive[12]. Additionally, given this association of the vector verb to the meaning of completion of the event, some scholars conceive it as a marker of perfectivity (Hook 1991, 1993). However, Butt and Geuder (2001) show that, while it is true that vector verbs supply an achievement or accomplishments reading, the semantics of complex predicates goes beyond the mere aspectual characterization. Butt and Geuder argue that V-V complex predicates may convey a wide range of meanings that appear to be related to the lexical meaning of the vector verb. In particular, they propose that the interaction between the vector verb and the main verb results in different interpretations of the event.

However, the semantics of V-V complex predicates is difficult to frame, as frequently the complex predicate can alternate with its simple verb counterpart with very little semantic difference. Two examples are given in sentences 90 and 91. In the first sentence, the complex predicate *band karnā* "close" is followed by the vector verb *denā*, which is absent in the second sentence. The contrasting meaning of these two sentences is difficult to convey in English, and it implies a nuance of perfectivity and responability/benefactivity expressed by the vector verb (Butt and Ramchand 2003, Butt and Geuder 2001).

90.	*us=ne*	*darvāz-ā*	*band*	*kar*	*diy-ā*
	3SG=ERG	door-M.SG.NOM	close	do	give.PRF-M.SG
	"He closed the door."				

[12] According to some scholars (Caracchi 1992:169), the use of vector verbs is ruled out with the progressive form.

91. *us=ne* *darvāz-ā* *band* *kiy-ā*
3SG.OBL=ERG door(M)-SG.NOM close do.PRF-M.SG
"He closed the door."

The vector verb *lenā* expresses benefactivity as well, but with a different orientation of the predicate: while *denā* implies that the action expressed by the main verb benefits a second Participant, *lenā* implies that the action has effects and consequences on the Agent[13] (Butt and Ramchand 2003, Slade 2013, Drocco and Tiwari 2020a, 2020b). An example is 92, where the verb *dekhnā* "look at, see" and the verb *socnā* "think" are followed by the vector verb *lenā*, which indicates that the outcome of the ecent expressed by the main verb benefits the agent. Similarly, the verb *jānā* and *ānā* besides the aspectual reading implying perfectivity (example 93) also add a nuance of dynamicity and directionality.

92. मैंने खूब सोच लिया है सब कगाज़ देख लिया।

main=ne *khūb* *soc* *liy-ā* *hai* *sab* *kāgāz*
1SG=ERG very_well think take.PRF-M.SG be.PRS.3SG all paper(M.PL.NOM)
dekh *liy-ā*
look_at take.PRF-M.SG
"I've thought about it thoroughly. I've examined (lit. look take) all the papers."

93. वसीयत की बात फिर उसे याद आ गई।

vasīyat=kī *bāt* *phir* *use* *yād*
will(M.SG.OBL)=GEN matter(F.SG.NOM) again 3SG.DAT memory(F.SG.NOM)
ā *ga-ī*
come go.PRF-F
"He remembered the will again."

Other frequent vector verbs are *baiṭhnā* "sit" which implies regret about the action expressed by the main verbs, *uṭhnā* "rise" which adds the semantics of suddenness, *ḍālnā* "throw" which expresses an action done violently and *paṛnā* "fall" which is used to express a lack of intentionality by the subject (Hook 1974). Besides these semantic and aspectual functions, some scholars (Hook 1974, Drocco and Tiwari 2020b) pointed out that the use of vector verbs is related also to the concept of mirativity

[13] Andrea Drocco (p.c.) proposes to identify this use of *lenā* with the term auto-benefactive.

and intersubjectivity: i.e., the use of a vector verbs implies some share knowledge between the speaker and the listener (Drocco and Tiwari 2020b:170).

3.2.5.2. Noun + Verb Complex predicates

Like many other languages of South Asia, Hindi makes extensive use of N + V complex predicates. In the literature in English, these types of construction are commonly referred to also as "conjunct verbs" (McGregor 1972, Shapiro 1984, Mohanan 1997, Kachru 2006, Kulkarni 2011). The phenomenon is widespread in Hindi and encompasses a wide range of semantic classes of verbs. Among these, experiential verbs are particularly productive, and it is therefore important to briefly examine how Hindi complex predicates function and how they interact with the way experiential situations are expressed in Hindi. These predicates consist of a nominal host and a light verb that act as a single predicate unit, whose semantics is different from that of the light verb alone (Mohanan 1994, Butt 1995). In these formations, the semantic core is expressed by the nominal, while the verb adds information on TAM properties and voice. Compare sentences in 94: sentence a) shows the complex predicate *dikhāī denā* "see" formed by the light verb *denā* "give" and the noun "*dikhāī*" "seeing", while sentence b) shows the verb *denā* when used as a simplex verb. A similar example is given in 95, that shows the verb *ānā* "come" used as a light verb in the complex predicate *yād ānā* "remember" (a), and the same verb *ānā* used as a simplex verb (b).

94.

a) पन्ना को कैरम और अंधेर जी दिखाई देता था।

pannā =ko	*cārom̐ or*	*am̐dher-ā*	*hī*	*dikhāī*
panna=DAT	all_four_directions	darkness(M)-SG.NOM	EMPH	seeing(F)SG.NOM
de-t-ā	*th-ā*			
give-IPRF-M.SG	be.PST.3SG			

"Panna could see only darkness all around."

b) करुणा गौओं को पानी देती है।

karūnā	*gau-om̐=ko*	*pānī*	*de-tī*	*hai*
karuna.NOM	cow(F)-PL.OBL=DAT	water(M.SG.NOM)	give-PRS.F.SG	be.PRS.3SG

"Karuna gives water to the cows."

95.

a) रतन को अपना वादा याद आया।

ratan=ko *apn-ā* *vād-ā* *yād* *āy-ā*
ratan=DAT REFL-M.SG promise(M)-SG.NOM memory(F.SG.NOM) come.PRF-M.SG
"Ratan remembered her promise."

b) अभी तुम्हारी सहेली रतन आई थी।

ab=hī *tumhārī* *sahel-ī* *ratan* *ā-ī*
now=EMPH 2PL.GEN friend(F)-SG.NOM ratan.NOM come-PRF.F.SG
th-ī
be.PST-F.SG
"Your friend Ratan had just come."

When they are translated into other languages, Hindi N-V complex predicates are usually translated with simple verbs, whose semantics differs from the semantics of the light verb. For this reason, some studies propose to detect complex predicates in the Hindi lexicon by adopting a translation equivalent approach (Mahesh and Sinha 2009). The class of nominals that can function as a host in Hindi is highly productive: any noun or adjective – and even some adverbs – can be used to form a complex predicate. In contrast, the class of verbs that can be used as light verbs is closed and only a dozen verbs can occur in this function. Among these, the verb *honā* "be" and the verb *karnā* "do" are the most frequent: the first is used to form intransitive predicates and the second to form transitive predicates. Few other verbs can be used in complex predicates: among these, there are verbs such as *ānā* "come", *lagnā* "adhere", *rahnā* "stay", *paṛnā* "fall", *uṭhnā* "rise", which form intransitive predicates, or verbs such as *denā* "give", *lenā* "take", *mārnā* "hit", *ḍālnā* "throw", which form transitive predicates[14]. Both the nominal host and the verb can be used independently.

Complex predicates constitute the bulk of Hindi verbal lexicon: indeed, they do not only occur in huge numbers, but they also express basic notions that do not have a corresponding simple verb. For example, notions like "work" (*kām karnā*), "love" (*pyār karnā*), "wait" (*intazār karnā*) and "remember" (*yād honā*/*karnā*) can only be expressed by complex predicates in Hindi. The high productivity of this phenomenon is also evident in the incorporation of verb borrowings into the lexicon. When a verb is borrowed from another language (such as English, Persian, Arabic, and Sanskrit), it becomes a part of the Hindi lexicon by forming a complex predicate consisting of the borrowed term functioning as the nominal host, followed by a Hindi light verb. For instance, the verb *print karnā,* meaning "print", is composed by the English borrowed noun *print* and the Hindi verb

[14] Except for the experiential predicates *dikhāī denā* and *sunāī denā* that I will discuss in the chapter dedicated to Perception verbs.

karnā “do”. Similarly, the verb *mālūm honā* “know” is formed by the Arabic adjective *mālūm* “known” and the Hindi verb *honā* “be” (see on this Kachru 2006 and Montaut 2004).

As mentioned above, in complex predicates, the nominal host and the light verb form a single predication. This means that the argument structures selected by these complex predicates are partly determined by the light verb requires and partly influenced by the semantics of the noun. The noun can contribute to the number of arguments of the entire predicate and their meaning and case marking (see on this also Mohanan 1997: 212-225). See, for example, sentence 96 in which the PP *mujh=par* “on me” cannot be licensed by the verb *karnā* “do” which typically only requires two arguments.

96. आप वह पहले आदमी हैं जिसने मुझपर विश्वास किया है।

āp	*vah*	*pahl-e*	*ādmī*	*haiṁ*
2SG.HON.NOM	CORR.PRN.SG.NOM	first-M.PL.NOM	man(M.PL.NOM)	be.PRS.3PL

jis=ne	*mujh=par*	*viśvās*	*kiy-ā*	*hai*
REL.PRN.OBL=ERG	1SG.OBL=on	trust(M.SG.NOM)	do.PRF-M.SG	be.3SG.PRS

“You are the first person that trusts me.”

Mohanan (1997: 212-225) provides further evidence for considering the sequence noun + verb as a single predication. She argues, for example, that gapping and ellipsis in answers to yes/no questions are not allowed either for the light verb alone or for the noun alone, while the entire complex predicate can be gapped. And she also points out that scrambling tests show that the nominal is part of the verbal constituent, as the noun in a complex predicate is not free to scramble.

Notably, however, the status of the nominal host is far from being stable (Mohanan 1997: 225-235). On the one hand, the nominal element is part of the predicate, on the other hand it behaves as an argument: it may be passivized (see Shapiro 1989) and it may agree with the verb. In example 97, the verb is in the perfective form and requires an ergative marking on the subject, thus the perfective participle of the light verb *karnā* (*kī*) is in the feminine and agrees with the nominal host *madad* “help”. The same happens in 96 above, in which the verb *karnā* agrees with the nominal host *vicār,* as both the arguments *jis=ne* and *mujh=par* are oblique. In example 98, the nominal host *vicār* is passivized and it is raised to subject status.

97.

admī=ne	*kitāb*	*paṛh-n-e=meṁ*	*apn-e*
man(M.SG.OBL)=ERG	book(F.SG.NOM)	read-INF-OBL=in	REFL-M.SG.OBL

bacc-e=kī	*madad*	*k-ī*
child(M)-SG.OB=GEN	help(F.SG.NOM)	do-PRF.F

“The man helped his son to read the book.”

98. ऐसी दलीलों पर विचार किया ही नहीं जा सकता है।

esī	*dalīl-oṁ=par*	*vicār*	*kiy-ā*	*hī*	*nahīṁ*
such	argument(F)-SG.OBL=on	thought(M.SG.NOM)	do.PRF-M.SG	EMPH	not
jā	*sak-t-ā*	hai			
go	can-IPRF-M.SG	be.3SG.PRS			

"Such arguments could not be considered."

The category of N - V complex predicates in Hindi is extremely heterogenous and many scholars tried to propose a taxonomy. Most of them focused mainly on the parameters of agreement and argument structure.[15] In the present work, I will use the classification proposed by Montaut (2016) based on three parameters: agreement, case marking and the licensing of external argument (Davison 2005 proposes a similar classification). Montaut (2016) distinguishes two classes, one of them further divided into two subclasses.

The first class consists of complex predicates with internal agreement, i.e. complex predicates in which the nominal host agrees with the light verb. In transitive predicates agreement is visible only in the ergative alignment, as the nominal host functions as the direct object of the verb and thus triggers agreement only when the subject is marked with the ergative. In contrast, in intransitive predicates, the nominal host always functions as the subject of the verb, thus it always triggers agreement. Predicates belonging to this class can be divided into two subcategories, according to the way they treat their external argument. In the first subtype, the external argument retains its status and it is marked standardly, with the postposition typically associated to that semantic role (as in 99). In the second type, the external argument is obligatorily marked with the genitive postposition *kā* (/*ke*/-*kī*), as in 100.

99.

tum=ne	*mujh=par*	*viśvās*	*kiy-ā*	*hai*
2PL=ERG	1SG.OBL=on	trust(M.SG.NOM)	do.PRF-M.SG	be.3SG.PRS

"You trusted me."

100.

admī=ne	*kitāb*	*paṛh-n-e= meṁ*	*apn-e*
man=ERG	book(F.SG.NOM)	read-INF-OBL=in	REFL-M.SG.OBL
bacc-e=kī	*madad*	*k-ī*	
child-M.SG.OBL=GEN	help.F.SG.NOM	do-PRF.F	

[15] For an overview of the classifications that have been proposed in the previous literature the reader is referred to Montaut (2016).

"The man helped his son read the book."

The second class consists of complex predicates in which the nominal host does not trigger agreement with the light verb. Verbs belonging to this class treat the external argument as a canonical direct object: it may be subject to DOM and it may agree with the verb in ergative alignments (as in 101).

101. *rānā=ne* *kāngres=kī* *sadasyatā* *grahaṇ*
rana=ERG congress(M.SG.OBL)=GEN membership(F.SG.NOM) grasping(M.SG.NOM)
k-ī
do-PRF.F.SG
"Rana took the Congress's membership."

Kulkarni (2011) analyses semantic and syntactic criteria governing agreement in Hindi complex predicates aiming to detect under what circumstances the verb agrees with its nominal host (internal agreement) and under what circumstances it does not (external agreement). According to her, the clue is in the semantics of the nominal host: if the nominal host is a "verbal noun" (i.e. an action noun) then the external argument takes the genitive, and the verb agrees with the nominal host (internal agreement). While if the noun is not a verbal noun, two options are available. In the first option, the external noun is oblique and followed by a postposition according to its semantic role, and in this case the nominal host is treated as an argument and triggers agreement with the verb (internal agreement). In the second option, the external argument is treated as a direct object, the verb agrees with it and the internal agreement with the nominal host is thus ruled out.

Notably, however, this classification is not clear-cut and that some complex predicates may take either external or internal agreement. Consider the contrastive examples with the verb *yād karnā* "remember" in 102. In 102a, the second argument is treated as a direct object (it is marked with the accusative *ko*) and the verb (*k-i**yā̃***) does not agree with the feminine nominal host *yād*, it is in the default masculine form because of the co-occurrence of ergativity and accusative. In 102b the second argument is marked with the genitive, and the verb (*k-**ī̃***) agrees with the feminine nominal host.

102.

a) *maiṁ=ne* *us* *laṛk-ī=**ko*** *yād* *ki**y-ā̃***
1SG=ERG that.OBL girl(F)-SG.OBL=**ACC** memory(F.SG.NOM) do.PRF-**M**.SG
"I remembered that girl."

b) *maiṁ=ne* *us* *laṛk-ī=**kī̃*** *yād* *k-**ī̃***
1SG=ERG that.OBL girl(F)-SG.OBL=**GEN** memory(F.SG.NOM) do.PRF-**F**.SG

"I remembered that girl."

3.2.5.2.1. Ligh verb alternations: actionality and valency

While the nominal host functions as the semantic core of the complex predicate, the light verb conveys grammatical information related to TAM properties. However, in some respects, it also conveys information on the lexico-semantic level, and in particular on Aktionsart properties. The same nominal host can occur with several light verbs, and the alternation of the light verb generates a difference in the lexical aspect. These alternations are very frequent with experiential complex predicates, and they contribute different construals of a given experiential situation. For example, they are extremely productive for the expression of remembering in Hindi. This experience is expressed by several complex predicates, some of theme already exemplified in sentence 102 above, all formed by the noun *yād* "memory". The light verbs alternating with the noun convey information regarding the semantic properties of the event and the participant and regarding Aktionsart properties. The verb *yād karnā* (*karnā* "do"), for example, implies that the situation is construed as an activity and the experiencer resembles an Agent in its semantic properties (as in 103a) and it is opposed to the verb *yād honā* (*honā* "be"), which implies lack of control by the Experiencer and construes the experience as a state (as in 103b). Besides these two basic forms, other forms may be used. The verb *yād rakhnā* (*rakhnā* "put"), for example, differs from the verb *yād karnā* as it implies the same agentive reading, but conveys a durative meaning (as in 103c). The verb *yād ānā* (*ānā* "come") contrasts with the verb *yād honā*, as it expresses the same lack of agentivity but also supplies an achievement reading (as in 103d).

103.

a) मेरी मोहब्बत की याद करो।

merī	*muhabbat= kī*	*yād*	*karo*
my	love(F.SG.OBL)=GEN	memory(F.SG.NOM)	do.IMP

"Remember my love."

b) मेरी माता को तुम्हारी सूरत याद है।

merī	*mātā= ko*	*tumhārī*	*sūrat*	*yād*	*hai*
my	mother=DAT	your	face	memory(F.SG.NOM)	be.3SG.PRS

"My mother remembers your face."

c) हमें हर क्षण इसे याद रखना चाहिए।

hameṁ	*har*	*kṣaṇ*	*ise*	*yād*	*rakh-nā cāhiye*
1PL.DAT	every	moment	this.ACC	memory(F.SG.NOM)	put-INF should

"We should remember this every moment."

d) सहसा उसे गोबर की याद आई।

sahsā	*use*	*gobar=kī*	*yād*	*ā-ī*
suddenly	3SG.DAT	gobar=GEN	memory(F.SG.NOM)	come-PRF.F.SG

"Suddenly he remembered Gobar."

Another example is the alternation between the verb *ānā* "come " (or *lagnā* "be attached") as opposed to the verb *honā* in the expression of mental experiences or physical sensations. As I will argue in more detail in the following chapters, these two verbs convey an inchoative meaning thus pointing toward an achievement interpretation (as in104), in contrast with the light verb *honā* that tend to construe the event as a state (as in 105). The verb *ānā* is typically used in complex predicates expressing cognitions or emotions, whereas the verb *lagnā* usually occurs with complex predicates referring to bodily sensations. As I will discuss in section 8.8.2, this semantic difference probably derives from the different meanings expressed by the two light verbs, as *lagnā* implies a physical contact that is absent in the meaning of *ānā*.

104. मुझे अपने आस-पास की हर चीज़ पर गुस्सा आता रहा।

mujhe	*apne*	*ās-pās=kī*	*har*	*cīz=par*	*guss-ā*
1SG.DAT	REFL	near=GEN	all	thing=on	anger(M)-SG.NOM
ā-t-ā	*rahā*				
come-IPRF.M.SG	stay-PRF.M.SG				

"I kept getting angry at everything around me."

105. मुझे अपने पर है।

mujhe	*apne= par*	*guss-ā*	*hai*
1SG.DAT	REFL=on	anger(M)-SG.NOM	be.PRS.3SG

"I am angry at myself."

The alternation of the verb used in the complex predicate may also bring about changes in valency patterns. For example, the intransitive anticausative vs. transitive causative alternation is realized through the alternation of the light verb: *honā* "be" vs *karnā* "do". In sentence 106, the complex predicate *band karnā* "lit. close do" expresses a causal event with an agentive first argument "*us ādmī=ne*", while in sentence 107, the agentive causer is removed, and the event is construed as spontaneous.

106. *us* *ādmī=ne* *khirk-ī* *band* *k-ī*
that.OBL man=ERG window(F)-SG.NOM close do.PRF-F.SG
"That man closed the window."

107. *khirk-ī* *band* *ho* *gay-ī*
window(F)-SG.NOM close be go.PRF-F.SG
"The window closed (by itself)."

As I will discuss in the following chapters, this mechanism is regular and highly productive for the expression of experiential events in Hindi, as it is systematically used to express two main ways of construing an experience. The transitive causative form is used to encode experiences in which the Experiencer is construed as endowed with control and volitionality, while the intransitive anticausative form typically occurs with dative constructions and is used to express non-agentive experiences.

3.3. Complementation in Hindi

Since in many experiential events (in particular in perceptions and cognitions), the Stimulus is very frequently expressed by complement clauses, in this section I briefly address how complementation is realized in Hindi. Clausal complementation can involve finite clauses or non-finite clauses. Finite clauses are always extraposed, while non-finite clauses typically appear pre-verbally.

Finite complementation in Hindi is always realized through a subordinating sentence following the main verb: the clause may be introduced by the particle *ki* "that" (as in sentence 108) or it may and be realized by a mere juxtaposition to the main clause (as in 109). This complementizer derives from a form of the verb *kahnā* "say" and in previous literature has often been labeled "quotative" (Subbarao 2012: 194). This origin of the complementizer is a shared feature of South Asian languages and it is interesting since, as Noonan (1985: 47) points out, complementizers rarely derive from verbs and typically origin from conjunctions, pronouns, case markers and adpositions. The complement clause may also be anticipated by a correlative in the main clause (see on this Kachru 2006: 215-220; Bhatt, Farudi and Rambow 2013: 88). This correlative can be instantiated by the pronoun *yah* "this" (110) or by a noun (111). In my analysis, I will treat all these forms of the finite complemental clause as different instantiations of the same construction. The complement clause may occur either in subject or object position. In the first case, we deal with subject complementation, while in the second case with object complementation. As I will discuss in chapter 8, this construction is particularly

frequent with verbs of cognition in Hindi, and it is used to express the propositional content of the mental state or activity.

108. अब मुझे सूझने लगा कि जीवन का लक्ष्य सुख –भोग ही है।

ab	*mujhe*	*sūjh-n-e*	*lag-ā*	*ki*
now	1SG.DAT	be_perceived-INF-OBL	attach-PRF.M.SG	that

jīvan kā lakshya sūkh-bhog hī hai.
the goal of life is happiness and enjoyment.
"Now I started to think/understand that the goal of life is happiness and enjoyment."

109. वह सोच रही थी, आदमी में स्वार्थ की मात्रा कितनी अधिक होती है।

vah	*soc*	*rah-ī*	*th-ī,*
3SG.NOM	think	PRG-F.SG	be.PST-F

ādmī meṁ svārth kī mātrā kitnī adhik hotī hai.
how much selfishness is there in a man?
She was thinking: "How much selfishness is there in a man?"

110. अगर तुम यह समझ रही हो कि संसार में तुम्हारा कोई नहीं है, तो यह तुम्हार भ्रम है।

agar	*tum*	*yah*	*samajh*	*rah-ī*	*ho*	*ki*
if	2PL.NOM	this.NOM	understand	PRG-F.SG	be.SBJV.PRS.3SG	that

saṁsār meṁ tumhārā koī nahīṁ hai, to yah tumhārā bhram hai.
you have no one in the world, then this is your misconception.
"If you are thinking that you have no one in the world, then you are wrong."

111. उसे बार-बार एक ही बात सूझ रही थी कि वह उसे बांहों से पकड़ ले और उसके मुंह पर हाथ रखकर उसका मुंह बन्द कर दे।

use	*bār-bār*	*ek*	*hī*	*bāt*	*sūjh*
3SG.DAT	time-time	one	EMPH	matter(F.SG.NOM)	be_perceived
rah-ī	*th-ī*	*ki*			
PRG-F.SG	be.PST-F	that			

vah use bāṁhoṁ se pakaṛ le aur uske muṁh par hāth rakhkar uskā muṁh band kar de.
he should hold her by the arms, put a hand over her mouth and shut her mouth.
"He kept thinking about one thing repeatedly: that he should hold her by the arms, put a hand over her mouth and shut her mouth."

Another complemental clause that appears very frequently with the verbs investigated in this dissertation is the predicative participle construction. This sentence-type is particularly relevant in the case of perception verbs and employs a participle in predicative relation with the Stimulus NP. This construction occurs both with the transitive construction (as in 112) and with the dative construction (as in 113). The Stimulus NP, which is an argument of the main verb, includes a secondary predication. In sentence 112, for instance the object of the first predication becomes the subject of the secondary predication and the event expressed by the participle is construed as a property of the Stimulus. As I will discuss in section 7.5, this construction is particularly frequent with verbs of perception, but it can also be found with some verbs of cognition.

112. सूयोर्दय हुआ तो उसने जलपक्षियों को नदी के किनारे एक पैर पर खड़े देखा।

sūyorday	*hu-ā*	*to*	*us=ne*	*jalapkṣiy-oṁ=ko*	
sunrise	be.PRF-M.SG	then	3SG.OBL=ERG	water-birds(M)-PL.OBL=ACC	
nadī=ke	*kināṛ-e*	*ek*	*pair=par*	*khaṛ-e*	*dekh-ā*
river=GEN	side-OBL	one	leg=on	stand-PRF.M.PL	see-PRF.M.SG

"When the sun rose, he saw water birds standing on one leg by the riverbank.

113. एक सौ पग भी न चला होगा कि उसे नदी के तट पर एक मनुष्य पाल्थी मारे बैठा दिखाई दिया।

ek sau pag bhī na calā hogā ki
He must not have taken even a hundred steps

use	*nadī=ke*	*taṭ=par*	*ek*	*manuṣy*
3SG.DAT	river(F)SG.NOM=GEN	bank(M.SG.OBL)=on	one	man(M.SG.NOM)

pālthī māre	*baiṭh-ā*	*dikhāī*	*di-yā*
cross-legged	sit-PRF.M.SG	seeing(F)	give-PRF.M.SG

he saw a man sitting cross-legged by the river-bank.
"He must not have taken even a hundred steps when he saw a man sitting cross-legged by the river-bank."

Hindi verbs of experience may also appear with non-finite complement clauses. Verbs requiring this complement type typically occur with a direct infinitive. This construction is not very frequent with the verbs analyzed here, but it appears with verbs expressing the possession of a skill or the acquisition of a skill. Since the subject of this complement clause is coreferential with the subject of the main clause, it undergoes obligatory deletion. Consider, for example, sentence 114, in which the verb *sikhnā* "learn" appears with an infinitive Stimulus.

114. बचपन से उसने गऊओं का पालना और घास छीलना सीखा था।

bacpan=se	*us=ne*	*gaū-oṁ=kā*	*pāl-nā*
childhood(M.SG.OBL)=from	3SG.OBL=ERG	cow(F)-OBL.PL=GEN	herd-INF
sīkh-ā	*th-ā.*		
learn-PRF.M.SG	be-PST.M.SG		

"Since childhood, he had learned to herd cows and peel grass."

4. The theoretical framework

The theoretical framework I refer to in this work is situated within the broad functional-cognitive paradigm (Croft and Cruse 2004). The basic tenets of this approach are:

a) Language is not an autonomous cognitive faculty.
b) Grammar is conceptualization.
c) Grammar emerges from usage.

Language is not autonomous in the sense that even though the specific configuration of cognitive abilities employed during linguistic communication (such as simultaneous perception and production of a sequence of structured symbolic units) is unique to language, the single cognitive skills required for such communicative ability are not (Croft and Cruse 2004: 2). This hypothesis emerged in the 1970s, in response to the generative grammar's theory that language is innate and separate from other non-linguistic cognitive modules. Saying that language is anchored to all other cognitive skills also implies that linguistic knowledge is represented in the same way as other conceptual structures in the mind. This applies not only to meaning, but also to the syntactic, morphological and phonological levels. The fact that linguistic knowledge is not autonomous from other cognitive modules of conceptualization also implies that grammar has a role in the conceptualization of the external world and that the study of grammar can give us insights on the way we categorize and conceptualize reality (Croft and Cruise 2004: 3). Finally, grammar is built through a process of abstraction and schematization of specific utterances (Croft and Cruse 2004: 4).

In this study, I will borrow concepts and tools from three main approaches to the study of the interaction between grammar and conceptual structure. These approaches are: Frame Semantics, which focuses on the organization of semantic-conceptual knowledge (Fillmore 1977a, 1977b, 1982; Fillmore and Baker 2001, 2010; Langacker 1987); Construction Grammar, which focuses on the organization of grammatical knowledge (Goldberg 1995, 2006; Hilpert 2014, Perek 2015); and the usage-based framework, which starts from the assumption that grammar emerges from language use (Langacker 1987, Barlow and Kemmer 2000, Bybee and Hopper 2001). The following section are dedicated to a brief overview of these theoretical approaches.

4.1. Frame semantics

The term frame semantics refers to a wide variety of approaches to the systematic description of natural language meanings. The one common feature of all these approaches is the following slogan by Charles Fillmore (1977a):

A. Meanings are relativized to scenes.

Fillmore's slogan implies that meanings are determined and interpretated according to the background frame or scene they are situated in. The term "frame" refers to a system of concepts which are related by some motivating context such that, without knowledge of the entire system, one does not have complete knowledge of any of the component concepts (Fillmore 1982: 111). One of the arguments presented in favor of a frame-based approach to lexical semantics is that many concepts cannot be understood without the knowledge of the social and cultural background in which the situation exists. For example, the concept word *vegetarian* makes sense only in a community of speakers in which meat-eating is common (Fillmore 1982:120). An often-cited example by Fillmore (1977) is the difference in meaning between 115 and 116 emerging from the different backgrounds generated by the words *land* and *ground*. Sentence 115 refers to a sea voyage while sentence 116 refers to an air travel.

115. *I spent three hours on land this afternoon.*
116. *I spent three hours on the ground this afternoon.*

The relationship between a word concept and the frame to which it belongs is called the profile-based relation (Langacker 1987; see also Croft and Cruse 2004). Langacker (1987) illustrates the profile-based relation with the word *radius*. He points out that we can understand the meaning of *radius* only if we have access to an understanding of the concept of a *circle*, as a *radius* is a line segment, but not any line segment: it is the line segment that joins the center of a circle with any point on its circumference. So, *radius* is defined in relation to the structure of the circle (see Figure 7).

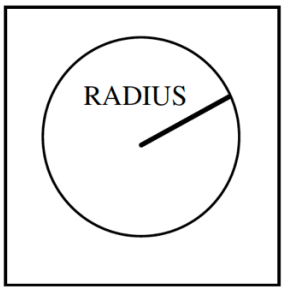

Figure 7. The profile-base relation: *radius* and *circle* (Croft and Cruse 2004: 15).

In Frame Semantics, lexical entries and phrases are assigned a specific semantic frame, i.e. a structured representation of the speakers' conceptualizations of reality, essentially blocks of encyclopedic information that contain knowledge about referential concepts or situations. Frames

consist of a set of interconnected elements such as participants, attributes, and relations, which together form a coherent representation.

Thus, it is not only relational nouns that represent a concept profile against a base: the process of profiling can also be applied to situations in general. When referring to a situation, a frame characterizes a small abstract scene which identifies (at least) the participants of the scene and thus the arguments of predicates and sentences describing the scene. These frame elements are called roles because they generalize over many potential situations and individuals. In specific instances of a frame, roles receive fillers. Not all participants are obligatorily expressed linguistically: the *profiling* of a situation indicates which participants need to be obligatorily expressed and which participants function as focal points in the profiling process (Langacker 1991, 2009).

Let us apply the concepts of frame and profile to the expression of possession in Hindi, in order to show how the relationship between semantics and syntax is represented in Frame Semantics. The possessive situation identifies a relation between two roles, a Possessor and a Possessee, which can be differently profiled. One fundamental distinction holds between *ascriptions of possession* (or *have*-constructions) and *predications of belonging* (or *belong*-constructions; see on this Langacker 1995, 2001; Heine 1997, 2001; Keidan 2008, Stassen 2009, Carnesale 2022). These two ways of construing the event are exemplified in 117 for ascription of possession, and in 118 for predication of belonging.

117. ASCRIPTION OF POSSESSION

लेकिन मेरे पास नगद नहीं है।

lekin	*mere pās*	*nagad*	*nahīṁ*	*hai*
but	1SG.LOC(beside)	cash.SG.NOM	not	be.3SG.PRS

"But I have no cash."

118. PREDICATION OF BELONGING

जिस मकान में रहता हूँ, वह अब मेरा नहीं है।

jis	*makān=meṁ*	*rah-t-ā*	*hūṁ*	*vah*
Rel.OBL	house(M.SG.OBL)=in	live-IPRF-M.SG	be.1SG.PRS	Crr.NOM
ab	*mer-ā*	*nahīṁ*	*hai*	
now	1SG.GEN-M	not	be.3SG.PRS	

"The house I am living in now does not belong to me."

The difference between these two sentences depends on the different ways they profile the same situation: an *ascription of possession* encodes the relation between the two *relata* from a possessor-oriented point of view: it takes the Possessor as the topical item, while the Possessee is the new information and has the role of focus. In *predications of belonging*, instead, the Possessee is the topic, while the Possessor adds new information: this construction encodes the relationship from a possessee-oriented point of view.

Another example given by Goldberg (1995) involves the English verbs *rob* and *steal* and illustrates how differences in the lexical profiling of participants lead verbs encoding the same event type to develop different arguments structures. The verb *rob* profiles the thief and the victim and shows the argument structure in 119 (profiled participant roles are represented in boldface), while the verb *steal* profiles the thief and the goods, and shows the argument structure in 120.

119. rob <**thief target** goods>
Mark robbed the millionaire (of all her money).
** Marked robbed a million dollars from the millionaire.*

120. steal <**thief** target **goods**>
Mark stole a million dollars from the millionaire.
** Mark stole the millionaire of her money.*

4.2. The constructional approach

Following recent language-specific studies on experiential constructions (Verhoeven 2007, Fedriani 2014, Luraghi 2020a), I adopt the theoretical framework of Cognitive Construction Grammar (Goldberg 1995, 2006, Hilpert 2014, Perek 2015). Construction Grammar (CxG) was initially proposed at Berkeley University, through the seminal work of Charles Fillmore, Paul Kay and Mary Catherine O'Connor (Fillmore and Kay 1993; Fillmore, Kay and O'Connor 1988, Kay and Fillmore 1999) as an alternative to the traditional generativist approaches (Chomksy 1957, 1965). Despite sharing some foundational ideas with the generative approach, such as the conception of natural language as a cognitive system and the focus on its creative productivity, Construction Grammar contrasts sharply with it. While generativists focus on the formal structures of language and mainly disregard semantic and pragmatic functions, constructionists focus on the inextricability of these two levels. Central to the constructional framework is the primacy of *constructions* in the grammar of a language, conceived as any linguistic sign consisting of a conventionalized pairing of form and

meaning. This way of intending constructions thus implies that the semantic level is as important as the syntactic one (Hilpert 2014, Goldberg 1995, 2006). As Golberg says (2005: 5):

> All levels of grammatical analysis involve constructions: learned pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully general phrasal patterns. […] Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency.

Figure 8 taken from Croft and Cruse (2004: 258) illustrates a schematic representation of a construction: the formal level contains information related to the syntactic, morphological and phonological level, whereas the functional level is associated to the meaning of the construction and specifies information related to its semantic and pragmatic properties, but also to its discourse-functional properties.

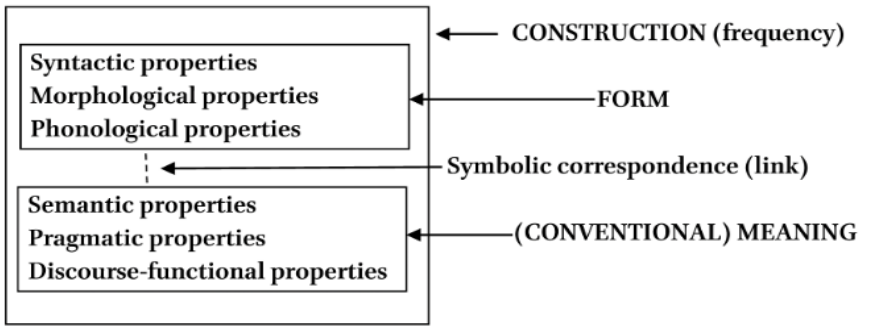

Figure 8: Schematic representation of a construction (Croft and Cruse 2004: 258).

Another important tenet of CxG that sharply differentiates it from the generativist framework is the emphasis on surface structures: generativist approaches aim to provide a universal description of languages that minimizes the number of linguistic items that need to be stored in the speaker mind and thus highlights the importance of transformational mechanisms and compositional semantics (Chomsky 1957, 1965). The main aim of the generativist approach is to explain how the linguistic mind can generate a vast array of surface constructions from a limited number of core items stored in the speaker's memory. From the constructional perspective there is no distinction between the surface and the core of the grammar, and every linguistic element, ranging from morphological affixes to argument structures, is treated as a construction (Hilpert 2013: 2). Consequently, the grammar is conceived as the repository of all constructions of a language: this repository is called *Constructicon* (Goldberg 1995, 2003). A corollary of this approach is the rejection of distinct and discrete modules

within the grammar and instead the positing of a continuum that encompasses constructions of varying degrees of abstraction, thus implying that there is no distinction between syntax and the lexicon.

This lexicon-syntax continuum where constructions are collocated is generated on two dimensions: from substantive to schematic and from atomic to complex (Croft 2001, Goldberg 2006). These two dimensions are not correlated: this means that a construction can be highly complex but not necessarily highly schematic. Table 11 represents the syntax-lexicon continuum as proposed by Croft (2001) and shows that the different combinations of the dimension of complexity and schematicity correspond to the various level of the traditional analysis of a language (lexicon, syntax, morphology, etc.). Table 12, adapted from Goldberg (2006: 5), shows some examples of constructions varying in size, schematicity and complexity.

Table 11: The syntax-lexicon continuum (adapted from Croft 2001: 17).

Construction type	**Traditional name**	**Examples**
Complex and (mostly) schematic	Syntax	[SBJ be-TNS VERB-en by OBL]
Complex and (mostly) specific	Idiom	[pull-TNS NP's leg]
Complex and bound	Morphology	[NOUN-s], [VERB-TNS]
Atomic and schematic	Syntactic category	[DEM], [ADJ]
Atomic and specific	Word/lexicon	[this], [green]

Table 12: Construction types (adapted from Goldberg 2006: 5).

Construction type	**Example**
Morpheme	e.g. *pre-*, *-ing*
Word	e.g. *avocado*, *anaconda*, *and*
Complex word	e.g. *daredevil*, *shoo-in*
Complex word (partially filled)	e.g. [N-s] (for regular plurals)
Idiom (filled)	e.g. *going great guns*, *give the Devil his due*
Idiom (partially filled)	e.g. *jog* <someone's> *memory*, *send* <someone> *to the cleaners*
Covariational Conditional	The Xer the Yer (e.g, *the more you think about it, the less you understand*)
Ditransitive	Subj V Obj1 Obj2 (e.g. *he gave her a fish taco; he baked her a muffin*)

Passive	Subj auxVPpp (PPby) (e.g. *the armadillo was hit by a car*)

As Masini (2018) points out, the schematicity of a construction does not correlate with its productivity. To be productive, a construction must possess a certain degree of schematicity, but this parameter alone does not solely determine its productivity. As Masini (2018: 53) says, the productivity of a construction is strictly connected to the variables and the restrictions that operate in it. For example, the construction *jog* <someone's> *memory* is unspecified, but it is not very productive, because the empty slot can only be filled by specific items. Rather, productivity correlates with the *type frequency* of the construction: i.e. to the count of unique specific linguistic items that can instantiate it: the higher the type frequency, the more productive the construction. In usage-based studies the distinction between token frequency and type frequency is central (see on this section 4.2.2 below).

In CxG, constructions are not derived from one another via transformations. Rather, different constructions merge with one another in order to form other constructions. In the *Constructicon*, actual expressions (also called *construct*) typically involve the combination of several different constructions. Consider example 121 from Goldberg (2006: 10): the sentence *What did Liza buy Zach?* involves the list of constructions below.

121. what did Liza buy Zach?
   a. *Liza*, *buy*, *Zach*, *what*, *do* constructions.
   b. Ditransitive construction
   c. Question construction
   d. Subject – Auxiliary inversion construction
   e. VP construction
   f. NP construction

Just like the generative approach, the constructional approach thus aims to explain the inherent creativity of languages: positing that constructions are free to combine with one another allows linguists to explain the limitless creative possibilities of a language. Constructions are combined freely to form actual expressions as long as they are not in conflict: unresolved conflicts result in judgments of ill-formedness. Conflicts may also result in a different meaning of the construction (see for example the opposition between the locative-adessive construction and the possessive construction in Hindi discussed in section 4.2.3).

The *Constructicon* is conceived as a structured net organized according to various types of relations linking the constructions. Goldberg (1995) lists the following four link types:

i. Instance link: a construction (C_2) is the instantiation of another abstract construction (C_1). C_2 is thus more specified than C_1 that is more abstract and more general.
ii. Polysemy link: a construction (C_2) shows a polysemic relation with another construction (C_1).
iii. Metaphorical Extension Link: a construction (C_2) is the metaphorical extension of a construction (C_1).
iv. Subpart Link: some properties of construction (C_2) are derived from another construction (C_1), but C_2 is independent from C_1 and is not its instantiation.

Notions such as polysemy and metaphor refer to semantic relations that are usually associated to the lexical level and not to the syntactic level. However, since in CxG every construction (also the most schematic one) is conceived as a conventionalized union of form and meaning, cognitive-semantic mechanisms such as metaphoric extensions can apply also to constructions. Hilpert (2014) proposes as an example the polysemic network of attributive possessive constructions in English, that can also be applied to the Hindi possessive postposition *kā (/ke/kī)*. When used attributively, the possessive postposition is also used to express other types of relation between two entities, such as kinship and meronymic relations, but also more semantically vague ones such as associative relations, temporal or spatial relations, and so on. A representation of the network generated from the Hindi ATTRIBUTIVE POSSESSIVE construction [NP] *kā/ke/kī* [NP] is given in Figure 9 (on agreement patterns of the genitive postposition see section 3.2.2.1).

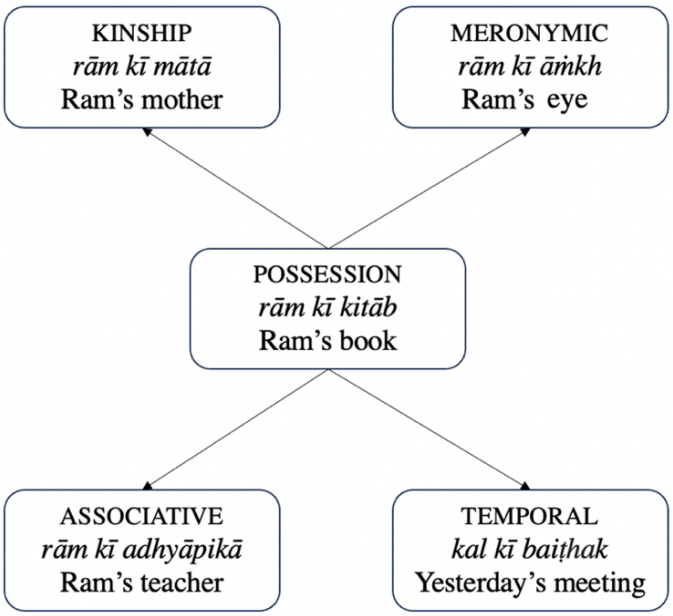

Figure 9: Representation of the network generated from the Hindi attributive possessive construction [NP] -*kā/ke/kī* [NP].

4.2.1. A constructional approach to argument structures

Goldberg (1995, 1999, 2006) employed this theoretical framework to examine argument structures, one of the most general and abstract mechanisms of a language. With her analysis, she offered an alternative perspective to the lexicalist approach (Levin 1993), which placed a strong emphasis on individual verbs as the primary determinants of argument structure. Following this approach, each verb has its own semantics and inherent valency pattern, which specifies the number and type of arguments it requires. These valency patterns are seen as part of the lexical knowledge associated with each verb, and the syntactic structure is derived from the verb's meaning. In contrast, in Goldberg's constructional framework, argument structures bear their own semantics, independently of the lexical elements that instantiate them and they are not licensed by the semantics of the individual verb; rather they contribute to the semantics of the whole construction. Once again, since argument structure constructions bear their own semantics, they are involved in mechanisms such as prototype effects, polysemes, metonymies and metaphors. Examples of English argument structure constructions discussed by Goldberg (2006: 73) include the constructions shown in Table 13:

Table 13: English argument structure constructions adapted from Goldberg (2006: 73).

Form/Example	Meaning	Construction Label
Subj V $Obl_{path/loc}$	X moves $Y_{path/loc}$	**Intransitive Motion**
e.g. The fly buzzed into the room.		
Subj V Obj $Obl_{path/loc}$	X causes Y to move $Z_{path/loc}$	**Caused Motion**
e.g. Pat sneezed the foam off the cappuccino.		
Subj V Obj Obj2	X causes Y to receive Z	**Ditransitive**
e.g. She faxed him a letter.		
Subj V Obj RP	X causes Y to become Z_{state}	**Resultative**
e.g. She kissed him unconscious.		

As already mentioned, one of the basic tenets of CxG is that argument structures are assumed to have meaning independently of the lexical items, and that the meaning of a construction arises from the fusion of the meaning of the lexical items and of the argument structure. The corollary of this is that certain classes of verbs correlate with certain constructions. Another corollary of CxG is that the same verb can have different interpretation according to the argument structure construction it occurs in. According to CxG, the semantics of a sentence thus emerges from the union of the semantics of the verb and the semantics of the argument structure. This interpretation is quite distant from the traditional analysis of valency proposed in Levin's pivotal work on verbal classes in English (1993).

In Levin's view, a verb has more than one meaning and this allows it to occur with different argument structures, since the inherent grammatical valency of a lexical item is stored within the lexical-semantic information in the verbal lexicon. In contrast, from a constructional perspective, the different meanings exhibited by the same verb should not be interpreted as contained in the lexical entry, but as derived from different constructions.

One of the most famous examples proposed by Goldberg (1995) to exemplify this mechanism is given in sentence 122.

122. *Pat sneezed the napkin off the table.*

Either we admit that the English verbal lexicon stores an entry of the verb *sneeze* featuring the semantics of "moving something by sneezing" and requiring a three arguments argument structure, or we explain the meaning of sentences such as 122 as emerging from the fusion of the semantic contribution of the verb and the semantics of the argument structure. The argument structure in question is what Goldberg (1995) defined CAUSED MOTION argument structure and which she describes as in 123.

123. *Pat put the ball on the table.* X CAUSES Y TO MOVE Z SUB V OBJ OBL

The formal representation of the CAUSED MOTION construction given by Goldberg (1995: 52) is proposed in Figure 10.[16]

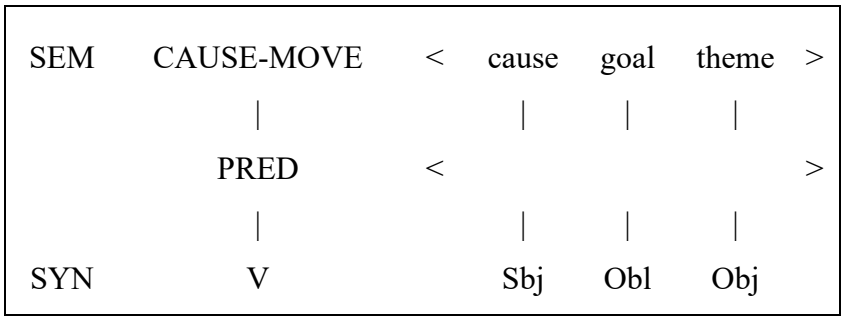

Figure 10: Representation of the CAUSED MOTION construction (adapted from Goldberg 1995: 52).

[16] The formalism in Figure 10 follows the guidelines traced by Goldberg (1995, 2006: 21): Solid lines indicate that the argument role of the construction must fuse with an independently existing participant role of the verb. Dashed lines indicate that the argument role of the construction may be contributed by the construction without corresponding role existing as part of the inherent verbal meaning.

In sentence 123, the CAUSED MOTION construction is instantiated by the verb *put* which generally requires three-arguments: the Agent, the Theme/Patient, and the Location. From a lexicalist perspective, the pattern is licensed by the verb by means of semantic and syntactic information stored in its lexical entry. However, this explanation does not account for cases such as 122, where the verb *sneeze* is intransitive and only requires a single-argument argument structure. According to Goldberg's explanation, the meaning of the construction emerges from the fusion of the CAUSED MOTION argument structure with the verb *sneeze*. This constructional explanation is more elegant because it offers a descriptive framework which allows us to generalize over a vast range of cases in which the same verb in different constructions shows systematic differences in meaning. Moreover, it allows us to generalize over cases in which different verbs in the same construction show similar meanings, such as the verb *boo* in 124 and *sneeze* in 122.

124. *They booed him off stage.*

Thus, in CxG, the interpretation of an expression is derived by integrating the semantics contributed by the argument structure construction with the semantics of the main verb and its various arguments. This amounts to saying that the participant roles profiled by the verb are distinct from the argument roles associated with the construction. Verbs display a richer frame semantic representation than constructions, which are associated with a more abstract semantics. This difference between verbal and constructional meaning derives from the fact that verbs display frame-specific roles (*participant roles*), while constructions refers to event types which are basic to human experience, such as those denoting that *something moved* or *someone experienced something*, etc., and are thus associated with more general roles (*argument roles*) such as agent, patient or goal (Goldberg 1995: 39-43; see also Langacker 1991). From the different semantic frames of verbs and constructions, it follows that systematic differences in meaning associated with the same verb in different constructions can be explained by a particular construction providing its own meaning (Goldberg 2006: 19-20).

According to Goldberg (1995: 50) there are two general principles that can be understood to constrain the ways in which the participant roles of a verb and the argument roles of a construction can merge:

1. The Semantic Coherence Principle, according to which the participant role of the verb and the argument role of the construction must be semantically compatible.
2. The Correspondence Principle, according to which "the profiled participant roles of the verb must be encoded by profiled argument roles of the construction, with the exception that if a

verb has three profiled roles, one can be represented by an unprofiled argument role (and realized as an oblique argument)."[17]

Every argument role profiled by the construction is then linked to a grammatical relation according to the linking generalizations elaborated by Dowty (1991). Dowty distinguishes two macro-roles, Proto Agent and Proto Patient, which he conceives as prototype concepts in the sense elaborated by Rosch and Mervis (1975). He lists a series of properties which define these two macro-roles (given in A) and B)) and then argues that "in predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject; the argument for which the predicate entails the greatest number of Proto-Patient properties will be lexicalized as the direct object" (Dowty 1991: 576).

A) Contributing properties for the Agent Proto-Role (Dowty 1991:572):
    1. volitional involvement in the event or state
    2. sentience (and/or perception)
    3. causing an event or change of state in another participant
    4. movement (relative to the position of another participant)
    5. exists independently of the event named by the verb

B) Contributing properties for the Patient Proto-Role (Dowty 1991:572):
    1. undergoes change of state
    2. incremental theme
    3. causally affected by another participant [roughly "force recipient"—BL]
    4. stationary relative to movement of another participant
    5. does not exist independently of the event, or not at all

In Construction Grammar mechanisms such as transformations, derivations and alternations are avoided: there is no distinction between deep level and surface as in generative syntax, and there are only "surface forms" (i.e. constructions) that create new forms through the process of unification. Constructionists emphasize the fact that generalizations involving surface forms are broader than

[17] The Correspondence Principle derives from the fact that lexical semantics and discourse pragmatics are typically aligned and requires that the semantically salient profiled participant roles are encoded by grammatical relations that provide them a sufficient degree of discourse prominence: i.e. by profiled argument roles (Goldberg 1995: 50).

generalizations captured by transformative mechanisms. Goldberg (2006: 25) sums this concept in what she labels the Surface Generalization Hypothesis.

> SURFACE GENERALIZATION HYPOTHESIS
> "There are typically broader syntactic and semantic generalizations associated with a surface argument structure form than exist between the same surface form and a distinct form that it is hypothesized to be syntactically or semantically derived from."

To sum up, in this work I frame my analysis of the argument structures of Hindi experiential verbs within the theoretical framework of Construction Grammar and in particular I will refer to the following basic tenets.

1. The grammar is interpreted as the repository of all constructions of a language. This repository is called *Constructicon* (Fillmore 1988, Croft 2001, Goldberg 1995, 2006) and it is shaped as a syntax-semantics continuum generated by a network of constructions. Constructions vary in relation to schematicity and complexity: new constructions are generated by the unification of already existing constructions.
2. Constructions are conceived as conventionalized pairings of form and meaning: the basic corollary of this is that constructions bear their own semantics and participate in cognitive-semantic mechanisms such as metaphoric extensions. Moreover, since constructions are meaningful, they can also contribute systematic semantic properties that are not generally associated with the verb they occur in (Perek 2015).
3. Constructions are subject to generalizations, for example they tend to correlate with specific classes of verb (Barðdal 1999, 2008, Perek 2015), in the sense that verbs that are closely related semantically tend to appear in the same argument structure constructions (Goldberg 1995, Pinker 1989, Levin 1993).
4. The overall interpretation of a construction is arrived at by integrating the argument structure construction with the main verb and its various arguments (Goldberg 2006: 38).

I believe that this theoretical approach suits well the aims of my work. One of my main purposes is to show that in Hindi the argument structure constructions of experiential verbs are related to the semantic foundations of the event expressed. Following this approach, I will not view case-markings in isolation but as occurring as part of a construction and I will try to demonstrate that the constructions analyzed in this work bear their own meanings and that certain constructions correlate with certain classes of verbs because of their similar semantics.

4.2.2. Usage-based approach

My interpretation of the functional distributions of constructions follows a usage-based approach. In usage based CxG, the categorization leading to the formation of the grammar of a language is a cognitive process based on experience and usage, and grammar is constituted by abstract patterns and concrete items. Usage-based approaches focus on the importance of grounding linguistic theory in actual instances of language usage and are sensitive to frequency and entrenchment. The frequency in use of concrete items determine the emergence of abstract patters in the speakers' mind (Langacker 1987, Barlow and Kemmer 2000, Bybee and Hopper 2001, Goldberg 2006, Perek 2015). Usage-based models differ from structuralist and generative models in their representation of the grammar of a language. In generative and structuralist approaches, only the structure of grammatical elements determines their representation in the speakers' mind. On the other hand, in usage-based approaches, the usage properties of the grammatical items contribute to the representation. In particular, two notions are central in usage-based approaches: token frequency and type frequency.

- Token frequency is the frequency in use of a specific **item** (e.g. word, constituent, construction).
- Type frequency is the frequency in use of a specific **pattern** (e.g. word+word, verb+construction).

Type frequency and token frequency have an influence on the consolidation of abstract schemes or concrete sequences and a difference in the frequency correlates with a difference in how these items are stored in the grammar (Bybee 2006: 719; see also Bybee and Thompson 1997):

1. A low frequency only leads to conventionalization.
2. A higher frequency can lead to the creation of a new construction.
3. A very high frequency leads to a grammaticalization process.

In my analysis of Hindi experiential constructions, I will follow Barðdal (1999) and Luraghi (2020) and I will define token frequency as the number of times a certain construction occurs with a specific verb, and type frequency as the number of different verbs that occur with a certain construction. A high type frequency of a given construction also correlates with the construction's productivity: argument structure constructions that occur with a large variety of verbs are more likely to be extended to new verbs than constructions that only appear with few verbs (Goldberg 1995, Barðdal 2008, Luraghi 2020a).

Nevertheless, the degree of semantic similarity between the new verbs and verbs already associated with a specific argument structure plays a crucial role in determining the productivity of

the argument structure construction. This means that constructions that are used with a wide variety of verbs tend to be extended more easily than constructions that are used with a semantically restricted set of types (Goldeberg 2006). The consequence of this is an inverse correlation between the productivity of a construction and its semantic coherence: the more a construction is extended to other classes of verbs, the less semantically coherent it is (see on this also Barðdal 2008). Figure 11, adapted from Barðdal (2008: 35) shows the inverse correlation between the type frequency and semantic coherence.

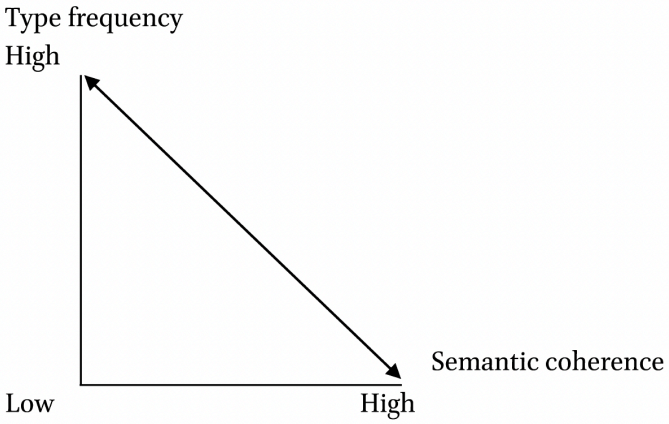

Figure 11: Inverse correlation between the type frequency and semantic coherence (taken from Barðdal 2008: 35).

4.2.3. A constructional approach to Hindi

In this section, I will present two examples of constructions in Hindi. I will first address the relation between the locative-adessive construction and the possessive construction in the languages. I will then discuss the verb *milnā* "meet" and the argument structures it occurs with, in order to show that constructions have their own semantics.

Figure 12 and Figure 13 show the locative adessive construction and the possessive construction respectively. The first construction is used to express proximity between two entities in Hindi, while the second construction is used to encode inalienable possessive relations. These two constructions are very similar. In the adessive construction, the Location is in the oblique case and followed by the postposition *ke pās* "next, near", while the Theme stands in the nominative. The verb *honā* "be" agrees with the Theme and has an existential function (Figure 12).

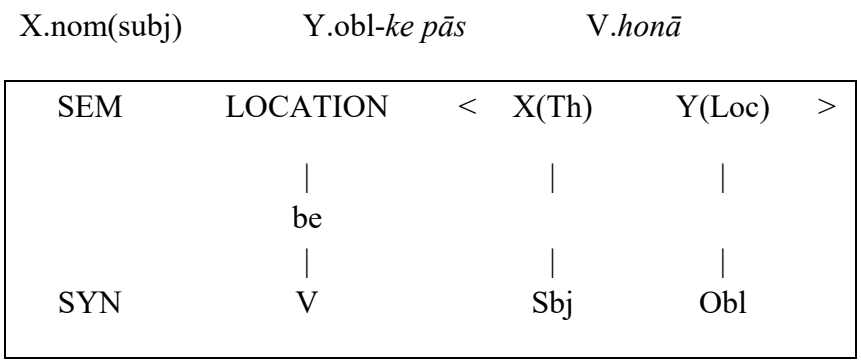

Figure 12: The Hindi locative-adessive construction.

In the possessive construction, the Possessor (Pr) is in initial position, marked in the oblique case and followed by the postposition *ke pās* "beside", next comes the Possessee (Pe), marked in the nominative case. The verb *honā* "to be" agrees with the Pe and has an existential function. Note that the possessive construction shows certain syntactic properties that the adessive construction does not feature. In particular, as I will discuss below in this section, the locative Possessor is a non-nominative subject (Figure 13).

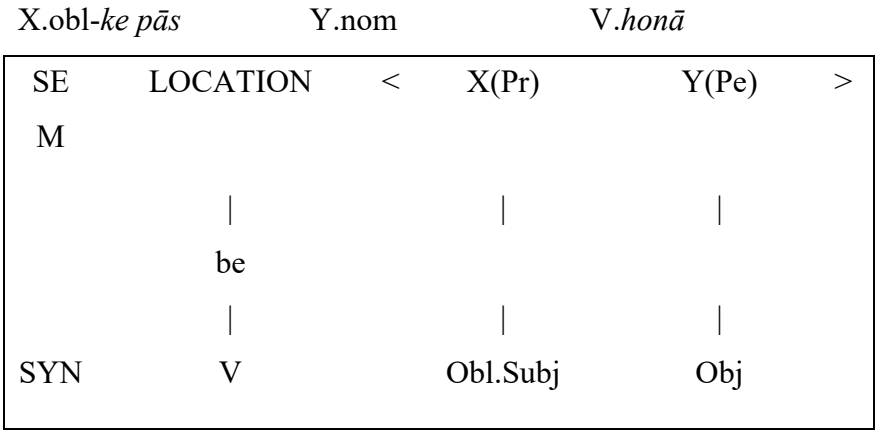

Figure 13: The Hindi possessive construction.

Let us briefly consider the semantic properties of a possessive relation in order to discuss the relation between these two constructions in Hindi. One of the most important properties that defines the semantic prototype of possession is proximity between the Pr and the Pe (Heine 1997, Stassen 2009). Other semantic properties typical of possession are the humanness of the Pr and the inanimateness of the Pe. It is reasonably to hypothesize that in Hindi the possessive construction is linked to the adessive construction through a metaphorical relation. This metaphorical extension led to the adessive construction, which typically expresses proximity, being used to encode the domain of possession in Hindi. This however occurs only when the construction is instantiated by specific semantic properties of the arguments (Carnesale 2022). In particular, when the argument preceding the postposition *ke*

pās is [+HUMAN] and the argument in the nominative is [-ANIMATE], the resulting construction is a possessive construction. This extension is quite common from a typological perspective: as Heine (1997: 75) points out, the locative construction is the most used strategy for the encoding the domain of possession cross-linguistically. This can be explained by means of two underlying metaphors very common typologically: POSSESSORS ARE PLACES and PHYSICAL VICINITY IS CONTROL (Luraghi 2014: 108). Remarkably, in this construction the semantics of possession is not lexicalized in a lexical item, it rather emerges from the instantiation of the locative construction through these specific semantic features. As a consequence, the possessive construction has a distribution which is complementary to that of the adessive construction. If the argument preceding the postposition *ke pās* is not [+HUMAN] and/or the second argument in the nominative is not [-ANIMATE], then the resulting construction has a locative interpretation. In all other cases, the construction encodes spatial proximity. These distributions are summed up in Table 14 and exemplified in sentences (125-128).

Table 14: Complementary functional distribution of the adessive and possessive construction.

		1st ARGUMENT	
		Human	**Non-human**
2st ARGUMENT	**Human**	ADESSIVE	ADESSIVE
	Non-human	POSSESSIVE	ADESSIVE

125. 1° argument: [+ HUMAN]; 2° argument: [-ANIMATE]: Possession
उसके पास आठ सौ रूपये थे।

us=ke pās	*āṭh*	*sau*	*rupay-e*	*th-e*
3SG.OBL=LOC(beside)	one	hundred	rupees(M)PL.NOM	be.PST-M.PL

"He had eight hundred rupees."

126. 1° argument: [- HUMAN]; 2° argument: [-ANIMATE]: Location
उसका घर हमारे घर के पास ही है।

us=kā	*ghar*	*hamāre*	*ghar=ke pās*
3SG.OBL=GEN	house(M.SG.NOM)	1PL.GEN	house(M.SG.OBL)=LOC(beside)

hī	*hai*
EMPH	be.3SG.PRS

"His house is near our house." (from hiTenTen21)

127. 1° argument: [- ANIMATE]; 2° argument: [+HUMAN]: Location
गाड़ी के पास तुम्हारी पत्नी है।

gāṛī=ke pās	*tumhārī*	*patnī*	*hai*
car=LOC(beside)	2PL.GEN	wife(M.SG.NOM)	be.3SG.PRS

"Near the car there is your wife."

128. 1° argument: [+ HUMAN]; 2° argument: [+HUMAN]: Location
गोबर के पास धनिया थी।

gobar=ke pās	*dhaniyā*	*th-ī*
Gobar=LOC(beside)	Dhaniya.NOM	be.PST-F

"Near Gobar, there is Dhaniya."

Some semantic and syntactic features differentiate the possessive construction from the adessive. For example, only in the adessive construction allows the postposition *ke pās* "beside" to alternate with other synonymic postpositions, like *-ke bagal meṁ* or *-ke nikaṭ* "next, near to". The possessive construction does not allow for the use of other postpositions; if another locative postposition is selected, the resulting construction acquires an adessive meaning. See the contrasting examples below (129 and 130).

129.

rām=ke pās	*na-ī*	*kitāb*	*hai*
Ram=LOC(beside)	new-F	book(F.SG.NOM)	be.3SG.PRS

"Ram has a new book."

130.

rām=ke bagal meṁ	*na-ī*	*kitāb*	*hai*
ram=LOC(beside)	new-F	book(F.SG.NOM)	be.3SG.PRS

"Next to Ram there is a new book. *Ram has a new book."

This seems to indicate that the adessive construction instantiated with a [+HUMAN] location and a [-ANIMATE] second argument has undergone a grammaticalization process causing the desemantization of the postposition *ke pās*, which, in this context, has lost its original lexical meaning (see Lehmann 2015). Additionally, in pragmatically unmarked possessive constructions, the Pr is in initial position, whereas in unmarked locative constructions, the Location is preverbal, and the subject is in initial position.

Let us now move on to the discussion of the verb *milnā,* in order to show how constructions contribute their semantics and points toward a specific construal of the event. This verb has the basic

meaning of “be mixed or be united with someone or something” and can variously be interpreted as “receive”, “find”, “meet” according to the context. In 131, the verb occurs in an oblique second argument construction: the first argument is in the nominative case while the second argument is followed by the postposition *se*, expressing here a comitative function. When occurring in this construction, the verb indicates a situation which implies volitionality and control on behalf of the first participant and the meeting event is interpreted as intentional. In 132, the same verb occurs in a dative construction, in which the first argument is encoded in the dative and the second one stands in the nominative. As I will discuss in the following chapters, the dative construction implies a specific semantics in Hindi, i.e. the dative participant must lack control and volitionality. When occurring in this construction, the verb *milnā* expresses an event that is not intentional for the first participant.

131.	*maiṁ*	*apn-ī*	*bahan=se*	*mil-ā*
	1SG=ERG	REFL-F	sister(F.SG.OBL)=COM	meet-PRF.M.SG
	“I met my sister (purposefully/ volitionally).”			

132.	*mujhe*	*apn-ī*	*bahan*	*mil- ī*
	1SG.DAT	REFL-F	sister(F.SG.NOM)	meet-PRF.F.SG
	“I met my sister (not purposefully/volitionally)”			

From a constructional perspective these two meanings are derived from the different argument structure construction that occurs with the verb and are not encoded in the lexical entry of the verb *milnā* as a lexicalist perspective would suggest. The verb expresses an event in which two participants meet each other, while the argument structure contributes by adding information about the volitionality of the first participant: the oblique second argument construction (example 131) implies control by the two participants, while the dative construction (example 132) expresses lack of agentivity and implies that the first participant receives the consequences of the event without intentionally taking part in it.

Another example of construction alternation can be found with the verb *kahnā* “tell” (133 and 134). In sentence 133, the verb occurs in an oblique second argument construction, with the first participant in the nominative and the second argument followed by the postposition *se*, once again marking a comitative function. The content of the communicative event is expressed with an objective clause introduced by the subordinating conjunction *ki*. While in sentence 134 the second argument is not marked with the comitative, but with the dative. These constructions are quite similar, but they express a subtle difference regarding the semantic properties of the second participant. In the first example, s/he participates in the event and is construed as a co-Agent in the communicative situation,

whereas in the second example s/he is construed as a Recipient. This semantic information is not contributed by the verb, but by the argument structure the verb occurs with.[18]

133.	*maiṁ*	*tum=se*	*kah rah-ā*	*hūṁ*	*ki*
	1SG.NOM	2SG=COM	say PRGR.M.SG	be-1SG.PRS	that
	"I am telling you that …"				

134.	*maiṁ*	*tumheṁ*	*kah rah-ā*	*hūṁ*	*ki*
	1SG.NOM	2SG.DAT	say PRGR.M.SG	be-1SG.PRS	that
	"I am telling that …"				

As I will show in Section 4.3, except for the transitive construction, Hindi constructions seem to have a clear semantics and occur only with certain classes of verbs. Moreover, when a verb alternates among different constructions, this alternation is often associated with a change in the semantic properties of the event, such as the properties of the participants. Additionally, some constructions carry aspectual nuances. For instance, in the expression of cognitions, the contrast between the genitive and the dative construction seems to indicate a difference in Aktionsart. As I will discuss in chapter 8, the dative construction may express both a stative and an achievement semantics, and it is used to encode mental states such as "know" or mental achievements such as "understand", while the genitive construction is used to indicate states resulting from actions and it is used to encode the result of mental activities such as "think".

4.2.3.1. A constructional approach to Hindi N-V complex predicates

Recall from section 3.2.5, that Hindi makes extensive use of N + V complex predicates consisting of a nominal host and a light verb that act as a single predicate unit. The semantic core is expressed by the nominal, while the verb supplies information with respect to TAM properties, valency and actionality. In N-V complex predicates, the argument structure is partly determined by the argument structure taken by the light verb, and partly influenced by the semantics of the noun (Mohanan 1994, Butt 1995, Butt and Geuder 2001). As I mentioned in chapter 3.2.5, this category is extremely heterogenous, and the status of the nominal host, and consequently of the whole complex predicate, is often ambiguous. Consider for instance the complex predicate *madad karnā* "help (lit. help do)" in 135. This predicate typically occurs in a construction in which the first participant is marked with the

[18] Note however that, as Andrea Drocco notices (p.c.), the dative marking on the second argument is very common in spoken Hindi, but is not present in standard Hindi.

nominative/ergative, and the second participant is marked with a genitive, as in 135. The nominal host is the direct object of the verb *karnā* "do", as it is indicated by verb agreement in ergative alignments. However, it cannot be considered as a standard direct object as it never occurs with the accusative *ko*. Moreover, the nominal host contributes to the argument structure of the complex predicate. Even if the status of the noun *madad* is ambiguous, it clearly cannot be considered as a normal argument of the verb.

135. मैं तुम्हारी क्या मदद कर सकता हुँ ?

maiṁ	*tumhārī*	*kyā*	*madad*	*kar*	*sak-t-ā*	*hūṁ ?*
1SG.NOM	2PL.GEN	what	help(F.SG.NOM)	do	can-IPRF-M.SG	be.1SG.PRS

"How can I help you?"

As I have mentioned, these mismatches are quite frequent in Hindi complex predicates as well. Some complex predicates show internal agreement (the nominal host agrees with the light verb) and an oblique external argument that is either marked with the genitive postposition *kā* (/*ke*/*-kī*) or with the postposition typically associated to the semantic role expressed by the argument. While other complex predicates display external agreement, as the external argument is treated as the subject (in intransitive predicates) or the direct object (in transitive predicates). This classification is not clear-cut, and some complex predicates may allow for both internal and external agreement.

Compare sentences 136 and 137 below, in which the complex predicate *yād ānā* "remember" appears in two distinct argument structures (for a detailed analysis of the expression of memory in Hindi see 8.7.1). This verb consistently features a dative Experiencer, while the Stimulus can be expressed in two different ways. In 136, the verb *yād ānā* acts as a single lexical unit: the Stimulus *nām* "noun" is the syntactic subject and the verb shows external agreement. In 137, the noun + verb sequence is not fully integrated, and the complex predicate appears in an internal agreement construction: the nominal host is treated as an argument, it is the syntactic subject, and it agrees with the verb; the Stimulus *garīboṁ* "poor" is linked to the nominal host with the genitive case.

136. हां, उसका नाम भी याद आ गया क्रिप या क्रुप...

hāṁ,	*(mujhe)*	*us=kā*	*nām*	*bhī*	*yād*
yes	(1SG.DAT)	3SG.OBL=GEN	name(M.SG.NOM)	too	memory(F.SG.NOM)

ā	*ga-yā*	*krip yā krup…*
come	go.PRF-M.SG	krip or krup

"Yes, I also remember his name, Krip or Krup…"

137. हमारे लिए तो यही बहुत है कि तुम्हें हम गरीबों की याद तो आयी।

hamāre lie to yahī bahut hai ki

For us it means a lot that

tumheṁ	*ham*	*garīb-oṁ=kī*	*yād*	*to*
2PL.DAT	1SG(OBL)	poor(M)-PL.OBL=GEN	memory(F.SG.NOM)	EMPH

ā-yī.

come-PRF.F.SG

"For us it means a lot that you remembered us, the poor."

As I will discuss in more detail in chapter **Errore. L'origine riferimento non è stata trovata.**, these complex predicates are widely used in the language, and their proliferation in the Hindi verbal lexicon was significantly influenced by prolonged contact with Persian (Montaut 2016). Interestingly, similar ambiguous properties are also shown by Persian complex predicates (Goldberg 2003). In Persian, complex predicates consist of a host that may be a noun, an adjective or a preposition, and a light verb. As in Hindi, a large part of the Persian verbal lexicon consists of complex predicates and these formations are highly productive, as the integration of verbs borrowed from other languages happens via complex predicates (Karimi-Doostan 1997).

Goldberg (2003) offers an account of Persian complex predicates and treats them as constructions represented in the lexicon. As it happens in Hindi, complex predicates in Persian show some ambiguities in how they treat their nominal host. In particular, Goldberg (2003: 83) notes that they display a mismatch of lexical and phrasal properties: they act in some ways as a single unit, and in other ways as more than one unit. They act as a single lexical item in the sense that they resist separation (for example by adverbs and arguments), and the host cannot appear with a determiner. Additionally, the argument structure of the complex predicate may differ from the argument structure that the light verb features when used as a simple verb, thus implying that the nominal host may contribute to the argument structure of the whole complex predicate. Compare examples 138 and 139 adapted from Goldberg (2009: 87-88). In 138, the verb "take" occurs as a simple verb, while in 139 it is used as a light verb in the complex predicate *arusi gereftan* "throw a wedding" and appears with a benefactive argument. Moreover, primary stress, which usually falls on the simple verb in Persian, in 139 falls on the nominal host.

138. *ketâb râ az man gereft*

book ACC from me took

"She/He took the book from me."

139. *barâye* *u* *arusi* *gereftam*
for her/him weeding took
"I threw a wedding for her/him."

Goldberg (2003) represents the construction involving complex predicate as in Figure 14. The formalization is intended to simultaneously represent the internal constituents and the external status of the construction. The external syntax of the complex predicate is that of a single lexical item, but the internal syntax includes two lexemes: a host (represented with the variable X0) and a simple verb (V0). The fact that the host precedes the verb is represented by the symbol <. This representation also works for the Hindi complex predicate.

Figure 14: The complex predicate construction in Persian (adapted from Goldberg 2003: 90).

As Goldberg (2003: 99) notes, "the preference for treating the complex predicate as a single syntactically integrated predicate is motivated by its status as a semantically integrated predicate. This can be seen to be a special case of a general iconic principle: namely a tight semantic bond between items tends to be represented by a correspondingly tight syntactic bond". However, this alignment between semantics and syntax does not always hold and complex predicates' constructions, in Persian as in Hindi, are sometimes subject to mismatches: the host and the light verb may act as two lexically distinct units. For example, Goldberg notes that the elements in a Persian complex predicate may be separated by a number of elements, including the auxiliary of the future, negative and imperfective prefixes, and direct object clitcs (2003: 91).

In established literature on Hindi complex predicates (Alsina 1996, Mohanan 1994, Butt 1995) complex predication occurs when two or more elements enter into a relationship of co-predication: meaning that each element contributes arguments to the argument structure. In this view, when two argument structures are combined, individual arguments contributed by the nominal host and by the light verb can be identified with one another and this leads to complex predication. Typically, the argument contributed by the nominal host combines with the most semantically similar argument contributed by the light verb. Consider for example the experiential complex predicate *yād ānā*

exemplified in (136). In this complex predicate, the argument structures of the two predicational elements in a) and b) are combined and give rise to the composed argument structure given in Figure 15.

a) *yād* "memory" is an experiential noun and licenses an Experiencer and a Stimulus.
b) *ānā* "come" licenses a Theme and a Goal.

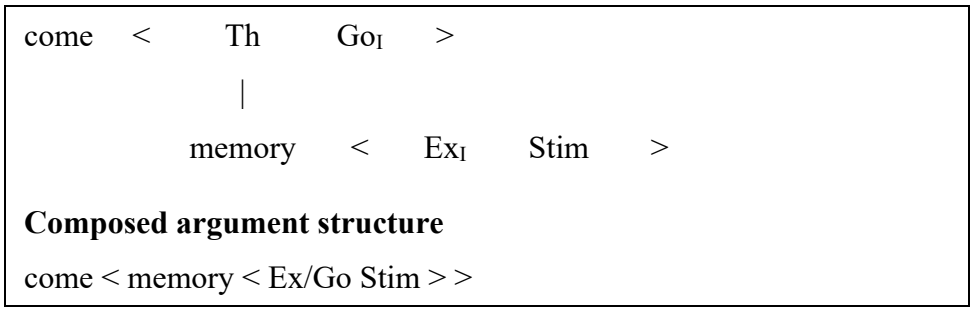

Figure 15: Composed argument structure of the complex predicate *yād ānā.*

When the nominal host is fully integrated with the light verb, the resulting argument structure construction is that shown in Figure 16 in which the noun *yād* is linked at the syntactic level to PRED, indicating that the verb and the noun are in a relationship of copredication (as in 136 above). However, when the nominal host is not fully integrated with the verb, then it does not result in a complex predication and the argument structure is as given in Figure 17, in which the nominal host memory is not linked to PRED, but to the syntactic function Obj, and the Stimulus is its nominal modifier (as in 137 above).

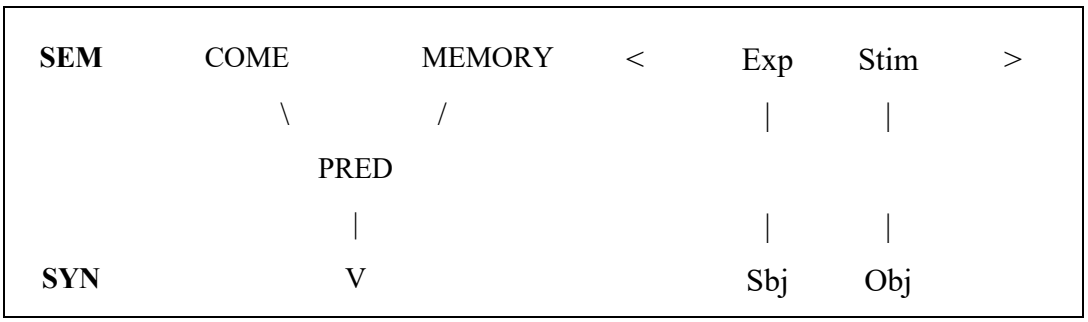

Figure 16: Argument structure construction of fully integrated complex predicate: example for *yād ānā.*

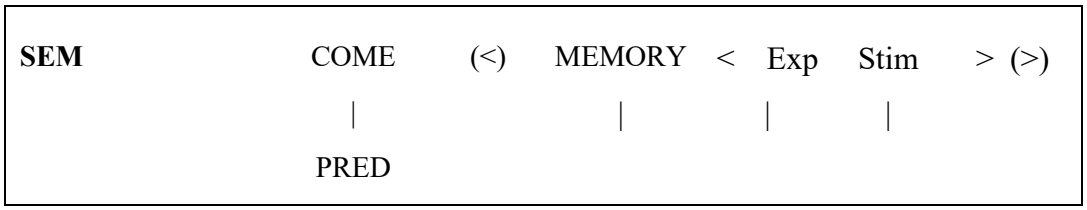

Figure 17: Argument structure construction of non-integrated complex predicates: example for *yād ānā.*

How should we address such alternations in a Construction Grammar (CxG) analysis aimed at exploring the underlying semantic motivations for using different constructions? I argue that in cases such as sentence 137, the genitive marking of the Stimulus is not connected to a different construal of the event and does not imply a different semantics. It is not the result of some specific semantic properties that trigger a genitive marking on the Stimulus, it is rather derived from the fact that the nominal host and the verb are not integrated, and the noun *yād* "memory" is treated as an argument and must be taken as an independent NP. In other words, the Stimulus NP takes a genitive marking because it is not the argument of a verb, but the argument of a noun: the genitive is the canonical marking of arguments of nouns in the language.

Complex predicates origin when independent syntactic forms begin showing properties such as joint meaning and a composite argument structure (Ackerman and LeSourd 1996). The atypical behaviors of Hindi complex predicates that I have discussed in this section might be motivated by the origin of these predicates: verbs occurring as light verbs in Hindi complex predicates may also appear as simple verbs in the language and they occur as simple verbs quite frequently. As mentioned above, complex predicates are new in the grammar of the language, and they were introduced after a long contact with Persian, during the 14th-16th century (Montaut 2016). This means that while simple verb constructions are well entrenched in the grammar, complex predicates' constructions are still a recent innovation. It is reasonable to hypothesize that originally these noun + verb complex predicates were interpreted as the sequence of two distinct units in the canonical SOV order. This interpretation was also supported by the fact that the verbs that were used in the complex predicates had a full semantics and were already used very frequently as main verbs in the grammar. The evolution from two lexically distinct elements to complex predicates was probably a consequence of the fact that, in this sequence, the nominal host and the verb are semantically very close. Their emerging syntactic integration can be seen as the result of the principle of iconicity (Givón 1991): entities tightly connected on the semantic level tend to be coded as tightly connected at the syntactic level. In a similar way, Goldberg (2003; 2006: 103) suggests that the strong preference in Persian for treating complex predicates as single syntactically integrated predicates is motivated iconically by the complex predicates' status as a semantically integrated predicate. For these reasons, in the following chapters, I will not treat alternations involving a genitive marking of the Stimulus in experiential complex predicates as

involving a different semantics in the construal of the event, but as an indication of the fact that the noun + verb sequence is not treated as a complex predication.

4.3. Transitivity marginality and iconicity in Hindi

Many typological studies investigated the use of the transitive pattern across languages in relation to the semantic class of the verbs it occurs with and pointed out that there seems to be a consistent cross-linguistic tendency to avoid the transitive frame for the encoding of certain classes of verbs (see Tsunoda 1981, 1985; Malchukov 2002, 2015; Haspelmath 2015). Tsunoda (1981, 1985) for example is one of the first study to investigate the extent to which a set of languages use the transitive frame to encode non-prototypical transitive events. As result of his investigation, the scholar draws a verb-type hierarchy (Table 15) illustrating the cross-linguistic tendency of specific semantic verb classes to be encoded with the transitive coding frame. The Implicational Hierarchy allows to predict the syntactic coding of certain classes of verbs according to their semantic properties. It states that, cross-linguistically, verbs expressing actions with a direct effect on Patients tend to be coded with a transitive frame, and that the more distant the semantic class is form the direct effect prototype in the hierarchy, the less probable it is that this class is encoded with the same transitive frame. For example, emotion verbs are less frequently encoded with transitive frames than perception verbs. This robust cross-linguistic tendencies highlight the intricate relationship between verb semantics and argument structure constructions.

Table 15: Tsunoda's Implicational Hierarchy of Transitivity (2015: 1598).

Type	1	2	3	4	5	6	7
Meaning	Direct effect on patient	Perception	Pursuit	Knowledge	Feeling	Relationship	Ability
Examples			search, wait, await	know, understand, remember, forget	love, like, want, need, fond, fear, afraid, angry, proud, boast	possess, have, lack, lacking, resemble, similar, correspond, consist	capable, proficient, good

Later, Malchukov (2005) suggested to split the hierarchy proposed by Tsunoda into a two-dimensional map, arguing that verbs can deviate from the transitive prototype according to two semantic hierarchies: one showing decreased agentivity and one showing decreased patienthood. The two-dimensional transitivity hierarchy elaborated by Malchukov (2005, 2015) is represented in Figure 18.

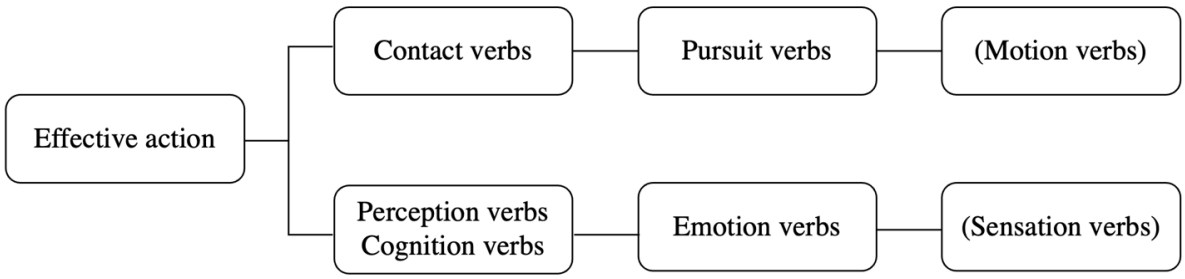

Figure 18: Two-dimensional transitivity hierarchy by Malchukov (2015: 82).

In the upper hierarchy, the deviation path from the effective action prototype (verbs of direct effect in Tsunoda's table) features the loss of the properties of patienthood. In effective verbs the second argument is affected by the event and undergoes a physical change. In contact verbs the second argument is not physically affected but it is still reached at by the Agent and there is contact between the two participants (such as for the English verb *hit*). In pursuit verbs contact between the two arguments is not involved (such as the English verb *look for*). While in most motion verbs the second argument is not even present, as this class requires only a single agentive argument (such as the English verb *run*).

In the lower hierarchy, the deviation path from the effective prototype features the loss of the agentive properties: the first argument gradually loses properties such as control and volitionality. The lower hierarchy is composed by the class of verbs analyzed in this dissertation (experiential verbs) and shows how variable this verb class is. Some experiential verbs are closer to the transitive prototype than others. For example, as mentioned in 2.1, verbs of cognition and verbs of perception may show some degree of agentivity on behalf of the Experiencer, while verbs of bodily sensation never do it. Moreover, as for the case of the last verb class of the upper hierarchy, i.e. motion verbs, bodily sensation verbs do not have a second argument and only require a one-place argument structure with the Experiencer as the only argument. Notably, however, there is an important difference between the upper and the lower hierarchy: while the upper hierarchy only displays the loss of patienthood features, preserving agenthood properties, the lower hierarchy is characterized by the loss of the control and volitionality of the Agent, as well as by the loss of the properties of the Patient. In verbs of perception, for example, the second argument, i.e. the Stimulus, is not physically affected by the event, even if the Experiencer shows some degree of agentivity. As I mentioned in section1.1, dedicated to the final discussion, the semantic variability of verbs of Experience makes this class of verbs particularly intriguing for an analysis focused on the interaction between semantics and syntax within a language.

Languages of the world differ in the way they encode these two semantic deviations from the transitive prototype; however, when languages do not extend the transitive construction, two main tendencies are found cross-linguistically. Verbs in the upper hierarchy tend to occur in constructions in which the first argument is marked as an Agent, while the second argument is marked with a case distinct from that used for Patients. Verbs in the lower hierarchy, instead, shows a higher frequency to occur in constructions in which the first argument is not marked as an Agent, but with the oblique case typically used to mark Recipients or Beneficiaries. Haspelmath (2001) notes that this tendency is particularly relevant with verbs of emotion in Standard Average European languages. On the other hand, Onishi (2001) argues that verbs of sensation lend themselves more easily to non-canonical subject constructions than other experiential verbs. However, I would like to point out that this is not always the case and that, in many languages of the world, verbs in the lower hierarchy may occur in constructions in which the first argument is marked as an Agent and the second argument is marked with cases distinct from that reserved to Patients (Aldai and Wichmann 2018, Luraghi 2020a: 22-23).

Another interesting typological insight on transitivity is given by Haspelmath (2015). Haspelmath proposes the first quantitative study to investigate how languages differ in relation to their transitivity prominence: i.e. the extent to which a language employs the transitive coding. His study is based on a sample of 35 languages and 80 verb meanings and relies on the database of the ValPal project, a broader project which aim is the typological investigation of argument structure properties of verbs belonging to different valency classes (Hartmann et al. 2013). In Haspelmath's study, transitivity prominence is measured by the percentage of verbs that occur with a transitive construction in a given language (Haspelmath, 2015: 139). In languages with high transitivity prominence, verbs deviating from the semantic transitive prototype tend to be encoded with the same transitive frame, whereas languages displaying a low transitivity prominence tend to marginalize the transitive coding frame to verb classes semantically close to the transitive prototype.

In this dissertation, I follow Haspelmath (2015) and Croft (2022) and I define the transitive construction as the coding-frame used in Hindi to encode the prototypical transitive event. The prototypical transitive event involves "a change of state in which an external volitional agent brings about a change in a patient, such that the patient enters a resulting state" (Croft 2022: 183). In other words, prototypical transitive events involve dynamic processes performed by an Agent starting and controlling the event that have consequences on a Patient physically affected by the action, as in 140 (see on this also Tsunoda 1981, 1985; Kittilä 2002). These are the verbs that are referred to as verbs of *direct effect on Patient* in Tsunoda's Transitivity Hierarchy, and as verbs of *effective action* in Malchukov's Two-Dimensional Hierarchy. Verbs expressing these agentive change-of-state events include, for example, verbs of breaking and verbs expressing changes in a physical property of the

Patient, such as *dry*. Haspelmath (2015: 138), for example, selects the verb "break" as "the yardstick by which to define transitivity".

140. *bacc-e=ne* *khidkī* *tod-ī*
child-M.SG.OBL=ERG window(F.SG.NOM) break-PRF.F.SG
"The child broke the window."

Recall from 3.2.4 that Hindi is a split ergative language with DOM. This means that in this language the transitive construction can be instantiated by four different coding-frames according to the combination of these two parameters. The four instantiations of the transitive frame in Hindi are reproposed in Table 16.

Table 16: Coding-frames instantiating the transitive construction in Hindi.

CODING FRAME	Ergative	Accusative
1-nom 2-nom V.subj[1]	×	×
1-erg 2-nom V.subj[2]	✓	×
1-nom 2-acc V.subj[1]	×	✓
1-erg 2-acc V[3SG.M]	✓	✓

The definition proposed by Haspelmath (2015) and Croft (2022) has the purpose to identify the transitive pattern as a comparative concept, in order to formulate typological generalizations on transitivity and transitivity-prominence across languages. However, it is quite useful in this study as well, because it allows to generalize over the specific instantiations of a Hindi transitive verb without selecting a specific coding frame. As mentioned in 3.2.4, in Hindi DOM and ergativity are not verb-sensitive and they occur with any transitive verb. For this reason, I will disregard variation in the coding-frame driven by the aspectual properties of the predicate and/or the properties of participants, because they do not depend on the verb and its semantics. I will consider the four patterns shown in Table 16 as instantiations of the same transitive frame.

In Hindi the transitive pattern does not only encode prototypical transitive events, but it may be used to encode verbs belonging to other semantic classes deviating from the transitive prototype in different ways. For example, pursuit verbs such as "search" may be encoded by a transitive pattern even though they deviate from the transitive prototype: the Patient of these verbs is not affected and does not undergo a change of state (see example 141 below).

141. वह उसी दिन से तुम्हें खोज रहे हैं।

vah *us=ī* *din=se* *tumheṁ* *khoj*
3SG.NOM that.OBL=EMPH day(M.SG.OBL)=INS 2PL.ACC search(F.SG.NOM)
rah-e *haiṁ*
stay-PRF.M.PL be.PRS.3PL
"He has been searching for you since that day."

In Indo-European languages, it is not uncommon for the transitive construction to be extended to the encoding of non-semantically transitive situations: two-place verbs encoding events which involves two participants, one of which sharing the properties of the Agent (mainly humanness), usually require a transitive construction (Luraghi 2020a). Hence, non-semantically transitive events are frequently expressed by syntactic transitivity. However, as Montaut (2004a) remarks, the transitive construction is marginal in Hindi. Besides this construction, Hindi exhibits various syntactic patterns that seem to be selected on the basis of semantic parameters. This peculiarity led Montaut (2004a, 2013) to define these patterns as semantic alignments (following the definition given in Donohue and Wichmann 2008) rather than syntactic ones. As I introduced in 3.2.4.3, in these patterns the less salient entity is encoded in the nominative and conceptualized as the starting point of the event, while the most salient argument is encoded by the case that is typically associated to the semantic role expressed by the argument.

Literature on case (Comrie 1989, de Hoop and Narasimhan 2005, Malchukov 2005, 2015, Butt 2006) generally distinguishes two main functions of case-marking: the so-called indexing and disambiguating function. The indexing function uses cases to express semantic roles (or specific semantic features of the argument), while the disambiguating function uses cases mostly or exclusively to mark core arguments and express grammatical relations. Following this distinction, Malchukov (2015) proposes two typological tendencies determining case-marking cross-linguistically:

a) Iconicity, which implies the "choice of the most semantically fitting frame" (Malchukov 2015: 85) when encoding semantic roles, thus favoring the indexing function.
b) Markedness, which implies the "choice of the transitive frame as a major default pattern" (Malchukov 2015: 85) for the expression of most events, thus favoring the distinguishing function.

The world's languages vary in the way they rank these two parameters. Languages that rank Iconicity over Markedness are more concerned with the faithful encoding of the semantic features of their

arguments: these languages are typically low in transitivity prominence and do not extend the use of transitive constructions to non-transitive events, because in such languages transitive constructions are typically constrained to prototypical transitivity. In contrast, languages that favor Markedness over Iconicity are more concerned with differentiating the main grammatical relations (subject and object) from peripheral elements, and therefore tend to use transitive patterns by default, regardless of the semantic properties of the event.

Hindi is relatively low in transitivity prominence and fits the definition of iconic language. A clear example of Hindi iconicity is the expression of possessive relations. In this language, possessive relations are encoded through different construction types depending on the semantic properties of the possession: inalienable possessions are expressed through a genitive construction (example 143), while alienable and temporary possessions are expressed through the adessive construction, already discussed in section 4.2.3 (example 142; Carnesale 2022, see also Sulger 2015).

142. "We have no straw to sell."
[HUMAN – INANIMATE]: [Ownership] → Locative construction
हमारे पास बेचने को भूसा नहीं है।

hamāre pās	*becn-e=ko*	*bhūsā*	*nahīṁ*	*hai*
1PL.LOC(beside)	sell-INF.OBL=ACC	straw(M.SG.NOM)	not	be.3SG.PRS

143. "He had three daughters."
[HUMAN – HUMAN]: [Inalienable possession: Kinship] → Genitive construction
उनकी तीन लड़कियाँ थीं।

un=k-ī	*tīn*	*laṛk-iyāṁ*	*th-īṁ*
3PL.OBL=GEN-F	three	daughter(F)-PL.NOM	be.PST-3PL.F

Other examples of Hindi iconicity are the instrumental and the dative constructions as opposed to the transitive pattern. The instrumental pattern is typically used with an involuntary or inefficient Agent (Pandharipande 1979, Mohanan 1994, Montaut 2004); while the dative construction always implies a human participant that receives the consequences of the event and/or is affected by it. The transitive pattern in this language is generally associated to agentive events with a prototypical Agent volitionally acting and controlling the scene. There are many contrastive examples available showing that when the Agent lacks these semantic properties, other patterns are available. Consider example 144, taken from Montuat (2004a: 211), which was already discussed in 3.2.2.1 and is repeated here. In the first sentence the Agent is marked with the ergative, the speaker here believes that the

interlocutor purposefully killed a person; in the second sentence, the Agent is marked with the instrumental: the speaker is saying that it was an accident.

144. Instrumental Agent vs Ergative Agent:

A:	*tum-hīṁ=ne*	*us-kā*	*khūn*		*ki-yā*
	2SG-EMPH=ERG	3SG.OBL-GEN	blood(M.SG.NOM)		do.PRF-M.SG

A: "It's you who murdered him."

B:	*sāhab maiṁ=ne*	*us-kā*	*khūn*	*nahīṁ*	*ki-yā,*
	sir 1SG=ERG	3SG.OBL-GEN	blood	not	do.PRF-M.SG
	mujh=se	*ho ga-yā*			
	1SG.OBL=INS	be go.PRF-M.SG			

B: "Sir, I did not kill him, it happened by myself (I did it unconsciously)."

As I will discuss in section 7.3.3, similar examples are found also in the domain of experiential events: for instance, when the perceiver of a visual or auditory perception is agentive and controls the perception, the transitive pattern is selected, and the perceiver is encoded as the Agent (in the nominative/ergative). Conversely, when the perception is not controlled by the perceiver the dative pattern becomes available, the choice thus evolves from the semantic parameters of the event. Remarkably, the transitive pattern may often be used as a default construction and can also encode non prototypically agentive events. However, while the transitive pattern can be extended regardless of the semantic properties and used to encode almost any situation, all other Hindi patterns cannot be used to express events lacking the semantic properties required.

Another interesting issue regarding the semantic-syntactic interplay in Hindi is the use of ergative marking. Many scholars pointed out that ergativity in Hindi is sensitive to both syntactic and semantic features. Indeed, even if the ergative construction in Hindi seems to be partly triggered by syntactic features (transitive verbs in perfective aspect), in many cases, ergative marking seems to be associated to agentive arguments (see among others Mohanan 1994, Butt and King 2002, De Hoop and Narashiman 2005).[19] For example, the default use of ergative marking only occurs when the ergative is used in a default transitive construction. However, intransitive patterns using ergative marking of the most salient participant frequently occur in Hindi and in these cases the ergative clearly expresses a high level of agentivity. For example, the single argument of many "body emission" verbs

[19] For a thorough investigation upon differential subject marking and indexing function in Hindi the reader can refer to de Hoop and Narasimhan (2005), Mohanan (1994). For a general discussion of case theory see Butt (2006). For a detailed overview of the interaction between ergativity and semantic transitivity in Hindi see Drocco (2008).

can optionally be marked with the ergative case. When this happens, ergative case-marking encodes a more Agent-like argument: volitional and in control of the event, as the contrasting examples 145 and 146 show.[20]

145. *laṛk-ī* *chīllā-ī*
girl-F.SG.NOM/OBL scream-PRF.F.SG
"The girl screamed."

146. *laṛk-ī=ne* *chīllā-yā*
girl-F.SG.OBL/NOM=ERG scream-PRF.M.SG
"The girl screamed (purposefully)."

As I hope I have demonstrated, there are valid reasons to consider Hindi as a language ranking iconicity over markedness. As said above in this section, one of the main claims of the constructional approach is that a construction bears its own semantics and a corollary of this is that certain constructions tend to correlate with certain classes of verbs: usually the classes whose verb meanings better align with the semantics of the construction. The assumption that constructions have their own semantics allows us to interpret the distribution of the pattern analyzed in this study across the exceptional variability of experiential events. In the following chapters, I will discuss the distribution of the constructions used in Hindi to encode experiential events and show how they correlate with specific semantic verb classes. To sum up the research question of the present work and the assumption it derives from are:

- ASSUMPTION: Hindi is a highly iconic language and constructions in Hindi are semantically constrained (except for the transitive pattern).
- HYPOTHESIS: the analysis of the functional distribution of different constructions across the experiential domain will show that constructions in Hindi are specialized for encoding of specific semantic properties and correlate with specific classes of verbs.

[20] This is a classical instance of the unergativity verb class first identified by Perlmutter in his seminal paper on the Unaccusative Hypothesis (Perlmutter 1978).

5. Methodology

This work is a corpus-based study on argument structure constructions of experiential verbs in Hindi. In the last decades it became clear that the investigation of a given linguistic phenomenon should not rely on a random selection of instances that the linguist happens to come across while searching the corpus for specific examples. Corpus linguistics offers a data-based approach to language analysis and provides a solid foundation for quantitative research, based on the systematic analysis of the distribution of linguistic phenomena across a corpus (Stefanowitsch 2020: 55). The use of corpora is supposed to provide representative samples of linguistic phenomena within a language. Hence, the most effective approach to linguistic investigation is to retrain a comprehensive sample of the phenomenon and then systematically analyze it. In corpus-based studies, the corpus is considered to be a valuable source of evidence for the analysis of linguistic phenomena, and it is described as a large body of linguistic data composed of samples of language use with specific properties. First, the instances of language use are authentic, which implies that the language samples contained in the corpus were produced for the purpose of communication, and not for linguistic analysis (Sinclair 1996). Moreover, the collection of examples is conceived to be representative of the language under investigation (Stefanowitsch 2020: 22-23).

As I mentioned in 1.1, one of the main aims of this study is to investigate the extent of the semantic-syntactic interplay in Hindi. One of the main theoretical tenets in this work is that given that constructions have their own semantics they tend to correlate with specific classes of verbs. This concept is based on the Distributional Hypothesis firstly introduced by Harris in the 70s. Harris (1970: 785) noted that "If we consider words or morphemes A and B to be more different in meaning than A and C, then we will often find that the distributions of A and B are more different that the distributions of A and C." This means that differences in the meaning of two elements correlate with differences in their distribution. The implications of the distributional hypothesis have led linguists to explore a corpus-based method known as the study of association phenomena (Gries 2003, Gries and Stefanowitsch 2006). These studies investigate into the tendencies of linguistic elements, such as morphemes, words, and syntactic constructions, to either co-occur or avoid co-occurrence. Building on these theoretical preliminaries, my intent is to identify which verbs tend to occur with a specific construction in order to establish the semantic coherence of the constructions used in Hindi to encode experiential events. In the following chapters, I analyze the Hindi encoding of experiences from an onomasiological point of view. Starting from the verbs that lexicalize the three semantic subdomains under investigation, I search for their constructions in the corpus. Since many experiential verbs (such as *dekhnā* "see, look at" or *socnā* "think") are highly frequent and occur thousands of times, a manual scrutiny of all their occurrences would have been extremely time-consuming, for this reason I decided

to select a random set of maximum 200 occurrences to analyse for each verb. Many times, the absolute frequency of the verb was higher than 200 occurrences, in such cases I did a manual scrutiny of a random sample of 200 occurrences. However, some verbs show an absolute frequency of less than 200, in such cases I scrutinized all the occurrences in the corpus (see the discussion on data saturation and sample representativeness in section 5.1.5 below). In chapter **Errore. L'origine riferimento non è stata trovata.**, dedicated to the final discussion, I draw my conclusions presenting the results from a semasiological perspective and I address the semantics of the constructions analysed in chapters 6, 7 and 8.

Table 17, Table 18 and Table 19 present the verbs I analyzed for this study for the three experiential subdomains: bodily sensations, perceptions, and cognitions. Each table lists the basic meanings of the Hindi verbs, their absolute frequencies in the corpus, and the number of occurrences that I manually scrutinized and annotated. It is important to note that when the absolute frequency exceeds 200 occurrences, the corresponding data in the corpus might be unreliable. As in such cases, I did not scrutinize all the occurrences of the verb in the corpus, the data could include instances of deverbal nouns. Therefore, the data on absolute frequency should be considered as an indicative value.

Table 17: Verbs of bodily sensation.

Meaning	MSH verb	Frequency in Corpus	Manual scrutiny
Be hungry	*bhūkh honā*	3	3
	bhūkh lagnā	24	24
	bhūkhā honā	33	33
Be thirsty	*pyās honā*	9	9
	pyās lagnā	21	21
	pyāsā honā	25	25
Be hot	*garmī honā*	1	1
	garmī lagnā	5	5
Be cold	*ṭhaṇḍ honā*	-	-
	ṭhaṇḍ lagnā	7	7
Be in pain	*dard honā*	75	75

Table 18: Verbs of perception.

Meaning	MSH verb	Frequency in Corpus	Manual scrutiny
appear/see	*dikhnā*	11	11
	dikhāī denā	197	197
	dikhāī paṛnā	3	3
	najar ānā	257	257
see/look at	*dekhnā*	4317	200
show	*dikhānā*	458	200
look at/stare	*tāknā*	230	230
be heard	*sunāī denā*	53	53
be heard	*sunāī paṛnā*	6	6
hear/listen	*sunnā*	902	200
touch	*chūnā*	-	-
	sparś karnā	-	-
taste	*cakhnā*	-	-
	svād lenā	-	-
	svād milnā	-	-
smell	*khuśbū ānā*	-	-
touch/smell/taste	*lagnā*	-	-

Table 19: Verbs of cognition.

Meaning	MSH verb	Frequency in Corpus[21]	Manual Scrutiny
Think	*socnā*	873	200
	sūjhnā	158	158
	vicār + light verbs	293	200

[21] When the absolute frequency is > 200, the data of the frequency in the corpus might be unreliable: I have not scrutinized all occurrences, so there might be cases of deverbal nouns, or participles with a subordinating adverbial function, etc. This number should be considered as an indicative value.

	khayāl + light verbs	64	64
Understand	*samajhnā*	1665	200
	samjhānā	464	200
	samajh ānā (1)	145	145
	samajh ānā (2)	18	18
Know	*jānnā*	1749	200
	jān paṛnā	234	200
	jñāt honā	59	59
	malūm honā	1527	200
	patā + light verbs	277	200
Forget	*bhūlnā*	351	200
Remember	*yād* + light verbs	694	200

5.1. The corpus

In this section I present the two corpora that I used to extract the linguistic data I analyzed for this dissertation. The main source is a corpus I collected specifically for this work consisting of literary texts of the 20th century (section5.1.2). On few occasions, when the data from the literary corpus was too scarce (for example in the case of verbs expressing perception through smell and taste), I relied on data taken from the hiTenTen corpus available on SketchEngine[22] (section 5.1.3). Since the literary corpus constitutes my primary and default source, when presenting examples from it I do not provide the reference to the source. While when using examples extracted from the hiTenTen corpus, I explicitly specify this beneath the sentence.

5.1.2. The literary corpus

My corpus consists of literary texts from the 20th century. I believe that a literary corpus is the most effective and convenient means to extract experiential constructions. Novels and short stories generally deal with human characters and how they emotionally and psychologically react to what happens to them. Additionally, literary texts are often deeply embedded in the cultural and historical contexts in which they are written and can provide valuable insights into how an experiential event

[22] https://www.sketchengine.eu/e

is conceptualized in a given community. I am aware that this study would benefit if the analysis of the corpus was complemented by the analysis of various more contemporary uses of modern Hindi. However, the decision to limit the study to literary Hindi was driven by the necessity to analyze texts with a well-defined, standardized, and homogeneous language, in order to avoid the introduction of sociolinguistic varieties and features that could have been too complicated to manage effectively. Table 20 shows the structure of the corpus used for the present study and gives information on the literary works composing it and their size in terms of tokens. The corpus was specifically collected for the purpose of this study, and it offers an adequate and manageable number of instances.

Table 20: Corpus structure and size.

	Author	Title	Tokens
Novels	Premchand (1880-1936)	*Alankār*	61.378
		Godān	167.429
		Gaban	109.878
		Karmabhūmi	129.637
	Dharmavir Bharti (1926-1997)	*Sūrāj kā sātvāṁ ghoṛā*	24.878
		Gunāhoṁ kā devtā	100.620
	Mohan Rakesh (1925-1972)	*Aṁdhere band kamre*	101.127
	Jaishankar Prasad (1889-1937)	*Kaṁkāl*	67.742
Short stories	Premchand (1880-1936)		420.000
	Raghuveer Sahay (1929-1990)		18.022
	Mahadevi Varma (1907-1987)		5.674
	Amritlal Nagar (1916-1990)		18.052
TOTAL: Tokens: 1,224,437; Words: 987.787; Sentences: 78.054 (from SketchEngine).			

As shown in Table 20, Premchand's production constitutes the bulk of the corpus. In the following, I would like to focus on the reasons that led me to choose him as my main source. Premchand, whose real name was Dhanpat Rai Srivastava, was one of the most significant authors of 20th-century, as he is usually considered the "father of modern Hindi/Urdu literature" (Chaudhuri 2004:133). He is considered as the first and main exponent of social realism in Hindi literature, as his works, which include novels, short stories and essays, often depict the struggles of the lower classes and the

injustices of the caste system in pre-independence India. Some of his most famous works included in my corpus are *Godān*, *Gaban*, *Karmabhūmi*, *Alankār*, and numerous short stories, in which he addresses complex social issues such as the problems of rural life, of the impoverished peasantry and of widowhood in North India. Premchand's use of Hindi is characterized by a high degree of simplicity and realism, as he opted for a language that mirrored the way people speak in real life, making his stories relatable and accessible to a broad audience.

Premchand was also deeply involved in the linguistic policies of his time, and he took a stance regarding the type of language that was to be chosen as the *rāṣṭrabhāṣā* (i.e. national language). His position stands distinctly apart from that of his contemporary nationalist extremists that aimed at a completely Sanskritized Hindi, which resulted in a variety distant from spoken language (see 3.1). Instead, Premchand aligned more with Gandhi, who believed that, in order to break free from the political, cultural and psychological dominion of the British, Indians had to reject the English language from the institutional context and replace it with Hindi (Lelyveld 2001: 71). The difference between Gandhi and the extremist nationalists lies in the fact that for him the term Hindi did not denote the extremists' *śuddh hindī*, the language "purified" from any Persian influence, but rather the language spoken by both Hindus and Muslims of Northern India, written in the Devanagari or the Persian script interchangeably.

Premchand was one of the strongest supporters of Gandhi's linguistic policies; in his novels and short stories, he used the Hindustani as outlined by Gandhi and, as Chaudhuri (2004: 133) notes, he "indignantly rejected the popular myth that Urdu was for Muslims while Hindi was for Hindus". According to Premchand, the future national language should transcend the literary, religious, or writing-related divisions and to represent the multifaceted aspects of Indian identity. His novels display a very diverse vocabulary, as he extensively uses both Arabic-Persian terms and *tadbhava* terms, the choice dictated by communicative and stylistic needs rather than ideological and identity-related factors.

Alongside Premchand's work, novels and short stories by other authors are also included. The corpus consists of a diverse compilation that spans several decades and that tries to reflect the rich range of the literary production in Hindi. In particular, the corpus includes novels from prominent writers such as Dharmavir Bharti, Mohan Rakesh, and Jaishankar Prasad. **Dharmavir Bharti's** works, *Sūrāj kā sātvāṁ ghoṛā* and *Gunāhoṁ kā devtā*, bring forth the intricacies of human emotions, the societal norms and the life of urban middle-class at the turn of pre- and post-independence India. The novel *Gunāhoṁ kā devtā* is one of his most famous works, which made him one of the most recognized figures of Hindi literature alongside Premchand. **Mohan Rakesh** was among the pioneers of the *Nai kahānī* (New Story) literary movement of the 1950's. The production of the writers

belonging to this movement mainly consisted of short stories and novels addressing the loss of values and the rising of insecurity, loneliness and anxiety in middle class families of the time (Singh 2016). *Aṁdhere band kamre* is one of the best-known novels by Mohan Rakesh, which explores the complexities of a dysfunctional relationship in the background of urban settings.

Jaishankar Prasad and **Mahadevi Verma**, on the other hand, are cosidered two Pillars of the *Chāyāvād* movement (Romanticism) in Hindi literature (Rubin 2002). In particular, Mahadevi Verma production, included in the corpus with a collection of her short stories, mainly addressed women's lives in North India, and the development of women's education at the beginning of the 20th century (Schomer 1983). Alongside Verma's production, the section of the corpus composed by short stories also includes a collection by **Raghuvir Sahay** and **Amritlal Nagar**.

5.1.3. The HiTenTen Corpus

The hiTenTen corpus belongs to the TenTen corpora family on SketchEngine.[23] This is a range of corpora collected from the Internet that are currently available in more than 40 languages. The name TenTen indicates the size of the corpora, which often aim to include roughly 10 billion words, although the actual size may vary. The current version of hiTenTen, and the one that I relied on for this dissertation, is the hiTenTen2021,[24] which consists of 901,352,786 tokens. The texts were downloaded between March and April 2021 from the web and in December 2020 from Hindi Wikipedia. All TenTen corpora are collected following the same criteria and can be regarded as comparable. The corpora are created using technology specialized in collecting only linguistically valuable content from the web and aim to be a large, representative sample of the language. This includes a wide variety of texts from different domains such as news articles, blog posts, forums, and other types of web content. The diversity of sources ensures that a particular genre or style is not overrepresented. However, the extremely variable sources of texts collected from the web make it challenging to manage all the variants, and as I am not a native speaker and I lack the intuitive understanding to discern what is standard in the language and what is not, I decided to rely on literary works and did not choose the hiTenTen corpus as the main source for my linguistic data.

5.1.4. Corpus creation and data extraction

The corpus was created and interrogated on SketchEngine, an advanced corpus management and query system that allows for in-depth linguistic analysis. SketchEngine offers the possibility to create a corpus by uploading user's own texts. While the hiTenTen corpus is a public corpus supplied by

[23] https://www.sketchengine.eu/documentation/tenten-corpora/.

[24] https://www.sketchengine.eu/hitenten-hindi-corpus/.

SketchEngine and accessible by all users that have a license, the Literary corpus I collected is a private one. Corpora uploaded by single users are stored in their personal space so that other users cannot access them. However, the owner of a private corpus can grant access to individually selected users who owns a license on SketchEngine by sharing the corpus.

I created the corpus by uploading the files in.txt format, which is among the formats best supported by SketchEngine. Once the corpus is uploaded and compiled, SketchEngine can calculate its size and the size of the lexicon. The lexicon size is obtained by calculating the number of unique items in the corpus: each item is counted only once even though it appears many times in the corpus. The lexicon size of the Literary corpus is given in Table 21. Metadata can be added on the uploaded files, so I included information on the author, the title of the novel or the short story collection and the year when they were written. Metadata allowed me to keep track of the origin of each example extracted from the corpus.

Table 21: Corpus size.

CORPUS SIZE	
Tokens	1,136, 573
Words	987,787
Sentences	78, 054
LEXICON SIZE	
Word	43, 517
Lemma	36,776

Once the corpus is compiled, it is automatically tokenized, lemmatized, and pos-tagged. Each token in the corpus is thus associated to its lemma and pos-tagged according to the tagset of that given language. A tagset is a list of labels used to indicate the part of speech of each token in a text and sometimes also other grammatical categories (case, tense, etc.). Since April 2022, Sketch Engine offers tools for lemmatization and POS-tagging for many Indian Languages[25], including Hindi. In Table 22 are given the most common POS-tags used with Hindi corpora on SketchEngine.

Table 22: Part-of-speech tagset for Indian Languages (most common tags).

COMMON TAGS	
noun	N.*|XC.*
verb	V.*

[25] https://www.sketchengine.eu/tagset-indian-languages/

adjective	J.*
pronoun	PRP.*
adverb	RB.*
postposition	PSP.*
interjection	INJ.*
conjunction	CC.*

To extract the occurrences of the verbs I analyzed, I used the *Concordance* tool on Sketch Engine, which allows the user to search for words, phrases, tags, syntactic patterns, etc., and displays the results in the KWIC format shown in Figure 19.

Details	Left context	KWIC	Right context
Alamkar_Pr...	वरन उसकी इतनी रक्षा भी न करते थे जो वर्तमानकाल में अनिवार्य	समझी	जाती है।</s><s>उनका विश्वास था कि देह को जितना कष्ट दिया
Alamkar_Pr...	दिनों में कृषकों की सहायता करते थे।</s><s>शहर के रहने वाले	समझते	थे कि यह चोरों और डाकुओं का गिरोह है, यह सब अरब के लुटेरों
Alamkar_Pr...	ते हैं।</s><s>किन्तु यह भरम था।</s><s>तपस्वी धन को तुच्छ	समझते	थे, आत्मोद्धार ही उनके जीवन का एकमात्र उद्देश्य था।</s><s>उ
Alamkar_Pr...	कहा-"भगवान तू साक्षी है कि मैं पापों को कितना घोर और घातक	समझ	रहा हूं।</s><s>" धीरेधीरे इस मूर्ति का मुख विकृत होने लगा, उस
Alamkar_Pr...	न्य है।</s><s>मुझे शक्ति दे कि तेरे जीवों को तेरी दया की ज्योति	समझाकर	परेम करुं, क्योंकि संसार में सब कुछ अनित्य है, एक तू ही नित्य, अ
Alamkar_Pr...	ार करुंगा।</s><s>यही मेरा संकल्प है।</s><s>आप इसे उचित	समझते	हैं ?</s><s>' पालम-'पिरय बन्धु, मैं एक अधम पराणी हूं किन्तु ह
Alamkar_Pr...	देखा और अपनी ज्ञानशक्ति से सहज में इसका आध्यात्मिक आशय	समझ	लिया।</s><s>तीतरी के रूप में थामस थी, जो पापजाल में फंसी
Alamkar_Pr...	हा।</s><s>उसने मानो कुछ सुना ही नहीं।</s><s>पापनाशी ने	समझा	कि वह ध्यान में मग्न है।</s><s>वह हाथ बांधकर उकडूं बैठ गया
Alamkar_Pr...	उसकी ओर बिना ताके ही उत्तर दिया- 'पथिक, मैं तुम्हारी बात नहीं	समझा	और न परभु मसीह को ही जानता हूं।</s><s>' पापनाशी ने विस्मि
Alamkar_Pr...	हो।</s><s>लेकिन मुझे इसकी भी चिन्ता नहीं कि कोई मुझे क्या	समझता	है, मेरा मान करता है या निन्दा।</s><s>' पापनाशी ने फिर शंका
Alamkar_Pr...	ापस्वियों की, जैसे अबोध बन्दर दीवार पर रंग पोतकर अपने मन में	समझता	है कि मैं चित्रकार हो गया।</s><s>इसका तुम्हारे पास क्या जवाब
Alamkar_Pr...	न्ते की कोई परवाह नहीं, और न इसकी परवाह है कि तुम मुझे क्या	समझते	हो।</s><s>मुझे न परेम है न घृणा।</s><s>बुद्धिमान पुरुष को
Alamkar_Pr...	हाथ से निकल जाने का भय होता है अथवा जिस चीज को वह बुरा	समझते	हैं उसका उन्हें सहन करना पड़ता है।</s><s>इन विचारों को चित्त
Alamkar_Pr...	बेल्कुल अर्थशून्य नहीं है।</s><s>संसार की धनसम्पत्ति को तुच्छ	समझना	बुद्धिमानों का काम है।</s><s>लेकिन अपने अनन्त सुख की उपेक्ष
Alamkar_Pr...	s><s>' पापनाशी को मानव चरित्र का पूरा ज्ञान था।</s><s>वह	समझ	गया कि इस मनुष्य पर ईश्वर की कृपादृष्टि नहीं हुई है और उसकी अ

Figure 19: The KWIC Concordance for the verbs *samajhnā*.

Each concordance is linked to the exact position it has in the corpus, so that the user has easily access to the wider context in which the example is contained. Concordances can then be sorted and filtered according to parameters defined by the user. Furthermore, they can be counted and processed for statistical purposes directly through the SketchEngine interface. For example, the system allows for the calculation of absolute and relative frequencies, as well as the analysis of the distribution of the searched element within the corpus. For extracting all the occurrences of a given verb or a given construction in the corpus, I mainly used the advanced search, employing specific queries that I formulated with the Corpus Query Language (CQL). The CQL is a special code that combines the search for specific tokens with the use of POS tags, corpus structures, and even unspecified tokens, all in one query. Hence, the CQL allows the user to search for syntactic or lexical patterns without

necessarily specify concrete words. The syntax is easy and transparent. Each token is represented as a value-attribute pair in square brackets, as follows:

- [attribute = "value"]

The attribute refers to the level of annotation (e.g. word, POS, lemma, etc) and the value refers to what the user is looking for. Two or more attribute-value pairs can also be combined inside the same pair of brackets to search for tokens characterized by more than one feature at different levels of annotation. For example, a query for the different Hindi complex predicates formed with the nominal host *jān* "knowledge" (जान) followed by an unspecified light verb would look as shown in 147.

147. [lemma=" जान " & tag="N.*|XC.*"] [tag!="V.*"]? [tag="V.*"]

The query above is meant to search for all the occurrences of sequences of three tokens: the first token is defined as the lemma *jān* (जान) tagged as a noun (the tag for nouns is: "N.*|XC.*"), the second element is an optional token that must not be tagged as a verb, while the third token is an unspecified lemma that must be tagged as a verb. The optional unspecified token between the nominal host and the light verb is necessary, as the components of a complex predicate in Hindi are not necessarily sequential. I added an unspecified token that must be different from a verb, because otherwise the query would give only results in which the light verb is followed by a tense auxiliary. As mentioned above, the results of the query are shown in context in the form of concordances, displayed in the KWIC (KeyWord In Context) format. This is the format commonly used in corpus linguistics to display concordance lines that show how a specific word or phrase appears in its surrounding context. This is particularly useful for analysing word usage, collocations, and linguistic patterns: the keyword is placed in the centre of each line, and the context words surrounding the keyword are displayed to the left and right.

Another example is the query given in 148 that I used to retrieve all the occurrences of the construction expressing an ability or knowing how to do something.

148. [word=".*ना" & tag="V.*"] []? [lemma="आ" & tag="V.*"]

As I will discuss in chapter 8.6, in order to express the possession of an ability or a skill, Hindi uses a particular construction in which the Experiencer is marked with the dative, the action that they are capable of performing is encoded by the infinitive of the verb, and the main verb is *ānā* "come" (as in 149).

149. मुझे लिखना आ जाए, तो मैं भी एक कहानी लिखूं।

mujhe	*likh-nā*	*ā*	*jāe,*	*to*	*maiṁ*	*bhī*
1SG.DAT	write-INF	come-IPRF-M.SG	go-SBJV.3SG	then	1SG.NOM	too

ek kahānī *likh-ūṁ.*
on story(F.SG.NOM) write.SBJV.1SG
"If I learned how to write, I would write a story too."

The query shown in 148 indicates that the pattern we are searching for is composed by three tokens: the first token must be a lemma ending in *nā* (ना), that is the ending marking the infinitive of every verb in Hindi) and tagged as a verb; while the last token is the lemma *ā-* (आ- is the root of the verb *ā-nā* "come") tagged as verb. Once again, since the components of the construction are not necessarily sequential, between the two exact lemmas I added an optional undefined token. I then manually examined and annotated all the sentences resulting from the query, and I excluded the instances of irrelevant constructions.

Note however that the pos-tagging and lemmatization offered by SketchEngine are not always free of problems. In the first place, lemmatization is not always perfect. Let us illustrate this by means of the Hindi verb *samajhnā* "understand, consider" (discussed in section 8.6). In order to search for all the occurrences of the verb *samajhnā* in the corpus, the query would be as the following:

150. [lemma= "समझ" & tag="V.*"]

This query is intended to look for a single token instantiating any form of the verb *samajhnā* represented in the query by the lemma *samajh* "समझ" tagged as a verb. The first fifteen lines resulting from this query are shown in the KWIC format in Figure 19. Figure 20 (glossed in 151) and Figure 21 (glossed in 152) shows two occurrences of the verb in the corpus. In the figures, the annotation of the lemma is given in the grey line below the concordance.

</s><s> मैंने तुम्हारा मतलब नहीं समझा । </s><s>

मैं तु मतलब नहीं समझा

Figure 20

</s><s> आप इसे उचित समझते हैं? </s><s>

आप इसे उचित समझते है

Figure 21

151. *maiṁ=ne* *to* *tumhār-a* *matlab* *nahīṁ* *samjh-ā*
1SG=ERG EMPH your-M.SG meaning(M.SG.NOM) not understand-PRF.M.SG
"I don't understand what you mean."

152. *āp* *ise* *ucit* *samajh-t-e* *haiṁ?*
2SG.HON.NOM this.ACC fair understand-IPRF-M.PL be.PRS.3PL
"Do you think this is fair?"

Sentence 152 shows a form of the lemma *samajhnā* "understand", appearing here in the masculine plural imperfective participle. This form is correctly lemmatized: as it is shown in the annotation below the concordance, the red keyword is interpreted as a form of the lemma *samajhnā* (समझ = *samajh*). Sentence 151 as well shows a form of the lemma *samajhnā*, in this case a masculine singular perfective participle (*samjhā*). Yet, in this example, the keyword is incorrectly lemmatized, and it is interpreted as a form of the lemma *samjhā* (समझा), that is the causative of *samajh*. This happens probably because the root of the causative, e.g. *samjhā*, is formally identical to the masculine singular perfective participle of the non-causative form *samajh*. Both the examples shown above should result from the query in 150; however, since the form in sentence 151 is wrongly lemmatized, it would not appear in the list of concordances resulting from this query. This and other issues can be fixed by modifying the query accordingly. For example, to address this issue I formulated the new query in 153.

153. [lemma= "समझ.*" & tag="V.*"]

The sequence *lemma="समझ.*"* means that the lemma I am searching for has a root formed by समझ and followed by an undefined number of undefined characters. This new query will bring problems of its own, in particular its results will not only show all the occurrences of the verb *samajhnā*, but also the occurrences of its causative *samjhānā*. But in this case the problem can be easily solved during the annotation process, by manually taking off the undesired results.

5.1.5. Sample representativeness

As I mentioned in the introduction to this chapter, many experiential verbs have a very high frequency and occur thousands of times, so that a manual scrutiny of all the occurrences would have been extremely time-consuming. Hence, I decided to select a random set of 200 occurrences for each verb.

Since the chance of a construction being chosen by a given verbs correlates positively with its frequency with that verb, we can assume that a random sampling method is a good way for collecting the 200 occurrences. When working with a large concordance set, SketchEngine offers the random sample function, which is meant to reduce the number of concordance lines while preserving the balance and representativeness of the sample, as the number of lines defined by the user are randomly selected from all parts of the corpus.

Corpus based approaches always rely on the main theoretical and methodological assumption that a corpus is representative of the actual use of the language (Raineri and Debras 2019). From a statistical point of view, samples are scaled down versions of larger corpora, hence a sample is assumed to be representative if the results of the analysis of the sample is also valid for the whole corpus (Yates 1965, Manning and Schütze 1999). One might object that 200 occurrences are not representative enough. Unfortunately, while there are saturation measures for the lexical representativeness of a sample, there are no such measures for syntactic constructions that I am aware of (see on this also Zanchi and Inglese 2022). However, we can assess whether a sample of 200 random occurrences reaches data saturation or not. In general, a sample reaches saturation when adding more data does not change the statistical distribution of its contents (McEnery and Wilson 2001: 166), in our case the frequency distributions of the constructions analyzed. At this point, the sample is considered large and diverse enough to be representative of the linguistic phenomenon it aims to model. While there may not be universally accepted measures for syntactic construction saturation, several criteria could signal that a sample is approaching saturation, in particular frequency stability is generally acknowledged as a good indicator of data saturation (McEnery, Xiao and Tono 2006). If the frequencies of the syntactic constructions stabilize and new additions cause minimal fluctuation regarding new information, the sample may be nearing saturation. In order to determine saturation, researchers often rely on iterative analysis, where they check the statistical properties of the sample as they add more data. When the addition of new texts does not provide new information or cause changes to the frequency distributions, data saturation may be assumed. As McEnery, Xiao and Tono (2006: 20) point out "in order to obtain a representative sample from a population, the first concern to be addressed is to define the sampling unit and the boundaries of the population. The population is the assembly of all sampling units while the list of sampling units is referred to as a sampling frame." In my case, the basic sampling unit is the construction, the population is the assembly of all the constructions occurring with a given verb in the corpus, while the sampling frame is the list of the occurrences randomly selected.

In order to assess whether or not a sample of 200 occurrences randomly selected from the corpus is representative of the assembly of all the constructions occurring with a given verb, we can

do an iterative analysis where we check the statistical properties of the sample as we add more data. Here, I present a case study for the verb *dekhnā* "see, look at". Given that, as I will discuss in section 7.3.1, this verb is one with the highest absolute frequency in the corpus and with the widest range of constructions, we can assume that, if 200 occurrences are representative in the case of *dekhnā*, then this also holds for the other verbs analyzed. I checked the relative frequencies of the constructions occurring with the verb *dekhnā* for four random samples of different size and I assessed whether the statistical distribution changed or remained constant as the sample size increased, and at which cut-off point the distribution of the frequency of the constructions stopped changing in a significant way. The sizes of the four random samples were: 60 occurrences, 100 occurrences 150 occurrences and 200 occurrences. Table 23 shows the relative frequencies of the constructions for each sample size. As the data in the table show, the relative frequency of each construction remains relatively stable across the four samples, the only difference is that as the sample size increases rare constructions enter the list. However, rare constructions are not of much interest for my investigation, as generalizations on the semantics of a construction cannot be made if this occurs few times with a given verb. However, I opted for widest sample size, and I collected 200 occurrences.

Table 23: The relative frequencies of the construction occurring with *dekhnā* for the four random samples.

	Sample size			
Construction	**60**	**100**	**150**	**200**
Transitive	58,3%	54%	58%	55,5%
Oblique Stimulus (*kī or/taraf*)	15%	13%	15,3%	12%
Finite complement clause	13%	12%	14%	13,5%
Predicative construction	11,7%	13%	10%	11,5%
Locative adverbial	1,67%	2%	1,3&	4%
Oblique Stimulus (*ke andar*)	0%	0%	0%	0,5%
Oblique Stimulus (*meṁ*)	0%	1%	0%	0,5%
Passive	0%	4%	1,3%	2,5%

6. Bodily sensations

In this chapter, I analyze the expression of bodily sensations in Hindi. Bodily sensation verbs are the most semantically distant from the transitive prototype and they differ from it in many respects. In the first place, unlike effective action verbs, verbs of bodily sensation typically profile only one participant, as they are the only type of experience which is typically not caused by an external element, and which is not directed toward a content, as argued by Verhoeven (2007: 47, see also Fedriani 2012: 56, Luraghi 2020a). This aspect distinguishes bodily feelings also from all other experiential subdomains. Secondly, bodily sensations typically denote states and not activities, even if they may also depict dynamic events and be construed as processes or as changes of state. Additionally, the Experiencer of a bodily sensation is never conceptualized as an Agent since it is never volitional and has never control over the experience (Bossong 1998, Luraghi 2020a). It is instead more similar to a Beneficiary/Recipient or a Patient. Like the prototypical Recipient it is always animate and non-volitional, and like the prototypical Patient it never controls the event, and it is affected by it. It is not surprising therefore that, as I will discuss more in detail in this chapter, bodily sensations are never encoded with a transitive construction in Hindi and that the typical case-marking for this Experiencer type is the dative case. This is quite common in the languages of the world, which usually encode bodily sensations via intransitive verbs with one argument either in the nominative/absolutive case or an oblique case or via experiential adjectives.

The subdomain of bodily sensation is concerned with a large variety of feelings related to the body of the Experiencer (Verhoeven 2007). In the following analysis, I will mainly focus on three different bodily sensation types:

a. Feelings concerning a state of saturation, such as “be hungry” or “be thirsty” (section 6.1).
b. Feelings of temperature, like “feel hot” or “feel cold” (section 6.3).
c. Feelings related to specific areas of the body such as “aching” and “itching” (section 6.2).

Note that, even though they are semantically related, here I keep distinct the following classes of verbs: verbs referring to bodily states (such as “be sick”), verbs referring to bodily sensations (such as “feel sick”) and verbs referring to bodily emission (such as “cough”), and I will only consider bodily sensations as belonging to the domain of experience. Bodily sensations are also distinct from bodily functions such as bleeding or sweating. The domains of bodily functions and that of bodily states border with the domain of bodily sensations and this may result in metonymical extensions. For example, in many languages, bodily states constructions pattern with constructions encoding

bodily sensations, as English *be cold*, which can refer either to a sensation (*Let's go inside, I am cold*) or to a state (*The tea is cold*).

From the point of view of the argument structure, Hindi verbs of bodily sensation display some peculiar patterns which are not found in the other experiential subdomains (or at least are very rare). For instance, the expression of pain employs a locative construction which is hardly found in the other experiential domains analyzed in this study. In this construction the Experiencer is encoded as the possessor of a body part where the sensation is located, and the Expertum is in the nominative (as in 155). Another construction typical of bodily sensations and marginalized in the other experiential domains is the copular constructions, in which the Experiencer is in the nominative and the Expertum is lexicalized in an adjective that agrees in gender and number with the Experiencer. The verb *honā* "be" functions as a copula (as in 156). Bodily sensations also display constructions that pattern with other experiential verbs and that link this verb class to other experiential classes. In particular, as already mentioned and as I will discuss in more details below, the most frequent construction used in Hindi to encode bodily sensation is the dative construction, in which the Experiencer is marked with the dative and the Expertum is in the nominative (as in 154), the predicate is expressed either by the verb *honā* "be" or *lagnā* "adhere". The dative marking on the Experiencer occurs in the encoding of all experiential subdomains in Hindi, thus providing a link between bodily sensations and the rest of the experiential domain. Additionally, other constructions with a very limited scope may be found in Hindi for the expression of bodily sensations, such as the inverse transitive construction occasionally used for the expression of pain that I discuss in section 6.1.

154. उसे भूख लगी है।

use	*bhūkh*	*lag-ī*	*hai*
3SG.DAT	hunger(F.SG.NOM)	attach-PRF.F	3SG.PRS

"He is hungry."

155. धनिया के सिर में दर्द था।

dhaniyā=ke	*sir=meṁ*	*dard*	*th-ā*
dhaniya=GEN	head(M.SG.OBL)=in	pain(M.SG.NOM)	be.PST-M.SG

"Dhaniya had a headache."

156. वह बच्चा प्यासा है।

vah	*bacc-ā*	*pyās-ā*	*hai*
that.NOM	child(M)-SG.NOM	thirsty-M.SG	be.3SG.PRS

"That child is thirsty."

6.1. Feelings of pain and itch

In Hindi, feelings of pain and itching can be expressed with a locative construction, a dative construction, and a transitive construction. As shown in Table 24, the locative construction is the most frequent in the corpus for the expression of this sensation type.

Table 24: Occurrences of constructions encoding feelings of pain and itch in Hindi.

Construction	Absolute Frequency
Transitive pattern	4
Dative pattern	19
Locative pattern	52

In the locative construction, the Experiencer is in the genitive and is conceptualized as the possessor of the body part where the sensation is located. The body part is marked with an inessive locative: the noun is in the oblique case and is followed by the postposition *meṁ* "in". The Expertum is a nominative NP. Note that the genitive Experiencer is frequently omitted, as in 157 (in the corpus it is absent 27 times over the 52 occurrences of this construction). The relation between the Experiencer and their body part is meronymic and ensures that when the body part is affected the Experiencer in his/her globality is affected as well, hence the genitive Experiencer can be omitted and reference to the Experiencer may be indicated only by his/her body part.

157. रीढ़ में दर्द था और थकान की वजह से बुख़ार-सा भी लग रहा था, इसलिए मैं इतनी देर सोया रह गया।

rīṛh=meṁ *dard* *th-ā* *aur*
spine(M.SG.OBL)=in pain(M.SG.NOM) be.PST-M.SG and
thakān=kī_vajah=se *bukhār-sā* *bhī* *lag* *rah-ā*
tiredness(F.SG.OBL)=because_of fever-like.M.SG also attach PRGR-M
th-ā *islie maiṁ itnī der soyā rah gayā.*
be.PST.M.SG that is why I overslept for so long
"My spine was hurting and, due to tiredness, I was also feeling feverish, that is why I overslept for so long."

Remarkably, this construction is never used to express feelings of temperature or feelings of saturation such as hunger or thirst. This is probably explained by the fact that such sensation types are related to a general condition of the Experiencer and not to a specific area of his/her body. It is reasonable to conjecture that this construction is specifically used for the expression of bodily

sensations that are localized on a specific area of the experiencer's body such as the head (as in 0 above) the stomach (as in 158) or the tongue (as in159). In this case, the body part is clearly conceived as a container, and the experiencer is conceptualized via the cross-linguistically quite common metaphors according to which THE BODY IS A CONTAINER and SENSATIONS ARE THE CONTAINED THINGS (Kövecses 2000).

158. शाम को उसके पेट में दर्द होने लगा।

śām=ko	*us=ke*	*peṭ=meṁ*	*dard*
evening(F.SG.OBL)=at	3SG.OBL=GEN	stomach(M.SG.NOM)=in	pain(M.SG.NOM)
ho-n-e	*lag-ā*		
be-INF-OBL	start-PRF.M.SG		

"In the evening he started having a stomachache."

159. खन्ना की जीभ में खुजली हो रही थी।

khannā=kī	*jībh=meṁ*	*khujlī*	*ho*	*rah-ī*
khanna=GEN	tongue(F.SG.OBL)=in	itching(F.SG.NOM)	be	PRGR-F
th-ī				
be.PST-F.SG				

"Khanna's tongue was itching."

It is quite common in the languages of the world to find constructions that put the Experiencer in the background and profile it through his/her body part (Verhoeven 2007, Croft 2022). In these constructions, the Experiencer is encoded either as an argument of the verb in an oblique case (external possessor constructions) or as a genitive nominal modifier of the body part NP (internal possessor constructions). An example of an external possessor construction is the expression of pain in Italian: consider sentence 160, in which the body part is encoded as the subject of the experiential predicate and the Experiencer is expressed by the dative form of the first-person singular pronoun *mi* "to me".

160.

Mi	*fa*	*male*	*la*	*gamba.*
1SG.DAT	do.3SG.PRS	pain	the.F	leg(F)

"My leg hurts".

Internal possessor constructions are found for example in English. Compare the translation of the Italian example in 160 "My leg hurts", in which the Experiencer is encoded as an attributive possessor of the body part and the latter is the subject of the single argument verb *hurt*. The English translation

in 160 shows that one should not assume that a genitive modifier is not a participant in the event, especially when the genitive modifier refers to the possessor of a body part. As Croft (2022: 32) rightly points out "there are some situations in which the semantic possessor of a participant is itself sufficiently salient, often in part because of its humanness, for it to be construed as a participant in the event expressed by the predicate, despite its indirect relationship to the event per se". Notably, however, the Hindi construction differs from the Italian and the English strategies as it does not foreground the body part, but the Expertum. As I will show, this is a common feature in the Hindi expression of bodily sensations. Among constructions discussed in this chapter, only the adjectival construction used for the expression of feelings of lacking encodes the Experiencer in the nominative. All other constructions reserve the nominative case marking to the Expertum. This correlates with the fact that in Hindi verbs of bodily sensations are mainly expressed through noun verb complex predicates in which the nominal host encodes the Expertum and are never expressed by single argument verbs with a nominative Experiencer.

The corpus also attests to a transitive construction expressing feelings of pain. This construction is somehow similar to the construction Verhoeven (2007: 80-81) refers to as reverse transitive construction (see also Belletti and Rizzi 1988 on the Italian verb *preoccupare* "worry"), in which the Stimulus or the Expertum are conceptualized as the Agent. Notably, however, in this construction the Agent-like argument is once again the body part, the Expertum is lexicalized in the complex predicate *dard karnā*, consisting of the noun *dard* "pain" and the transitive light verb *karnā* "do". The Experiencer is present as a genitive that is part of the NP headed by the noun encoding the body part, hence it is encoded as an attributive possessor (as in 161). This construction is quite rare, and it occurs only four times in the corpus, against the 52 occurrences of the locative construction and the 18 occurrences of the dative one.

161. उस रोज जानवर पर बड़ी मार पड़ी। मेरे हाथ दर्द करने लगे।

us roj jānvar par baṛī mār paṛī
I hit that animal hard that day

mere	*hāth*	*dard*	*kar-n-e*	*lag-e*
1SG.GEN	hand(M.PL.NOM)	pain(M.SG.NOM)	do-INF-OBL	start-PRF.M.PL

"I hit that animal so hard that day. My hands started to hurt."

Feelings of pain and itching can also be expressed by a dative construction in which the Experiencer is in the dative case and the Expertum is lexicalized in a nominative NP which agrees with the verb *honā* "be". Interestingly, unlike the locative construction, which is used to express only localized sensations, the dative construction can be employed either to encode a localized feeling or to encode

a bodily sensation that is not related to any specific area of the body (see example 162). The difference between the dative and the locative construction is that the latter is used in the Hindi expression of bodily sensation in order to constrain the circumscribed area within which the sensation is perceived. Notably, however, in this construction type the Expertum is usually expressed by a noun that lexicalizes the body part where the sensation is felt (like *sir-dard* "headache" in 163) or a locative adverb occurs in the sentence (see example 164). Sentence 163 might seem similar to the locative constructions exemplified in 158 or 159, but it differs from them with regard to the encoding of the Experiencer. In the locative constructions discussed above, the Experiencer is encoded as a genitive adjunct, while here it is marked with the dative. This latter construction resembles the external possessor construction typical of some Indo-European languages (Haspelmath 1999, Luraghi 2020b). The main difference between these two patterns is that when the Experiencer is marked with the genitive it is constructed as a modifier of the body part NP and not as an argument, while when it is encoded in the dative it is an argument that depends on the verb (see also Croft 1985).

162. रग्घू को इस समय मर्मान्तक पीड़ा हो रही थी।

ragghū=ko	*is*	*samay*	*marmāntak*	*pīṛā*	*ho*
ragghu=DAT	this.OBL	time(M.SG.OBL)	piercing	pain(F.SG.NOM)	be
rah-ī	*th-ī.*				
PRGR-F	be.PST-F.SG				

"Ragghu was in piercing pain at that moment."

163. ऐसा सिर-दर्द मुझे आज तक नहीं हुआ था, मगर तुम्हारे हाथ रखते ही सिर ऐसा हल्का हो गया है, मानो दर्द था ही नहीं।

esā	*sir-dard*	*mujhe*	*āj=tak*	*nahīṁ*	*hu-ā*	*th- ā*
such	head-ache(M.SG.NOM)	1SG.DAT	today=till	not	be.PRF-M	be.PST-M.SG
magar	*tumhāre*	*hāth*	*rakh-t-e*		*hī*	*sir*
but	2PL.GEN	hand(M.PL.NOM)	take-IPRF-M.PL		EMPH	head(M.SG.NOM)
aisā	*halk-ā*	*ho*	*gayā*	*hai,*	*māno*	
such-M.SG	light-M.SG.NOM	be	go.PRF.M.SG	be.3SG.PRS	like	
dard	*th-ā*	*hī*	*nahīṁ.*			
pain(M.SG.NOM)	be.PST-M.SG	EMPH	not			

"I had never had such a headache till date, but as soon as I held your hand, my head became so light, as if there was no pain."

164. तुम्हें कहीं दर्द तो नहीं है।

tumheṁ	*kahīṁ*	*dard*	*to*	*nahīṁ*	*hai?*
2SG.DAT	somewhere	pain(M.SG.NOM)	then	not	be.PRS.3SG

"Do you have any pain (somewhere)?"

6.2. Feelings of temperature

Referring to the observations made by the physiologist Hensel (1981), Koptjevskaja-Tamm (2015: 8) notes that the perception of temperature basically involves two types of situations: temperature sensations and thermal comfort. These two situations differ from each other in that according to Hensel (1981: 168) "temperature sensation is a rational experience that can be described as being directed towards an objective world […]" as in *It is very cold today*, while "thermal comfort is an emotional or affective experience referring to the subjective state of the observer", as in *I am cold*. From a semantic point of view, previous scholars (Plank 2003, Goddard and Wierzbicka 2007, Koptjevskaja-Tamm 2011, 2015, Luraghi 2015) distinguish three different temperature related situations which may be expressed cross-linguistically: these are tactile temperature (165), ambient temperature (166) and personal-feeling temperature (167).

165. पानी ठंडा है।

pānī	*ṭhaṁḍ-ā*	*hai*
water(M.SG.NOM)	cold-M.SG	be.3SG.PRS

"The water is cold."

166. बाहर ठंड है।

bāhar	*ṭhaṁḍ*	*hai*
outside	cold(F.SG.NOM)	be.3SG.PRS

"It is cold outside."

167. मुझे ठंड लग रही है।

mujhe	*ṭhaṁḍ*	*lag*	*rah-ī*	*hai*
1SG.DAT	cold(F.SG.NOM)	attach	PRGR-F	be.3SG.PRS

"I am cold."

Personal feelings are fundamentally different from the other two types, as they are the only type of situation that explicitly involves a person experiencing a bodily sensation. In this section, I will only

focus on this type of temperature related situations. Of course, ambient temperature and tactile temperature imply that someone feels a temperature in relation to ambient circumstances or specific entities, but in such cases the experiencer is not necessarily linguistically expressed. As Koptjevskajae-Tamm (2011: 394) points out "languages vary considerably as to whether or to what extent the three-fold distinction is made explicit, either by morphosyntactic or by lexical means or by a combination of both." In some languages, the three subdomains are kept distinct. Table 25 exemplifies three different ways of encoding and distinguishing these three subdomains linguistically, presenting data for Hindi, Italian and English.

Table 25: Constructions used in Hindi, Italian and English to encode the three temperature-related domains.

	TACTILE	AMBIENT	PERSONAL FEELING
Hindi	Predic. adj construction	Single argument construction	Dative construction
Italian	Predic. adj construction	Impersonal verb construction	Possessive construction
English	Predic. adj construction	Predic. adj construction	Predic. adj construction

As Luraghi (2015) discusses, in Italian three different constructions are used for the three fields: a copular construction is used to express tactile temperature, an impersonal verbal construction employing the support verb *fare* "do" together with a temperature noun is used to encode ambient temperature, while a possessive construction with the verb *avere* "have" followed by a temperature noun is used to encode personal-feeling temperatures. In contrast, other languages do not show such distinction. In English, for example, a copular construction is used to express all three situation types: *It's cold today*, *The water is cold*, *I am cold*. Hindi uses constructional ways to keep these three temperature related situations. Tactile temperature is expressed by a copular construction (as in 165), ambient temperature is expressed by a single argument construction (as in 166), while personal feeling temperature is expressed by a dative construction (as in167). In Hindi, feelings of temperature are encoded in the same way as other bodily sensations: the Experiencer is in the dative and the Expertum is lexicalized in the nominal host of a noun verb complex predicate. The basic term for "cold" in Hindi is *ṭhaṇḍ*, while the basic term for "hot" is *garm*. Both terms are quite opaque with regard to their etymology. Turner (1971:240) traces the term *garm* back to the Sanskrit word *gharma* "heat", but other scholars propose to consider it as a borrowing from Persian (see Molesworth et al 2005: 226, Liljegren and Haider 2015: 459). As regard the term for cold, *ṭhaṇḍ*, Turner (1971: 778) argues that it comes from the Sanskrit *stabdha* "firmly fixed" via the semantic shift: firm → sluggish → cold. The light verb may be instantiated either by the verb *honā* "be" or by the verb *lagnā* "adhere,

be attached". Unfortunately, verbs expressing feelings of temperature are scarcely attested in the corpus, with only 13 occurrences. An example is given in sentence 168.

168. तुम्हें गर्मी लग रही है अम्मांजी? मैं तो ठंड के मारे कांप रही हुँ।

tumheṁ	*garm-ī*	*lag*	*rah-ī*	*hai*	*Ammāṁ jī?*	
2SG.DAT	hot(F)-SG.NOM	adhere	PRGR-F	be.PRS.3SG	Ammāṁ jī?	
maiṁ	*to*	*ṭhaṇḍ=ke māre*		*kāmp*	*rah-ī*	*hūṁ.*
1SG.NOM	EMPH	cold(F.SG.OBL)=because_of		shiver	PRGR-F	be.PRS.1SG

"Are you hot, Ammāṁ jī? I am trembling with cold."

Notably, in Hindi the copular construction is not allowed for the encoding of feelings of temperature. As Verhoeven (2007: 43) notes, an important point to bear in mind when analyzing bodily sensations is that these experiential types are related to bodily states, but they are not bodily states. This distinction is important because only bodily sensations are related to the domain of experience. Moreover, a bodily sensation may occur regardless of the bodily state it is usually triggered by (as Verhoeven points out, one can feel sick without being sick). While many languages do not necessarily distinguish these two notions formally (like English), Hindi keeps them distinct by using different constructions. The dative construction is used to encode temperature related feelings, while the copular construction is used to express temperature related states. This latter construction is specified to expresses states (such as *to be tall,* as in 169) and not feelings of states and for this reason, it cannot be used to express sensations such as *to be cold*: in Hindi, a sentence such as 170 refers to the body temperature of the person and not their bodily feeling.

169. *vah* *lark-ā* *lamb-ā* *hai*
that.NOM boy(M)-SG.NOM tall-M.SG be.3SG.PRS
"That boy is tall."

170. *vah* *lark-ā* *ṭhaṇḍ -ā* *hai*
that.NOM boy(M)-SG.NOM cold-M.SG be.3SG.PRS
"That boy is cold. (Meaning that his body temperature is cold.)"

6.3. Feelings of saturation or lack (hunger and thirst)

In Hindi, feelings of saturation or lack (like hunger and thirst) can be expressed with a dative construction or with a copular construction. In the first construction, the Experiencer is in the dative, the Expertum is in the nominative and the Stimulus is usually missing. The predicate can take two

different verbs: the verb *honā* "be" or the verb *lagnā* lit. "adhere". The difference between these two light verbs lies on lexical aspect they contribute to the whole complex predicate: the verb *honā* indicates a state (example 171), while the verb *lagnā* usually indicates an inchoative event (example 172). In other words, the verb *honā* simply profiles the experiential state, whereas *lagnā* profiles the entering into the experiential state. However, when *lagnā* is in the perfective aspect it tends to construe the bodily sensation as a state resulting from a change of state (as in 173). The inchoative semantics of the construction with *lagnā* is evident in example 172, where the temporal subordinate clause distinctly depicts a prior moment in which the Experiencer is not hungry, followed by a subsequent moment (after seeing "a delicious meal") when they become hungry.

171. पीने की इच्छा होने पर भी अमर ने कहा-अभी तो प्यास नहीं है, मुन्नी।

pīne kī icchā hone par bhī amar ne kahā:
Even if he wanted to drink, Amar said:

abhī	*to*	*pyās*	*nahīṁ*	*hai*	*munnī*
now	then	thirst(F.SG.NOM)	not	be.PRS.3SG	munni

"Now I am not thirsty, Munni."

172. इन दिनों रूचिकर भोजन देखकर ही उसे भूख लगती थी।

in	*din-oṁ*	*rūcikar*	*bhojan*	*dekh=kar*
those.OBL	day(M)-PL.OBL	delicious	meal(M.SG.NOM)	see=Cè

hī	*use*	*bhūkh*	*lag-t-ī*	*th-ī*
EMPH	3SG.DAT	hunger(F.SG.NOM)	attach-IPRF-F.SG	be.PST-F.SG

"In those days, as soon as he saw a delicious meal, he got hungry."

173. उसने पूछा - पानी लाऊं ? मेहता ने कहा - हां, प्यास तो लगी है।

usne pūchā "pānī lāūṁ?" Mehtā ne kahā –
He asked - should I bring water? Mehta said –

hāṁ	*pyās*	*to*	*lag-ī*	*hai*
yes	thirst(F.SG.NOM)	EMPH	adhere-PRF.F.SG	be.PRS.3SG

"Yes, I am thirsty. / Yes, I got thirsty."

In the copular construction, the Experiencer is in the nominative and the Expertum is lexicalized in an adjective (*bhūkhā* "hungry", *pyāsā* "thirsty") that agrees in gender and number with the Experiencer (as in 174 and 175). As mentioned in the introduction to this chapter, when analyzing

bodily sensations, we need to keep distinct verbs that denote bodily states from verbs that denote bodily feelings, because despite their frequent association they do not refer to the same situations. Someone can be physically hot without feeling hot. Indeed, someone can be physically hot but feel cold at the same time (for example because of fever). This indicates that temperature states and temperature feelings are not directly linked. However, in the case of hunger and thirst this distinction becomes irrelevant, as these sensation types are intrinsically connected to bodily states. Someone cannot be hungry without also feeling hungry and, similarly, when someone feels thirsty is because s/he is thirsty. As mentioned in section 6.2 above, copular constructions are typically used to encode states. This unique intrinsic connection between states and feelings allows for the use of the adjectival construction for the expression of hunger and thirst in Hindi. In other words, this construction can be metonymically extended to the expression of this sensation type since there is no distinction between the bodily feeling and the bodily state when referring to hunger and thirst.

174. मुबारक हैं वे लोग जो प्यासे रहते हैं, क्योंकि वह स्वर्ग की निर्मल नदियों का जल पियेंगे।

mubārak	*haiṁ*	*ve*	*log*	*jo*
blessed	be.PRS.3PL	3PL.NOM	people.NOM	REL.PRN.NOM
pyās-e	*rah-t-e*	*haiṁ*		
thirsty-M.PL	stay-IPRF-M.PL	be.3PL.PRS		

kyoṁki vah svarg kī nirmal nadiyoṁ kā jal piyeṁge

because they will drink the water of the pure rivers of heaven

"Blessed are those who will be thirsty (lit. who will remain thirsty), because they will drink the water of the pure rivers of heaven."

175. बहन, मुझे कुछ खाने को दो, भूखी हुँ।

bahan, mujhe kuch khāne ko do,	*bhūkh-ī*	*hūṁ*
Sister, give me something to eat,	hungry-F.SG	be.1SG.PRS

"Sister, give me something to eat, I am hungry."

6.4. Metaphorical uses

As I have diffusely discussed in the previous sections, the copular construction can only be used for the expression of hunger and thirst in Hindi. Remarkably, however a copular construction instantiated by temperature related adjectives such as *garm* "hot" can be metonymically used for the expression of emotions such as anger. Consider sentence 176 and 177. In the first sentence (176), anger is metaphorically expressed by the adjective *garm* "hot" followed by the conjunctive participle of the

verb *honā* (*ho=kar*), a similar metaphor is expressed in sentence 177 in which the verb *honā* occurs in a V-V complex predicate followed by the light verb *jānā*. This light verb supplies an inchoative reading to the emotion situation, the whole complex predicate *garm ho jānā* means "become angry".

176. फिर पति से गर्म हो कर कहा - तुम भी वहाँ से कमाई करके लौटे तो खेत में पहुंच गए।

phir	*pati=se*	*garm*	*ho=kar*	*kah-ā*
then	husband=INS	hot	be=CP	say-PRF.M.SG

tum bhī vahāṁ=se kamāī kar=ke lauṭe to khet=meṁ pahuṁc gae.

"Then she said to her husband angrily (lit. being angry) - When you returned after taking the money from there, you went to the fields."

177. यह गर्म मिजाज का आदमी है लेकिन इतनी जल्दी इतना गर्म हो जाएगा, इसकी उसे आशा न थी।

yah	*garm*	*mijāj=kā*				*ādmī*	*hai*
this.NOM	hot	temperament(M.SG.OBL)=GEN				man(M.SG.NOM)	be.3SG.PRS
lekin	*itn-ī*	*jaldī*	*itn-ā*	*garm*	*ho*	*jāegā,*	
but	as_much-F	quick	as_much-M	hot	be	go-3SG-FUT-M	

is=kī use āśā na thī.

he did not expect this.

"He is a hot-tempered man, but he did not expect that he would get so angry (lit. hot) so quickly."

Metaphorical extensions of this type are quite common, and the expression of emotions is typically realized through metaphors and metonymies in many languages of the world thus offering one of the most interesting areas of study for linguists interested in investigating how conceptual metaphors operate cross-linguistically (Lakoff and Johnson 1980). Several works for example have been devoted to the study of the basic emotion of anger (see among others Kövecses 1986, 1990, 2000, Lakoff 1987, Gibbs 1994), but other emotions such as fear have been investigated as well (Kövecses 2000, Esenova 2011, Ho 2016, Csillag 2018, Namrata, Abhijeet and Gosh 2023). These semantic extensions from bodily sensations to emotions are explained in cognitive linguistics through the general metonymic principle according to which THE PHYSIOLOGICAL EFFECTS OF AN EMOTION STAND FOR THE EMOTION (Lakoff and Kövecses 1987, Kövecses 2020). Since these physiological effects are universal as they are associated to anatomical and physical properties of human beings, many metonymic and metaphorical extensions are common cross-linguistically.

Anger, as a universal human emotion, manifests itself through various psychological and physiological effects. Among these effects, Lakoff and Kövecses (1987) observe that the most typical are increased body heat, heightened internal pressure (encompassing blood pressure and muscular tension) and agitation. These effects resonate with common experiences of anger, where individuals often feel a heat coursing through their bodies. The semantic extension operating in the Hindi examples 176 and 177 is based on the metonymy ANGER IS HEAT, which is in turn based on the folk theory according to which increased body heat is the principal effect of anger (Lakoff and Kövecses 1987: 203). This is a quite common metonymy drawing a parallel between an abstract emotional state and a tangible physical sensation, and it is not a specific feature of Hindi, but can be found quite frequently also in other languages. The conceptual metonymy ANGER IS HEAT, and the associated metonymy ANGER IS A HOT FLUID IN A CONTAINER for example are found in Italian where a sentence such as *Perché ti scaldi tanto?* "Why are you getting so hot?" actually means "Why are you getting so angry?". The same semantic extension exists also in English (Lakoff and Kövecses 1987), in Chinese (Yu 1995), in Spanish (Barcelona 1989), in Japanese (Vasiljevic and Graham 2020) and in many other languages (Koptjevskaja-Tamm 2015). A similar metaphor is also found in other Indo-Aryan languages; for example, Liljegren and Haider (2015) note that in Palula, the expression *téet-i dimaáy,* which literarily means "a hot/warm brain", is used to indicate someone who is "easily angered". The domain of emotion is not the only one to be expressed metaphorically via temperature related sensations or states. Another interesting semantic extension is discussed by Reznikova et al. (2012), who in a typological study on the expression of pain found that the feeling of heat (for example burning) is one of the main semantic fields used as sources for metonymic and metaphorical extensions for the expression of pain in many languages, thus indicating that there is a cross-cultural tendency to associate feelings of temperature to the feeling of pain (see on this also Viberg 2015: 117). Moreover, feelings of temperature may also be used to express other emotions when associated to cold. For example, Namrata, Abhijeet and Gosh (2023) note that the expression of fear in Hindi (and Bangla) is based on the conceptual metonymy according to which FEAR IS COLD (as in the Hindi expression *hāth-pāṁv ṭhaṇḍā honā* "hands and feet getting cold in fear").

Another metaphorical extension that frequently occurs in the corpus is the use of the verbs *bhūkhā honā* "lit. hungry be" and *pyāsā honā* "lit. thirsty be" to express desire. As said before, verbs of bodily sensation usually have a single argument structure. Among them, only verbs that express hunger or thirst may take a Stimulus, even if this second argument is rarely expressed. Half of the occurrences of the copular construction shows a Stimulus, while dative construction occurs with a Stimulus only once in the corpus. In both the dative and the single argument construction, the Stimulus is marked with the genitive. Remarkably, all cases in which hunger and thirst occur with a

Stimulus show a metaphoric meaning and are used to express emotions / volitions such as “desire something, crave for something”. In these constructions, the Experiencer of an intense emotion is metaphorically conceptualized as the Experiencer of a physical need, and the Stimulus that triggers the emotion is expressed by the genitive Stimulus of the bodily sensation verb (as in 178, 179 and 180).

178. वह केवल उसके स्नेह की भूखी है।

vah	*keval*	*us=ke*	*sneh=kī*	*bhūkh-ī*
3SG.NOM	only	3SG.OBL=GEN	affection(M.SG.OBL)=GEN	hungry-F.SG
hai				
be.PRS.3SG				

“She only wants his affection. (Lit. She is only hungry for his affection).”

179. मैं उनके बखान की भूखी नहीं हुँ, अपना बखान धरे रहें।

maiṁ	*unke*	*bakhān=kī*	*bhūkh-ī*	*nahīṁ*	*hūṁ*
1SG.NOM	3PL.OBL=GEN	praise(M.SG.OBL)=GEN	hungry-F.SG	not	be.PRS.1SG

“I don't want his praise (Lit. I'm not hungry for his praise).”

180. मगर फिर भी हम एक-दूसरे के खून के प्यासे थे।

magar phir	*bhī*	*ham*	*ek-dūsre=ke*	*khūn=ke*
but again	also	1PL.NOM	one-other=GEN	blood(M.SG.OBL)=GEN
pyās-e	*th-e*			
thirsty-M.PL	be.PST-M.PL			

“But still we were thirsty for each other’s blood (meaning ‘We wanted to kill each other’)”

The frequency of constructions expressing feelings of hunger and thirst attested in the corpus is given in Table 26. As the data show, Hindi does not display a preference toward one of the two constructions, but the presence of a Stimulus clearly correlates with the copular construction.

Table 26: Frequencies of the constructions encoding hunger and thirst in Hindi.

Construction	**No Stimulus**	**With Stimulus**	**Tot**
Dative construction	55	1	56
Copular construction	24	34	58
Total	79	35	114

This metaphorical mapping of volitions into sensations is frequent cross-linguistically. For example, in Chagga, a Bantu language of Tanzania (Emanatian 1995), the expression of desire, in particular sexual desire, is systematically linked to hunger (181). The same metaphorical linking is also present in Brazilian Portuguese (see example in 182 taken from Gibbs, Lima and Francozo 2004: 1200), English and in Italian.

181. *ngi'ichuo* *njàa* *(ia mndu mka)*
I feel hunger (for a woman)
"I'm desirous (of a woman)" (Taken from Emanatian 1995: 167).

182. *tenho* *sede* *de* *saber*
have-1SG.PRS thirst of knowledge
"I am thirsy for knowledge" (Taken from Gibbs, Lima and Francozo 2004: 1200).

This metaphor is interesting because the subdomains of bodily sensations and volitions are generally conceived as very different from a semantic point of view. They can be seen as constituting the opposite edges of the experiential domain: on the one hand, bodily sensations are the least controlled experiential types, on the other hand volitions are prototypically associated to intentionality and consequently they are typically conceptualized as showing some degree of control (Verhoeven 2007; Luraghi 2020a). However, these two subdomains can be conceived contiguous to the domain of bodily needs (Reh and Simon 1998: 42, Luraghi 2020a: 100). Bodily needs border with bodily sensations and partly with volitions and provide a conceptual link between these two experiential areas. A bodily need is based on a sensation and ultimately results in a desire: for instance, the sensation of hunger derives from the need for food and results in the desire to eat. Thus, a physical need can be understood as a desire based on a physical sensation and this allows bodily sensations to be used as source domains for the expression of the target domain of emotions.

6.5. Discussion

Bodily sensations are never expressed in Hindi by a transitive construction in which the Experiencer is encoded in the nominative/ergative. This is explained by the semantic properties of this experiential types which, as I already discussed in the introduction to this chapter, never share the features of the Agent. Moreover, the transitive construction displays two arguments while bodily sensations typically only profile one participant, that is the Experiencer. As a consequence, the most common constructions (Verhoeven 2007: 74-75) used by the languages of the world for the expression of

bodily sensations consist either of experiential adjectives in predicative functions with the Experiencer (or his/her body part) in the subject function or of intransitive constructions with a single argument in which the Expertum is lexicalized in the verb. In Hindi, bodily sensations can be expressed by various constructions, but only the dative construction is used for the encoding of every type of bodily sensations. This suggests that the dative construction is the most productive for the expression of sensations in the language, as it shows the highest type frequency in the corpus. The other constructions are more semantically constrained, as they seem to be used for the expression of specific bodily sensation types. The copular construction is only used for the expression of sensations of hunger and thirst, while the locative construction seems to be used to encode bodily sensations typically localized on a specific area of the body. The reverse transitive construction has a very limited scope and can encode only feelings of pain. Table 27 summarizes the constructions used to encode the sensation types investigated in this chapter and gives their absolute frequencies. The semantic maps in Figure 22 are intended to illustrate the functional distribution of these constructions over the experiential subdomains of bodily sensations. I would like to briefly focus on the constructional distribution of the expression of localized sensations (pain and itching) as I think that it can give us interesting insights on the interplay between semantics and syntax in the language and it allows me to introduce some tendencies that will appear as more evident in the following chapters. In particular, data in Table 27 show that, when we look at the expression of single sensation types from an onomasiological perspective, the construction with the highest type frequency does not necessarily correspond to the construction with the highest token frequency. This means that the productivity of a construction in the grammar is not necessarily related to the productivity of that construction with a single verb or set of verbs. In the case of pain and itching the higher frequency of the locative construction is explained by the fact that this is the most semantically fitting frame for the expression of localized sensations, while the dative construction shows a vaguer semantics generically associated to experience. As I will discuss in the following chapters this is not an isolated case and it frequently happens that constructions showing a lower productivity in the grammar show a higher token frequency with specific verbs. As I will argue, this is a consequence of the iconic encoding strategies of the language which results in the tendency to use frames that appear as more semantically coherent with the event expressed by the verb.

Table 27: Frequency of Hindi constructions covering the domain of bodily sensations.

	Pain/Itch	Temperature	Hunger/Thirst	TOT
Reverse transitive construction	4/75	-	-	**4/203**
Locative construction	52/75	-	-	**52/203**
Dative construction	19/75	13/13	56/114	**89/203**
Copular construction	-	-	58/114	**58/203**
TOTAL	**75**	**13**	**114**	**203**

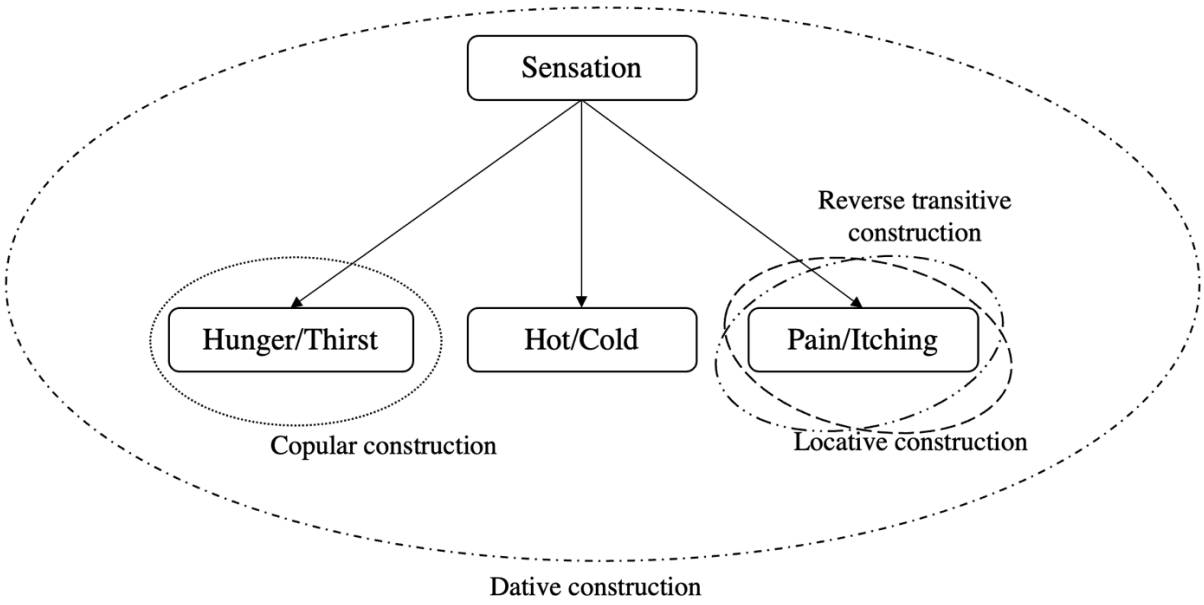

Figure 22: Functional distribution of Hindi constructions encoding bodily sensations.

All verbs occurring with the dative construction analyzed in this chapter are complex predicates formed by a nominal host which lexicalizes the Expertum and a light verb. The noun exhibits ambiguous behaviors (section 4.2.3.1): on the one hand, it forms a single constituent with the verb, contributing, for instance, to profiling the argument structure. On the other hand, it functions as an argument, as it agrees with the light verb. The light verb typically alternates between *honā* "be" and *lagnā* "adhere, be attached". The alternation reflects semantic differences about the way the event is construed (see on this Mohanan 1994, Butt and Geuder 2001, Montaut 2016): the verb *honā* contributes a stative semantics and it typically profiles the bodily sensation as a durative atelic condition in which the Experiencer experiences a state, while *lagnā* supply an inchoative reading, and it construes the event as telic and dynamic, hence as a change of state.

As I will discuss further in the next chapters, the verb *lagnā* appears in numerous experiential complex predicates. Besides the domain of bodily sensations, this verb is also found in subdomains related to emotions, and to a lesser extent it can also be used to encode perceptions and cognitions. This verb consistently appears with a dative Experiencer, as it always profiles non-agentive participants, and construes the experience as an event happening to the Experiencer. It may also encode a generic experience, and it is usually used in contexts in which an English speaker would use the verb *feel*. Consider for example sentence 157 above, repeated here in 183, in which *lagnā* expresses a generic physical feeling associated to sickness and fever.

183. रीढ़ में दर्द था और थकान की वजह से बुख़ार-सा भी लग रहा था, इसलिए मैं इतनी देर सोया रह गया।

rīṛh=meṁ *dard* *th-ā* *aur*
spine(M.SG.OBL)=in pain(M.SG.NOM) be.PST-M.SG and
thakān=kī_vajah=se *bukhār-sā* *bhī* *lag* *rah-ā*
tiredness(F.SG.OBL)=because_of fever-like.M.SG also attach PRGR-M
th-ā *islie maiṁ itnī der soyā rah gayā.*
be.PST.M.SG that is why I overslept for so long
"My spine was hurting and, due to tiredness, I was also feeling feverish, that is why I overslept for so long."

The Hindi verb *lagnā* is particularly interesting because of the wide range of meanings associated to it even when used outside the functional domain of experience. The basic meaning of the verb is "adhere" or "be attached" (McGregor 1994). However, depending on the context in which it occurs it can appear with significantly variable readings: as a consequence, dictionaries generally list dozens of different meanings associated to the entry *lagnā* (see for example the Oxford English Dictionary by McGregor 1984). Many meanings reported by Hindi dictionaries are triggered by lexical restrictions (Shapiro 1987), as they occur only in specific collocations and, therefore, give rise to complex predicates. When used as a full verb with its original semantics, *lagnā* is intransitive and occurs with a nominative first argument and an oblique second argument typically marked with a locative case. In sentence 184, *lagnā* is used with a metaphorical extension of its original meaning "attach" and it appears with an oblique second argument followed by the postposition *par* "on". In sentence 185, the verb is followed by the postposition *meṁ* "in". Sentence 186, which shows a dative first argument, is clear evidence of the extent of the semantic range of this verb.

184. गोबर की आंखें उसी गाय पर लगी हुई थी।
gobar=kī *ankh-eṁ* *us=ī* *gāy=par*

gobar=GEN eye(F)-PL.NOM that.OBL=EMPH cow(F.SG.OBL)=on

lag-ī *hu-ī* *th-īṁ*

attach-PRF.F be-PRF.F be.PST-3PL.F

"Lit. Gobar's eyes were attached to that cow."

185. शहर में आग लगी, हमें बंगला से निकलना पड़ा।

śahar=meṁ *āg* *lag-ī* *hameṁ*

town(M.SG.OBL)=in fire(F.SG.NOM) attach-PRF.F 1PL.DAT

bangl-oṁ=se *nikal-nā* *paṛ-ā*

bungalow(M)-PL.OBL=INS go_out-INF fall-PRF.M.SG

"The town was on fire (Lit. The fire attached in the town) We had to go out of the bungalow".

186. *samajh-n-e=meṁ* *mujhe* *das* *minaṭ* *lag-e*

understand-INF-OBL=in 1SG.DAT ten minute(M.PL.NOM) attach-PRF.M.PL

"It took me ten minutes to understand."

Additionally, the same verb may also be used with an auxiliary function (3.2.3.2). In this case, it follows the main verb, which is in the oblique infinitive form, and it expresses an inchoative reading, as in 187.

187. लज्जित नजरों से उसकी ओर ताकने लगा।

lajjit *najroṁ=se* *us=kī or* *tāk-n-e* *lag-ā*

ashamed eye(F)PL.OBL=INS 3SG.OBL=towards stare-INF-OBL attach-PRF.M.SG

"He started staring at him with ashamed eyes."

As mentioned above, the verb *lagnā* refers to a very broad semantic area with a central meaning, i.e. "physically attach" or "adhere". This meaning is extended to the encoding of emotions and bodily sensations based on the conceptual metaphor according to which sensations and emotions are entities that attach themselves to the Experiencer. From the physical proximity and contact between two material objects, the verb's meaning extended to the psychological and perceptual proximity between the Experiencer and the Expertum. From a diachronic point of view, it is difficult to trace the path of the semantic evolution of this verb from Old Indo-Aryan to modern Hindi. Shapiro (1987) was able to trace the root *lag-* back to Old Indo-Aryan, where it still retained its original meaning "attach" but was rarely used to encode experiences. Shapiro only found a case of experiential use in which the sequence *lag-* + *hṛdi* "heart" + genitive is used to express liking. This construction seems to foreshadow the use of *lagnā* for the expression of experiences in modern Indo-Aryan, however it

seems to be very marginal, as similar constructions do not seem to be attested in that period and in the later Middle Indo-Aryan phase. Shapiro hypothesizes that Hindi started making such extensive use of *lagnā* only after the restructuring of the Hindi verbal lexicon, which occurred after contact with Persian, when a huge number of noun-verb complex predicates entered the verbal lexicon (see on this also section 9.1.3). Today, the verb *lagnā* is used across various experiential subdomains, and, depending on the nominal host with which it occurs, it can encode different experiential types. As I discussed in this chapter, it is used in complex predicates expressing bodily sensations, such as "be hungry" *bhūkh lagnā* or "be thirsty" *pyās lagnā*, but also feeling of temperatures as "feel hot" *garmī lagnā* or "feel cold" *ṭhaṇḍ lagnā*. In these cases, the verb is mainly used to construe the event as an achievement, as opposed to the light verb *honā* alternating with it. As I will discuss more in detail in the next chapter (7.6), the same verb may also be used as a simple verb in dative Experiencer constructions for the expression of non-agentive perceptions, in particular when the perception refers to the senses of touch, taste and smell. Moreover, some cognitive situations as well may be expressed by this verb. For example, the noun verb complex predicate *patā lagnā* means "find out (with a non-agentive reading)". Additionally, the verb may also be used as a simple verb with a dative Experiencer followed by a complement clause and in this case, it encodes situations such as "It seems to me that…" or "It occurred to me that…". All these uses of the verb share the feature that the Experiencer is conceived as being completely non-agentive. Besides the experiential subdomains of sensations, perceptions and cognitions that are analyzed in this study, the verb *lagnā* may also be used to encode emotions. Consider for example sentence 188 and 189 in which the verb *ḍar lagnā* and *acchā lagnā* encode fear and liking respectively.

188. मुझे मौत से बहुत डर लगता है।

mujhe	*maut=se*	*bahut*	*ḍar*	*lag-t-ā*	*hai.*
1SG.DAT	death=INS	much	fear(M.SG.NOM)	adhere-IPRF-M.SG	be.PRS.3SG

"I am very afraid of death."

189. अब तुम्हारी यही बात मुझे अच्छी नहीं लगती।

ab	*tumhār-ī*	*yah=ī*	*bāt*	*mujhe*	*acch-ī*
now	2PL.GEN-F	this.NOM=EMPH	words(F.SG.NOM)	1SG.DAT	good-F
nahīṁ	*lag-t-ī.*				
not	adhere-IPRF-F.SG				

"Now, I don't like your words."

In section 8.8.2, I will try to give an overview of the productivity of this verb across the different experiential subdomains and argue that, even though it can be used to express situations referring to all subdomains, it is prototypically associated to bodily sensations. I will show that the distribution of this verb might be explained when considering the verb as referring to a very broad semantic area with the central meaning of physical attachment between two entities. I will discuss that this specific semantic component referring to physical contact makes the verb more suitable for the encoding of experiences in which the Experiencer is somehow physically affected, as in bodily sensations, and less suitable for situations in which the involvement of the Experiencer is less anchored to its physical feeling and is more mental.

7. Perceptions

7.1. The typology of the lexicalization of perceptions

Perceptive situations are related to the intake of external information through the five sense modalities: sight, hearing, touch, taste, and smell. The expression of perceptive situations has been the subject of many typological studies, focused in particular on the universality of patterns of polysemy and of mechanisms of metaphorical and metonymic extension in the semantic domain of perception verbs. Two main universals have been proposed. The first is Viberg's (1984) hierarchy of sense modalities that places the sense of sight at the top and that develops into a unidirectional path for semantic extension and lexicalization across the five sense modalities. The second is Sweetser's (1991) hypothesis according to which there is a crosslinguistic tendency to extend perceptive verbs to the expression of cognitive events. In particular, Sweetser suggests that typologically verbs expressing visual perceptions tend to be extended to the cognitive domain and to the expression of higher intellectual situations such as knowing and thinking more than verbs referring to other sense modalities. As has been pointed out by Evans, Nicholas and Wilkins (2000) these two theories refer to two distinct types of semantic extension. Viberg's theory focuses on intrafield extensions: meaning that both the source meaning and the target meaning of the metaphorical extension belong to the same domain (perceptive situations). Whereas Sweetser's theory involves extrafield extensions in which the target meaning belongs to a different semantic domain, i.e. cognitions (see on this also Matisoff 1978). Since the extrafield theory involves cognitions as the target domain, I will discuss Sweetser's theory in more detailed in the chapter dedicated to the expression of cognitive situations (see the introduction to chapter 8). In this chapter, I will address Viberg's intrafield proposal on the five sense modalities' hierarchy and discuss how the Hindi expression of perceptive events aligns with his typological remarks. I will limit my discussion on the extension from perception to cognition to the domain of evidentiality (see section 7.5).

Viberg (1983, 1984) investigated the lexicalization patterns within the semantic domain of perception verbs from a typological point of view. He examined data taken from a questionnaire on perception verbs of 53 languages representing 14 different linguistic groups and elaborated the modality hierarchy represented in Figure 23.

SIGHT > HEARING > TOUCH > SMELL/TASTE

Figure 23: Viberg's lexicalization pattern of verbs of perception.

The hierarchy should be interpreted as follows: a verb originally expressing a perception through a sense modality higher (to the left) in the hierarchy can be extended to express perceptions acquired through sense modalities lower in the hierarchy. So, verbs referring to sight can be used to express hearing and verbs referring to hearing can be used to express touch and so on. The hierarchy works only unidirectionally. This means for example that, as touch is before smell in the hierarchy, verbs referring to touch can extend their meaning to smell, but the opposite is not possible. In Italian, for example the same verb *sentire* "hear, feel" is used to encode perceptions referring to hearing, touch, smell and taste, but not to sight.

As Viberg remarks, this hierarchy may not always be applied contiguously, which means that a particular sense modality can be omitted in some languages. The non-contiguity also implies that some intermediate sense modalities can be skipped in the semantic extension still maintaining the directionality of the hierarchy. For this reason, Viberg elaborates a more complex network of shifts, which is represented in Figure 24 (1984: 147).

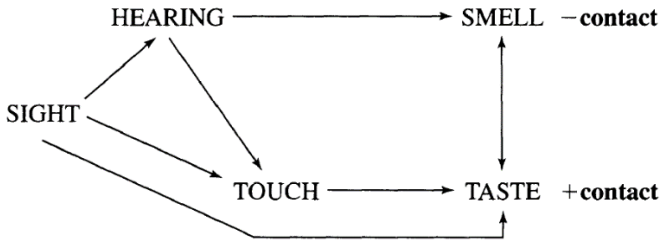

Figure 24: Viberg's lexicalization network of verbs of perception (1984: 147).

The sense of sight has a primacy over the other senses not only because it is the most prone to extending its usage to the expression of other sense modalities, but also because it is the most lexically complex one across languages. For example, it is the most susceptible to the proliferation of synonyms and hyponyms. This complexity is not restricted to the lexico-semantic level, but also involves the syntactic level. As I will show, Hindi tends to be highly specific in its lexicalization patterns and presents different verbs for each sense modalities (see Table 29 below). However, as we will see, verbs of visual perception show the widest constructional variation among perception verbs. Hindi thus seems to align with Viberg's typological observations according to which visual perception shows the highest complexity at all linguistic levels.

7.2. The semantic properties of perceptions

Perceptions concern two participants: the Experiencer that perceives the situation and the Stimulus that is both the trigger of the experience and its content. Following previous scholars (Viberg 1984,

Usoniene 1999), I distinguish two types of perceptive verbs: agentive and non-agentive. In non-agentive perceptions, the Experiencer does not control the event and is not volitional. In this type of situation, the perceiver is completely passive, and it is affected by the experience because of humans' involuntary predisposition to perceive physical triggers through senses. For this reason, this Experiencer resembles a Recipient/Beneficiary rather than an Agent: it lacks control and volitionality, and at the same time receives the consequences of the event (as in the English example 190). In agentive perceptions the Experiencer shares the semantic properties of a prototypical Agent: it controls the event and is typically volitional. These situations feature an Experiencer that starts the event and controls it, as in 191, and are on the border between experiential events and activities.

190. *She was driving when she saw a deer crossing the street.*

191. *She was looking at the dogs playing outside.*

Besides a higher degree of agentivity, agentive perceptions also imply a lower degree of affectedness on the Experiencer in comparison with other experiential verbs. As Kemmer (1993: 137) points out, "simply perceiving an object has a relatively small effect on the perceiver, as compared to thinking about the object or being affected by emotions aroused by it". Given this semantic characterization, this verb class is typically located quite close to the transitive prototype in the transitivity hierarchy (see Tsunoda 1985, 2015 or Malchukov 2005, 2015). Indeed, agentive perceptions are semantically similar to some verb classes in the upper hierarchy of Malchukov's transitivity scale. In particular, they share many semantic properties with pursuit verbs, such as *follow*, *search for*, and so on. Both verb classes imply a volitional first participant that controls the event and a not affected second participant. This led many scholars (Dixon 1991, Lazard 1989) to emphasize a similarity between the argument structure of verbs of agentive perception and that of pursuit verbs. Non-agentive perceptions, in contrast, are more similar to emotions and bodily sensations and can be associated to the verb classes belonging to the lower hierarchy in Malchukov's scale.

As Luraghi (2021: 25) points out, besides the parameter of control, another parameter contributes to distinguish agentive and non-agentive perceptions: i.e, actionality. Generally, two groups of perceptions can be singled out, *uncontrolled states* (such as *see* or *hear*) and *controlled activities* (such as *look at* and *listen to*). The same distinction is made by Croft (2012) who defines experiential predicates such as *see* and *hear* as mental state verbs, and verbs such as *listen to* or *look at* as inactive actions, i.e. "atelic actions that imply that the Agent experiences somethings, but they also emphasize the fact that s/he volitionally attends the Stimulus" (Croft 2012: 156). Notably, however, verbs encoding uncontrolled states such as *see* and *hear* can also construe the event as an achievement. This seems to be supported by the aspectual distribution of some Hindi verbs. As we

will see in the following sections (7.3), specific semantic properties of the Experiencer seem to correlate with a specific aspectual characterization with some verbs. For example, *tāknā* "observe, stare", which encodes highly agentive visual perceptions, correlates with the imperfective aspect, thus indicating that it basically refers to atelic actions (see section 7.3.1). In contrast, the verb *dikhāī denā* "see", which refers to non-agentive perceptions, occurs both in the perfective and imperfective aspect with a comparatively similar distribution, thus indicating that this verb allows for both an achievement and a stative interpretation.

Languages of the world vary in the way they encode these semantic differences (Usoniene 1999). In some languages, lexical means seem to dominate while other languages prefer to use different morpho-syntactic encodings. In Italian, for example, different semantic properties are lexicalized by different verbs. The verb *guardare* "look at" encodes agentive visual perceptions, as in 192, while the verb *vedere* "see" encodes non-agentive visual perceptions, as in 193. Similarly, the verb *ascoltare* "listen to" refers to agentive auditory perception, while the verb *sentire* "hear" expresses non-agentive auditory perceptions. Italian does not use constructional means to distinguish between different degrees of agentivity on the Experiencer. As I will discuss, Hindi uses both strategies: it uses different verbs for the expression of different degrees of agentivity, but it also associates them with different constructions.

192.	*la*	*madre*	*guarda*	*il*	*figli-o*	*gioc-are.*
	The.F.SG	mother (F)	look_at-3SG.PRS	the.M.SG	son(M)-SG	play-INF
	"The mom watches her son play."					

193.	*Sara*	*ha*	*visto*	*un*	*topo*	*in*	*camera.*
	Sara	have.3SG.PRS	see.PRF.PRT	a	mouse	in	room
	"Sara saw a mouse in her room."						

A classification similar to those given by Luraghi and Croft was proposed in 1984 by Viberg, who adds a third class (i.e. *sensory copulas*; see on this also Viberg 2008, 2015). He distinguished between verbs of controlled activities, verbs of uncontrolled experiences and source-based copulas. The first type (controlled activities) profiles the perceptive event as an unbounded process controlled by an agentive Experiencer (as in *look at* or *listen to*). The second type refers to a state or inchoative achievement that is not controlled (uncontrolled states or achievements such as *see* or *hear*). The first and the second types belong to a macro-class that Viberg defines Experiencer-based perceptions, because they consist of verbs in which the "syntactic base" (i.e. the subject) refers to the Experiencer. The third verb class is defined as Source-based, as it consists of verbs profiling the event starting

from the Stimulus/Expertum and leaving the Experiencer in the background. Following Biber at al. (1999), Viberg (2015: 99) calls verbs belonging to this class *sensory copulas* since they take subject predicative complements and, at the same time, they refer to one of the sense modalities, as in "The soup smells good". By combining the five sense modalities and the three event-types based on agentivity and actionality, Viberg elaborates the three-by-five grid represented in Table 28 (adapted from Viberg 1984). In this chapter, I will mainly refer to this classification while discussing Hindi lexicalization of perception events.

Table 28: The basic classification of verbs of perception (Viberg 1984).

	Base selection		
	Experiencer-based		**Source-based**
Dynamic system	**Experience (state/inchoative)**		**Copulative (state)**
Sense modality			
sight	Peter looked at the birds	Peter saw the birds	Peter looked happy
hearing	Peter listened to the birds	Peter heard the birds	Peter sounded happy
touch	Peter felt the cloth	Peter felt a stone under his foot	The cloth felt soft
taste	Peter tasted the food	Peter tasted garlic in the food	The food tasted good/bad
smell	Peter smelled the cigar	Peter smelled cigars in the room	Peter smelled good/bad

In addition to the verb classes given in the table, Hindi displays another group of verbs referring to a Source-based way of profiling the event. These verbs in Hindi refer to the possibility to perceive something and they require a single argument encoding the Stimulus, while the Experiencer is left in the background. These verbs are present also in other languages, for example in Swedish (Viberg 2008) and in Finnish (Huumo 2010) and are usually referred to as perceptibility verbs (see also the class of Light, Sound and Smell emission verbs in Levin 1993: 233). Typologically perceptibility verbs are related to Experiencer-based perception verbs, as they are often derived morphologically from the same lexical roots of Experiencer-based verbs. This is actually a quite common phenomenon: in English, for example, there is no lexical differentiation among controlled activities, uncontrolled experiences and sensory copulative states with respect to taste, as 194, 195 and 196 clearly show.

194. Agentive Experiencer-based perception
Taste the soup and tell me if there is enough salt.

195. Non-agentive Experiencer-based perception
I can taste ginger in the soup.

196. Sensory copulative state
Your soup tastes exactly like my grandmother's.

Table 29 shows the lexicalization grid of perception verbs in Hindi. The columns represent the four perception-types discussed above, while the rows show the five sense modalities. As shown in the Table, the verbs expressing generic visual perception (*dekhnā*) and generic auditory perception (*sunnā*) can be used to encode both controlled activities and uncontrolled experiences. However, they show a different complexity from a constructional perspective. As I will discuss, *sunnā* does not use constructionist ways to distinguish these two semantics; in contrast, *dekhnā* may occur in two distinct constructions that differ with respect to the semantic properties they contribute. As Table 29 shows, Hindi also exhibits verbs that can only allow for non-agentive reading, such as *dikhāī denā* and *sunāī denā*. These verbs can also be used to encode copulative states and perceptibility events. The sense modality referring to touch does not show a specific verb for the expression of uncontrolled experiences, copulative states and perceptibility events. Such perceptions are typically expressed by the verb *lagnā* "adhere, be attached". As I will discuss in section 6.7. *lagnā* refers to generic feelings and is also used to express non-agentive perceptions referring to the smell and taste.

Table 29: The lexicalization grid of perception verbs in Hindi.

	Controlled activities	**Uncontrolled experiences**	**Copulative states**	**Perceptibility events**
Sight	*dekhnā / tāknā*	*dekhnā /dikhnā / dikhāī denā / najar ānā*	*dikhnā / dikhāī denā / najar ānā*	*dikhnā / dikhāī denā / najar ānā*
Hearing	*sunnā*	*sunnā /sunāī denā*	*sunāī denā*	*sunāī denā*
Touch	*chūnā / sparś karnā*	*lagnā*	*lagnā*	*lagnā*
Taste	*cakhnā/ svād lenā*	*svād milnā/ lagnā*	*svād honā / lagnā*	*svād honā / lagnā*
Smell	*sūṁghnā*	*khuśbū ānā / lagnā*	*mahknā / lagnā*	*khuśbū ānā / lagnā*

7.3. Presentation of Hindi data

In the following sections, I will present data extracted from the corpus and analyze it, focusing primarily on the argument structure of verbs of perception and their aspectual characterization. The

corpus shows uneven attestation of verbs of perception. Visual and auditory perceptions are highly frequent, with verbs related to visual perception displaying the highest absolute frequency. While verbs referring to the other three sense modalities appear less frequently.

Since verbs referring to sight and hearing are the most frequent and exhibit the gratest complexity, both lexically and constructionally, I will first present data for these two verb classes. In section 7.3.1, I address the verb expressing generic visual perception in Hindi, while in section 7.3.2, I will deal with the verb expressing generic auditory perception. Section 7.3.3 focuses on the expression of non-agentive perceptions, perceptibility, and sensory copulative states in Hindi. As perception verbs referring to sight and hearing frequently encode their stimuli through complement clauses, I dedicate section 7.5 to the analysis of complementation with this verb class. In section 7.6, I then move on to discussion of verbs referring to the other three sense modalities – touch, taste and smell.

Table 30 shows the Hindi verbs of perception analyzed in this study with their absolute frequency in the corpus and the number of occurrences I manually scrutinized for each verb. Given that the data for the sense modalities of touch, taste and smell were too scarce in the literary corpus, I had to rely on the hiTenTen corpus as well for my analysis. As the data were scarce and sourced from different corpora, I decided to limit my discussion to a qualitative analysis, for this reason I do not provide quantitative data for these three sense modalities.

Table 30: Hindi verbs of perception analyzed in this study and their absolute frequency.

Meaning	MSH verb	Frequency in Corpus	Manual scrutiny
appear/see	*dikhnā*	11	11
	dikhāī denā	197	197
	dikhāī paṛnā	3	3
	najar ānā	257	257
see/look at	*dekhnā*	4317	200
show	*dikhānā*	458	200
look at/stare	*tāknā*	230	230
be heard	*sunāī denā*	53	53
be heard	*sunāī paṛnā*	6	6
hear/listen	*sunnā*	902	200

touch	*chūnā*	-	-
	sparś karnā	-	-
taste	*cakhnā*	-	-
	svād lenā	-	-
	svād milnā	-	-
smell	*khuśbū ānā*	-	-
touch/smell/taste	*lagnā*	-	-

7.3.1. Dekhnā: the generic verbs for visual perceptions

In Hindi visual perception is expressed by several verbs. The verb most frequently found in the corpus is *dekhnā,* which can be translated either as "look at" or as "see" depending on the context. This verb is used in Hindi to encode a general visual perception and it is not specified either according to the properties of the Experiencer (agentivity) or according to the properties of the event (lexical aspect). Table 31 shows the distribution of Hindi aspectual forms over the occurrences of the verb *dekhnā* that I manually scrutinized: as the data shows, there is no specific aspectual characterization and the verb can appear with all aspectual forms.

Table 31: Aspectual distribution of the verb *dekhnā*.

	IPRF							
Construction	**IPRF**	**PRGR**	**CONT**	**PRF**	**CP**	**FUT**	**IMP**	**TOT**
Transitive	22	20	8	47	3	4	7	111
Oblique Stimulus	1	0	4	17	2	0	2	26
Finite Compl. clause	9	3	2	10	0	2	1	27
Predicative participle	4	0	0	14	3	1	0	22
Predicative adjective	0	0	0	0	0	1	0	1
Locative adverbial	0	0	1	6	0	0	1	8
Passive	4	0	0	0	0	1	0	5
Total	**40**	**23**	**15**	**94**	**8**	**9**	**11**	**200**

As the data in Table 31 show, *dekhnā* most frequently appears with a transitive construction, in which the Experiencer is marked with the ergative/nominative, while the Stimulus is a direct object. Sentences 197 and 199 below are two examples of this construction. In 197, the verb is in the

imperfective, hence the Experiencer is in the nominative, while the Stimulus is marked with the accusative *ko*. In sentence 198, the verb is in the perfective aspect and the Experiencer is marked with the ergative *ne*. As the Stimulus is expressed by the third person singular pronoun in the accusative, the verb cannot agree with any element in the sentence, and it is in the default masculine singular form.

197. मैं चुपचाप बैठा सडक़ से गुज़रती हुई गाडिय़ों को देखता रहा।

maiṁ	*cupcāp*	*baiṭh-ā*	*saṛak=se*	*guzar-t-ī*
1SG.NOM	silently	sit-PRF.M.SG	street(F.SG.OBL)=INS	go_by-IPRF-F

hu-ī	*gāṛi-yoṁ=ko*	*dekh*	*rah-ā*	*th-ā*
be.PRF.F	car(F)-PL.OBL=ACC	look_at	PRG-M.SG	be.PST-M.SG

"I kept looking at the cars running down the street sitting in silence."

198. जमादार ने उसे सिर से पांव तक देखा।

jamādār=ne	*use*	*sir=se*	*pāṁv=tak*
jamadar=ERG	3SG.ACC	head(M.SG.OBL)=INS	foot(M.SG.OBL)=to

dekh-ā.
look_at-PRF.M.SG

"Jamadar looked at him from head to foot."

This verb is also found in an oblique Stimulus construction in which the Experiencer is consistently marked with the nominative/ergative, while the Stimulus is in the oblique case and followed by the postposition *kī or* or *kī taraf* meaning "towards" (199). When in the perfective aspect, the verb cannot agree with either the ergative Experiencer or the oblique Stimulus and consequently it appears in the default masculine singular form, as in 200.

199. रतन एक क्षण तक छत की ओर देखती रही ।

ratan	*ek*	*kṣan=tak*	*chat=kī or*	*dekh*
Ratan.NOM	one	moment=till	ceiling(F.SG.OBL)=towards	look_at

rahī
PRG-F.SG

"Ratan was staring at the ceiling."

200. ओंकारनाथ ने दु:खी आंखों से पत्नी की ओर देखा।

oṁkārnāth=ne *du:kh-ī* *ānkh-oṁ=se* *patnī=kī_or*
Onkarnath=ERG sad-F eyes(F)-PL.OBL=INS wife(F.SG.OBL)=towards
dekh-ā.
look_at-PRF.M.SG
"Onkarnath looked at his wife with sad eyes."

As mentioned above, the verb *dekhnā* is not specified with respect to the agentivity of the Experiencer. However, interestingly, the non-agentive reading is only allowed when *dekhnā* occurs with a transitive construction (as in 201) while it is ruled out when the verb occurs in the oblique Stimulus construction.

201. उनमें से एक ने कहा-'हम लोग इधर से होकर जा रहे थे तो हमने इस कबर से चिल्लाने की आवाज निकलती हुई सुनी, और जब अन्दर आये तो तुम्हें पृथ्वी पर अचेत पड़े देखा।

us meṁ se ek ne kahā– "ham log idhar se hokar jā rahe the to
*"*One of them said «We were passing by, when"
ham=ne *is* *kabar=se* *cillā-n-e=kī*
1PL=ERG this.OBL tomb(M.SG.OBL)=INS scream-INF-OBL=GEN
āvāj *nikal-t-ī* *hu-ī* *sun-ī,* *(aur jab andar āye to)*
sound(F.SG.NOM) exit-IPRF-F.SG be.PRF-F hear-PRF.F.SG (and when we came inside)
tumheṁ *pr̥thvī=par* *acet* *paṛ-e* *dekh-ā".*
2SG.ACC ground(F.SG.OBL)=on unconscious lay-PRF.M.PL see-PRF.M.SG
"One of them said 'We were passing by, when we heard screams coming from this tomb and when we entered, we saw you lying unconscious on the ground.'"

In sentence 201 two perceptions are encoded: a visual perception expressed by the verb *dekhnā*, and an auditory perception expressed by the verb *sunnā*. Both verbs occur in a transitive construction. The Experiencer *ham=ne*, shared by the two verbs is in the ergative, while the two Stimuli are encoded as direct objects. The Stimulus of the verb *sunnā* (*āvāj* "sound, voice") is in the nominative, while the Stimulus of *dekhnā* is in the accusative because of DOM (*tumheṁ* "you.ACC"). The semantic property of non-agentivity is evident from the context. The clause *ham log idhar se hokar jā rahe the to* "we were passing through here when", which precedes the experiential constructions, implies that the Experiencer did not expect to hear screaming and find a person lying on the floor.

When occurring with an oblique Stimulus, *dekhnā* can only express agentive perceptions, as the postpositions *kī or* and *kī taraf* "towards" imply directionality and consequently always construe a controlled activity. Interestingly, in this construction, a difference in the semantic properties of the

Experiencer (agentivity) is marked on the Stimulus. As mentioned above in this chapter, several scholars (Dixon 1991, Lazard 1989) pointed out that perception verbs can be assimilated to the semantic class of pursuit verbs. In Tsunoda's Implicational Hierarchy of Transitivity (Tsunoda 1985, 2015), for example, perception verbs are located between the prototype (Direct effect on patient) and the class of pursuit verbs. Remarkably, when discussing the patterns typically associated to pursuit verbs cross-linguistically, Malchukov (2015) concludes that the nominative-oblique pattern (or the ergative-oblique depending on the alignment type of the language) is the most semantically fitting frame for the encoding of this verb class. The fact that agentive perception verbs and pursuit verbs patterns with respect to the semantically based argument structures support the assumption that these two classes are semantically similar.

In sum, even if the verb *dekhnā* does not distinguish between agentive and non-agentive perceptions, the grammar offers a constructional way to keep them distinct. The transitive construction is not semantically specified and is used to express both agentive and non-agentive perceptions, while the oblique Stimulus construction is used to express agentive perceptions. Notably, the oblique Stimulus construction is also frequently found with verbs that lexicalize the agentivity of the Experiencer on the root, such as *tāknā* "observe, stare". Table 32 reports the frequencies of the two constructions occurring with this verb: as the data shows, *tāknā* strongly prefers the oblique Stimulus construction. The fact that this construction correlates with highly agentive verbs supports the hypothesis that this pattern is semantically motivated and implies control by the Experiencer.

Table 32: Relative frequencies of the transitive construction and the oblique Stimulus construction with the verb *tāknā*.

Construction	**Frequency**
Transitive construction	82
Oblique Stimulus construction	148
Total	**230**

In 202 *tāknā* occurs in a transitive construction, while in 203 it appears in an oblique Stimulus construction. Sentence 203 clearly expresses a highly agentive perception as shown by the presence of the instrumental *āṁsū bharī āṁkhoṁ=se* "with eyes full of tears", as the presence of instrumentals is typical of prototypical agentive events (see on this Croft 2022). Interestingly the encoding of an instrument seems to correlate with the oblique Stimulus construction, as 66% of the occurrences of an instrumental adjunct with the verb *tāknā* co-occur with an oblique Stimulus.

202. ठाकुर हैरत में आकर उसका मुँह ताकने लगा।

ṭhākur *hairat=meṁ* *ā=kar* *us=kā* *muṁh*
thakur.NOM wonder=in come=CP 3SG.OBL=GEN face(M.SG.NOM)
tāk-n-e *lag-ā*
stare-INF-OBL start-PRF.M.SG
"The thakur was surprised and started staring at his face."

203. चौंककर आँखें खोलीं तो रानी सचमुच सामने खड़ी उसकी तरफ आँसू भरी ऑखों से ताक रही थीं।

cauṁk=kar āṁkheṁ kholīṁ to
When he opened his eyes with surprise
rānī *sacmuc* *sāmne* *khaṛ-ī* *us=kī taraf*
queen(F.SG.NOM) really ahead stand-PRF.F 3SG.OBL=towards filled-F
āṁsū *bhar-ī* *āṁkh-oṁ=se* *tāk* *rah-ī* *th-īṁ.*
tear filled-F eye(F)-PL.OBL=INST observe PRG-F.SG be.PST-F.PL
"When he opened his eyes with surprise, the queen was actually standing in front of him, and she was staring at him with eyes full of tears."

Interestingly, unlike *dekhnā* which, is used to express a generic visual perception and does not show preferences with respect to aspectuality, *tāknā* seems to have a specific aspectual characterization. Table 33 shows the distribution of the aspectual forms over the occurrences of the verb *tāknā* that I manually scrutinized. This verb occurs very rarely with the perfective aspect (only three occurrences over 178 sentences) and it mostly appears in the progressive (204), the durative (205) and the imperfective aspect (206). The data suggests that the agentivity of the Experiencer lexicalized in the verb *tāknā* correlates with a durative reading. The semantic properties of the participants and the aspectual characterization of the verb thus contribute to construe visual perception as an atelic controlled action.

204. मेहता उसकी ओर भक्तिपूर्ण नेत्रों से ताक रहे थे।

mehtā *us=kī or* *bhaktipūrṇ* *netr-oṁ=se* *tāk* *rah-e*
mehta.NOM 3SG.OBL=towards devotion_full eye(M)-PL.OBL=INS stare PRG-M.PL
th-e.
be.PST-M.PL
"Mehta was looking at her with devotional eyes."

205. रतन उसके मुंह की ओर अपेक्षा के भाव से ताकती रही, मानो कुछ कहना चाहती है और संकोचवश नहीं कह सकती।

ratan *us=ke* *muṁh=kī or* *apekṣā=ke* *bhāv=se*
ratan.NOM 3SG.OBL=GEN face=towards expectation=GEN expression=INS
tāk-t-ī *rah-ī,*
stare-IPRF-F PRG-F
māno kuch kahnā cāhtī hai aur saṁkocavaś nahīṁ kah saktī.
as if she wanted to say something and couldn't due to hesitation
"Ratan kept looking at his face with expectation, as if she wanted to say something and couldn't due to hesitation."

206. इतनी ही एकाग्रता से वह कदाचित आकाश की काली, अभेध मेघ-राशि की ओर ताकता !

itnī *hī* *ekāgrtā=se* *vah* *kadācit* *ākāś=kī*
as_much EMPH concentration=INS 3SG.NOM sometimes sky=GEN
kālī, *abhedh* *megh-rāśi=kī or t* *tāk-t-ā!*
black-F impenetrable cloud=towards stare-IPRF-M.SG
"With the same concentration he probably gazes towards the black, impenetrable clouds in the sky!"

Table 33: Aspectual distribution of the verb *tāknā*.

Construction	IPRF			INC	PRF	FUT	CP	Tot
	Dur	Hab	Prgr					
Oblique Stimulus	26	22	28	29	2	1	1	149
Transitive	6	12	1	10	1	1	0	34
No Stimulus	5	6	1	1	0	0	1	21
Locative adverbial	2	4	4	0	0	0	1	14
Total	**39**	**44**	**34**	**40**	**3**	**2**	**3**	**178**

7.3.2. Sunnā: the generic verbs for auditory perceptions

The verb for generic perception in Hindi is *sunnā* which can be translated as both "hear" and "listen to". Like the generic verb expressing visual perception (*dekhnā*), *sunnā* does not seem to be associated to a specific aspectual distribution and it occurs in the corpus with all Hindi aspectual forms. When the Stimulus in expressed by an NP this verb only features a transitive construction and does not allow an oblique Stimulus of the type employed by the verb *dekhnā*. In 207 *sunnā* is in the

imperfective aspect, while in 208 it is in the perfective aspect, the Experiencer is marked with the ergative and the verb agrees in gender and number with the feminine noun *bāt* “words, discourse”.

207. गाँव-वालो की फरियाद कौन सुनता!

gāṁv-vāl-oṁ=kī	*phariyād*	*kaun*	*sun-t-ā!*
villager-PL.OBL=GEN	complaint(F.SG.NOM)	who.NOM	listen-IPRF-M.SG

“Who listens to the complaints of the villagers?”

208. देवीदीन ने दरोगा की बात सुनी।

devīdīn=ne	*dārogā=kī*	*bāt*	*sun-ī*
devidin=ERG	inspector(M.SG.OBL)=GEN	words(F.SG.NOM)	listen-PRF.F.SG

“Devidin listened to the inspector's words.”

Interestingly, this verb can appear sometimes in the corpus with a genitive Stimulus. Consider for example sentence 209 and 210, in which Stimulus refers to a human participant and it is encoded with a genitive pronoun: *merī* “my (1SG.GEN-F)” in 209, and *us=k-ī* “her (3SG.OBL=GEN-F)” in 210.

209. उसके सामने मेरी कौन सुनेगा ?

us=ke sāmne	*mer-ī*	*kaun*	*sun-e-g-ā.*
3SG.OBL=in front of	1SG.GEN-F	who.NOM	listen_to-3SG-FUT-M

“Who will listen to me in front of him.”

210. मुन्नी मना कर रही है पर कोई उसकी सुन नहीं रहा।

munnī	*manā*	*kar*	*rah-ī*	*hai*	*par*	*koī*
munni.NOM	refuse	do	PRG-F.SG	be.PRS.3SG	but	INDF.PRN.NOM
us=k-ī		*sun*	*nahīṁ*	*rah-ā.*		
3SG.OBL=GEN-F		listen	not	PRG-M.SG		

“Munni is refusing but no one is listening to her.”

An apparently similar construction occurs in Old Indo-Aryan and in many ancient Indo-European languages in general, where verbs expressing hearing allow an alternating marking on the Stimulus between the accusative and the genitive. However, as I will discuss in more details below in this section, this similarity is only apparent, as in ancient languages the opposition between the genitive

and the accusative bring semantic differences that are not associated to the Hindi genitive Stimulus. In example 211 (adapted from Dahl 2009: 42), the Early Vedic verb *śrav-* "hear, listen to" appears twice with a genitive Stimulus: *śyāvāśuvasya* "Śyāvāśuva" and *atreḥ* "Atri". Luraghi, Caviglia and Pinelli (2014) and Dahl (2014) note that such verbs of perception primarily tend to select the genitive Stimulus construction when the Stimulus refers to an animate or human participant, while when the Stimulus refers to an inanimate or abstract entity, these verbs usually receive the accusative case marking. Other perception verbs that can take a genitive stimulus in Early Vedic are *cet-* "perceive, take notice of" and *ved-* "find, learn, know" (Dahl and Fedriani 2012).

211. *śyāvāśuvasya* *sunvatas* *tathā* *śr̥ṇu* *yathā*
śyāvāśuva.GEN extract.PRS.PRT.GEN thus hear.2SG.PRS.IMP like
aśr̥ṇor *atreḥ* *karmāṇi* *kr̥ṇvataḥ*
listen.2SG.IPRF Atri.GEN sacred_deeds.ACC perform.PRS.PRT.GEN
"Listen thus to Śyāvāśuva, who is extracting (soma), as you listened to Atri, who was performing sacred deeds." (RV VIII 36.7 adapted from Dahl 2009: 42)

The genitive vs accusative case marking alternation on the second argument is quite common in ancient Indo-European languages and, according to previous literature (see Jamison 1976, Hettrich 2014), the genitive marking on the second argument indicates different degrees of affectedness as it implies that the action affects the Patient only partially (as in 212 and 213 taken from Jamison 1976: 129).

212. *pibā* *somam*
drink.Imp soma.Acc
'Drink soma.'
(Rgveda VIII.36.1, from Jamison 1976: 129)

213. *pibā* *somasya*
drink.Imp soma.Gen
'Drink (of) soma.'
(Rgveda VIII.37.1, from Jamison 1976: 129)

Luraghi (2020) notes that the same alternation is very frequent with verb of auditory perception in Homeric Greek as well and discusses how the partitive meaning of the genitive can function with verbs of perception, which do not imply an affected Patient. She concludes that in the case of experiential verbs the genitive marking points toward less agency and control by the first participant

and it is opposed to the transitive construction that expresses a full degree of agency. Fedriani (2012: 84) discusses the same genitive marking on the Stimulus of hearing verbs in Latin and she concludes that "the use of the genitive case is neither triggered by inherent properties of the Stimulus or by relational determinants such as the degree of control or agentivity projected by the Experiencer, but rather by event-based properties to do with the conceptualization of the state of affairs in terms of temporal boundaries and the (ir)realis character of the verbal process". A similar conclusion is proposed by Dahl (2009), who discusses the genitive marking on the Stimulus of auditory perception verbs in Vedic and argues that the genitive is semantically in contrast with the accusative, as the first is used to convey atelicity and imperfectivity, while the latter characterizes the event as more telic and perfective. According to Dahl, this semantic difference correlates with the fact that the genitive is typically used to mark human Stimuli, hence atelicity would derive from the fact that the act of listening to a person does not generally imply an endpoint to the event, as opposed to the act of listening or hearing something.

However, as I mentioned above, I do not believe that the genitive Stimulus construction in Hindi can be assimilated to the semantically based alternations typical of ancient Indo-European languages and in particular Vedic. The genitive Stimulus construction occurs rarely in the corpus (only three occurrences over the 200 occurrences scrutinized) and is actually quite peculiar in the language. Hindi simple verbs usually do not allow one of their arguments to be marked with the genitive. Moreover, interestingly, the genitive postposition here is in the feminine, even if there is no feminine noun to trigger agreement. Cases in which a simple verb occurs with a feminine agreement in appearance inexplicable are found elsewhere in the language but seems to be marginal and restricted to specific verb classes. In particular, as I will discuss widely in Chapter 8, this behavior seems to be associated with verbs that tend to occur with Stimuli instantiated by propositional contents, and for this reason it also occurs with verbs of cognitions. As I will argue the genitive postposition is probably explained by the presence of the not-overtly expressed feminine noun *bāt* "words, speech, discourse, matter": which is frequently used with verbs such as *listen* (as in 214), *think* (as in 215) and *remember* (as in 216).

214. मैं तुम्हारी बात सुन रहा हुँ।

maiṁ	*tumhār-ī*	*bāt*	*sun*	*rah-ā*	*hūṁ*
1SG.NOM	2PL.GEN-F	words(F.SG.NOM)	listen	PRG-M.SG	be.1SG.PRS

"I am listening to your words."

215. तब वह पालम सन्त की बातों पर विचार करता हुआ धीरे धीरे अपनी कुटी की ओर चला।

tab	*vah*	*pālam*	*sant=kī*	*bātoṁ=par*	*vicār*
then	3SG.NOM	palam	sant=GEN	words(F.PL.OBL)=on	thought(M.SG.NOM)
kar-t-ā		*hu-ā*		*dhīre dhīre apnī kuṭī kī=or calā.*	
do-IPRF-M.SG		be.PRF-M.SG		he slowly started walking towards his hut.	

"Then, he slowly started walking towards his hut thinking about the words of Saint Palam."

216. मुझे इससे सुरजीत की कही हुई बात याद हो आयी।

mujhe	*is=se*	*surjīt=kī*	*kah-ī*	*hu-ī*
1SG.DAT	this.OBL=INS	surjit=GEN	say-PRF.F	be.PRF-F
bāt	*yād*	*ho*	*ā-yī*	
words(F.SG.NOM)	memory.F.SG.NOM	be	come-PRF.F	

"This reminded me of something Surjit said."

This assumption is supported by the feminine agreement of the verb, which otherwise would remain unexplained. Additionally, it can also explain why verbs of visual perception such as *dekhnā* "see, look at" or *tāknā* "stare, observe", which typically do not occur with propositional arguments, are not found with the feminine genitive construction in the corpus. In sum, I will not treat occurrences with a genitive marking of the Stimulus as a distinct construction, but I will consider them as instantiations of the transitive pattern with an unexpressed object.

7.3.3. Non agentive visual and auditory perceptions

Besides the verbs analyzed in the previous sections, non-agentive perceptions can also be expressed by other verbs in Hindi. Visual non-agentive perceptions can be expressed with two distinct, but etymologically related, verbs: the simple verb *dikhnā* "appear, see" and the noun-verb complex predicate *dikhāī denā* lit. "give seeing" (formed by the noun *dikhāī* "seeing" and the light verb *denā* "give"). Auditory non agentive perceptions, on the other hand, are expressed by the complex predicate *sunāī denā* lit. "give hearing" (formed by the noun *sunāī* "hearing" and the verb *denā* "give"). As I will discuss in more detail in section 7.4 below, *dikhnā* is the morphological anticausative of the verb *dekhnā*. The nouns *dikhāī* and *sunāī*, instead, each consist of a verb stem (*dikh* 'appear to' and *sun* 'hear') that is causativized via the addition of the causative morpheme *-ā* and is further nominalized via the feminine nominalization affix *-ī* (Butt, Carnesale and Ahmed 2023). Both the causative and the nominalization morphemes are productive, yet the nominalized causative is no longer productive in the language. Some fixed expressions still exist but are rare: see for example *cāṛh-āī* "climb, ascent", *līp-āī* "painting", *laṛ-āī* "fight", *lūṭ-āī* "plundering", *pāṛh-āī* (Kachru 1980, Saksena 1982).

Besides non-agentive perceptions, these verbs can also be used to encode perceptibility events and sensory copulative states; the interpretation depends on the construction they occur with. In 217 and 218, for example, the verbs *sunāī denā* and *dikhāī denā* occur in a single argument construction featuring only a nominative Stimulus. In such cases, these verbs express the possibility of perceiving something and they can be translated as verbs of appearance, such as "be visible, appear". When the single nominative argument is followed by a predicative element, as the adjective *prasann* "happy" in 219, they are interpreted as sensory copulas. Lastly, when they occur with a dative argument as in 220 and 221, they encoded non-agentive perceptions.

217. मन की एक दशा वह भी होती है, जब (…) कान खुले रहते हैं और कुछ नहीं सुनाई देता।

man kī ek daśā vah bhī hotī hai, jab
There is a condition of the mind when (…)

kān	*khul-e*	*rah-t-e*	*haiṁ*	*aur*
ear(M.PL.NOM)	open-M.PL	stay-IPRF-M.PL	be.3PL.PRS	and
kuch	*nahīṁ*	*sunāī*	*de-t-ā.*	
something(M.SG.NOM)	not	hearing(F)	give-IPRF-M.SG	

"There is a condition of the mind when (…) the ears are open and nothing is heard. "

218. नगर में सूर्य का परकाश फैल चुका था। गलियाँ अभी खाली पड़ी हुई थीं। गली के दोनों तरफ भवनों के ऊंचेऊंचे सतून दिखाई देते थे।

nagar meṁ sūry kā parkāś phail cukā thā. galiyāṁ abhī khālī paṛī huī thīṁ.
"The light of the sun had spread in the city. The streets were still empty."

galī=ke	*donoṁ*	*taraf*	*bhavan-oṁ=ke*	*ūṁce-ūṁce*	*satūn*
street=GEN	both	sides	building-PL.OBL=GEN	tall-tall	
satūn	*dikhāī*		*de-t-e*	*th-e.*	
column(M.PL.NOM)	seeing(F)		give-IPRF-M.PL	be.PST-M.PL	

"The tall pillars of the buildings were visible on both sides of the street."

219. दारोगा जी प्रसन्न दिखाई दे रहे थे।

dārog-ā	*jī*	*prasann*	*dikhāī*	*de*	*rah-e*	*th-e*
inspector(M)-SG.NOM	HON	happy	seeing.F	give	PRG-M.PL	be.PST-M.PL

"The inspector looked happy."

220. और वहां मुझे एक बहुत बड़ा मेंहदी का वृक्ष दिखाई दिया।

aur	*vahāṁ*	*mujhe*	*ek*	*bahut*	*baṛā*	*meṁhdī=kā*	*vṛkṣa*
and	there	1SG.DAT	one	very	big	rosemary=GEN	tree(M.SG.NOM)
dikhāī		*di-yā*					
seeing(F)		give-PRF.M.SG					

"And there I saw a huge rosemary tree."

221. एकाएक उन्हें गोगी की महीन आवाज सुनाई दी।

ekāek	*unheṁ*	*gogī=kī*	*mahīn*	*āvāj*	*sunāī*
suddenly	3PL.DAT	gogī= GEN	sweet	voice(F.SG.NOM)	hearing(F)
d-ī					
give-PRF.F.SG					

"Suddenly, they heard Gogi's sweet voice."

When they occur with an Experiencer, the complex predicates *dikhāī denā* and *sunāī denā* consistently features a dative Experiencer. As is well known by scholars of Hindi, the dative construction is the prototypical construction for the encoding of experiential events in Hindi (Verma and Mohanan 1990, Kachru 1990, Mohanan 1994, Montaut 2004a). This pattern is semantically constrained and only occurs when the Experiencer is not volitional and does not control the perception. Consider for example sentence 222, in which the Experiencer's low control is made clear by the context: the man is trying not to think of a woman, Thayas, but he cannot help seeing her image everywhere. Even when contemplating God, when his mind should be clear from earthly thoughts, he sees her sitting in front of him staring at him.

222. उससे दूर रहकर भी थायस नितय उसके साथ रहती थी। जब वह कुछ पूता था, ईश्वर का ध्यान करता था तो वह सामने बैठी उसकी ओर ताकती रहती, वह जिधर निगाह डालता, उसे उसी की मूर्ति दिखाई देती।

usse dūr rahkar bhī thāyas nitay uske sāth rahtī thī. jab vah īśvar kā dhyān kartā thā to vah sāmne baiṭhī us kī or tāktī rahtī

"Although Thayas was far away, she was always with him. When he would contemplate God, she sat in front of him staring at him"

vah	*jidhar*	*nigāh*	*ḍāl-t-ā*	*use*
3SG.NOM	anywhere	gaze(M.SG.NOM)	throw-IPFR-M.SG	3SG.DAT
us-kī	*mūrti*	*dikhāī*	*de-t-ī*	
3SG.OBL-GEN	image(F.SG.NOM)	seeing(F)	give-IPRF-F.SG	

"wherever he turned his gaze he saw her image".

Consider now sentence 223, which clearly exemplifies the semantic contrast between the ergative Experiencer of the verb *sunnā* and the dative Experiencer of the verb *sunāī denā*. In the first sentence the dative Experiencer does not control the event and is not intentional, as the occurrence of the adverb *sahsā* "suddenly" indicates. In the following sentence, the conjunctive participle *kān lagā=kar* (lit. "attaching the ear") makes it clear that the Experiencer has acquired control over the perception through the action of stretching his ears. Hence, the second sentence does not allow a dative Experiencer, and the perceiver is marked with the ergative.

223. सहसा उसे मंड़ैया के सामने चूड़ियों की झंकार सुनाई दी। उसने कान लगाकर सुना। हाँ, कोई है।

sahsā	*use*	*maṁṛaiyā=ke_sāmne*	*cūṛi-yoṁ= kī*	
suddenly	3SG.DAT	shed= in_front_of	bracelet-OPL.OBL=GEN	
jhaṁkār	*sunāī*	*d-ī*	*us=ne*	*kān*
tinkle(F.SG.NOM)	hearing(F)	give-PRF.F.SG	3SG.OBL=ERG	ear(M.SG.NOM)
lagā=kar	*sun-ā*	*hāṁ, koī hai.*		
attach=CP	listen-PRF.M.SG	Yes, there was someone.		

"Suddenly he heard the tinkle of bracelets outside the shed. He strained his ears and listened. Yes, there was someone."

Even if this construction displays an oblique Experiencer, it clearly construes the event as Experiencer based. Subjecthood tests show that the dative element here is functioning as a subject and indicate that the Experiencer is not left in the background. Consider for example sentence 224, in which the reflexive *apne* is co-referent with the dative *mujhe* "to me" and not with the nominative stimulus *aurat* "woman". In Hindi the reflexive is subject-oriented (see Mohanan 1994, Montaut 2004; see section 3.2.4.3) and here it is oriented toward the dative, indicating that the dative Experiencer is functioning as a non-nominative subject. In example 225, the unexpressed subject of the conjunctive participle *jākar* "go=CP" is co-referent with the dative element and not with the Stimulus *bāg* "garden". The unexpressed subject of a conjunctive participle in Hindi is controlled by a subject thus showing that here the subject is the element in the dative.

224. मुझे अपने घर में से एक बूढ़ी औरत बाहर निकलती हुई दिखाई दी।

mujhe	*apne*	*ghar=meṁ=se*	*ek*	*būṛh-ī*	*aurat*	*bāhar*
1SG.DAT	REFL	house=in=INS	one	old-F	woman(F.SG.NOM)	out

nikal-t-ī	*hu-ī*	*dikhāī*	*d-ī*
come_out-IPRF-F.SG	be.PRF-F.SG	seeing(F)	give-PRF.F.SG

"I saw an old woman coming out of my house."

225. आगे जाकर उन्हें राण के पास एक खूबसूरत बाग दिखाई देता है।

āge	*jā=kar*	*unheṁ*	*rāṇ=ke pās*	*ek*	*khūbasūrat*
forward	go=CP	3PL.DAT	ran=LOC(beside)	a	beautiful

bāg	*dikhāī*	*de-t-ā*	*hai*
garden(M.SG.NOM)	seeing(F)	give-IPRF-M.SG	be.3SG.PRS

"They continue forward and they see a beautiful garden next to Ran." **(**from hiTenTen21**)**

The nominal hosts *dikhāī* and *sunāī* may also occur a different light verb, i.e. *paṛnā* "fall" (as in 226). These complex predicates occur with the same dative construction and convey a similar meaning, however, the verb *paṛnā* contributes a slightly different semantic interpretation.

226. एक दिन रमानाथ वाचनालय में बैठा हुआ पत्र पढ़रहा था कि एकाएक उसे रतन दिखाई पड़ गई।

ek din ramānāth vācnālay meṁ baithā huā patr paṛh rahā thā
One day, Ramanath was reading some documents in his office

ki	*ekāek*	*use*	*Ratan*	*dikhāī*	*paṛ*	*ga-ī*
when	suddenly	3SG.DAT	Ratan.NOM	seeing(F)	fall	go.PRF-F.SG

"One day, Ramanath was reading some documents in his office, when suddenly he saw Ratan.

The verb *paṛnā* can be used with a number of functions in Hindi. It can be used as a main verb meaning "fall" or in constructions with a deontic function. It can also be used as a light verb in V-V complex predicates, and in this case it indicates that an event happens suddenly and unexpectedly. The same nuance of suddenness and unpredictability is present also in the complex predicates *dikhāī paṛnā* and *sunāī paṛnā* and it is what distinguishes these complex predicates from those formed with the light verb denā. This semantics is evident in 226, in which abruptness is expressed by the adverb *ekāek*, "suddenly". Moreover, in 226 the *dikhāī paṛnā* is embedded in a temporal subordinate clause introduced by the conjunction *ki* "that", and it is preceded by a main clause in which the verb is in the progressive aspect. The contrast between the progressive aspect of the main verb and the perfect aspect of *dikhāī paṛnā* depicts a scene in which something was going on when suddenly and unexpectedly something else happened. Notably, *dikhāī denā* and *sunāī denā* do not seem to be aspectually characterized and do not seem to belong to a specific actionality class. Obviously, they can never be interpreted as activities as they never allow an agentive reading, however they lend

themselves to both a stative and an achievement interpretation. As the data in Table 34 show, these verbs may occur with both the imperfective and the perfective aspect, although *dikhāī denā* shows a pronounced preference for the perfective aspect. When *dikhāī denā* and *sunāī denā* occur with the imperfective, both habitual (227) and progressive (228), they favor an atelic stative reading, while they point toward an achievement interpretation when appearing with the perfective aspect (as in 220 and 221 above).

227. पन्ना को चारों ओर अंधेरा ही दिखाई देता था।

pannā=ko	*cārom̐*	*or*	*am̐dher-ā*	*hī*	*dikhāī*
panna=DAT	all_four	direction	darkness(M)-SG.NOM	EMPH	seeing(F)
de-t-ā	*th-ā*				
give-IPRF-M.SG	be.PST-M.SG				

"Panna could see only darkness all around."

228. उसका शरीर एक लम्बी चौड़ी चादर से का हुआ था, जिससे उसका मुंह भी छिप गया था केवल दो आंखें दिखाई दे रही थीं।

uskā śarīr ek lambī cauṛī cādar se kā huā thā, jisse uskā muṁh bhī chip gayā thā
His body was covered with a long wide sheet, which hid his face

(use)	*keval*	*do*	*āṁkh-eṁ*	*dikhāī*	*de*	*rah-ī*	*th-īṁ,*
3SG.DAT	only	two	eyes-F.PL.NOM	seeing(F)	give	PRG-F.PL	be.PST-F.PL

"His body was covered with a long wide sheet, which hid his face, he could see only two eyes/only two eyes were visible."

Table 34: Aspectual distribution across the occurrences of *dikhāī denā* and *sunāī denā* in the corpus.

	Iprf		**Prf**	**Other**			**Tot**
	Hab	Progr		Inc	Fut	Subj	
dikhāī denā	13	74	118	9	4	2	220
sunāī denā	23	4	24	1	0	0	52
Total	36	78	142	10	4	1	272

Interestingly, *dikhāī denā* and *sunāī denā* show some peculiarities in their argument structure and case-markings. As discussed above, these two complex predicates mainly occur with a dative Experiencer and a nominative Stimulus agreeing with the light verb *denā* "give". These constructions are peculiar because, when these verbs are in the perfective aspect, there is no ergative element in the

sentence (as in 220 and 221 above). Yet, as Butt, Carnesale and Ahmed (2023) point out, when in the perfective the verb *denā* always takes an ergative subject elsewhere in the language. This verb is used in Hindi at least with four different functions: as a main verb, as a permissive light verb (Butt 1995), as a vector verb in V-V complex predicates and as a light verb in N-V complex predicates. When used as a main verb it typically occurs in a three-arguments argument structure, with an Agent, a Theme and a Recipient, respectively encoded with a nominative/ergative, a nominative/accusative and a dative (as in 229). The verb occurs with a nominative/ergative subject also when it is used with a permissive function (example 230, adapted from Butt 1995: 34). When used in V-V complex predicates, the light verb *denā* conveys completion of the action (Hook 1974, 1991, Butt 1995, Drocco 2020) and responsibility (Butt and Geuder 2001) and it requires an ergative subject regardless of the verb it follows as in 231. Lastly, even if verb *denā* is not used very often in N-V complex predicates (obviously with the exception of the experiential predicates I am discussing here), when it occurs in such a construction, the verb features a nominative/ergative subject. See, for instance, sentence 232 in which *vividhtā=par* "on the diversity" is profiled by the semantics of the noun *dhyān* thus indicating complex predication and the Agent is marked with the ergative *ne*.

229. *(maim)* *kal* *ratan=ko* *rupye* *de* *d-ūṁ-g-ā*
(1SG.NOM) tomorrow ratan=DAT rupee(M.PL.NOM) give give-1SG-FUT-M.SG
"I will give the rupees to Ratan tomorrow."

230. *anjum=ne* *saddaf=ko* *hār* *banā-n-e* *di-yā*
Anjum=ERG Saddaf=DAT necklace(M.SG.NOM) make-INF-OBL give-PRF.M.SG
"Anjum let Saddaf make a necklace."

231. हिला ने केवल बंगले का नंबर बतला दिया था।

hilā=ne *keval* *bangle=kā* *nambar* *batlā* *di-yā*
hila=ERG only bungalow=GEN number(M.SG.NOM) tell give.PRF-M.SG
th-ā
be.PST-M.SG
"Hila had only told the bungalow's number."

232. भाषा के विविधता पर हम ने आरम्भ से ध्यान दिया।

bhāshā=ke *vividhatā=par* *ham=ne* *ārambh=se*
language=GEN diversity(M.SG.OBL)=on 1.PL=ERG beginning=INS
dhyān *di-yā*

attention(M.SG.NOM) give.PRF-M.SG

“From the beginning, we paid attention to the diversity of languages.” (from hiTenTen21)

As the examples above show, in all uses the verb *denā* occurs with an ergative subject and never a dative one. Butt, Carnesale and Ahmed (2023) did a corpus investigation of the argument structure of these two verbs, and they found no trace of an agentive/ergative argument in any of the examples with *dikhāī denā* and *sunāī denā*. Moreover, they note that the addition of an agentive argument these verbs is judged as severely ungrammatical by native speakers. Since the light verb *denā* “give” does not combine with any other experiential nouns in the language, they conclude that it exceptionally does not occur with agentive argument in this construction and that the construction is very limited and not productive.

What is even more unclear is why the verb *denā* should be involved rather than some other semantically light verb like *jānā* “go” or *ānā* “come”. In particular, the verb *ānā* is frequently used as a light verb in complex predicates expressing experiences. Consider for example sentence 233 in which the visual non-agentive perception is expressed by the complex predicate *najar ānā*, consisting of the nominal host *najar* “seeing” and the light verb *ānā* “come” (see below in this section).

233. बहुत तलाश के बाद उसे अपना पुराना घर नजर आया।

bahut	*talāś=ke bād*	*use*	*apnā*	*purān-ā*	*ghar*	*najar*
much	search=after	1SG.DAT	REFL	old-M.SG	house(M.SG.NOM)	seeing(F)

ā-yā.

come-PRF.M.SG

“After much searching, he saw his old house.”

Butt, Ahmed and Carnesale (2023) suggest that the lack of ergative marking with *dikhāī denā* and *sunāī denā* is explained when considering the argument structure construction as emerging from the reanalysis of a complex predication, in which a former Recipient was reinterpreted as an Experiencer, leading ultimately to the Experiencer prototypical construction. This accounts for the fact that the light verb *denā* is not productive anymore with experiential nominal hosts (it only appears with *dikhāī* and *sunāī*). Since the Recipient originally profiled by the verb *denā* is reinterpreted as an Experiencer, the dative argument acquires the most salient properties and is thus linked to the subject role, thus becoming a non-nominative subject (as examples 224 and 225 clearly show). A similar explanation also applies to the complex predicates *dikhāī paṛnā* and *sunāī paṛnā*. The dative Experiencer of these complex predicates would originally be derived from the locative argument contributed by the verb *paṛnā* “fall”.

Another verb which occurs frequently to express non-agentive visual perception is the complex predicate *najar ānā*, consisting of the noun *najar* "seeing" and the light verb *ānā* "come". This verb shows a semantic characterization very similar to *dikhāī denā*. According to the construction in which it occurs, it may express perceptibility, sensory copulative states or non-agentive perceptions. The most frequent construction is the single nominative argument construction, which may encode both achievements (as in 234) and states (as in 235). The verb can also be used to encode perceptions (as in 236), in this case it occurs with a dative construction and the semantics implies lack of agentivity by the Experiencer.

234. कभी-कभी तुलिया स्वप्न की एक झलक-सी नजर आ जाती, और स्वप्न ही की भांति विलीन भी हो जाती।

kabhī-kabhī	*tuliyā*	*svapn=kī*	*ek*	*jhalak-sī*	*najar*
Sometimes	Tulia.NOM	dream=GEN	one	glimpse-like	sight(F.SG.NOM)
ā	*jā-t-ī*	*aur svapn=hī=kī bhāṁti vilīn bhī ho jātī.*			
come	go-IPRF-F.SG	and would also disappear like a dream			

"Sometimes Tulia would appear like a glimpse of a dream and would also disappear like a dream."

235. शक्कर-मिल की चिमनी यहाँ से साफ नजर आती थी।

śakkar-mil=kī	*cimnī*	*yahāṁ=se*	*sāph*	*najar*
sugar-mill=GEN	chimney(F.SG.NOM)	here=from	clear	sight(F.SG.NOM)
ā-t-ī	*th-ī*			
come-IMPRF-F.SG	be.PST-F.SG			

"The chimney of the sugar mill was clearly visible from here."

236. वे जब कभी मुझसे मिलते, उन्हें मेरी आखें नशे से लाल नजर आतीं।

ve	*jab*	*kabhī*	*mujh=se*	*mil-t-e,*	*unheṁ*
3PL.NOM	when	sometimes	1SG.OBL=COM	meet-IMPRF-M.PL	3PL.DAT
merī	*ākh-eṁ*		*naś=se*	*lāl*	*najar*
1SG.GEN	eye-F.PL.NOM		intoxication=INS	red	sight(F.SG.NOM)
ā-t-īṁ.					
come-IPRF-F.PL					

"Whenever they met me, they would see my eyes red with intoxication."

7.4. (Anti)causative alternation in the expression of Hindi perceptive events

The four different perception event-types singled out by Viberg can be expressed in Hindi by cognate verbs belonging to the causative paradigm. This is particularly true in the case of visual perceptions. These morphologically related verbs correspond to different ways of construing the event and correlate with different constructions. For example, Hindi exhibits three different verbs related to the semantics of visual perception: these verbs are all related to the basic verb referring to sight in the language, i.e. *dekhnā*. As I will discuss below in this section, *dikhnā* is formally the anticausative of *dekhnā* and it is translated as "appear", "seem, look as" or "see" in English, according to the argument structure construction in which it occurs. The verb *dikhānā* is derived by the addition of a suffix *-ā-* to the root *dikh-* and is the causative form of the verb *dekhnā*. The original meaning of the verb was "make someone see", but today it has been lexicalized with the meaning "show". As discussed in section 7.3.3 above, the verbs *dikhāī denā* and *dikhāī paṛnā* are also related to this causative paradigm, as they are formed by the nominal host *dikhāī* "seeing", which is the nominalization (via the feminine suffix *-ī*) of the causative verb *dikhānā*. In this section, I will discuss how this causative paradigm takes part to the expression of perceptive events in Hindi.

As is well known by scholars of Hindi, the phenomenon of the causative/anticausative verb pairs is well entrenched in the language: these alternations occur with most Hindi verbs, and they are morphologically overt. This phenomenon is not typical of the Hindi language alone, but of South Asian Languages in general, and it is often considered one of the features characterizing South Asia as a linguistic area (Masica 1976). The literature abounds with studies devoted to the causative system in Hindi (see Masica 1976, Saksena 1980, 1982, Montaut 2004, Butt 2002) for this reason in this section I do not aim to provide a comprehensive survey of the phenomenon, but simply to give a short overview. I will then move on to discuss the role of (anti)causative alternations in the expression of events of perception in Hindi.

In the case of Hindi simple verbs, the (anti)causatives are realized by changes in the root vowel and the addition of derivational suffixes. The base verb (which may be both transitive and intransitive) can take a suffix *-ā*, and a suffix *–vā* (sometimes accompanied by the shortening of the root vowel, more rarely by changes in the root consonant). See for example: *parh-* "study" > *parh-ā-* "teach" >. *parh-vā-* "make/have s.o. study". Some verbs do not form the causative by adding a derivational suffix, but via the strengthening of the root (see *mar-* "die" > *mār-* "hit/kill"). Previous literature proposed to interpret the two derivational morphemes as encoding two levels of causation: each increasing the valency of one argument (cfr. Masica 1991). Kachru (1980), for example, suggests that the suffix *-ā* adds an argument marked with the postposition *ko* (as in 237 vs. 238), and suffix –vā adds an argument marked with the postposition *se* (as in 238 vs. 239).

237. *rita=ne* *angūr* *khā-e*
Rita=ERG grape(M.SG.NOM) eat-PRF.M.PL
"Rita ate some grapes."

238. *rita=ne* *sima=ko* *angūr* *khil-ā-e*
Rita=ERG Sima= ACC/DAT grape(M.SG.NOM) eat-CAUS1-PRF.M.PL
"Rita fed Sima some grapes."

239. *rita=ne* *kala=se* *sima=ko* *angūr*
Rita=ERG kala=INS Sima= ACC/DAT grape(M.SG.NOM)
khil-vā-e
eat-CAUS2-PRF.M.PL
"Rita made Kala feed Sima some grapes."

However, Saksena (1980, 1982) and Butt (2002) note that there is no evidence for assuming the existence of different layers of causatives, as this classification does not really explain the behavior of many causatives in the language, such as different patterns of case marking with the same causative morpheme (consider 240, 241 and 242, adapted from Butt 2002: 3). Moreover, many verbs may appear with either - *ā*- or -*vā*- without an increase in the valency and with the same pattern of case marking, but with different meanings (see 243 vs. 244, adapted from Saksena 1982: 57). These examples suggest the different causative morphemes may contribute different semantic components.

240. *saddaf=ne* *māsālā* *cakh-ā*
Saddaf=ERG spice(M.SG.NOM) taste-PRF.M.SG
"Saddaf tasted the seasoning".

241. *anjum=ne* *saddaf=ko* *māsālā* *cakh-vā-yā*
Anjum=ERG Saddaf= ACC/DAT spice(M.SG.NOM) taste-CAUS.2-PRF.M.SG
"Anjum had Saddaf taste the seasoning".

242. *anjum=ne* *saddaf=se* *māsālā* *cakh-vā-yā*
Anjum=ERG Saddaf=INS spice(M.SG.NOM) taste-CAUS.2-PRF.M.SG
"Anjum had the seasoning tasted by Saddaf"

243. *maiṁ=ne* *lark-e=ko* *parh-ā-yā.*
1SG=ERG boy-M.SG.OBL=ACC/DAT study-CAUS.1-PRF.M.SG
"I taught the boy"

244. *maiṁ=ne* *larke=ko* *parh-vā-yā.*
1SG=ERG boy-M.SG.OBL= ACC/DAT study-CAUS.2-PRF.M.SG
"I had the boy study."

Saksena argues that the two derivational morphemes do not correspond to different levels of causation, but to two different semantic characterizations with respect to the degree of involvement of the causee: *-ā* suffixation encodes direct causation (either contactive or non-contactive) while the suffix *-vā* expresses indirect non-contactive causation. She also argues that differential marking of the argument results in an opposition between an affected and a non-affected causee: the postposition *ko* implies an affected causee, while the postposition *se* implies a non-affected causee. Table 35 summarizes how the semantic characterization of the two morphemes interact with the different case marking of the causees according to Saksena. In sentence 241 and 242 above the verbs are derived by adding the causative morpheme to the root *cakh-* "taste". In 241, the morpheme *-ā-* occurs with a causee marked with the postposition *ko* and the meaning of the sentence implies an involved causee (construed as affected by the event). In sentence 242 the morpheme *-vā-* is used and the causee is marked with the postposition *se*: the meaning of the sentence implies a non-involved causee construed more as a second Agent than as a Patient.

Table 35: Interaction between causative morphemes and case marking in Hindi (adapted from Saksena 1982: 86).

		Causation	
Affectedness		**Direct, contactive**	**Indirect, nonontactive**
	+ affected causee	*-ko* and *-ā-*	*-ko* and *-vā-*
	- affected causee	*-se* and *-ā-*	*-se* and *-vā-*

In addition to the two causative morphemes *-ā-* and *-vā-*, some transitive verb can derive a corresponding intransitive anticausative via the shortening of the root vowel. See for example, *toṛnā* "break tr" in 245 vs. *ṭūṭnā* "break intr" in 246".

245. पापनाशी की नींद टूट गयी।

pāpnāśī=kī *nīṁd* *ṭūṭ* *ga-yī*
papnashi=GEN sleep(F.SG.NOM) break go.PRF-F-SG
"Papnashi lost his sleep. Lit. Papnashi's sleep broke."

246. पुरूषों ने इस आर्दश को तोड़ा है, स्त्रियों ने निबाहा है।

purūṣ-oṁ=ne *is* *ārdaś=ko* *toḍ-ā* *hai,*

man(M)SG.OBL=ERG	this.OBL	ideal=ACC	break-PRF.M.SG	be.3SG.PRS
striy-oṁ=ne	*nibāh-ā*		*hai.*	
Woman(F)SG.OBL=ERG	accomplish-PRF.M.SG		be.3SG.PRS	

"Men have always broken this ideal, women have always lived up to it."

Before going into detail, a few remarks on my use of the term "causative" are in order. Different scholars use the term causative with different meanings, but two main uses of the term can be singled out: traditional grammarians have generally used the term to refer to an overtly marked verbal category thus addressing a morphosyntactic definition, while modern usage refers to the semantic level and to the presence of a causative component in the verb's meaning, whether or not this is overtly marked (Masica 1976: 41). In this latter sense, for example, the English verb *break* is a causative verb when it is used in a sentence like "I broke the pencil", while it is not causative in a sentence like "The pencil broke". Haspelmath (1993: 90), for instance, defines an inchoative/causative verb pair semantically and says that "it is a pair of verbs which express the same basic situation (generally a change of state, more rarely a going on) and differ only in that the causative verb meaning includes an agent participant who causes the situation, whereas the inchoative verb meaning excludes a causing agent and presents the situation as occurring spontaneously". In Hindi the (anti)causative form can express a number of semantic differences between morphologically related verb stems and, as it happens in the case of experiential verbs, these differences do not necessarily imply a causative meaning (see below in this section). For this reason, in the following pages I will use the term causative to refer to the overtly marking of a verbal category and not to a semantic component.

7.4.1. The (anti)causative paradigm and verbs of perceptions

As mentioned above, the causative paradigm operates with verbs of perceptions resulting in a series of cognates expressing the same perception but construing it differently. I will illustrate this with verbs referring to visual perception, as they show the most complex paradigm. This is given in Table 36.

Table 36: The (anti)causative paradigm of the verb *dekhnā*.

anticausative	**base verb**	**causative in *-ā***	**causative in *-vā***
dikhnā "appear"	*dekhnā* "see, look at"	*dikhānā* "show"	*dikhvānā*

The forms shown in Table 36 differ in the way they profile the event as they distribute prominence among the participants differently. The intransitive anticausative *dikhnā* for instance only profiles the

Stimulus and leaves the Experiencer in the background, while the transitive verb *dekhnā* includes the Experience and profiles it as the starting point of the event. Notably, depending on lexical-semantic verb classes, the (anti)causative morphology can express a number of semantic differences that do not necessarily imply an (anti)causative meaning. This is at odds with the semantic characterization of the (anti)causative paradigm given in the literature. As many scholars (Dixon and Aikhenvald 2000, Haspelmath 1993, Zúñiga and Kittilä 2019), (anti)causative derivations typologically correlate with the semantics of causation. Dixon and Aikhenvald (2000: 14), for example, say that "In a causative derivation a new argument is introduced […], it has constant meaning, that of being a causer. That is, causative derivations all have common semantics, of causation." A similar observation is made by Zúñiga and Kittilä (2019: 15) who point out that "The events coded by causatives differ semantically from their non-causative counterparts in that the caused event includes a causer/agent that is lacking in the non-caused event. Speakers use causatives when they need, or want, to introduce to the clause an external agent, the causer". See for example sentence 246 above, in which the transitive causal verb *toṛnā* construes the event as caused by an Agent, which is marked with the ergative, as opposed to 245 in which the non-causal form *ṭūṭnā* construes the event as spontaneously happening. However, when applied to experiential verbs, these alternations do not really add a causer bringing about the event, but an agentive Experiencer perceiving some Stimulus. Sentences 247 and 248 are two examples showing the causal transitive verb *dekhnā* "see" (248) and its non-causal intransitive cognate *dikhnā* "appear" (247). In 248, the argument structure of *dekhnā* is increased by one, however this additional argument is not an Agent that brings about the event, it is rather an Experiencer. Additionally, the semantic contribution of the alternation does not include a causative component, as the meaning of *dekhnā* is "see, look at something" and not "make something appear".

247. उसके ऊपर तनाव साफ दिखता था।

us=ke upar	*tanāv*	*sāph*	*dikh-t-ā*	*th-ā*
3SG.OBL=on	tension(M.SG.NOM)	clear	appear-IPRF-M.SG	be.PST-M.SG

"The tension was clearly visible on him."

248. मैं सडक़ से गुज़रती हुई गाडिय़ों को देखता रहा।

maiṁ	*saṛak=se*	*guzar-t-ī*	*hu-ī*	*gāṛi-yaṁ*
1SG.NOM	street=INS	go.by.IPRF.F	be.PRF.F	car(F)-PL.NOM
dekhtā	*rah-ā*			
look_at-IPRF-M.SG	PRG-M.SG			

"I was looking at the cars running down the street sitting in silence."

Notably the intransitive anticausative verb *dikhnā* "appear" may also occur in a dative construction (249) in which the Experiencer is reintroduced in the argument structure with a dative marking. In this construction, the dative Experiencer is construed as the Beneficiary/Recipient of the event expressed by *dikhnā.* The fact that the Experiencer is an argument in this construction, and not an adjunct, is shown by the fact that it behaves as a non-canonical subject.

249. मुझे तो तुम्हारे स्वास्थ्य में अभी कोई तबदीली नहीं दिखती।

mujhe	*to*	*tumhār-e*	*svāsthy=meṁ*	*ab=hī*	*koī*
1SG.DAT	then	2PL.GEN	health=in	now=EMPH	INDF.ADJ
tabdīlī		*nahīṁ*	*dikh-t-ī.*		
change(F.SG.NOM)		not	appear-IPRF-F.SG		

"I don't see any change in your health right know."

The causatives in -*ā*- and -*vā*- construe the perceptive situation as caused by an external Agent as shown in 250 and 251. Notably the interpretation of sentences 250 and 251 support Saksena's assumption that the two causative morphemes do not correspond to different layers of causativization but to two different semantic characterizations. The - *ā* - morpheme implies direct causation and the - *vā* – morpheme implies indirect causation. This is shown also by the different English translations of the two Hindi sentences. The verb *dikhānā* in 250 can only be translated with "show" in English, while *dikhvānā* in 251 is better translated with "have me seen". However, since in both cases the causee is an Experiencer, it is always affected and this results in the same pattern of case marking: both *dikhānā* and *dikhvānā* occur with a dative/accusative Experiencer and not with an instrumental one (see Table 35 above).

250. मैं तुम्हें अपने कमरे दिखाउंगी।

maiṁ	*tumheṁ*	*apnā*	*kamr-ā*	*dikh-ā-ūṁ-g-ī*
1SG.NOM	2PL.DAT/ACC	REFL	room(M).SG.NOM	appear-CAUS1-1SG-FUT-F

"I will show you my room."

251. मुझे किसी अच्छे डॉक्टर को दिखवा दिजिएगा।

mujhe	*kisī*	*acch-e*	*ḍākṭar=ko*	*dikh-vā*
1SG.ACC	IND.ADJ	good-M.SG.OBL	doctor=DAT/ACC	appear-CAUS2

d-ijie-g-ā

give-IMP.HON-FUT-M.SG

"Please have me checked (lit. seen) by a good doctor."

Notable, the system of anticausative verb pairs is no longer transparent nor productive and verbs belonging to this paradigm are usually perceived as lexically distinct units, although correlated (Montaut 2016). As a consequence, Hindi simple verbs' causative system is characterized by pervasive idiosyncrasies and complex predicates morphology usually fills the gaps characterizing the defective morphology of simple verbs. This happens also for verbs of perceptions. For instance, as shown in Table 37 the verb *sunnā* "hear, listen to" does not have the anticausative form, but features a suppletive anticausative realized by the complex predicate *sunāī denā* "be audible" (discussed in 7.3.3).

Table 37: The (anti)causative paradigms of *dekhnā* and *sunnā*.

	anticausative	**base verb**	**causative -*ā*-**	**causative -v*ā*-**
See, look at	*dikhnā* "appear"	*dekhnā* "see, look at"	*dikhānā* "show"	*dikhvānā*
Hear, listen to	----	*sunnā* "hear, listen to"	*sunānā* "narrate"	*sunvānā*

That the simple verbs' anticausative system is no longer productive and is falling into disuse in Hindi is also suggested by a quantitative analysis of the corpus. For instance, even if visual perceptions still have a simple anticausative verb, this verb occurs very rarely in the corpus (only 11 times), as opposed to the complex predicate *dikhāī denā* which occurs hundreds of times (197). This predominance of complex predicates over simple verbs is a consequence of the restructuring of the Hindi verbal lexicon which occurred in the 18th/19th century because of the contact with the Persian language first and the English language later (see on this section 9.1.3).

The fact that some of the verbs belonging to this (anti)causative paradigm are felt as independent roots by native speakers is proved also by the analysis of the argument structure constructions in which these verbs occur. This is particularly true for the verb *sunānā*, the causative form of *sunnā* "hear, listen to", which originally meant "make someone hear something" and that today has been lexicalized with the meaning "narrate". Saksena (1982, see also Shibatani 1957) suggests that reanalysis of this type typically involve direct causativization with the suffix *-ā-* and affected causees, as affected causees are generally associated to Patients or Recipients and thus "potentially invite reanalysis as a single activity expressed by a [di]transitive verb" (Saksena 1982: 52). Since the causee here is an Experiencer (affected), it should always be marked with the postposition *ko* (see Table 35). However, in the case of highly lexicalized causatives other markings seem to be allowed. The

examples below show two occurrences of the verb *sunānā* in the corpus. In 252, the verb occurs with the causee (the Experiencer) marked with the postposition *ko*. Sentence 253 shows a construction in which the causee is followed by the postposition *se*. Recall from section 3.2.2.1 that this postposition can express the instrumental but also the comitative. It is reasonable to conjecture that in sentence 253, the postposition *se* does not mark an instrumental agent, which would imply non-affectedness, but a comitative argument. The comitative marking on the Experiencer shows the advanced stage in the lexicalization process, as it implies that *sunānā* is felt today as an independent root belonging to the class of verbs of communication. These verbs typically allow the second argument to be marked both with the dative and with the comitative.

252. तुम्हें हिन्दू-देवियों की कथा सुनाऊं।

tumheṁ	*hindū-dev-iyoṁ=kī*	*kathā*	*sun-ā-ūṁ*
2PL.DAT	hindu-goddess-F.PL.OB=GEN	story(F.SG.NOM)	hear-CAUS1-SUBJ.1SG

"I'll tell you the story of the Hindu goddesses."

253. उसने जाकर पति से यह सारी कथा सुनायी।

us=ne	*pati=se*	*yah*	*sar-ī*	*kathā*
3SG.OBL=ERG	husband=INS	this.NOM	whole-F	story(F.SG.NOM)

sun-ā-yī

hear-CAUS1-PRF.F.SG

"She narrated the whole story to her husband."

254. मिर्जा साहब से अंग्रेजी में बोले "अब क्या करना चाहिए?"

mirjā	*sāhab=se*	*aṁgrezī=meṁ*	*bol-e*	*"ab kyā karnā cāhie?"*
mirja	sahab=INS	English=in	say-PRF.M.PL	"what should we do now?"

"He said to Mirja Sahab: 'What should we do now?'

7.5. Complement clause constructions

This section is dedicated to the discussion of clause complementation with verbs of perception in Hindi. In the present work, I follow Noonan's (2007: 41) definition, according to which complementation "arises when a notional sentence or predication is an argument of a predicate" and it functions as the subject or object of that predicate. As I will discuss, perception verbs in Hindi occur

mainly with two types of complemental clause: a predicative participle construction and with a finite complement clause.

The predicative participle construction is typical of visual perception verbs and involves a participle encoding a secondary predication. In this construction the participle is in a predicative relation with the Stimulus NP and is used to refer to a situation perceived by the Experiencer in which another participant is involved. This participial construction may appear with all verbs of visual perception discussed above and with both a nominative/ergative Experiencer and a dative Experiencer. Sentence 255 shows the predicative participle construction occurring with the verb *dekhnā* and an ergative Experiencer, and 256 shows the predicative participle occurring with the verb *dikhāī denā* and a dative Experiencer.

255. उसने जलपक्षियों को नदी के किनारे एक पैर पर खड़े देखा।

us=ne	*jalpakṣi-yoṁ=ko*	*nadī=ke*	*kināŕ-e*
3SG.OBL=ERG	water_birds(M)PL.OBL=ACC	river=GEN	bank.OBL

ek	*pair=par*	*khaṛ-e*	*dekh-ā*
one	leg=on	standing-M.PL	see-PRF.M.PL

"He saw water birds standing on one leg on the bank of the river."

256. वहाँ उसे बाल, वृद्ध, नरनारियों का एक बड़ा समूह एक समाधि के सामने सिजदा करता हुआ दिखाई दिया।

vahāṁ	*use*	*bāl,*	*vr̥ddha,*	*narnār-iyoṁ=kā*	*ek*	*baṛā*
there	3SG-DAT	child	old_man	woman(F)-SG.NOM=GEN	one	big-M

samūh	*ek*	*samādhi=ke_sāmne*	*sijdā*	*kar-t-ā*
group(M.SG.NOM)	one	mausoleum=in_front_of	bow	do-IPRF-M.SG

hu-ā	*dikhāī*	*di-yā*
be.PRF-M.SG	seeig.F	give.PRF-M.SG

"There he saw a large group of children, old men and women prostrating in front of a mausoleum."

As shown in the examples, in predicative participle constructions, the perception verb and the verb in the complementation predicate share one argument. Consider sentence 255, in which the accusative NP expressing the Stimulus of the visual perception, i.e. *jalpakṣi-yoṁ=ko* "birds=ACC", is also the subject of the perfective participle *khaṛe* "standing" and agrees with it in gender and number. Similarly, in sentence 256, the nominative NP *baṛā samūh* "big group" is both the Stimulus of the perception verb and the head of the imperfective participle *sijdā kartā* "prostrating lit. bow do" and agrees with the participle in gender and number.

Since the main function of participle is either encoding attributes of nouns (when used with attributive function) or circumstantial adjuncts of verbs (when used with adverbial function), participles have the distinctive property to express simultaneity between two events. This makes this construction quite suitable for perception verbs but also unsuitable for most other verbs that typically allow complement clauses (see Noonan 2007: 85). This holds true for Hindi as well, in which the predicative participle construction is marginal and seems to be restricted to the domain of visual perception, in which it is the most frequent complement type. These participles usually indicate direct perception as is shown by the fact that this construction mostly features imperfective participles in the corpus, indicating coincidence in time between the event they encode and the event in the main clause. However, it may also occur with a perfective participle and in this case, it usually encodes the result of an action: a state which co-occurs with the perceptive event (as in 257).

257. एक सौ पग भी न चला होगा कि उसे नदी के तट पर एक मनुष्य पाल्थी मारे बैठा दिखाई दिया।

ek sau pag bhī na calā hogā ki
He had not walked a hundred steps that
use nadī=ke taṭ=par ek manuṣy pālthī māre
3SG.DAT river=GEN bank=on one man(M.SG.NOM) cross-legged
baiṭh-ā dikhāī di-yā
sit-PRF.M.SG seeing(F) give.PRF-M.SG
"He had not walked a hundred steps that he saw a man sitting cross-legged on the riverbank."

Predicative participle complements are very frequent with verbs of visual perception, but I did not find traces of this construction with perception verbs referring to hearing in the corpus. Finite complement clauses, on the other hand, are used both with verbs of visual (258) and auditory (259) perception. In this construction, the Stimulus argument is realized by a clause introduced by the complementizer conjunction *ki* "that", the verb is in its finite form and inflected according to TAM properties. Unlike participial clauses, finite complement clauses are not associated to a specific interpretation and for this reason they are far more pervasive across the experiential domains. In Hindi, for example, they are the complement construction typically used with verbs of cognition as well.

258. एक दिन मैंने आंखों से देखा कि उस मूर्ति ने जीवित पराणी के समान अपना सिर फेर लिया।

ek din maiṁ=ne āṁkh-oṁ=se dekh-ā ki us
one day 1SG=ERG eye(F)-PL.OBL=INS see-PRF.M.SG that that.OBL
mūrti=ne jīvit parāṇī=ke samān apnā sir pher liyā

statue(F)=ERG living creature=like REFL head.M.SG.NOM turn take.PRF.M.SG
"One day I saw with my eyes that idol turned its head like a living creature."

259. खन्ना हंस देता है, "यह मैं नयी बात सुन रहा हुँ कि ज़िन्दगी में कुछ अर्थ भी है।"

khannā haṁsa detā hai,
yah *maiṁ* *nay-ī* *bāt* *sun* *rah-ā* *hūṁ*
this.NOM 1SG.NOM new-F words(F.SG.NOM) hear PRGR-M.SG be.1SG.PRS
ki jindagī meṁ kuch arth bhī hai.
that there is some meaning in life.
"Khanna laughs: "It's a new thing I'm hearing that there is some meaning in life"".

Dik and Hengeveld (1991) propose a classification of perception verb complementation in English and argue that different complements for perception verbs correspond to differences in meaning. They distinguish different event types associated with verbs of perception (given in i- iii) and they argue that each type is encoded by different constructions in English.

i. Immediate perception of individuals: typically expressed by transitive constructions, as in *I saw him*.
ii. Immediate perception of state of affairs: typically associated to participle complements as in *I saw him walking down the street*.
iii. Indirect perception and mental perception of propositional content: this reading concerns the acquisition of knowledge through one of the senses and it is not necessarily associated to simultaneity. This perception type is generally expressed by finite complement clauses as in *I saw that Mark has been crying*.

On the same line of Noonan's observations, Dik and Hengevald (1991: 240) claim that the participial construction in English implies simultaneity between the event expressed by the complement and the perception. Moreover, they also note that the participle does not allow the complement to be negated, while finite complemental clauses do. They conclude that the main semantic difference between these two English complement types is that finite complements emphasize the way the participant has acquired the information expressed by the finite clause, while participle complement clauses describe the perception event as such. Hindi shows a similar opposition between the predicative participle construction expressing immediate perception of state of affairs and the finite complemental clause construction, mainly expressing indirect perception. In sentences 255-257, the predicative construction expresses immediate perception and obligatorily implies simultaneity between the event expressed by the participle and the main verb. Hindi finite complement clauses, instead, tend to

describe the way the Experiencer acquires an information through perception. This is evident, for example, in sentence 260, in which the content expressed by the complement clause is in the subjunctive mood, ruling out a direct perception interpretation. Notably, in the same example the complement clause is also negated, thus rejecting a direct perception interpretation even more.

260. द्वार पर कान लगा कर सुना, चारों ओर ताक रहे थे कि कहीं कोई देख न ले।

dvār=par	*kān*	*lagā=kar*	*sun-ā,*	*cārom̐*	*or*
door=on	ear(M)SG.NOM	attach=CP	listen-PRF.M.SG	all_four	direction

tāk	*rah-e*	*th-e*	*ki*	*kahīm̐*	*koī*	*dekh*	*na*
stare	PRG.M.PL	be.PST-M.PL	that	somewhere	INDF.PRN	look	not

le.
take.SUBJ.3SG
"He listened with his ear at the door, looking around to see if anyone saw him."

The semantic difference between these two construction types also accounts for certain distributional properties observed in some perception verbs. For instance, I found very few occurrences of finite complement clauses with the non-agentive verbs *dikhāī denā* and *sunāī denā* in the corpus (as in 261). However, the non-agentive verb *dikhāī denā* frequently occurs with the predicative participle construction: this might be explained by the fact that the semantics of the verb better fits in a participial complement, as non-agentivity is often associated to a sudden perception, and suddenness is basically connected to the semantics of simultaneity.

261. उन्हें क्यों नहीं दिखाई देता कि अब समय बदल गया है?

unhem̐	*kyom̐*	*nahīm̐*	*dikhāī*	*de-t-ā*	*ki*	*ab*
3PL.DAT	why	not	seeing.F	give-IPRF-M.SG	that	now

samay	*badal*	*gay-ā*	*hai?*
time.M.SG.NOM	change	go.PRF-M.SG	be.3SG.PRS

"Why don't they see that times have changed?"

As Dik and Hengevald point out, on the basis of this distinction one would expect that, in languages that encode the difference between direct and indirect perception with distinct constructions, the construction used to encode indirect perception is also used as complement clause of verbs referring to knowledge or acquisition of knowledge (see on this also Noonan 1985). In Hindi this is exactly the case. As I will discuss in Chapter 8 verbs of cognition in this language most frequently occur with sentential complements and rarely with nominals. These complements are realized as finite

complement clauses with most verbs, and the predicative clause, which appears to be so prolific in the encoding of visual perceptions, never occurs with cognition verbs.

Notably, finite complement clauses can also be extended to the domain of evidentiality in Hindi, while predicative participle construction cannot, thus supporting the hypothesis that constructions encoding indirect perception may be extended to the expression of cognitive situations. Evidentiality consists in the indication by the speaker of the source of evidence for a statement (Aikhenvald 2004: 273-274). As mentioned in Chapter 2 and at the beginning of this chapter, a cross-linguistically common extension, which reflects the relation between perception and cognition, is the use of verbs referring to perception as markers of evidentiality (Sweetser 1990, Evans and Wilkins 2000, Vanhove 2008). This semantic extension especially concerns verbs of sight and hearing, but may also involve other sense modalities, and derives from the pragmatic inference that having perceived something means knowing it (Luraghi 2020a: 280).

In Hindi, when occurring with a finite complement clause, verbs of perception referring to sight and hearing can express the acquisition of knowledge. An example is 262, in which *dekhnā* is used not to express visual perception, but the acquisition of knowledge through the reading and studying of some papers. That the verb *dekhnā* here is used to express some controlled mental activity with the purpose of acquiring knowledge and not mere visual perception is also shown by the fact that it occurs right after a coordinate sentence expressing the mental activity of thinking (here encoded by the verb *socnā*, that is the basic verb for thought in Hindi, see section 8.4.1). Moreover, the sentence following the verb *dekhnā*, i.e. *is=meṁ koī begunāh nahīṁ hai* "there is no innocents in this", clearly expresses the inferred opinion that the Experiencer acquired by reading/studying the papers.

262. मैंने ख़ूब सोच लिया है दादा, सब काग़ज़ देख लिए हैं, इसमें कोई बेगुनाह नहीं है।

maiṁne	*khūb*	*soc liyā hai dādā,*	*sab*	*kāgaz*	*dekh*
I thought about it carefully, Dada.			all	paper(M.PL.NOM)	look_at
li-e	*haiṁ,*	*is=meṁ koī begunāh nahīṁ hai.*			
take.PRF-M.PL	be.3PL.PRS	There is no innocent person in this matter.			

"I thought about it carefully, Dada. I looked at all the cards. There is no innocent person in this matter."

Another clear example of the shift from sight to acquisition of a mental content is given in sentence 263, in which the Experiencer is inferring information on the opinion that others have about him based on direct perception of the surrounding context.

263. दयाशंकर ने जब देखा कि जल्द ही मुझपर बीवी का गुलाम होने की फबतियॉं कसी जाने वाली हैं, जिससे ज्यादा अपमानजनक बात मर्द की शान में कोई दूसरी नहीं कही जा सकती, तो उन्होंने बचाव की कोई सूरत न देखकर वापसी मुल्तवी कर दी।

dayāśaṁkar=ne jab dekh-ā ki jald hī mujh=par bīvī=kā
dayāśaṁkar=ERG when see-PRF.M.SG that soon EMPH 1SG.OBL=on wife=GEN
gulām ho-n-e=kī phabat-iyāṁ kas-ī jā-n-e=vālī
slave.NOM be-INF-OBL=GEN joke(F)-PL.NOM tie-PRF.F.SG go-INF-OBL=one.F
haiṁ,
be.3PL.PRS
jis=se jyādā apmānjanak bāt mard=kī śān=meṁ koī dūsrī nahīṁ kahī jā saktī, to unhoṁ=ne bacāv=kī koī sūrat na dekhkar vāpsī multavī kar dī.
"When Dayashankar saw that he would soon be accused of being a slave to his wife, an accusation that could not be more insulting to a man's honor, he delayed his return, seeing no possibility of defending himself."

The verbs *sunnā* "hear" can be used as a marker for evidentiality as well and can express situations in which the Experiencer acquires knowledge from a third party through the reception of the propositional content of a speech act (see on this also Dik and Hengeveld 1991: 238). Consider sentence 264, in which the verb *sunnā* cannot indicate an ongoing perception event but must refer to some information acquired previously to the time of utterance and clearly means "know by hearsay". This is made clear by the fact that the propositional content in the finite complemental clause "he wanted to run away from here" does not encode a state of affairs that can be physically perceived. Moreover, the verb *sunnā* appears in the present even if the perception clearly does not take place in the time of utterance, and this indicates that content of the complement clause is presented as fully acquired.

264. वह समझता था संसार में सब मनुष्य भलामानस है। हमको उपदेश करता था। अब उसकी आँख खुल गई है।अकेला घर में बैठा रहता है ! किसी को मुँह नहीं दिखाता। हम सुनता है, वह यहाँ से भाग जाना चाहता था।

vah samajhtā thā saṁsār meṁ sab manuṣya bhalāmānas hai. hamko upadeś kartā thā. ab uskī āṁkh khul gaī hai. akelā ghar meṁ baiṭhā rahtā hai ! kisī ko muṁh nahīṁ dikhātā.
He used to think that all humans in the world are good. Used to preach us. Now his eyes have opened. He sits alone at home! He doesn't show his face to anyone.

ham	*sun-t-ā*	*hai*	*vah yahāṁ se bhāga jānā cāhtā thā.*
1.PL.NOM	hear-IPRF-M.SG	be.3SG.PRS	he wanted to run away from here.

"We hear that he wanted to run away from here."

7.6. Other sense modalities: touch, taste and smell

In this section, I will discuss the Hindi expression of perception through the other three sense modalities: touch, taste and smell. As mentioned in section 7.3, verbs referring to these three senses are underrepresented in the corpus, and this does not allow any generalization. Due to the scarcity of data from the Literary corpus, I relied also on other sources for extracting linguistic data for these verbs. In particular, I mainly referred to the hiTenTen21 corpus (section 5.1.3). In presenting the data, I will proceed following the order in Viberg's Hierarchy, so I will first discuss the expression of touch, and then I will move on to taste and smell.

Two verbs are used in Hindi to express agentive perceptions referring to touch: the simple verb *chūnā* and the complex predicate *sparś karnā*. Note, however, that most of the times it is not clear whether these verbs indicate a perception or simply the action of coming into contact with someone or something. This is not a specific feature of Hindi, much to the contrary this ambiguity occurs also in other languages. Verbs of touching often tend to be associated to high agentivity and to be interpreted as referring to contact events rather than to experiences. This happens for example also in Italian and in English, in which the verbs expressing tough show the same semantic ambiguity. See, for example, sentence 265, in which the Italian verb *toccare* "touch" and the English verb *touch* in the translation clearly do not refer to an experiential event.

265.	*Tocc-are*	*lo*	*schermo*	*per*	*accendere*	*il*	*dispositivo.*
	touch-INF	the.M	screen(M)	to	turn_on-INF	the.M	device(M)
	"Touch the screen to turn on the device."						

As shown in 266, *chūnā* occurs in a transitive construction in which the Experiencer is marked with the nominative/ergative and the Stimulus with the nominative/accusative. *Sparś karnā* is a complex predicate formed by the noun *sparś* "touch, feel" (McGregor 1994) and the light verb *karnā* "do". This verb occurs in a transitive construction, as in 267.

266. धनिया ने होरी की देह छुई, तो उसका कलेजा सन से हो गया।

dhaniyā=ne	*horī=kī*	*deh*	*chu-ī,*
Dhaniya=ERG	hori=GEN	body(F.SG.NOM)	touch-PRF.F.SG
to us=kā kalejā san=se ho gayā.			
and her heart began to tremble			

“Dhaniya touched Hori’s body, and her heart began to tremble.”.

267. पापनाशी ने बारांगना के सुन्दर ललाट को अपने होंठों से स्पर्श किया।

pāpanāśī=ne	*bārāṁgnā=ke*	*sundar*	*lalāṭ=ko*	*apne*
papnashi=ERG	prostitute=GEN	beautiful	forehead=ACC	REFL
hoṁṭh-oṁ=se	*sparś*	*ki-yā.*		
Lips-PL.OBL=INS	touch(M.SG.NOM)	do.PRF-M.SG		

“Papnashi touched the beautiful forehead of the prostitute with his lips.”

In example 268, the two verbs co-occur and show that there is not much semantic difference between them. It is noteworthy, however, that the verb *chūnā* carries a semantic nuance due to its cultural significance related to the concept of caste purity. The same etymological root can be found, for instance, in the noun *chūt*, which denotes the touch of something ritually impure or the ritual contamination that ensues from it. On the other hand, the noun *sparś* is not linked to the concept of caste purity but rather indicates mere physical contact between two entities. Consider example 269, in which the verb *chūnā* takes on a heavily negative connotation.

268. खबरदार जो उसकी देह को स्पर्श किया। वह अब ईश्वर की है, मनुष्य उसे नहीं छू सकता।

khabardār	*jo*	*us=kī*	*deh=ko*	*sparś*
careful	REL.PRN.NOM	3SG.OBL=GEN	body=ACC	touch(M.SG.NOM)
kiyā.	*vah ab īśvar=kī hai,*			
do.PRF.M.SG	she now belongs to the Lord			
manuṣy	*use*	*nahīṁ*	*chū*	*sak-t-ā.*
man(M.SG.NOM)	3SG.ACC	not	touch	can-IPRF-M.SG

“Those who touch her body should be careful. She now belongs to the Lord; no man can touch her.”

269. ये दुष्ट रोज यहाँ आते थे सब-को छूते थे। इनका छुआ हुआ प्रसाद लोग रोज खाते थे, (इससे बढ़कर अनर्थ क्या हो सकता है)।

ye	*duṣṭ*	*roj*	*yahāṁ*	*ā-t-e*	*th-e*
this.PL	wicked(M.PL.NOM)	daily	here	come-IPRF.M.PL	be.PST-M.PL
sab=ko	*chū-t-e*	*th-e.*			
everything=ACC	touch-IPRF.M.PL	be.PST-M.PL			

in=kā chuā huā prasād log roj khāte the, (isase baḍhkar anarth kyā ho saktā hai).

"These wicked ones used to come here daily and touch everything. People used to eat their touched prasad every day."

Notably, both *chūnā* (270) and *sparś karnā* (271) may be used in metaphorical extensions. The metaphor in (271) is particularly interesting in this discussion because it connects two different experiential domains: emotion is metaphorically represented through perception.

270. जेल में शराब छुई तक नहीं।

jel=meṁ	*śarāb*	*chu-ī*	*tak*	*nahīṁ*
prison=in	alcohol(F.SG.NOM)	touch-PRF.F.SG	EMPH	not

"In prison, he had never touched alcohol."

271. लाला समरकान्त की गिरफ्तारी ने नैना के हृदय में उसी मर्मस्थल को स्पर्श किया।

lālā samarakānt=kī	*giraphtārī=ne*	*nainā=ke*	*hṛday=meṁ*
Lala Samarkant=GEN	arrest(F.SG.OBL)=ERG	Naina=GEN	heart=in
us=ī	*marmasthal=ko*	*sparś*	*ki-yā.*
that=EMPH	vital_point=ACC	touch(M.SG.NOM)	do.PRF-M.SG

"The arrest of Lala Samarkant had touched that vital point in Naina's heart."

In Hindi, taste is expressed through the verb *cakhnā* or through various complex predicates formed by the nominal host *svād* "taste". *Cakhnā* consistently appears in a transitive construction, and it implies an agentive reading (as in 272 and 273). In addition to *cakhnā*, a range of complex predicates formed with the noun *svād* "taste" are used in Hindi to describe taste. According to the light verb following the noun, the complex predicate construes the event differently. In 274, the noun is followed by the light verb *lenā* "take" and thus expresses an agentive perception in which the Experiencer is volitional and attentive. Interestingly, the light verb used here is not the verb *karnā* "do" which is typically used in complex predicates to encode agentive events, but *lenā*. This verb construes the Experiencer as agentive but also adds a benefactive interpretation implying that the Experiencer receives the consequences of the action the s/he performs (see section 3.2.5.1). This complex predicate appears with a genitive Stimulus in the corpus, however, as I have already discussed in 4.2.3.1, the genitive marking of the Stimulus in experiential complex predicates indicates that the noun + verb sequence is not treated as a complex predication, and it should not be interpretated as bearing a specific semantic interpretation.

272. बुढ़िया ने कहा, 'बडे मीठ संतरे लाई हुँ, एक लेकर चखो तो।'

buṛhiyā=ne	*kah-ā,*		*baṛ-e*	*mīṭh*	*saṁtar-e*
old_woman=ERG	say-PRF.M.SG		big-M.PL.OBL	sweet	orange(M)-PL.NOM
lā-ī	*hūṁ,*	*ek*	*le=kar*	*cakh-o*	*to*
bring-PRF.F.SG	be.1SG.PRS	one	take=CP	taste-IMP	then

"The old woman said: I've brought some very sweet oranges, try one.'"

273. आओ, आज तुम्हें गरीबों का खाना खिलाऊँ, इसका मजा भी चख लो।

āo, āja tumheṁ garīboṁ kā khānā khilāūṁ,
Come, today I will feed you a meal for the poor,

is=kā	*majā*	*bhī*	*cakh*	*lo.*
this.OBL=GEN	flavour(M)-SG.NOM	also	taste	take.IMP.

"Come, today I will feed you a meal for the poor, and I you will taste its flavor."

274. मैं सभी प्रकार के प्रेमों का स्वाद ले चुकी हुँ।

maiṁ	*sab=hī*	*prakār=ke*	*prem-oṁ=kā*	*svād*
1SG.NOM	all=EMPH	type=GEN	lovePL.OBL=GEN	taste(M.SG.NOM)

le	*cuk-ī*	*hūṁ.*
take	be_completed-PRF.F.SG	1SG.PRS.

"I have tasted all kinds of love."

The noun *svād* also appears with the light verbs *milnā* "meet, get, receive" and in this case it expresses a non-agentive perception (as in 275 and 276). The verb *milnā* occurs in a dative construction and contributes to construe the event as a prototypical experience in which the Experiencer does not control the perception and simply receives the consequences of the event. Sentence 276 exemplifies how differently *cakhnā* and *svād milnā* construe the perception of taste: the verb *cakhnā* refers to a controlled action, while the verb *svād milnā* clearly depicts an experiential achievement.

275. बहुत दिनों बाद आज उसे स्नेह की मधुरता का स्वाद मिला।

bahut dinoṁ bād āj	*use*	*sneh=kī*	*madhurtā=kā*
After a long time, today	3SG.DAT	affection=GEN	sweetness=GEN

svād	*mil-ā.*
taste(M.SG.NOM)	meet-PRF.M.SG

"After a long time, today, s/he tasted the sweetness of affection."

276. वहाँ बैलराम का रातिब था, साफ पानी, दली हुई अरहर की दाल और भूसे के साथ खली, और यही नहीं, कभी-कभी घी का स्वाद भी चखने को मिल जाता था।

vahāṁ bailrām=kā rātib thā, sāph pānī, dalī huī arahar=kī dāl aur bhūse=ke sāth khalī, aur yahī nahīṁ,

kabhī-kabhī	*ghī=kā*	*svād*	*bhī*	*cakh-n-e=ko*	*mil*
sometimes	ghee=GEN	taste(M.SG.NOM)	EMPH	taste-INF-OBL=ACC	meet

jā-t-ā	*th-ā.*
go-IPRF-M.SG	be-PST-M.SG

sometimes he would also get the flavor of ghee while tasting.

"There was Bailram's allowance: plain water, coarsely grinded lentils, oil cake with straw, and not only that, sometimes he would also get the flavor of ghee while tasting."

Smell is mainly expressed by two verbs in Hindi: these are *sūṁghnā* and *mahknā*. These two verbs occurs with different meanings in the corpus and seem to profile the perception differently. *Sūṁghnā* is an Experiencer-based verb and occurs in a transitive construction as shown in 277, while *mahknā* is a phenomenon-based verb, it means "smell, be fragrant" and occurs in a single-argument construction with the Stimulus in the nominative (278). When referring to a non-agentive perception, smell may also be expressed by a N-V complex predicate formed by the noun *khuśbū* "fragrance, scent" and the light verb *ānā* "come" occurring with a dative Experiencer and construing the event as non-agentive (as in 279).

277. मलका की फौज यह संजीवनी सुगन्ध सूंघते ही मतवाली हो गयी।

malkā=kī	*phauj*	*yah*	*saṁjīvnī*	*sugandh*
Malka=GEN	army(F.SG.NOM)	this.NOM	rejuvenating	fragrance

sūṁgh-t-e	*hī*	*matvālī*	*ho*	*ga-yī.*
smell-IPRF-M.PL	EMPH	intoxicated	be	go.PRF-F.SG

"As soon as Malka's army smelled this rejuvenating fragrance, they became intoxicated."

278. तुम्हारी रोटियाँ महक रही हैं काकी! मुझे बाजरे की रोटियाँ बड़ी अच्छी लगती हैं।

tumhārī	*roṭ-iyāṁ*	*mahak rahī*	*haiṁ*	*kākī!*
2PL.GEN	roti-F.PL.NOM	smell PRGR-F	be.2PL.PRS	aunt

mujhe bājre=kī roṭiyāṁ baṛī acchī lagtī haiṁ.

"Your rotis smell wonderful, Aunt! (I really like millet rotis a lot!)".

279. मुझे पायसम की खुशबू आ रही है।

mujhe	*pāyasam=kī*	*khuśbū*	*ā*	*rah-ī*	*hai.*
1SG.DAT	payasam=GEN	scent(F.SG.NOM)	come	PRG-F	be.PRS.3SG

"I smell payasam." (from hiTenTen21)

Non-agentive perceptions referring to touch (280), taste (281) and smell (282) may also be expressed with complex predicates formed by the light verb *lagnā*. As discussed in section 6.5, this verb literally means "be attached, adhere" and is typical of verbs referring to bodily sensations. When *lagnā* is used to express non-agentive perceptions it may occur in two constructions: either with a noun lexicalizing the Expertum (as in 282) or with an adjective in a predicative relation with the nominative Stimulus (as in 280 and 281). In both cases, the Experiencer is marked with the dative, as we would expect from non-agentive experiential predicates. Notably, this verb does not lexicalize a specific sense modality, it is semantically flexible, and it encodes a generic non-agentive perception that broadly corresponds to the English verb "feel" and whose semantics is clarified by the context. See for example sentence 283 in which the verb clearly expresses a non-agentive auditory perception.

280. जब टिन की किनारियों से केक छूटने लगे और छूने पर मुलायम लगे, तब केक पूरी तरह से तैयार है।

jab ṭin=kī kināriyoṁ=se kek chūṭ-n-e lage aur
When the cake starts leaving the sides of the tin and

chūne=par	*mulāyam*	*lag-e,*	*tab kek pūrī tarah=se taiyār hai.*
touch-INF-OBL=on	soft	adhere-SBJV.3SG	then the cake is completely ready

"When the cake starts leaving the sides of the tin and it feels soft to touch, then it is completely ready." (taken from hiTenTen21)

281. धीरे-धीरे मुझे उस गाय का दूध मीठा लगने लगा।

dhīre-dhīre	*mujhe*	*us*	*gāy=kā*	*dūdh*
slowly-slowly	1SG.DAT	that.OBL	cow=GEN	milk(M.SG.NOM)
mīṭh-ā	*lag-n-e*	*lag-ā.*		
sweet-M.SG	adhere-INF-OBL	star-PRF.M.SG		

"Gradually I started finding that cow's milk sweet." (taken from hiTenTen21)

282. सरकारी अधिकारियों ने हिन्दू मंदिरों का नियंत्रण अपने हाथों में ले रखा है, क्योंकि उन्हें इसमें पैसों की गंध लगती है।

sarkārī adhikār-iyoṁ=ne hindū maṁdir-oṁ=kā niyaṁtraṇ apne hāth-oṁ=meṁ le rakhā hai,
Government officials have taken control of Hindu temples in their hands,

kyoṁki	*unheṁ*	*is=meṁ*	*pais-oṁ=kī*	*gaṁdh*
because	3PL.DAT	this.OBL=in	money-(M)PL.OBL=GEN	smell(F.SG.NOM)
lag-t-ī	*hai*			
adhere-IPRF-F.SG	be.3SG.PRS			

"Government officials have taken control of Hindu temples in their hands because they smell money in it." (from hiTenTen21)

283. इसलिए फ़िल्मी संगीतकारों ने इस बात का ध्यान रखा है कि लोरियों के लिए ऐसे गायकों को चुना जाये जिनकी आवाज़ या तो मखमली हो या फिर वो ऐसे अंदाज़ में गायें कि सुनने में मुलायम लगे।

islie philmī saṁgītkāroṁ ne isa bāt=kā dhyān rakhā hai ki loriyoṁ=ke lie aise gāykoṁ=ko cunā jāye jin=kī āvāj yā to makhmalī ho yā phir
Therefore, film composers have taken care to choose singers for lullabies whose voice is either velvety or

vo	*aise*	*aṁdāj=meṁ*	*gā-yeṁ*	*ki*	*sun-n-e=meṁ*
3PL.NOM	such	style=in	sing.SBJV.3SG	that	listen-INF-OBL=in
mulāyam	*lag-e.*				
soft	adhere-SBJV.3SG				

"Therefore, film music composers have paid attention to choosing singers for lullabies whose voice is either velvety or sung in such a manner that it is soft to hear." (taken from hiTenTen21)

Table 38 represents the constructions investigated in this chapter and shows which construction can be used to refer to each of the five sense modalities. Remarkably, the constructions used by verbs referring to the sense modalities of touch, taste and smell only partly mirror the constructions used by verbs referring to sight and hearing in Hindi. In particular, the verbs analyzed in this section never occur with an oblique stimulus, and in this sense, they are similar to verbs expressing hearing. The oblique Stimulus construction only occurs with verbs of visual perception (*dekhnā* and *tāknā*), indicating that higher level of volitionality and control distinguishes sight from other sense modalities. However, all perception types, regardless of the sense modality, can be expressed by two distinct constructions: the transitive construction, which does not specify the semantic properties of the Experiencer, and the dative construction, which encodes only non-agentive perceptions.

Table 38: Constructions used in Hindi to refer to the five sense modalities.

Construction	Sight	Hearing	Touch	Taste	Smell
Single-nom argument	✓	✓	✓	✓	✓
Single-nom argument + pred participle	✓	✓			
Transitive construction	✓	✓	✓	✓	✓
Transitive construction + pred participle	✓				
Transitive construction + finite clause	✓	✓			
Oblique Stimulus construction	✓				
Dative construction	✓	✓	✓	✓	✓
Dative construction + pred participle	✓				

7.7. Discussion

Recall from 4 that one of the basic tenets of Construction Grammar is that the overall interpretation of a construction is reached at by integrating the argument structure construction with the verb and that this allows us to generalize over a vast range of cases in which the same verb shows systematic different meanings in different constructions. In this section I will argue that this is exactly the case for the verbs analyzed in this chapter, as when different argument structure constructions are combined with the same verb, we arrive at different interpretations of the verb. I will illustrate this with the verbs expressing visual perceptions, which as I have shown are the most complex from a constructional perspective. Figure 25 shows the argument structure construction network of *dikhāī denā* "appear, see" and Figure 26 shows the argument structure construction network of *dekhnā* "see, look at". When the verb *dikhāī denā* appears with a single nominative argument it encodes perceptibility events, while when it occurs with a single nominative argument and a predicative participle it encodes sensory copulative states. When it occurs with a dative argument it may encode either non-agentive perception of individuals or, when accompanied by a predicative participle, non-agentive perceptions of states of affairs.

Since constructions have their own semantics, not all verbs are allowed to appear with all constructions discussed in this chapter. In particular, the oblique Stimulus construction implies directionality and agentivity by the Experiencer and cannot occur with the verb *dikhāī denā* which implies a non-agentive Experiencer. The dative construction, on the other hand, is specifically used for the expression of uncontrolled states/achievements and is not allowed with agentive perceptions.

Lastly, the transitive construction does not seem to be semantically based and it may be used for both controlled activities and uncontrolled experiences. Table 38 summarizes the constructions used by verbs of perceptions in Hindi and gives the form and the semantics of each construction.

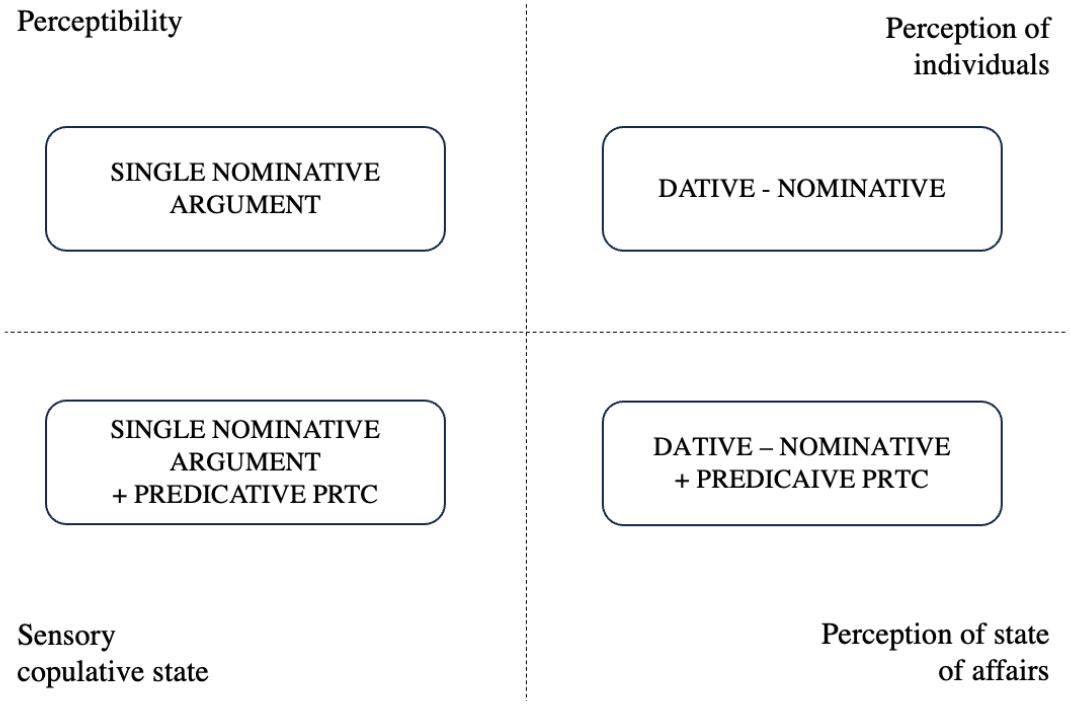

Figure 25: Argument structure construction network of *dikhāī denā* in Hindi

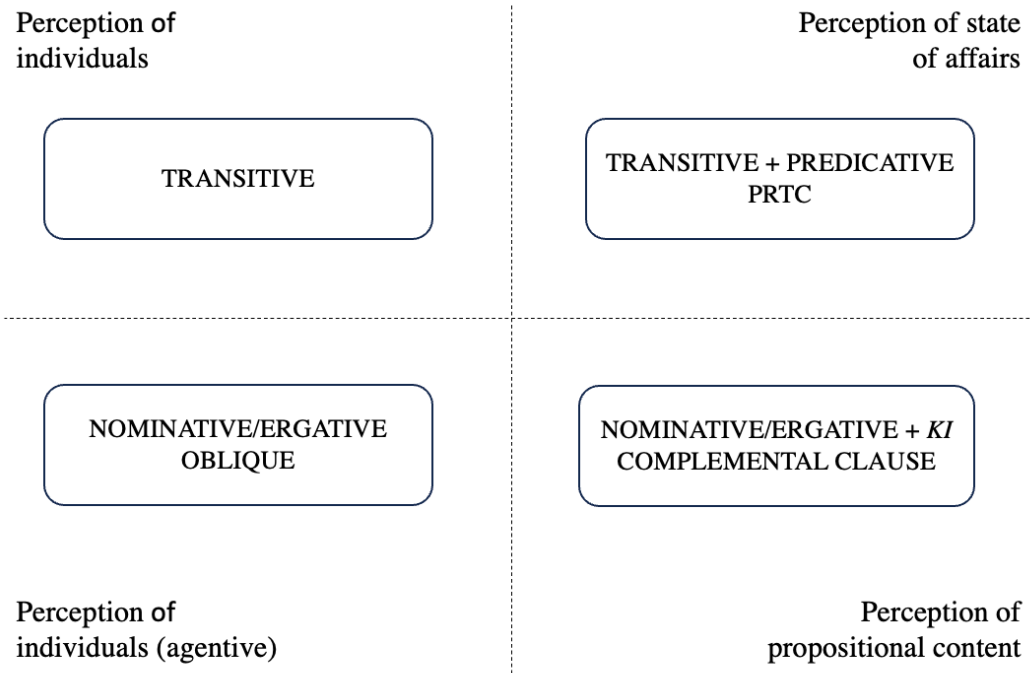

Figure 26: Argument structure construction network of *dekhnā* in Hindi

The fact that verbs expressing perceptibility and sensory copulative states (such as *dikhnā* "appear", *dikhāī denā* "appearing give" and *sunāī denā* "hearing give") can also occur in a dative construction and be used to express non-agentive perceptions gives us valuable insights on how non agentive perceptions are construed in Hindi. I will illustrate this with the verb *dikhnā* "appear". This verb lexically profiles a single participant role that is the appearer and it occurs in a single-nominative argument construction (228) as represented in Figure 27.

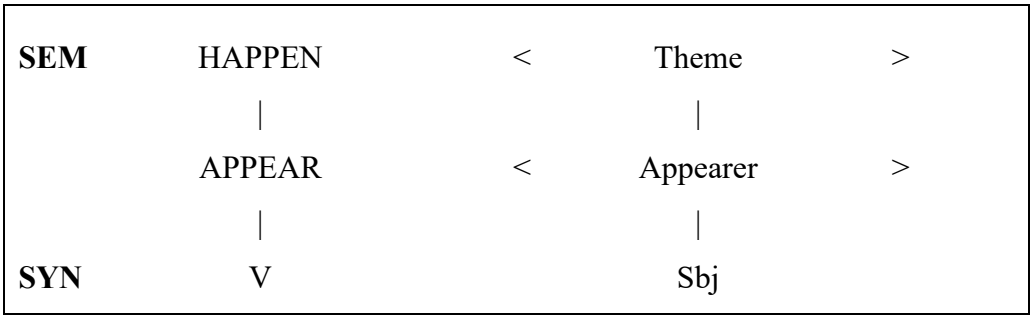

Figure 27. Single nominative construction with *dikhnā.*

When the verb *dikhnā* occurs in a dative construction (in Figure 28), the dative Experiencer is contributed by the dative construction which is the prototypical experiential construction in Hindi (Goldberg 1995: 53-54). The Experiencer is marked as a Recipient/Benificiary and it is linked to the subject relation as it is the most salient argument (Keenan 1976, Langacker 1987a, 1991) and the argument with most Proto-Agent properties (Dowty 1991, Goldber 1995: 116). When they are expressed by a dative construction, non-agentive events are construed as happenstances in which the Stimulus becomes perceivable.

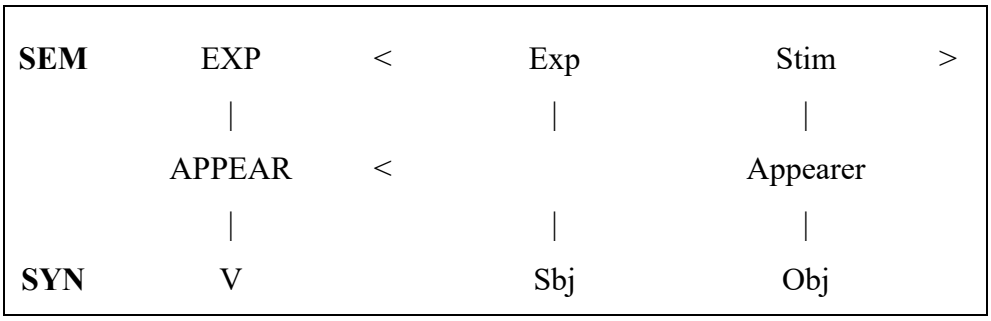

Figure 28: Dative construction with *dikhnā.*

In the oblique Stimulus argument construction used to encode agentive perception the Stimulus is followed by the postpositions *-ki or* or-*ki taraf* "towards". This construction is represented in Figure 29 and it is quite similar to the English Conative Construction (Goldberg 1995: 63) which can be used to express agentive perceptions when occurring with *look at* (see 9.1.5.4).

Sem	DIRECT-ACTION-AT	<	Agent	Theme	>
	|		|	|	
	SEE	<	Seer	Seen	
	|		|	|	
Syn	V		Sbj	Obl*"kī_or"*	

Figure 29: Oblique Stimulus construction with *dekhnā*.

This oblique Stimulus argument construction belongs to a class of constructions that can be used in Hindi to encode situations involving a participant sharing the properties of an Agent (it is human, volitional and in control) and a second participant that is not a prototypical Patient. Constructions belonging to this set change according to the semantic properties associated with the second participant. In this set of constructions, the first argument is marked with the nominative/ergative case. The case marking of the second argument is semantically based and requires the oblique form of the noun followed by a postposition: which postposition is to be employed depends on the properties of the event. For example, an event involving two agentive human participants interacting with each other (such as *marry someone*) typically requires the instrumental/comitative postposition *se* (as shown in 284).

284. अगर तुमसे शादी कर भी लूं तो नतीजा क्या होगा?

agar	*tum=se*	*śādī*	*kar*	*bhī*	*lūṁ*	*to*
if	2SG=COM	wedding(F.SG.NOM)	do	also	take.SUBJ.1SG	then

natījā	*kyā*	*hogā*
outcome(M.SG.NOM)	what	be-FUT-M.SG

"Even if I marry you, what will be the outcome?"

The postpositions used in the oblique Stimulus construction are also used with verbs of change-of-position, in particular when the entity towards which the movement is directed is an object or a person (as in 285). The semantics of the oblique Stimulus construction thus implies directionality of the event and control on behalf of the perceiver, for this reason it is only allowed with verbs of agentive perception.

285. *vah* *bacc-e=kī taraf* *cal* *ga-yī*
3SG.NOM child-M.SG.OBL=towards walk go.PRF-F
"She walked towards the child."

In section 4.3, I defined the transitive construction (Figure 30) as the construction used in Hindi to encode prototypically transitive events: dynamic processes performed by an Agent starting and controlling the event that have consequences on a Patient physically affected by the action. As I mentioned in that occasion, in Indo-European languages, it is not uncommon that the transitive construction gets extended to encode non-semantically transitive situations. Two-place verbs encoding events that involve two participants, one of which sharing some Agent properties (mainly humanness), usually require a transitive construction. As a consequence, non-semantically transitive events are frequently expressed by syntactic transitivity. It is not surprising then that the Hindi transitive construction can be used to encode experiential situations like agentive perceptions, since as I mentioned in section 7.2 many scholars consider agentive perceptions more similar to actions that to experiential states. More surprising is the fact that Hindi, which I assumed to be a highly iconic language, allows for the use of the transitive construction for the encoding of *non*-agentive events as well.

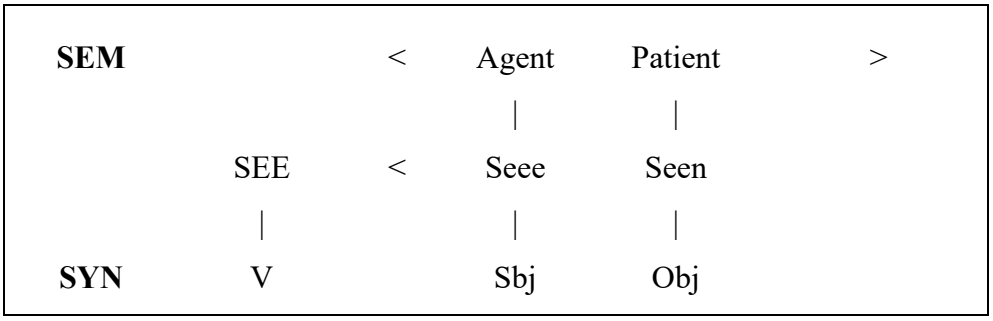

Figure 30: Transitive construction with *dekhnā.*

A possible explanation to this may be found in the fact that non-agentive perceptions and agentive perceptions belong to the same experiential types and that the boundaries between these two types of events are quite fuzzy, so that distinguishing them can be difficult sometimes. This has consequences on the linguistic level. The absence of a distinction between agentive and non-agentive perceptions is quite common cross-linguistically and frequently results in the presence in the language of a construction that can be used to encode both perception types. Consider for example English, which features verbs that lexicalize the agentivity of the perceiver, such as *look at* and *watch*, but that frequently uses the verb *see* to encode clearly agentive events such as "I saw a film yesterday evening". The same happens in Italian, where the verb *vedere* "see" – the only allowed to encode non-agentive visual perceptions in this language – can also express agentive events, as in *Ieri sera ho visto un film* "I saw a film yesterday evening". Like English and Italian, Hindi developed an unmarked construction to express both perception types. The dative construction is highly iconic and can be selected only on semantic basis, thus it cannot be extended to the expression of agentive events, and

it is not eligible to be the unmarked construction. The same happens for the oblique Stimulus construction that is semantically based and can only express agentive perceptions. The transitive construction, that is less iconic and is usually employed through a default mechanism, can be used to encode non-agentive events and can be selected as the unmarked one.

Table 39: Constructions used by verbs of perceptions in Hindi.

Vibger's class	**Construction**	**Form**	**Semantics**
Perceptibility verb	**Single-nominative argument**	2-nom V.subj[2]	Perceptibility
Sensory copula	**Single-nominative argument + predicative participle**	2-nom 3[2] V.subj[2]	Copulative states
Experiencer-based	**Transitive construction**	1-nom 2-nom V.subj[1]	Perception of individual
Experiencer-based (non/)agentive	**Transitive construction + predicative participle**	1-nom 2-nom Prtp[2] V.subj[1]	Direct perception of state of affair
Experiencer-based (non/)agentive	**Transitive construction + finite clause**	1-nom V.subj[1] + (*ki*) finite clause	Direct perception of state of affair Indirect perception Expression of evidentiality
Experiencer-based (agentive)	**Oblique Stimulus construction**	1-nom 2-obl V.subj[-1]	Perception of individual (agentive)
Experiencer-based (non-agentive)	**Dative construction**	1-dat 2-nom V.subj[2]	Perception of individual (non-agentive)
Experiencer-based (non-agentive)	**Dative construction + predicative participle**	1-dat 2-nom Prtp[2] V.subj[2]	Direct perception of state of affairs (non-agentive)

8. Cognition

Verbs of cognition refer to mental processes and states such as *think*, *believe*, *understand*, and *remember*. This class of verbs differs from verbs of bodily sensation and perception in that it implies that the Experiencer is not just animate, but specifically human, as cognitive situations necessarily require self-awareness by the Experiencer (Fortescue 2001, Luraghi 2020a). The semantic-functional domain of cognition borders with the domain of other experiential situations, in particular with perception, but also with bodily sensation (see section 8.5 below). As a consequence, it is typologically common that some perception verbs metaphorically develop into verbs referring to purely mental activity. As we will see, this also happens in Hindi, in which a verb such as *sūjhnā*, originally meaning "be perceived", has evolved to denote a mental achievement (see section 8.4.1). Cognitive linguistics generally explains the polysemy among perception verbs and verbs of cognition as resulting from the MIND-AS-BODY metaphor, according to which our mental structures are anchored in and are determined by the way we perceive the world through physical senses (Sweetser 1990, Ibarretxe-Antuñano 1999, 2002). This conceptual metaphor (Lakoff and Johnson 1980, Lakoff 1993) accounts for correspondences between our perception of external world and our internal states and activities and explains our conceptualization of the domain of cognition in terms of the domain of perception.

Sweetser (1990) argues that the connection between the domains of perception and cognition is particularly evident in the first place for sight verbs. As she remarks "[the] intellectual side of our mental life seems to be regularly linked with the sense of vision, although other senses [...] occasionally take on intellectual meanings as well. There are major similarities in our general linguistic treatments of vision and intellection" (1990: 37) and she adds that "hearing is connected with the specifically communicative aspects of understanding, rather than with intellection at large" (1990: 43). As I will discuss, Hindi seems to align with these observations as it does not extend verbs expressing hearing to verbs expressing cognitive event. Moreover, also from a constructional point of view Hindi seems to show a tighter connection between cognition and visual perception than between cognition and auditory perception. As discussed in chapter 7.3.1, verbs referring to agentive perception frequently occur with an oblique Stimulus construction, in which the Stimulus is marked with a locative postposition. This construction seems to be restricted to verbs referring to sight as I did not find traces of its occurrence with verbs referring to hearing. As I will argue, verbs of cognition allow a very similar oblique Stimulus construction, which semantically points toward a higher degree of control by the Experiencer (see for example the verb *vicār karnā* "think about, reflect on" in section

8.4.2). It should be pointed out, however, that some linguistic and anthropological studies are at odds with Sweetser's observation. For instance, Vanhove (2008) compared data from 25 different languages and found that the semantic extension from verbs referring to hearing to verbs encoding understanding and knowledge occurs more frequently than the extension from verbs referring to seeing. Moreover, Evans and Wilkins (2000) pointed out that Sweetser's hypothesis was predominantly based on Indo-European data and that, in contrast with her theory, Australian languages regularly recruit verbs of cognition like "think" and "know" from verbs of auditory perceptions rather than visual perceptions. They conclude that the paths of metaphorical extensions from perceptions to cognitions are culture-specific (Evans and Wilkins 2000: 546). Notably, the anthropological literature also seems to attest to a connection between auditory perceptions and cognitions. Mayer (1982:246) for instance reports that, in Papua New Guinea, intellectual activities, knowledge and memory are associated with the ear and both hearing and understanding / knowing are referred to by the same verb *iero*.

Languages around the world vary in the way they structure and lexicalize the domain of cognitions and some languages split this domain more finely than others. This variability may be connected with the fact that this class of verbs expresses internal states and activities that are not visible from the outside, and the categorization and construal of cognitions is thus subject to much cultural variation. However, three main semantic areas in this domain can be singled out: namely, *thought*, *knowledge* and *memory*/*forget*.

In Hindi, verbs referring to the same cognition type may construe the same event in different ways and accordingly tend to occur with specific aspectual properties and argument structures. These differences in the aspectual properties indicate a specific characterization of some verbs and suggests that they are connected to specific actionality classes. In previous literature (Luraghi 2020a: 150), cognition verbs are usually classified according to the four actionality categories singled out by Vendler (1967), and include states, such as "know", achievements, such as "understand" or "realize", activities such as "think", accomplishments such as "learn", and more complex situations, such as "remember" or "forget", which can be construed differently with respect to their actionality. As I will argue in this chapter, the specific aspectual distributions that characterize some Hindi verbs can give us valuable insights on the actionality class the verb belongs to. For instance, the Hindi verb referring to generic knowledge, *jannā*, occurs in the perfective aspect only once over the 200 random occurrences scrutinized from the corpus, thus implying that it typically points toward an atelic durative reading, and it tends to construe knowledge as a state rather than as an achievement. Other verbs referring to knowledge such as *mālūm honā* "know, lit. known be", on the other hand, are more flexible and frequently appear in the perfective aspect and construe the event as an achievement.

Argument structure constructions may add various specifications to the construal of cognitive situations as well. Different argument structures may highlight differences in the semantic interpretation of the same verb. As I will show, this is particularly true for specific verbs discussed in this chapter. For example, the Hindi verb *samajhnā*, which is typically associated with the meaning "understand", changes its semantics according to the construction it occurs in (see section). Another interesting case is the omission of the Stimulus in some Hindi verbs referring to thinking, which contributes to construe the event as atelic (see section 8.4). Before going into detail with my discussion on Hindi verbs of cognition, I would like to discuss the different type of constructions that involve the omission of an argument in Hindi and the different consequences that these constructions have in the interpretation of the event (section 8.1). Moreover, given the relevance that aspect and actionality have in my discussion, section 8.2 is dedicated to addressing how these two notions are understood in this study. In section 8.3, I proceed with the presentation of Hindi data.

8.1. Constructions with unexpressed roles

There are many conditions that lead to the omission of a participant in an argument structure construction. Since these have different consequences in the way the event expressed by the verb is construed a quick digression is in order here. Scholars (Fillmore 1986, Goldberg 1995, Haspelmath 2022) generally distinguish two types of argument omission. In the first type, the verb occurs in a construction that shades the argument or cuts it, thus resulting in a different construal of the event. The terms *shading* and *cutting* are taken from Fisher et al. (1991) who propose to metaphorically consider the profiling of a situation as a movie camera that focuses on certain participants of the scene rather than others. The difference between argument shading and argument cutting is that the shaded participant may still be expressed by an adjunct while the cut participant cannot (Goldberg 1995: 57).[26] A typical example of argument shading is the passive construction, which demotes the highest argument of the verb, as in 286. An example of argument cutting is the anticausative construction in 287, in which the argument is removed and not profiled by the construction. The different argument structures resulting from these types of argument removal are represented in Figure 31.

[26] Previous literature also calls these types of argument omission role removal and role demotion (see for example Malchukov 2015, Haspelmath 2022: 32).

Transitive construction		
SEM	$Role_1$	$Role_2$
	|	|
SYN	Subj	Obj

Shaded argument → Passive construction		
SEM	($Role_1$)	$Role_2$
	|	|
SYN	(Obl)	Subj

Cut argument → Anticausative construction	
SEM	$Role_2$
	|
SYN	Subj

Figure 31: Representation of the different types of argument removal

286. किसानों की मुसीबत उनसे नहीं देखी जाती।

kisān-oṁ=kī *musībat* *un=se* *nahīṁ* *dekh-ī*
farmers.PL.OBL=GEN problem(F.PL.NOM) 3PL.OBL=INS not see-PRF.F
jā-t-ī.
go-IPRF-F
"The problems of the farmers are not seen by them."

287. किसी निपुण प्रबंधक के हस्तकौशल, सुविचार और सुरूचि के चिन्ह दिखते थे।

kisī *nipuṇ* *prabandhak=ke* *hastakauśal* *aur*
INDF.OBL skilfull manger(M.SG.OBL)=GEN handicraft(M.SG.NOM) and
surūci=ke *cinh* *dikh-t-e* *th-e.*
flair(F.SG.OBL)=GEN sign(M.PL.NOM) appear.IPRF-M.PL be.PST- M.PL
"There were visible signs of the handicraft and flair of a skillful manager."

The second type of argument omission has been referred to in previous literature with the term null argument (Goldberg 1995). Previous scholars (Fillmore 1986, Haspelmath 2022) distinguish between two types of null arguments: indefinite and definite. Indefinite null arguments, also called argument withholding (see Haspelmath 2022: 30), refer to cases in which the unexpressed argument receives

an indefinite reading because the referent is either unknown or irrelevant, as in 288. In contrast, definite null arguments, also referred to as zero anaphora (Haspelmath 2022: 30) or latency (Croft 2001: 273), refer to cases in which the argument need not to be expressed because its referent is easily recoverable from the context (289 and 290). As Haspelmath points out, it is best to not use the same terminology for these two phenomena as they are semantically rather different. Therefore, I adopt Haspelmath's terminology, and I will use the term argument withholding for cases like in 288, and the term zero anaphora for cases like 289 and 290.

288. हरबंस इस तरह मेरी तरफ़ देखने लगा जैसे नक़ाब हटाने से उसे मेरा कोई और ही चेहरा नज़र आ रहा हो। कुछ देर वह चुपचाप सोचता रहा।

harbaṁs is tarah merī taraf dekhne lagā jaise nakāb haṭāne se use merā koī aur hī cehrā najar ā rahā ho.

Harbans began to look at me as if by removing a mask he could see a different aspect of my personality.

kuch	*der*	*vah*	*cupcāp*	*soc-t-ā*	*rah-ā*
INDF	lapse_of_time	3SG.NOM	silently	think-IPRF-M.SG	stay-PRF.M.SG

"He kept on thinking silently for a while."

289. मैं जानता हूँ कि मैं ज़िन्दगी में क्या चुन सकता हूँ और मुझे क्या चुनना चाहिए। यह भी जानता हूँ कि मेरे आसपास की दुनिया में किसे मेरी ज़रूरत है और किसे नहीं है। मैं सब कुछ जानता हूँ। खूब अच्छी तरह जानता हूँ।

maiṁ jāntā hūṁ ki maiṁ jindagī meṁ kyā cun saktā hūṁ aur mujhe kyā cunnā cāhie. Yah bhī jāntā hūṁ ki mere āspās kī duniyā meṁ kise merī jarūrat hai aur kise nahīṁ hai. maiṁ sab kuch jāntā hūṁ.

"I know what I can choose in life and what I should choose. I also know who, in the world around me, needs me and who doesn't. I know everything."

khūb	*acch-ī*	*tarah*	*jān-t-ā*	*hūṁ.*
well	good-F.SG	way(F.SG.OBL)	know-IPRF-M.SG	be.PRS.1SG

"I know what I can choose in life and what I should choose. I also know who in the world around me needs me and who doesn't. I know everything. I know it very well."

290. देवीदीन से पूछा, 'यह तीसरी औरत कौन है? ' देवीदीन ने कहा , 'मैं नहीं जानता।

Devīdīn se pūchā 'yah tīsrī aurat kaun hai?' Devīdīn ne kahā

(He) asked Devidin, "Who is this third woman?" Devidin said

maiṁ *nahīṁ* *jān-t-ā*
1SG.NOM not know-IPRF-M.SG
"He asked Devidin, 'Who is this third woman?' Devideen said, 'I don't know'."

Notably, zero anaphora does not point toward a different construal of the event and is rather triggered by the fact that the argument that is omitted is already given and it is highly salient in the context. In contrast, cases of argument withholding are interesting in our discussion as they seem to contribute a different construal of the event. This type of argument omission implies an indefinite and unspecified reading and seems to change the way in which the event is construed, pointing toward an atelic interpretation of the event expressed by the verb. In example 288, the verb *socnā*, the basic verb expressing thinking in Hindi, occurs in an indefinite argument construction, in which the Stimulus is omitted. In this case the verb is better translated as "ponder, reflect" rather than "think". Interestingly, not all verbs of cognition allow for indefinite argument withholding in Hindi and this phenomenon seems to be verb-specific. Compare the sentence in 288 with the sentence in 289, in which the verb *jānnā* "know", which typically occurs with two arguments, appears without the Stimulus. In 289, the context suggests that the omission is due to the fact that the Stimulus is highly definite, and its referent can be inferred from the context. Verbs referring to knowing generally only allow a definite reading of argument omission and never an indefinite one. The variation between how the argument omission is interpreted with *socnā* and *jānnā* can be explained through the nature of the cognitive event these verbs refer to. In particular, verbs denoting knowledge typically never allow an indefinite reading, since knowing is inherently definite and specific. Fillmore (1986) provides a test to distinguish these two types of omitted arguments. He notes that admitting ignorance of the referent is acceptable with argument withholding, while it is not acceptable with definite zero anaphora. Consequently, while adding the sentence "Who knows what he was thinking about" to example 288 is perfectly acceptable, adding "I wonder what he knows" to sentence 289 is not. In Hindi (and in South Asian languages in general) pro-drop phenomena are rampant, and they lead to the frequent omission of definite arguments of any type in the language (see on this Butt and King 1997, Butt 2001). Hence, the omission of either a definite Experiencer or a definite Stimulus is quite common in the corpus.

8.2. Actionality and aspect in verbs of cognition

In this study, I will distinguish between aspect and actionality, thus following the so called bidimensional approach (Sasse 2002, Tatevosov 2002, Bertinetto and Delfitto 2002).[27] In this

[27] Note that this approach is not universally accepted and not all scholars agree in keeping the notions aspect and actionality distinct (see among many others Bybee 1985, Bybee, Perkins and Pagliuca 1994, Croft 2012).

approach, verbal aspect is defined as an inflectional category of verbs. According to Bertinetto and Delfitto's (2000: 190) definition, verbal aspect corresponds to "the specific perspective adopted by the speaker/writer", which specifies whether the event is considered from a global or a partial point of view. Verbal aspect mainly corresponds to the distinction between perfective (global point of view) and imperfective (partial point of view). Actionality, on the other hand, is a feature linked to the lexical meaning of individual verbs, and reflects distinctions based on telicity, dynamicity and durativity.

There are studies (Napoli 2006, Luraghi 2020a) devoted to the interplay between aspect and actionality that argue that the tendency of verbs to feature certain TAM properties is heavily conditioned by their meaning and that there is a strong correlation between aspect and actionality. This seems to be particularly true for verbs of cognition. Luraghi (2020: 94), for example, points out that, among verbs of experience in Homeric Greek, "cognitive verbs are those that most consistently feature different lexemes in connection with different verbal aspects and show a higher connection of aspect with actionality". As I mentioned in the introduction to this chapter, this is true for Hindi as well, where some verbs exhibit a strong connection with specific aspectual forms and the choice of verbal aspect may change the actionality class of the verb.

The actionality of a verb can be analyzed following the model elaborated by Vendler (1967, see also Dowty 1979). Based on the parameters of telicity, dynamicity and durativity , this model singles out four classes: states, achievements, accomplishments, and activities. The characterization of each class is reported in Table 40, adapted from Vendler (1967). States and activities are atelic durative events that do not have a specific endpoint and that extend thorough time. On the other hand, achievements and accomplishments are telic and have an endpoint (Comrie 1976). States differ from the other three classes as they denote a situation that remains constant over time, while activities, achievements, and accomplishments involve a change. However, unlike activities which are unbounded, accomplishments and achievements are bounded: in particular, accomplishments indicate the conclusion of an activity hence are also durative, whereas achievements indicate only the point of conclusion and are punctual.[28]

[28]Aspect and actionality are also closely linked to the verbal category of tense. For instance, there is a general connection between the present tense and the imperfective aspect, as the present tense typically pertains to ongoing or habitual events and does not naturally lend itself to expressing an event from its endpoint (see on this Comrie 1976: 66).

Table 40: Vendler's (1967) classification of actionality.

	STATES	ACTIVITIES	ACHIEVEMENTS	ACCOMPLISHMENTS
Telicity	–	–	+	+
Dynamicity	–	+	+	+
Durativity	+	+	–	+

Even though verbs typically belong to specific actionality classes, when occurring with specific aspectual markers they can deviate from their semantic characterization and construe the event differently. Stative verbs in the perfective aspect, for instance, tend to be interpreted as achievements, and never as accomplishments, since states lack the temporal dimension that is implied in the durative interpretation of accomplishments. Activities, on the other hand, similarly to accomplishments, are ongoing processes and imply durativity. Thus, when featuring the perfective aspect, activities tend to be interpreted as accomplishments (see on this Luraghi 2020a: 91). See examples 291 and 292 showing the complex predicate *mālūm honā* (lit. "known be"), which in Hindi is one of the main verbs used to refer to knowledge. The verb is interpreted as state when in the present (291), and as an achievement when occurring in the perfective form (292).

291. उन्हें खूब मालूम है कि रायसाहब बड़े प्रभावशाली जीव हैं।

unheṁ	*khūb*	*mālūm*	*hai*	*ki*		
3SPL.DAT	well	known	be.PRS.3SG	that		
rāysāhab		*baṛ-e*	*prabhāvśālī*	*jīv*		*haiṁ.*
raisaheb(M.SG.NOM)		big-M.PL	influential	creature(M.PL.NOM)		be.PRS.3PL

"He knows very well that the Raisaheb is a very influential creature."

292. आज मुझे मालूम हुआ कि यह कितने क्रोधी हैं।

āj	*mujhe*	*mālūm*	*hu-ā*	*ki*	*yah*	*kitn-e*
today	1SG.DAT	known	be.PRF-M.SG	that	this.NOM	how_much-M.PL
krodhī	*haiṁ*					
hot-tempered	be.3PL.PRS					

"Today I learned/understood how hot-tempered he is."

Interestingly, some verbs of cognition only occur with specific aspectual forms, indicating that there is a clear connection between the meaning of a verb and its aspectual characterization. For example, verbs that denote states, such as *know*, do not generally occur in the progressive (Lakoff 1970: 121, Chung and Timberlake 1985, Tatevosov 2002: 348–349), as this aspectual form generally construes

the event expressed by the verb as an ongoing process (Dowty 1979: 145). Stevanovisch and Greis (2003) propose an analysis of the progressive construction in English in which they investigate the verbs that most correlate with this aspect. Interestingly, their analysis shows that among verbs that are strongly repelled by the progressive, cognition verbs are particularly prominent. This is interesting because it seems to suggest that in English cognitive situations tend to be construed either as states (when in the imperfective past or present) or as results (when in the perfective). Table 41, adapted from Stevanowitsch and Greis (2003: 231), shows the first 15 lexemes most strongly attracted to and most strongly repelled by the progressive construction in English (I have highlighted in bold the verbs of experience).

Table 41: Verbs most strongly attached to and most strongly repelled by the progressive construction in English (adapted from Stefanowitsch and Gries 2003: 231).

Attracted verbs	Repelled verbs
Talk	Be
Go	**Know**
Try	**Think**
Look	**See**
Work	Have
Sit	**Want**
Wait	Mean
Do	**Need**
Use	**Seem**
Come	**Believe**
Run	Call
Move	Put
Live	**Remember**
Deal	Find
Walk	Include

Notably, however, the correspondence between actionality and aspect is far from being perfect and it cannot be generalized for verbs with similar meanings both inter-linguistically and cross-linguistically (see on this Tatevosov 2002, Luraghi 2020a). For example, as I will discuss in more detail in 8.5.1, the Hindi verbs *mālūm honā* and *jānnā* both express knowing in Hindi, but they differ with respect to their aspectual distribution as *mālūm honā* can occur in the perfective aspect and can

be used to construe achievements, whereas *jānnā* does not lend itself to such interpretation. Another example is the Hindi verb *samajhnā* which is generally associated with the inchoative meaning "understand" (see for example McGregor's dictionary (1994)), thus indicating a change of state. Interestingly, however, this verb most frequently occurs in the imperfective and may also allow the progressive aspect. As I will discuss below in this chapter, its aspectual characterization (and its argument structure) seems to suggest that its primary meaning is "consider, think, deem" and implies a stative reading (see section 8.6). That the aspectual characterization may contribute to an interpretation that differ from the verb's default construal, as in the cases illustrated above, is acknowledged in discussions on aspect and actionality, and it is a phenomenon that some scholars refer to as aspectual hybridism (Bertinetto 1986, 1991). This aspectual hybridism is not a specific feature of verbs of cognition, and it also occurs with other classes of experiential verbs. See the following examples (adapted from Narasimhan 1998: 61) referring to the bodily sensation of thirst. In (293), the event is construed as an achievement, with the light verb *lagnā* "lit. adhere, be attached" in the perfective aspect. Sentence 294 shows that the same aspectual form can have a durative reading if the context favors it. Example 295 shows that the same verb can also appear in the progressive and that this aspect contributes to construe the event as durative.

293. *laṛkī=ko* *pyās* *lag-ī*
girl(F)SG.OBL=DAT thirst(F.SG.NOM) attach-PRF.F.SG
"The girl got thirsty."

294. *laṛkī=ko* *ghant-oṁ=ke liye* *pyās* *lag-ī*
girl(F)SG.OBL=DAT hour-PL.OBL=for thirst(F.SG.NOM) attach-PRF.F.SG
"The girl felt thirsty for hours."

295. *laṛkī=ko* *pyās* *lag* *rah-ī* *hai*
girl(F)SG.OBL=DAT thirst(F.SG.NOM) attach PRGR-PRF.F.SG be-3SG.PRS
"The girl is feeling thirsty."

Besides aspectual characterizations, other features allow us to diagnostic the actionality class a given verb belongs to. Dowty (1979) proposes a series of syntactic and semantic tests. For example, if a verb allows manner adverbials such as *attentively* or *steadily* it most probably encodes either an activity or an accomplishment. As Narasimhan (1998) notes, Hindi verbs occurring with dative subjects (which typically belong to the experiential class) do not allow adverbs such as *dhyān=se* "attentively" or *sthir rūp=se* "stably, steadily". She concludes that verbs occurring in dative constructions are neither activities nor accomplishments, they are either states or achievements (see

on this also Kachru 1990 and Abbi 1990). Moreover, verbs that allow adverbs supplying a temporal dimension such as "for an hour" probably refer to activities; in contrast, verbs that allow adverbs such as "in an hour" are most probably achievements or accomplishments. Narasimhan uses these tests show that the dative construction usually encodes achievements and not activities (see 296 and 297 adapted from Narasimhan 1998: 61, see on this also Kachru 2006: 84-85).

296. **Mohan=ko ek ghaṁ-ṭe=ke liye māmlā patā*
Mohan=DAT one hour(M)-SG.OBL=for matter(M)SG.NOM news(M)SG.NOM
cal-ā.
walk-PRF.M.SG
"Mohan got to know the matter for an hour."

297. *Mohan=ko ek ghante=meṁ māmlā patā*
Mohan=DAT one hour(M)-SG.OBL=in matter(M)SG.NOM news(M)SG.NOM
cal-ā.
walk-PRF.M.SG
"Mohan got to know the matter in an hour."

In the annotation of aspectual markers, I relied on a formal approach. Many verb forms in Hindi are periphrastic, and their aspect is overtly marked by a participle or an auxiliary (section 3.2.3.2). This made the annotation of the aspectual property of a given verb form unproblematic most of the times. Verb forms containing a perfective participle were annotated as perfective, while verb forms containing an imperfective participle were annotated as imperfective (habitual). The imperfective progressive and durative forms realized via the auxiliary *rahnā* "stay" were accounted for separately from the habitual imperfective. I also accounted for the inceptive form with the auxiliary *lagnā* separately. This form clearly shows an ambiguous status, as it lies in the middle between the two categories of actionality and aspect: it expresses the beginning of an action, thus pointing toward a telic construal of the event (as in 298), but it may occur also in imperfective aspect (example 299).

298. लेकिन फिर मैं बीती हुई बातें सोचने लगा।
lekin phir maiṁ bīt-ī hu-ī bāt-eṁ
but then 1SG.NOM pass-PRF.F.SG be.PRF-F matter(F)-PL.NOM
soc-n-e lag-ā.
think-INF-OBL start-PRF.M.SG
"But then I began to think about the past."

299. कभीकभी शारीरिक विकारों से बुद्धिशक्तियां परखर हो जाती हैं, बड़े वेग से उनका विकास होने लगता है।

kabhīkabhī śārīrik vikāroṁ se buddhiśaktiyāṁ parkhar ho jātī haiṁ,
Sometimes due to physical disorders, intellectual powers become sharp,

baḍ-e	*veg=se*	*un=kā*	
big-M.SG.OBL	speed(M.SG.OBL)=INS	3PL.OBL=GEN	
vikās	*ho-n-e*	*lag-t-ā*	*hai*
development(M.SG.NOM)	be-INF-OBL	start-IPRF-M.SG	be.3SG.PRS

"Sometimes due to physical disorders, intellectual powers become sharp, and they start developing at a great pace."

The Hindi forms that do not show a participle or an auxiliary are more problematic to annotate because they do not fit into the aspectual system of the periphrastic forms. These are the future, the present subjunctive and the imperative. Even though they do not show an overtly marking for their aspect, these forms can be classified as imperfective from a semantic point of view (see on this also Van Olphen 1975). However, since I decided to rely on a formal approach, I will account for these verb forms separately.

Some remarks are in order about the annotation of the aspectual forms of the verb *honā* "be", as this verb does not fit into the periphrastic paradigm of other Hindi verbs. The verb *honā* exhibits a present, a past and a future form. These are the same forms that are used as tense auxiliaries in the other verbs' paradigms to collocate the event on a deictic timeline. The present and the future are derived from the root *ho-* via inflection for the present (see *hūṁ* "I am", *ho* "you are", *hai* "you.SG are" but also "s/he/it is" and so on), and inflection plus suffixation for the future (see *hūṁ-g-ā* "I will be (M)" *hūṁ-g-ī* "I will be (F)" and so on). Past forms are instead derived form a suppletive paradigm realized by the root *th-*, which is etymologically related to the Sanskrit verb *sthā-* "stand". The past forms of *honā* are marked according to gender and number because they originally derive from a past participle (Butt and Rizvi 2008) and they are as follows: *th-ā* (be.PST-M.SG), *th-e* (be.PST-M.PL), *th-ī* (be.PST-F.SG), *th-īṁ* (be.PST-F.PL). The present and the simple past forms do not bear any overt aspectual marker, and they simply link the state expressed by the verb to a specific deictic point in the timeline. Since the verb *honā* "be" always implies the semantic properties of atelicity and durativity, it always refers to some ongoing or habitual event and it does not lend itself to encode an event from its endpoint. So, even though the imperfective aspect is not overtly expressed by a participle, I accounted for the present and the past forms of *honā* as imperfective. In addition to the imperfective past form, Hindi also exhibits a perfective past form, which is marked for gender and number and is as follows: *hu-ā* "be.PRF-M.SG", *hu-e* "be.PRF-M.PL", *hu-ī* "be.PRF-F.SG/PL". This

form is explicitly marked in relation to aspect as it is formed by the perfective participle of the verb and it clearly construes the event as an achievement, and for this reason it is often glossed as "become" (see, for example, Ahmed and Butt 2010). I accounted for this form separately as a perfective aspect.

8.3. Presentation of Hindi data

Table 42 shows the verbs of cognition that I analyzed for this study, together with their absolute frequencies in the corpus. In the next sections, I will discuss these focusing in particular on their TAM properties and the argument structures they occur with. I will divide my discussion in three main sections, according to three verb classes: verbs of thinking (section 8.4), verbs of knowing (section 8.5), verbs of remembering and forgetting (section 8.7).

Table 42: The Hindi verbs of cognition analyzed for the present study, their frequency in the corpus and the number of occurrences manually scrutinized and annotated.

Meaning	MSH verb	Frequency in Corpus	Manual Scrutiny
Think	*socnā*	873	200
	sūjhnā	158	158
	vicār + light verbs	293	200
	khayāl + light verbs	64	64
Understand	*samajhnā*	1665	200
	samjhānā	464	200
	samajh ānā (1)	145	145
	samajh ānā (2)	18	18
Know	*jānnā*	1749	200
	jān paṛnā	234	200
	jñāt honā	59	59
	malūm honā	1527	200
	patā + light verbs	277	200
Forget	*bhūlnā*	351	200
Remember	*yād* + light verbs	694	200

8.4. Think

Even though thinking is commonly considered a cognitive activity, this experiential type exhibits a much more complex characterization and there are different ways to conceptualize it. In particular, thinking may also refer to states and to mental achievements (Vendler 1967: 110, Goddard 2003). Recall from section 8.2 that *think* is among the verbs that most reject the progressive construction in English (Stefanowitsch and Gries 2003), a syntactic behavior quite atypical for a verb mainly expressing ongoing activities. This indicates that thinking has a more complex semantic characterization. Vendler (1967: 110), for example, distinguishes at least two different situations referred to by the English verb *think*: an unbounded activity controlled by the Experiencer, as in *think about*, and a state as in *think that*. Goddard (2003) points out that even though Vendler's classification seems accurate, it is somehow limited and does not account for the much more varied set of situations expressed by the verb *think* in English. He notes for example that *think that* may also be used to express a punctual thought as in "I think that someone is knocking at the door" (Goddard 2003: 111-112). He also notes that English displays a constructional way to keep distinct all the meaning expressed by the verb *think*. In particular, he singles out the following four constructions:

A. *think about*, expressing an ongoing activity controlled by the experiencer.
B. *think something about something*, which indicates considering or judging and usually refers to a state.
C. *think like this* followed by a complement clause expressing a propositional content, which usually refers to a state.
D. *think that*, which usually refers to a state.

To further complicate matter, Fortescue (2001: 20) notes that verbs referring to though show complex patterns of polysemy and languages of the world to do not necessarily align with one another. In particular, he argues that thinking may cover at least three main types of mental situations: thinking as believing, thinking as considering/judging and thinking as unspecified mental activity. Fortescue points out that the polysemy we deal with in the case of verbs of thought is not only related to the different ways a given language divides the semantic domain, but it also arises via metaphor and metonymy. Among the most typologically common metaphorical extensions of thinking he mentions thinking as weighing, thinking as observing, thinking as wanting and thinking as calculating.

In this section, I will focus on three main verbs used in Hindi to express thinking. First, I discuss the verb *socnā*, which typically refer to thinking as an unspecified mental activity, together with its morphological anticausative form *sūjhnā*, which depicts an inchoative situation and refers to a sudden

onset of a state of awareness. I then move on to complex predicates formed with a nominal host lexicalizing the cognition and a light verb that contributes to the construal of the situation in different ways, mainly related to aspectual properties and to the semantic properties of the Experiencer. I mainly focus on the complex predicates formed by the nouns *vicār* and *khayāl* (section8.4.2). As I will show, these two verbs refer to a different type of mental activity: *vicār* generally highlights a component of volitionality and control by the Experiencer and it is more appropriately translated as "thought, reflection", while *khayāl* typically refers to a lower degree of agentivity and it seems to refer to the mere possession of an opinion or an idea, hence it shows a more stative characterization than *vicār*.

8.4.1. Socnā

The Hindi verb most frequently used to express thinking in Hindi is *socnā*. The etymology of this verb can probably be traced back to the combination of the Sanskrit upasarga *su*, meaning "good, excellent, right" and the verb root *cint*- meaning "thinking, to have a though, consider" (Monier-Williams 1899: 398). Such a combination is attested in Sanskrit for example in the deverbal noun *sucintana* referring to "the act of thinking well, deliberate consideration" (see Monier-Williams 1899: 1223). This verb corresponds to the English *think,* in the sense that it expresses events of generical thinking and does not correspond to a specific construal of the situation. Fortescue (2001: 31) points out that it is typologically quite common that a given language shows a basic verb for expressing a general mental activity. Typically, these generic verbs express a range of polysemy that can be associated with different syntactic patterns. Besides a general verb for thinking, languages across the world usually also display other verbs, lexicalizing more semantically specific mental activities and states. As I will show, the verb *vicār karnā* "lit. thought do" lexicalizes a specific semantics and it differ from *socnā* as it is associated with a higher degree of agentivity.

The polysemy of *socnā* is associated also with the fact that this verb does not display a specific aspectual characterization and, as the data in Table 43 show, it may occur with any aspectual form. Notably, verbal aspect may point toward a specific construal of the event expressed by the verb. When it occurs in the progressive aspect, this verb undoubtedly encodes an activity (as in 300), while when in the perfective aspect it is usually interpreted as expressing an achievement (as in 301). When it is in the imperfective it typically encodes a state, but not always: see for example sentence 302, in which the adverbial *kabhī-kabhī* "sometimes" provides a temporal dimension thus contributing to construe the event as habitual. Lastly, example 303 is an occurrence of the verb in the imperative, showing that *socnā* can be also used to encode controlled activities.

300. रमेश इसका कोई जवाब सोच ही रहे थे।

rameś *is=kā* *koī* *javāb* *soc* *hī* *rah-e*
ramesh.NOM this.OBL=GEN INDF answer(M.SG.NOM) think EMPH PRGR-PRF.M.PL
th-e
be.PST-M.PL
"Ramesh was thinking of an answer to this."

301. मैंने सोचा कि तुम भी वहां बैठी-बैठी ऊब गयी होगी।

maiṁ=ne *soc-ā* *ki* *tum* *bhī* *baiṭh-ī-baiṭh-ī*
1SG=ERG think-PRF.M.SG that 2SG.NOM EMPH sit-PRF.F.PL - sit-PRF.F.PL
ūb *gay-ī* *ho-g-ī*
be_bored go.PRF-F be-FUT-F
"I thought that you must be bored sitting there."

302. मैं तो कभी-कभी सोचती हुँ, तुम यहां न आते, तो अच्छा होता।

maiṁ *to* *kabhī-kabhī* *soc-t-ī* *hūṁ,* *tum*
1SG.NOM then sometimes think-IPRF-F.SG be.1SG.PRS 2PL.NOM
yahāṁ na *ā-te,* *to* *acch-ā* *ho-t-ā.*
here not come-IPRF-M.PL then good-M.SG be-IPRF-MSG
"I sometimes think it would have been better if you had not come here".

303. यह तो तुम्ही सोचो कि ईश्वर की भुल के लिए मुझे दंड दे रहे हो।

yah *to* *tum=hī* *soc-o* *ki* *īśvar kī bhul ke lie mujhe daṇḍ de rahe ho*
this then 2SG=EMPH think-IMP that you are punishing me for the mistake of God
"You should think that you are punishing me for the mistake of God."

Table 43: Aspectual distribution of *socnā.*

	Iprf			**Prf**	**Fut**	**Imp**	**Inch**	**Sbjv**	**CP**	**Tot**
	Hab	**Cont**	**Prgr**							
Finite compl clause	31	3	15	52	4	11	6	2	1	125
Transitive	14	2	8	16	1	15	3	1	9	69
Passive	0	0	1	1	0	0	0	0	0	2
Infinitive clause	1	0	1	0	0	0	2	0	0	4

Locative Stimulus	0	0	0	1	0	0	0	0	0	1
Tot	**46**	**5**	**25**	**70**	**5**	**26**	**11**	**3**	**10**	**201**

Table 44 shows the argument structure constructions occurring with *socnā* in the corpus, together with their relative frequencies. The Experiencer is consistently marked with the nominative or the ergative, while the Stimulus may be encoded in different ways. In particular, *socnā* can combine both with NP Stimuli and complement clauses which propositionally express the specific content of the mental activity. However, this verb clearly favors complement clauses. As mentioned in section 3.3, four distinct main clause-types may realize finite complementation in Hindi. The clause may be introduced by the conjunction *ki* "that" (301) or it may be merely juxtaposed to the main clause without any conjunction (302). Moreover, the complement clause may also be anticipated in the main clause by a correlative NP instantiated by the pronoun *yah* "this" (303) or by a noun. In this study, I account for all four construction-types as *finite complement clauses*.

Table 44: Constructions occurring with the verb *socnā* in the corpus and their relative frequencies (1= Experiencer; 2=Stimulus).

Construction	**Form**	**Tot**
Transitive	1-nom 2-nom V.subj[1]	**72/200**
Passive	(1-instr) 2-nom V.subj[2]	**2/200**
Finite complement clause	1-nom (2-nom) V.subj[1] + (*ki*)obj clause	**122/200**
Non-finite complement clause	1-nom inf-gen V.subj[1]	**3/200**
Oblique Sitmulus construction	1-nom 2-loc V.subj[1]	**1/200**

As shown in Table 44, *socnā* also allows for other argument structures, which occur more rarely in the corpus, and which indicate that this verb may express a range of different meanings. One is the oblique Stimulus construction marking the second argument with the locative postposition *meṁ* "in" (304). As I will discuss in section 8.4.2 below, verbs that lexicalize a high degree of agentivity of the Experiencer, such as the complex predicate *vicār karnā*, tend to occur with an oblique Stimulus construction in Hindi. This is a differential marking on the Stimulus that I have addressed in the previous chapter while discussing verbs of perceptions (such as *dekhnā* or *tāknā*), which is triggered by the semantic characterization of the Experiencer rather than the Stimulus. Since the oblique Stimulus construction contributes an agentive reading, when *socnā* occurs with this construction it is better translated as "think about, reflect on" (304). The fact that the locative construction is so rare

with *socnā* is indicative of how thinking is construed when encoded by this verb. In particular, it indicates that *socnā* does not explicitly construe the Experiencer as volitional and in control of the situation. Since this verb is not semantically specified, it favors the transitive construction, which is the default construction in Hindi, in the sense that it is not associated with specific semantic properties of the participants (see section 4.3).

304. मगर एक-डेढ़ साल से जब से मैंने इस विषय में सोचना शुरू किया है, तब से मुझे लगता रहा है कि...

magar	*ek-ḍeṛh*	*sāl=se*	*jab=se*	*maiṁ=ne*	*is*
but	one-half	year(M.SG.OBL)=from	when=from	1SG=ERG	this.OBL
viṣay=meṁ	*soc-nā*	*śurū*	*kiy-ā*	*hai*	
theme(M.SG.OBL)=in	think-INF	start(M.SG.NOM)	do.PRF-M.SG	be.3SG.PRS	

tab se mujhe lagtā rahā hai ki…

I have been feeling that...

"But ever since I started thinking about this a year and a half ago, I have been feeling that..."

The verb *sūjhnā* "be perceived, be thought" is morphologically the anticausative of *socnā*. It is realized through the shortening of the root vowel (*-o-* > *-u-*) accompanied by a change in the root consonant (*-c-* > *-jh-*). The basic original meaning of this anticausative form is "be perceived" and it typically refer to a state or a change-of-state in which an entity is perceivable or becomes perceived (305). Perception is not really specified in relation to which sense modality is involved in the perceptive event. Most frequently, however, we can infer from the context that this verb is connected with visual perception. In sentence 306, for example, the verb *sūjhnā* co-occurs with the verb *sunāī denā* expressing auditory perception. The co-occurrence with a verb explicitly referring to hearing points toward a specific semantics of the verb *sūjhnā* in this sentence, which can be easily interpreted as referring to sight.

305. माघ के दिन थे। महावट लगी हुई थी। घटाटोप अंधेरा छाया हुआ था। मौत का सा-सन्नाटा छाया हुआ था। अंधेरा तक न सूझता था।

māgh ke din the. mahāvaṭ lagī huī thī. ghaṭāṭop aṁdherā chāyā huā thā. maut kā sā-snnāṭā chāyā huā thā.

It was a day in the month of Magh. There was a dense fog. Darkness had enveloped the surroundings. There was an eerie silence of death.

andher-ā	*tak*	*na*	*sujh-t-ā*	*th-ā*
darkness(M)-SG.NOM	even	not	be_perceived-IPRF-M.SG	be.PST-M.SG

"Not even darkness could be perceived."

"It was a day in the month of Magh. There was a dense fog. Darkness had enveloped the surroundings. There was an eerie silence of death. Not even darkness could be perceived."

306. मन की एक दशा वह भी होती है, जब आंखें खुली होती हैं और कुछ नहीं सूझता, कान खुले रहते हैं और कुछ नहीं सुनाई देता।

man kī ek daś vah bhī hotī hai, jab
There is a condition of the mind when

āṁkh-eṁ	*khul-ī*	*ho-t-ī*	*haiṁ*	*aur*
eye(F)-PL.NOM	open-F	be-IPRF-F	be.PRS.3PL	and
kuch	*nahīṁ*	*sūjh-t-ā*		*(hai)*
something(M.SG.NOM)	not	be perceivable-IMPF-M.SG		(be.PRS.3SG)

"The eyes are opened, and nothing is perceived,"

kān	*khul-e*	*rah-t-e*	*haiṁ*	*aur*
ear(M).PL.NOM	open-M.PL	remain-IPRF-M.PL	be.PRS.3PL	and
kuch	*nahīṁ*	*sunāī*	*de-t-ā.*	
something(M.SG.NOM)	not	hearing(F)SG.NOM	give-IMPF-M.SG	

"the ears are opened and nothing is heard."

"There is a condition of the mind when the eyes are opened and nothing is seen, the ears are opened and nothing is heard."

As is typical of Hindi anticausative verbs (see section 7.4.1), the participant that is suppressed by the anticausative is most frequently reintroduced as a dative argument. As the data in Table 45 show, this verb most frequently appears with a dative subject (116 times over 156).[29] When it occurs with a dative Experiencer *sūjhnā* can still express a perceptive situation (as in 307), but it mostly indicates a mental change of state (as in 308), a sudden rise of awareness that may be sometimes triggered by some vague and unspecified perception of an external entity.

307. उसे न कुछ सूझता था, न कुछ सुनाई देता था।

use	*na*	*kuch*	*sūjh-t-ā*
3SG.DAT	not	something(M.SG.NOM)	be_perceived-IPRF-M.SG

[29] Note, also, that not all the occurrences that appear without a dative Experiencer are to be considered as single nominative constructions, as among them there are also cases of zero anaphora, in which the Experiencer does not need to be expressed because already introduced in the context before.

th-ā, *na kuch sunāī detā thā.*
be.PST-M.SG nor hear anything
"He could neither perceive/see nor hear anything."

308. इस विषय में उसने ख़ुद कभी विचार न किया था, मगर तुरंत ही उसे जवाब सूझ गया।

is	*viṣay=meṁ*		*us=ne*	*khud*	*kabhī*
this	theme(M.SG.OBL)=in		3SG.OBL=ERG	self	sometimes
vicār	*na*	*ki-yā*	*th-ā,*	*magar*	*turaṁt*
thought(M.SG.NOM)	not	do.PRF-M.SG	be.PST-M.SG	but	immediately
hī	*use*	*javāb*	*sūjh*	*ga-yā.*	
EMPH	3SG.DAT	answer(M.SG.NOM)	be_perceived	go.PRF-M.SG	

"He himself had never thought about this, but immediately he understood the answer."

As already discussed in the introduction to this chapter, perception verbs are often linked to cognition verbs and this pattern of polysemy is cross-linguistically common. Frequently words that are originally used to refer to the experience of outward situations perceivable by senses are extended via metaphors to the expression of purely internal mental activity. A case frequently cited in the literature is represented by perceptual verbs used to express knowing and understanding in ancient Indo-European. In some Indo-European languages, the perfective form of the root **wid-* "see" underwent a semantic shift that resulted in the expression of knowing, via the metaphorical assumption that seeing something results in knowing it. This happens for example in Ancient Greek with the verb *oîda* and in Sanskrit with the verb *veda* (Sweetser 1990: 32–37; Luraghi 2020a: 169). *Sujhnā* displays a high variable semantics, and it can be variously translated according to the context it occurs in as "think", "understand" or "perceive". See for example sentence 309, in which it expresses a cognitive achievement that could also be translated as "occur to someone's mind". When it occurs in a particular context, it can recover its original semantics and be interpreted as a perception verb. See example 310 in which the verb co-occurs with *sunāī denā* pointing toward a perceptive reading, which is also indicated by the presence of an instrumental body part (*āṁkhoṁ=se* "with the eyes").

309. मदारी बाबू को अपनी प्राणरक्षा का कोई उपाय न सूझता था।

madārī bābū=ko	*apnī*	*prāṇ-rakṣā=kā*	*koī*	*upāy*
Madari Babu=DAT	REFL-F	life-saving(F.SG.OBL)=GEN	some	way(M.SG.NOM)
na	*sūjh-t-ā*	*th-ā.*		
not	be_perceived-IPRF-M.SG	be.PST-M.SG		

"Madari Babu could see no way to save his life."

310. ऑंखों से सूझता न था, कानों से सुनाई न देता था, जगह से हिलना मुश्किल था।

āṁkh-oṁ=se	*sūjh-t-ā*	*na*	*th-ā,*
eye(F)-PL.OBL=INS	perceive-IPRF-M.SG	not	be.PST-M.SG

kānoṁ se sunāī na detā thā, jagah se hilnā muśkil thā.

he could not hear with his ears, it was difficult to move from that place

"He could not see anything with his eyes, he could not hear with his ears, it was difficult to move from that place."

Evidence for considering this verb as belonging to the class of verbs of cognition rather than perception is provided by the types of complement clauses with which it occurs. Complementation with *sūjhnā* is typically realized through a finite clause introduced by the conjunction *ki* (311), which is the typical complement type of verbs of cognition in Hindi. In contrast, this verb never occurs in the corpus with a participle clause which, as I have discussed in section 7.5, is the typical complement clause of verbs of visual perception.

311. अब मुझे सूझने लगा कि जीवन का लक्ष्य सुख – भोग ही है।

ab	*mujhe*	*sūjh-n-e*	*lag-ā*	*ki*
now	1SG.DAT	be_perceived-INF-OBL	attach-PRF.M.SG	that

jīvan kā lakshya sūkh-bhog hī hai.

the goal of life is happiness and enjoyment.

"Now I started to think/understand that the goal of life is happiness and enjoyment."

Recall from the discussion on complementation of perception verbs (7.5) that the participle complement clause is used to indicate direct perception of state of affairs in Hindi. In contrast, when occurring with verbs of perception, finite complement clauses indicate indirect perception and mental perception of propositional content and thus can be extended to the domain of evidentiality, suggesting a connection with the cognitive domain. The fact that *sūjhnā* does not occur with a participle complement in the corpus suggests that this verb is not used to express direct perceptions of states of affairs, and it typically encodes the acquisition of some knowledge (through perceptions). The semantic shift from perception to cognition might be explained by the implication that when someone perceives a situation or an entity, he/she becomes aware of it and starts thinking about it. This shift is typologically quite common, and many languages all around the world use verbs that originally referred to the domain of perception to encode thought. For example, in Germanic languages thinking is ultimately anchored in feeling, as the verb for "think" is linked to the Indo-European root **tong*- encoding the meaning "feel " as well as "think" (Fortescue 2001: 30, see also

Pokorny 1959: 1088). Other examples are the Greek verb *noéō* "realize", which according to some scholars originally indicated an unspecified perception (Bertolín Cebrián 1996), or the Danish verb *synes,* which literally means "it seems/appears to me" (Fortescue 2001: 29).

Table 45: Constructions used with the verb *sūjhnā* in the corpus and their frequencies (1= Experiencer; 2=Stimulus).

Construction		**Form**	**Frequency**
Dat Exp	Nominative Stimulus	1-dat 2-nom V.subj[2]	84/156
	Finite complemental clause	1-dat V.[M(/F).SG] obj clause	21/156
	Non-finite complemental clause	1-dat INF-gen V.subj[F.SG]	7/156
	Genitve Stimulus	1-dat 2-gen V.subj[F.SG]	4/156
			116/156
NO Dat Exp	Finite complemental clause	V.[M(/F).SG] ki obj clause	2/156
	Non-finite complemental clause	INF-gen V.subj[F.SG]	4/156
	Nominative Stimulus	2-nom V.subj[2]	34/156
			40/156

Moreover, *sūjhnā* differs from verbs of perception and patterns with verbs of cognition also in that it displays a non-finite complement type, in which the Stimulus is expressed by an oblique infinitive followed by the genitive postposition (as in 312).

312. तुम्हें इतनी जल्द मांगने की क्यों सूझी?

tumheṁ	*itnī*	*jald*	*māng-n-e=k-ī*	*kyoṁ*	*sūjh-ī?*
2PL.DAT	so	soon	ask-INF-OBL=GEN-F	why	be_perceived-PRF.F.SG

"Why did you think to ask so soon?"

The meaning of this construction differs from that of the finite complement clause. This latter construction is used to describe that a propositional thought or opinion is activated or present in the mind of the Experiencer and it broadly corresponds to the *that*-construction in English (Wierzbicka 1988, Goddard 2003). The non-finite complement clause, on the other hand, does not represent a thought or opinion as it is in the mind of the Experiencer, it rather implies the intention of the Experiencer to do something, and it corresponds to the English *to*-construction (Wierzbicka 1988, Goddard 2003). The semantics of the non-finite complement clause with infinitives thus explains why this construction does not occur with verbs of perception in Hindi.

Notably, I found some occurrences in the corpus in which the genitive Stimulus is not expressed by an infinitive, but by a genitive NP (as in 313). This construction is apparently very similar to what

previous scholars (Verhoeven 2007, Fedriani 2012, Luraghi 2020a) call impersonal construction, in which both the Experiencer and the Stimulus are expressed by an oblique case marking and there is no nominative element.

313. क्यों उसे रोज़ सैर - सपाटे की सूझती थी?

kyoṁ	*us-e*	*roz*	*sair-sapāte=k-ī*	*sūjh-t-ī*
why	3SG-DAT	daily	walk(M.SG.OBL)=GEN-F	be_perceived-IPRF-F.SG
th-ī?				
be.PST-F.SG				

"Why did he think of going for a walk every day?"

Interestingly, however, this construction shows many peculiarities that might shed light on the real nature of this case pattern. First, the genitive marking on the second argument is quite rare in Hindi, as the genitive case typically occurs with arguments of nouns and not with arguments of verbs. Moreover, in this construction the genitive postposition is always in the feminine form even though there is no feminine NP to trigger agreement in the sentence. This is interesting because generally the default agreement form in Hindi is the masculine singular. In example 312 above, the Stimulus is expressed by the oblique form of the infinitive of the verb *māngnā* "ask", followed by the postposition *k-ī* "GEN-F". As discussed in section 3.2.2.1, the genitive postposition in Hindi must show agreement with some NP in the sentence. When there is no element to agree with, the genitive postposition should be in the default form, which in Hindi is the masculine direct singular, so that one would expect the masculine form of the genitive postposition *kā* "of" (M.SG). This is not the only case in which I found a construction displaying a feminine agreement but lacking a corresponding feminine NP that triggers such agreement. Notably, this happens also with the verb *socnā* (as in 314). Moreover, I addressed a similar construction when discussing the argument structure of the perception verb *sunnā* "hear" (see examples 209 and 210 from chapter 7.3.2).

314. मुझे प्रेम ने बताया था कि तुम बहुत जल्दी बम्बई छोडक़र दिल्ली आने की सोच रहे हो।

mujhe prem ne batāyā thā ki

"Prem told me that:

tum	*bahut*	*jaldī*	*bambaī*	*choṛ=kar*	*dillī*	*ā-n-e=kī*
2PL.NOM	very	soon	Bombay	leave=CP	Delhi	come-INF-OBL=GEN-F
soc	*rah-e*					
think	PRGR-PRF.M.PL					

"Prem told me that you are thinking to leave Bombay very soon and come to Delhi."

It is reasonable to conjecture that the feminine form is explained by postulating the presence of an unexpressed NP in the sentence, probably the noun *bāt,* which means "something said, word, speech" but also "matter, topic, subject, fact". This noun denotes an abstract notion involving or deriving from mental activities and it is often used with a vague semantics related to various speech or cognitive situations. See for example sentence 315.

315. *sacc-ī=kī* *bāt* *hai.*
truth(F)-SG.OBL=GEN matter(F.SG.NOM) be.PRS.3SG
"It is true (Lit. It is a matter/discussion of truth)."

Interestingly, a similar phenomenon occurs with complement clause constructions. Recall that *sūjhnā* occurs in constructions in which the Experiencer stands in the dative and the verb generally agrees in gender and number with the nominative Stimulus NP. However, when the Stimulus is a complement clause, the verb cannot agree with anything in the sentence and, in these cases, it should show the default agreement pattern. See, for example, sentence 316 in which the Experiencer is in the dative case and cannot trigger agreement with the verb, while the Stimulus is expressed by a complement clause. This configuration results in the verb standing in the (perfective) masculine singular form *sūjh-ā* (think-PRF.M.SG). However, sometimes in the corpus the verb appears in the feminine form even when there is no feminine noun in the sentence (as in 317). Note that this happens also when the complement clause is preceded by the pronoun/adjective *yah* "this", which elsewhere in the language stands in the default masculine form (as shown in 318). Then, why the feminine agreement in such sentences? Once again, I think that the explanation here is reached at by positing the presence of the unexpressed feminine noun *bāt* "matter, question". This hypothesis is supported by the fact that the same noun is also often found overtly expressed in the corpus (see sentence 319).

316. उन्हें एकाएक सूझा कि तीनों बच्चों को सिखाना चाहिए कि आदमी से वे सिर्फ एक हद तक रिश्ता रखें।

unheṁ *ekāek* *sūjh-ā* *ki* *tīnoṁ*
3PL.DAT suddenly be_perceived-PRF.M.SG that all_three
bacc-oṁ=ko *sikhā-nā* *cāhie* *ki* *ādmī=se*
child(M)-PL.OBL=ACC teach-INF should that man(M.SG.NOM)=INS
ve *sirph* *ek* *had=tak* *riśtā* *rakh-eṁ.*
3PL.NOM only one extent=till relation(M)SG.NOM keep-SBJV.3PL
"They suddenly thought that all three children should be taught that they should have only a limited relationship with man."

317. उसे इस बुढ़ापे में क्यों यह सूझेगी कि अपना घर गिरवाये...

use	*is*	*buṛhāp-e=meṁ*	*kyoṁ*	*yah*
3SG.DAT	this.OBL	old_age(M)-SG.OBL=in	why	this.NOM

sūjh-e-g-ī	*ki*	*apn-ā*	*ghar*	*girvā-ye…*
be_perceived-3SG-FUT.F	that	REFL-M	house(M.SG.NOM)	demolish-SBJV.3SG

"Why would she think in her old age to demolish her house?"

318. किसी को यह नहीं सूझता कि भारतीय खेल खिलाएँ, जो बिना दाम-कौड़ी के खेले जाते हैं।

kisī=ko	*yah*	*nahīṁ*	*sūjh-t-ā*	*ki*
INDF.OBL=DAT	this.NOM	not	be_perceived-IPRF-M.SG	that

bhārtīy khel khilāeṁ, jo binā dām-kauḍī ke khele jāte haiṁ.
Indians should play sports which are played without any cost.
"No one understands that Indians should play sports which are played without any cost."

319. उसे बार-बार एक ही बात सूझ रही थी कि वह उसे बांहों से पकड़ ले और उसके मुंह पर हाथ रखकर उसका मुंह बन्द कर दे।

us-e	*bār-bār*	*ek*	*hī*	*bāt*	*sūjh*
3SG-DAT	time-time	one	EMPH	matter(F.SG.NOM)	be_perceived

rah-ī	*th-ī*	*ki*
PRGR-PRF.F.S	be.PST-F.SG	that

vah use bāṁhoṁ se pakaṛ le aur uske muṁh par hāth rakhkar uskā muṁh band kar de.
he should hold her by the arms, put a hand over her mouth and shut her mouth.
"He kept thinking about one thing repeatedly: that he should hold her by the arms."

8.4.2. Complex predicates expressing thinking: *vicār* and *khayāl*

Another verb frequently used to encode thought in Hindi is the complex predicate *vicār karnā*, formed by the nominal host *vicār* "thought, reflection" and the light verb *karnā* "do". As I will show, this verb differs from *socnā* and *sūjhnā* as it explicitly lexicalizes the agentivity of the Experiencer and should be translated as "reflect, give consideration" (see also McGregor 1994). This complex predicate consistently occurs with a nominative/ergative Experiencer and most frequently features an oblique Stimulus marked with the postposition *par* "on" (320 and 321).

320. पंचों ने तुम्हारे मामले पर अच्छी तरह विचार किया।

pancoṁ=ne *tumhār-e* *māmle=par* *acchī* *tarah*
council=ERG your-M.SG.OBL affair.OBL=on good way
vicār *kiy-ā*
thought(M.SG.NOM) do.PRF-M.SG
"The council reflected on your case thoroughly."

321. लेकिन होरी इस प्रश्न पर जितना ही विचार करता, उतना ही उसका दुराग्रह कम होता जाता था।

lekin *horī* *is* *praśn=par* *jitnā* *hī*
but hori.NOM this.OBL question(M.SG.OBL)=on as_much EMPH
vicār *kar-t-ā,* *utnā hī uskā durāgrah kam hotā jātā thā.*
thought(M.SG.NOM) do-IPRF-M.SG the more his obstinacy reduced.
"But the more Hori thought about this question, the more his obstinacy reduced."

Recall from chapter 3.2.2.1 that the postposition *par* usually indicates a spatial relationship between two entities, and it is translated as "on, upon", but it may also encode a specific collocation in both space (*ghar=par* "at home", *skūl=par* "at school") and time (*samay=par* "on time"). Additionally, the same postposition may also be used to encode the Stimulus toward which an emotion is directed.

The same verb also occurs in oblique Stimulus constructions in which the Stimulus is followed by other spatial postpositions, namely *meṁ* "in" and *kī or* "toward" (322).

322. आकर्षण क्या वस्तु है और कैसे उत्पन्न हो सकता है, इसकी ओर उसने कभी विचार नहीं किया।

ākarṣaṇ kyā vastu hai aur kaise utpann ho saktā hai,
what attraction is and how it can arise
is=kī or *us=ne* *kabhī* *vicār*
this.OBL=toward 3SG.OBL=ERG sometimes thought(M.SG.NOM)
nahīṁ *ki-yā*
not do.PRF-M.SG
"He never thought about what attraction is and how it can arise.

These postpositions occur more rarely than *par* and seem to be synonymic with it, as the constructions they occur with imply an agentive Experiencer consciously bringing about the mental activity and controlling it. The fact that *vicār karnā* mainly occurs with oblique Stimuli marked with spatial postpositions gives us insights on how thinking and reflection are conceptualized when expressed by

this verb. All postpositions marking the Stimulus with *vicār karnā* typically occur with caused motion verbs in Hindi. The postposition *par* occurs when the change of location implies a superessive meaning, while the postposition *meṁ* implies an inessive meaning. The postposition *kī or* "toward" implies directionality. As I discussed in section 7.3.1, *kī or* is also typical of verbs expressing visual perceptions with high agentive Experiencers, such as *tāknā* "observe, look at, stare". The oblique Stimulus suggests that reflection in Hindi is metaphorically construed as a controlled movement of the mind toward the content of the mental activity. The use of spatial postpositions for the expression of abstract situations is a cross-linguistic phenomenon and it derives from the embodiment of experience and from the human tendency to metaphorically conceptualize and categorize the world according to our bodily perception of external entities (Lakoff and Johnson 2980, Tyler and Evans 2003). This analysis of postpositions as deriving from metaphorical and metonymical extensions has been supported by many scholars both in typological studies and in language specific studies (Niemeier and Dirven 1997, Fedriani 2012, Luraghi 2014). The Hindi oblique Stimulus construction can be viewed as derived from the Lakovian metaphor PURPOSES ARE DESTINATIONS, according to which we conceptualized purposes as destinations and the means we use to achieve them as paths. The postpositions *kī or*, *par* and *meṁ* are used to encode mental activities that are intentionally directed toward some cognitive content metaphorically conceived as a destination and, as a consequence, the Experiencer in this construction is always construed as agentive. The higher level of intentionality and awareness by the Experiencer is contributed by the semantics of the postpositions that imply purpose and directionality of the action.

The agentivity of the Experiencer expressed by *vicār karnā* is clearly shown in 323, in which the verb co-occurs with the simple verb *sūjhnā*. Example 323 clearly shows that different verbs can highlight different aspects of the same event and that this different construal is mirrored in Hindi by a different choice in the argument structure construction. In the first sentence, the verb *vicār karnā* expresses a controlled mental activity and occurs with an oblique Stimulus construction, while in the second sentence the verb *sūjhnā* construes thinking as a non-agentive mental change-of-state and occurs with a dative construction. This change of state is spontaneously activated in the mind of the Experiencer, as the adverb *turaṁt* "immediately" suggests.

323. इस विषय में उसने ख़ुद कभी विचार न किया था, मगर तुरंत ही उसे जवाब सूझ गया।

is	*viṣay=meṁ*	*us=ne*		*khud*	*kabhī*
this.OBL	theme(M.SG.OBL)=in	3SG.OBL=ERG		self	sometimes
vicār	*na*	*kiy-ā*	*th-ā,*	*magar*	*turaṁt*
thought(M.SG.NOM)	not	do.PRF-M.SG	be.PST-M.SG	but	immediately

hī *use* *javāb* *sūjh* *gay-ā.*
EMPH 3SG.DAT answer(M.SG.NOM) be perceived go.PRF-M.SG
"He himself had never reflected on this, but he immediately had the answer."

Interestingly, *vicār karnā* differ from *socnā* in that it rarely occurs with complement clauses and seems to prefer NP Stimuli. However, the Stimulus may also be encoded by a complement clause either finite, as in 324, or non-finite, as in 325, in which an oblique infinitive is followed by the genitive postposition. When occurring with *vicār karnā* these two different complement types imply the same semantic difference discussed above for *socnā*: the finite complement clause depicts a propositional thought or opinion as it is in the mind of the Experiencer, while the infinitive complement describes the intention of the Experiencer to do something.

324. मैंने विचार किया था कि ईश्वर ने उसे यह दण्ड दिया होगा।

maiṁ=ne *vicār* *ki-yā* *th-ā* *ki* *īśvar=ne*
1SG=ERG thought(M.SG.NOM) do.PRF-M.SG be.PST-M.SG that god=ERG
use *yah* *daṇḍ* *di-yā*
3SG.DAT that.NOM punishment(M.SG.NOM) give.PRF.M.SG
h-o-g-ā.
be-3SG-FUT-M.SG
"I thought that God must have given him this punishment."

325. उस मनुष्य की-सी दशा हो गयी, जो किसी नदी के तट पर बैठा उसमें कूदने का विचार कर रहा हो।

us manuṣy=kī-sī daśā ho gayī,
"It became like the condition of a man
jo *kisī* *nadī=ke* *taṭ=par* *baiṭhā*
REFL.NOM IND river(F.SG.OBL)=GEN bank(M.SG.OBL)=on sit-PRF.M.SG
us=meṁ *kūd-n-e=kā* *vicār* *kar* *rah-ā* *ho.*
3SG.OBL=in jump-INF-OBL=GEN thought(M.SG.NOM) do PRGR-M be.SBJV.3SG
"It became like the condition of a man who sits on the bank of a river and thinks of jumping into it."

In addition to *karnā*, the nominal host *vicār* can occur with other light verbs, and the alternation of the verb generates a difference in the construal of the event. Recall from–section 3.2.5.2.1, that changes in the light verb may bring about changes in valency patterns. For example, the intransitive anticausative vs. transitive causative alternation is realized by alternating between the light verb *honā*

"be" and *karnā* "do". This valency changing system resembles the anticausative system of simple verbs (Montaut 2016, Kachru 2006). This is shown by the fact that complex predicates occurring with the anticausative light verb *honā* pattern with constructions occurring with the anticausative simple verbs. In particular, they typically occur with a two arguments construction in which the Experiencer appears with an oblique marking, just like anticausative intransitive simple verbs. Additionally, the light verb in complex predicates usually conveys information at the lexico-semantic level (such as agentivity or directionality) and it may also express actionality properties (see Butt and Guder 2001).

Besides the light verb *karnā*, the noun *vicār* may occur with the light verb *honā* "be" and the light verb *ānā* "come". Notably, however, these two light verbs occur more rarely than *karnā*. When *vicār* appears with *honā*, it frequently does not feature an Experiencer. As is evident from example 326, in this case the light verb *honā* is interpreted as bearing an agentless passive reading and not an anticausative one. This is interesting and might be explained by the properties of the events expressed by the noun *vicār*. Indeed, agentive cognitions cannot be conceptualized as spontaneously happening out of the mind of the Experiencer, as it happens for example for perceptive situations such as "The Taj Mahal is visible from the opposite bank of the Yamuna" or "The Taj Mahal appeared on the horizon".

326. अर्जी पेश कर आया हुँ। उस पर विचार हो रहा है।

arjī	*peś=kar*	*ā-yā*	*hūṁ*
application(F.SG.NOM)	present=CP	come-PRF.M.SG	be.1SG.PRS
us=par	*vicār*	*ho*	*rah-ā*	*hai*
3SG.OBL=on	thought(M.SG.NOM)	be	PRGR-M	be.3SG.PRS

"I have presented the application. That is being considered."

The complex predicate *vicār honā* may also occur with an expressed Experiencer and remarkably when this happens the Experiencer is marked with the genitive (327) and rarely with the dative.

327. उनका विचार था कि तीनों को इंग्लैंड भेज कर शिक्षा के शिखर पर पहुँचा दें।

un=kā	*vicār*	*th-ā*	*ki*	*tīnoṁ=ko*	*iṁglēṁḍ*
3PL.OBL=GEN	thought(M.SG.NOM)	be.PST-M.SG	that	all_three=ACC	England
bhej=kar	*śikṣā=ke*	*śikhar=par*	*pahuṁcā*	*d-eṁ*
send=CP	education(F.SG.OBL)=GEN	peak(SG.OBL)=on	deliver	give-SBJV.3PL

"He thought that by sending all three [girls] to England he would provide them with the best education. Lit. His thought was that …"

The genitive marking on the Experiencer is interesting because only a limited number of Experiential verbs in Hindi allow it. While dative subjects are pervasive in the language, genitive subjects are rare and are typically confined to the expression of inalienable possession. However, the Experiencer of few complex predicates in Hindi can alternate between the dative and the genitive case marking. Previous studies have observed that the opposition between the genitive and the dative seems to imply also different semantic properties of the Experiencer. Montaut (2016) argues that the dative construction emphasizes the subject's affectedness and transience of the state, whereas the genitive construction emphasizes the stative nature of the process. She argues that the genitive marking on the subject in complex predicates "generally occur[s] with very weakly dynamic notions" (2016: 166). This seems to be supported by the distribution of the alternation between genitive and dative subjects across the Hindi verbal lexicon. Table 46 represents other experiential complex predicates that can alternate between a genitive and a dative Experiencer in the language (see on this also Caracchi 1992: 181). As it is evident from the table, these predicates usually refer to volitions and cognitions which can be construed as featuring an Experiencer endowed with control and volitionality, such as *irādā honā* "intend, decide", *vicār honā* "think, reflect"*, viśvās honā* "believe, trust" and *icchā honā* "want".

The genitive construction basically shows a stative semantics in Hindi, as it is suggested by its association with inalienable possession. It is reasonable to conjecture that the semantics of the genitive Experiencer results from the combination of the stative semantics of the genitive construction and the semantic component of control and volitionality implied by the nominal host in the complex predicate. In sum, when occurring with verbs of cognition, the genitive case profiles the cognition as a state resulting from a mental activity volitionally initiated and, through a typologically quite common metaphor (Luraghi 2014: 113), it construes the Experiencer as the possessor of this resultant state. This would be supported by the fact that verbs referring to cognitive states which do not imply an agentive Experiencer and an activity interpretation, such as *khabar honā* or *patā honā* "know, lit. information be" do not allow a genitive marking.

Table 46: Experiential complex predicates that can alternate between a genitive and a dative Experiencer.

VERB	MEANING	CONSTRUCTION	
		Dative Exp	**Genitive Exp**
icchā honā	want	✓	✓
irādā honā	decide	(✓)[30]	✓
āśā honā	hope	✓	

[30] The brackets signify that the verb may occasionally appear with dative marking on the Experiencer, but these instances are very rare, these could be explained by the highly agentive nature of the situations that these verbs denote.

vicār honā	think	(✓)	✓
khayāl honā	think	✓	✓
patā honā	know	✓	
yād honā	remember	✓	
viśvās honā	believe	✓	✓
pasand honā	like	✓	
dayā honā	have mercy	✓	
bharosā honā	trust	✓	
cintā honā	worry	✓	

The light verb *ānā* means "come" and contributes a non-causal achievement reading as is evident from example 328 and 329. In both sentences, the verb does not encode a controlled mental process; rather, it indicates a change of state in the mind of the Experiencer, like the sudden rise of an idea. When occurring with this light verb the complex predicate *vicār ānā* does not allow a genitive marking on the Experiencer. This verb may occur either with a dative subject (as in 328) or with a locative construction in which a body part is marked with the inessive case and the Experiencer is encoded as a genitive modifier and construed as the possessor of the body part, as in 329. Example 44 is interesting because the PP *dil=meṁ* "in (his) heart" suggests that verbs of cognitions, especially when referring to non-agentive changes-of-state such as *vicār ānā*, border with verbs of volitions. Here the sentence clearly points toward a volition reading of the verb *vicār ānā*, as the PP locates the experience in the heart of the Experiencer, which is the body part that is typically associated with emotions and volitions, while cognitions are typically associated with the mind.

328. एक क्षण के लिए उसे दिल में यह औपन्यासिक विचार आया भी कि वह सब कुछ छोड़कर चिकित्सा-विद्या पढ़े।

ek *kṣaṇ*(M.SG.OBL)=*ke lie* *use* *dil=meṁ* *yah*
one moment=for 3SG.DAT heart(M.SG.OBL)=in this.NOM
vicār *āyā* *bhī*
thought(M.SG.NOM) come-PRF.M.SG also
ki vah sab kuch choḍ=kar cikitsā-vidyā paḍhe.
that he should leave everything and study medicine
"For a moment she thought in her heart (lit. had the thought in her heart) that she should leave everything and study medicine."

329. उनकी व्यावहारिक बुद्धि में यह विचार ही न आता था कि जालपा किसी और चीज से अधिक प्रसन्न हो सकती है।

un=kī	*vyāvahārik*	*buddhi=meṁ*	*yah*	*vicār*
3PL.OBL=GEN	practical	mind(F.SG.OBL)=in	this	thought(M.SG.NOM)
hī	*na*	*ā-t-ā*	*th-ā*	
EMPH	not	come-IPRF-M.SG	be.PST-M.SG	

ki jālapā kisī aura cīja se adhika prasanna ho saktī hai.
that Jalpa could be happier than anything else
"It did not occur to his practical mind that Jalpa could be happier than anything else."

Another complex predicate used to encode thinking in Hindi is formed by the nominal host *khayāl*, a noun borrowed from Arabic that means "thought, idea, opinion". Like *vicār,* this nominal host may alternate among *honā* "be", *ānā* "come" and *karnā* "do" and accordingly occur with different constructions. The constructional distribution of this nominal host resembles that of *vicār* When occurring with the light verb *honā* the complex predicate may appear either with a genitive Experiencer (330) or with a dative one (331), even though it strongly prefers the genitive marking. When it occurs with the light verb *ānā* it consistently features a dative Experiencer (332) and never a genitive one. Lastly with the light verb *karnā* it occurs with a nominative/ergative Experiencer (333).

330. लोगों का ख़याल था कि सुबह होने से पहले ही वह चल बसेगा।

log-oṁ=kā	*khayāl*	*th-ā*	*ki*
people-PL.OBL=GEN	thought(M.SG.NOM)	be.PST-M.SG	that

subah hone=se pahle hī vah cal basegā.
that he would pass away before morning
"People thought that he would pass away before morning."

331. हरबंस को शायद ख़याल था उसकी इस भूख को मैं मिटा सकता हू हुँ।

harabaṁs=ko	*śāyad*	*khayāl*	*th-ā*
Haramb=DAT	maybe	thought(M.SG.NOM)	be.PST-M.SG

us=kī is bhūkh=ko maiṁ miṭā saktā hūṁ
that I could erase his hunger
"Maybe Haramb thought that I could erase his hunger."

332. सहसा उसे ख़याल आया, गुप्त पुलिस वाले सादे कपड़े पहने इधर-उधर घूमा करते हैं।

sahsā *use* *khayāl* *ā-yā*
suddenly 3SG.DAT thought(M.SG.NOM) come-PRF.M.SG
gupt pulis-vāle sāde kapḍe pahne idhar-udhar ghūmā karte haiṁ
policemen in disguise dressed in plain clothes roam around here and there
"It suddenly occurred to him that policemen in disguise might be roaming around here and there dressed in plain clothes."

333. बेहतर हो कि आप मेरे फायदे का इतना ख़याल न करें।

behtar ho *ki* *āp* *mer-e* *phāyad-e=kā*
better be.3SG.SBJV that 2HON.NOM 1GEN-M.SG.OBL benefit(M)-SG.OBL=GEN
itnā *khayāl* *na* *kar-eṁ*
so_much thought(M.SG.NOM) not do-3PL.SUBJ
"You should not think so much about my benefit."

Despite their similarities, *vicār* and *khayāl* also differ in many respects. Notably, they show a quite different distribution in relation to the frequency of light verbs. *Vicār* clearly favors the light verb *karnā*, while *khayāl* mainly occurs with the light verb *honā*. As a matter of fact, *khayāl* also favors the light verb *ānā* to *karnā*, which is very rare with *vicār* in the corpus. This seems to point toward a quite different semantic characterization of these nominal hosts. As discussed above, *vicār* indicates an ongoing, intentional mental activity. This is supported both by its tendency to occur with an oblique Stimulus and by the fact that it favors the light verb *karnā* "do" in the corpus, thus implying control and volitionality of the Experiencer. *Khayāl,* on the other hand, seems to point toward a lower degree of agency and a more static interpretation. Table 47 represents the aspectual forms appearing with *khayāl* complex predicates in the corpus and their distribution among the different light verbs. The light verb *honā* is the most frequent and clearly correlates with the imperfective aspect, while the light verb *ānā* almost always appears in the perfective form. The light verb *karnā* appears so rarely in the corpus with the noun *khayāl* that it is impossible to make any generalization. Data from Table 47 seem to suggest that *khayāl* construes the event as stative when occurring with the light verb *honā*. In contrast, when it appears with the light verb *ānā,* it construes thinking as an achievement in which the Experiencer lacks control and intentionality. In 332 above, for example, the complex predicate *khayāl ānā* refers to a sudden rise of awareness in the mind of the participant, the suddenness of the event is explicitly expressed by the adverb *sahsā* "suddenly".

Table 47: Aspectual distribution of complex predicates formed by the nominal host *khayāl.*

Light verb	PRF	IPRF	IMP	SBJV	CP	TOT
ānā	14	1	0	0	0	15
honā	1	44	0	0	0	45
karnā	1	0	1	1	1	4
TOT	**16**	**45**	**1**	**1**	**1**	**64**

8.5. Know

In this section, I address the expression of knowledge in Hindi. Knowledge is typically construed in two main ways: as a state, i.e. as the presence of some information in the experience's mind, or as an achievement/accomplishment, i.e. as a change-of-state in the experience's mind resulting from the acquisition of some piece of information that was previously unknown. In the next section (8.5.1). I deal with the Hindi expression of knowing as state. And in the following section (8.6), I will focus on verbs expressing the acquisition of knowledge in Hindi.

8.5.1. Knowing as a state

Knowing as a state can refer either to the possession of intellectual knowledge or to the possession of skills. Additionally, languages of the world may separate and lexicalize in different ways "knowing a person or a thing" and "knowing a fact" (Fortescue 2001). As I will discuss, Hindi does not lexically distinguish these two latter situations and encodes them with the same verb, namely *jannā* "know", but it uses constructional ways to keep them distinct. In particular when referring to the knowledge of a fact, the Stimulus is typically encoded with a finite complement clause. On the other hand, Hindi lexically distinguishes the possession of skills, and expresses this situation with a construction featuring the auxiliary verb *ānā* "come" (see below in this section). Notably, however, this distinction only appears when knowledge is construed as a state. When encoding the acquisition of an ability or a skill Hindi uses the verb *sikhnā* "learn", which can also be used to encode the acquisition of some piece of information through some mental effort. *Jānnā* is the most frequent verb used for the expression of knowledge in Hindi. This verb is connected to the Sanskrit root *jñā-* "know" (Turner 1971), which is etymologically related to the Indo-European root **gno-* quite common in European languages (see for example English *know*, French *connatre*, Italian *conoscere*, German *kennen*, Swedish *känna* etc.). This verb is used to express both "know something" (as in 334 and 335) and "know someone" (as in 336). In sentence 336, the Stimulus is instantiated by a person (*Ganda Sing*) and for this reason the object is marked by the postposition *ko*.

334. जालपा पति की आर्थिक दशा अच्छी तरह जानती थी।

jālpā *pati=kī* *ārthik* *daśā* *acch-ī*
Jalpa.NOM husband(M.SG.OBL)=GEN economic condition(F.SG.NOM) good-F.SG
tarah *jān-t-ī* *th-ī.*
way know-IPRF-F.SG be.PST-F.SG
"Jalpa knew very well the economic condition of her husband."

335. रतन ने ख़त का जवाब क्यों नहीं दिया- मामूली शिष्टाचार भी नहीं जानती?

ratan=ne *khat=kā* *javāb* *kyoṁ* *nahīṁ*
ratan=ERG letter(M.SG.OBL)=GEN answer(NOM.M.SG) why not
diyā *māmūlī* *shishtācār* *bhī* *nahīṁ* *jān-t-ī*
give.PRF.M.SG ordinary good_conduct(M.SG.NOM) even not know-IPRF-F.SG
"Why didn't Ratan reply to the letter - she doesn't know even simple manners?"

336. और गंडासिंह को जानते हो?

aur *gaṁḍā siṁh=ko* *jān-t-e* *ho?*
and Ganda Sing=ACC know-IPRF-F.SG be.PRS-2PL
"And do you know Ganda Sing? "

Jānnā consistently features a transitive construction in which the Experiencer is encoded with the nominative/ergative and the Stimulus NP stands in the nominative/accusative. Besides NP Stimuli, *jānnā* also occurs with complement clauses (as in 337) with a comparatively similar frequency (see Table 48).

337. मैं बस इतना जानती हुँ कि मैं अपने वर्तमान से बाहर आना चाहती हुँ।

maiṁ *bas* *itnā* *jān-t-ī* *hūṁ* *ki*
1SG.NOM just as_much know-IPRF-F.SG be.PRS.1SG that
maiṁ apne vartamān se bāhar ānā cāhtī hūṁ
I want to come out of my present.
"All I know is that I want to come out of my present."

Table 48 shows the distribution of the aspectual forms over the random occurrences I have manually scrutinized.

Table 48: Aspectual characterization of *jānnā.*

Construction	IPRF	PRF	SBJ	CP	Tot
Non finite complement clause	1	0	0	0	1
Finite complement clause	82	0	28	0	110
Transitive	78	1	4	1	89
Total	**166**	**1**	**32**	**1**	**200**

Notably this verb mostly occurs with the imperfective, and it appears in the perfective aspect only once (338). This correlates with the fact that "know" is usually construed as an unbounded state and not as a telic dynamic event. As mentioned in section 8.2, the perfective aspect profiles the event expressed from a global point of view and it is used to indicate that an action or an event is completed, often with a specific focus on the result or outcome of that action. For this reason, verbs that encode stative unbounded situations (such as *jānnā*) tend to reject this aspectual characterization. As a consequence, when the verb *jānnā* features the perfective aspect, it takes on a different reading: the event is construed as bounded, and the verb is no longer interpretated as a state but as an achievement. This can be seen in 338, in which the verb is better translated as "understand, get to know" rather than as "know".

338. जिसने अभी जीवन का मर्म नहीं जाना, वह मरना क्या जाने?

jis=ne	*ab=hī*	*jīvan=kā*		*marm*
REL.OBL=ERG	now=EMPH	life(M.SG.OBL)=GEN		essence(M)SG.NOM
nahīṁ jān-ā,	*vah*	*mar-nā*	*kyā*	*jā-ne?*
not know-PRF.M.SG	3SG.NOM	die-INF	what	know-SBJV.3SG

"The one who has not yet known/understood the essence of life, what does he know about death?"

The verb *jānnā* also occurs very frequently in a highly schematic construction, which I labelled the subjunctive construct (given in 339). The verb is in the subjunctive mood and stands in the third person singular; the subject is unexpressed and there is always the negation particle *na*. When expressing doubt or nescience, the subjunctive construction is often used to emphasize uncertainty in Hindi. This construction only occurs in the corpus either with finite complement clauses introduced either by the subordinating conjunction *ki* or by an interrogative particle (*kyā* "what", *kab* "when", *kaise* "how", *kyoṁ* "why" and so on). Notably, this construction has been grammaticalized, it has partly lost its original semantics and the verb seems to have evolved into a sort of discourse marker

of nescience or lack of understanding: it always implies an impersonal reading, and it is translated into English with “no one knows, who knows”.

339. लेकिन न जाने क्यों मुझे उनकी गुस्ताखी बुरी लग रही थी।

lekin	*na*	*jān-e*	*kyoṁ*	*mujhe*	*un=kī*
but	not	know-SBJV.3SG	why	1SG.DAT	3PL.OBL=GEN
gustākhī		*bur-ī*	*lag*	*rah-ī*	*th-ī.*
impudence(F)SG.NOM		bad-F	attach	PRGR-F	be.PST-F.SG

“But no one knows why I was feeling bad about his impudence / But who knows why I was feeling bad about his impudence.”

Other constructions occur more rarely in the corpus. Sentence 340 is an example of a non-finite complement clause in which the Stimulus is expressed by an infinitive. This construction is quite rare, and I only found two occurrences out of 200 scrutinized from the corpus. The low frequency is probably explained by the fact that, as 340 shows, the semantics of this construction basically implies the knowledge of a skill. However, Hindi displays a specific construction that is specialized to express possession of abilities and that employs the verb *ānā* “come”. In this construction, the Experiencer stands in the dative, while the ability or the skill is expressed by the infinitive of the verb. The verb is *ānā* “come” and agrees with the infinitive and stands in the masculine third person singular (see sentence 341 and 342)[31].

340. पूरा लट्ठ मारना जानता था, पर अपनी रक्षा करना न जानता था, जो लड़ाई में मारने से ज्यादा महत्व की बात है।

pūr-ā	*laṭṭh*	*mār-nā*	*jān-t-ā*		*th-ā,*	*par*
full-M.SG	stick(M.SG.NOM)	hit-INF	know-IPRF-M.SG		be.PST-M.SG	but
apn-ī	*rakṣā*		*kar-nā na*	*jān-t-ā*		
REFL-F.SG	protection(F)SG.NOM		do-INF not	know-IPRF-M.SG		
th-ā,	*jo laṛāī meṁ mārne se jyādā mahatv kī bāt hai.*					
be.PST-M.SG	which is more important than hitting in a fight.					

[31] Note that when the infinitive occurs with an object, the verb may show agreement with it as in the example below:

mujhe	*hindī*	*bol-n-ī*	*ā-t-ī*	*hai.*
1SG.DAT	hindi.F.SG.NOM	speak-INF-F	come-IPRF-F	be.PRS.3SG

“I can speak Hindi.”

"He knew how to swing the bat, but he didn't know how to protect himself, which is more important than hitting in a fight."

341. लेख लिखना मुझे आता ही है, पत्रों मे आयुर्वेद-महत्व पर दो-चार लेख लिख दूंगा, और लिखूगां भी जरा चटपटी भाषा मे।

lekh	*likh-nā*	*mujhe*	*ā-t-ā*	*hī*	*hai.*
article(M.PL.NOM)	write-INF	1SG.DAT	come-IPRF-M.SG	EMPH	be.PRS.3SG

"I know how to write articles,
āyurved-mahatv par do-cār lekh likh dūṁgā,
"I know how to write articles, I will write one or two articles on the importance of Ayurveda."

342. उसकी कमीज़ के आधे बटन खुले थे। इतनी बड़ी होकर भी उसे शरीर का होश रखना नहीं आया था।

uskī kamīj ke ādhe baṭan khule the. itnī baḍī hokar bhī

use	*śarīr=kā*	*hoś*	*rakh-nā*
1SG.DAT	body(M.SG.OBL)=GEN	consciousness(M.SG.NOM)	keep-INF

nahīṁ	*ā-yā*	*th-ā.*
not	come-PRF.M.SG	be.PST-M.SG

"Half the buttons of her shirt were open. Despite being so old, she did not know how to be aware of her body."

In Hindi knowledge can also be expressed by several complex predicates. The following part of this section is dedicated to these verbs. The verb *jān paṛnā* is a complex predicate formed by the nominal host *jān* "knowledge, acquaintance, understanding, opinion" which is clearly related to the verb *jānnā* "know", and the light verb *paṛnā* meaning "fall". Even though the noun *jān* is connected to the generic verb expressing knowledge in Hindi, the verb *jān paṛnā* shows a quite different semantic characterization and accordingly it occurs with different constructions and different aspectual distributions than *jānnā*. In the first place, unlike *jānnā,* it may occur in a construction with a single argument expressed by a nominative NP and an adjective or a participle in a predicative function (343 and 344). In these cases, it is translated as "seem, look like". The same verb may also appear with a dative Experiencer (as in 345 and 346). In such cases, it conveys a vaguely characterized experience that lies at the intersection of perceptions and cognitions. This experience can be translated in various ways, including "feel", "find something as", "know" or "seem to someone".

343. डॉक्टर ने कहा, 'मैं तुम्हारी और तुम्हारे घर के लोगों की जांच करना चाहता हुँ। बच्चे का रोग पुश्तैनी जान पड़ता है।

ḍākṭar ne kahā, 'maiṁ tumhārī aur tumhāre ghar ke logoṁ kī jāṁc karnā cāhtā hūṁ.
The doctor said: "I want to test both you and the members of your family.

bacc-e=kā	*rog*	*puśtainī*	*jān*
child(M)-SG.OBL=GEN	disease(M.SG.NOM)	hereditary	knowledge(F.SG.NOM)
paṛ-t-ā	*hai*		
fall-IPRF-M.SG	be.PRS.3SG		

"The doctor said: 'I want to test both you and the members of your family. The child's disease seems to be genetic."

344. बड़े घर की स्त्री जान पड़ती हैं।

baṛ-e	*ghar=kī*	*strī*	
big(M)-SG.OBL	house(M)SG.OBL=GEN	woman(F.SG.NOM)	
jān	*paṛ-t-ā*	*hai*	
knowledge(M.SG.NOM)	fall-IPRF-M.SG	be.PRS.3SG	

"She looks like the matron of a big house."

345. उस फटी हुई दरी पर बैठना रमा को अपमानजनक जान पड़ा।

us	*phaṭī*	*huī*	*darī=par*	*baiṭh-nā*	*ramā=ko*
that.OBL	torn	be.PRF-F	carpet(F.SG.OBL)=on	sit-INF	rama=DAT
apmānjanak	*jān*		*paṛ-ā.*		
humiliating	knowledge(M.SG.NOM)		fall-PRF.M.SG		

"Rama found it humiliating to sit on that torn carpet."

346. यह सभी चीजें न जाने क्यों मुझे अपरिचितसी जान पड़ती हैं।

yah	*sabhī*	*cīj-eṁ*	*na*	*jān-e*	*kyoṁ*
this.NOM	all=EMPH	things(F)-SG.NOM	not	know-.3SG.SBJV	why
mujhe	*aparicitsī*	*jān*		*paṛ-t-ī*	*haiṁ*
1SG.DAT	unfamiliar	knowledge(M.SG.NOM)		fall-IPRF.M.SG	be.PRS.3PL

"I don't know why all these things seem unfamiliar to me."

Notably, *jān paṛnā* always occurs either with a finite complement clause or with a predicative clause. In the predicative clause construction, the Stimulus is encoded by a nominative NP linked to a

secondary predication, which may be instantiated by a noun (344), an adjective (346) or a participle (347). The NP Stimulus agrees with the participle or the adjective and is included into the argument structure of the verb. Table 49 shows the frequency in the corpus of constructions occurring with this verb and their aspectual distribution. As the data shows, this verb occurs with a predicative construction more than half of the times, and this gives us valuable insights on its semantics. Predicative (participle) clauses are typical of verbs of perceptions, and they express the direct perception of state of affairs (section 7.5). Since predicative participle complements have the distinctive property to express simultaneity between two events, they are not only typical of verbs of perceptions, but they tend to be marginalized and restricted to this semantic class. The fact that *jān paṛnā* allows predicative participle complements indicates that the semantics of this verb has some connection with the domain of perception. However, unlike perception verbs, *jān paṛnā* also frequently occurs with adjectives and nouns as predicative elements, and it never occurs with an isolated NP Stimulus (not linked to a predicative element), thus suggesting that although sharing some properties with perception verbs it is also semantically distinct form this class. In particular, this verb seems to be on the border between perception and cognition and its meaning seems to be "feel, seem (to someone), look like (to someone)" and to imply a feeling or an opinion on an external situation based on perception (as in 347).

347. जालपा को इन शब्दों में स्नेह और सहानुभूति का एक सागर उमड़ता हुआ जान पड़ा।

jālpā=ko	*in*	*śabd-oṁ=meṁ*	*sneh*	*aur*
japla=DAT	these.OBL	word(F)-PL.OBL=in	affection(M.SG.OBL)	and
sahānubhūti=kā	*ek*	*sāgar*	*umaṛ-t-ā*	
sympathy(F.SG.OBL)=GEN	one	ocean(M.SG.NOM)	rise-IPRF-M.SG	
hu-ā	*jān*	*paṛ-ā.*		
be.PRF-M.SG	knowledge(M.SG.NOM)	fall-PRF.M.SG		

"Jalpa felt an ocean of affection and sympathy rising in these words."

As this verb is on the border with verbs of cognitions, it also frequently appears with finite complement clauses. Sentence 348 shows a sentential complement instantiated by an objective clause with no subordinating conjunction and anticipated by the correlative *esā* "this" in the main clause. Notably, in this example, the complement clause that expresses the content of the cognition resembles a direct speech: in the main clause the Experiencer is expressed by *use* "to him", which is the third person singular pronoun in the dative form. In the complement clause, the subject, which is coreferential with the Experiencer in the main clause, is in the first-person singular. The content of

the cognition is represented literally as it is propositionally expressed in the mind of the Experiencer. Sentence 349 shows an objective clause introduced by the subordinating conjunction *ki* "that".

348. उसे ऐसा जान पड़ा, मेरे पैरों के नीचे की जमीन धंस गयी।

use	*ais-ā*	*jān*	*paṛ-ā*
3SG.DAT	such-M.SG	knowledge(M.SG.NOM)	fall-PRF.M.SG

mere pairoṁ=ke nīce kī jamīn dhaṁs gayī.

"He felt as if the ground beneath his feet sank. Lit. He felt like this/it seemed to him like this: the ground beneath my feet sank."

349. उसे ऐसा जान पड़ता था कि कोई दैवी शक्ति मेरी मदद कर रही है।

use	*aisā*	*jān*	*paṛ-ṭ-ā*
3SG.DAT	such-M.SG	knowledge(M.SG.NOM)	fall-IPRF.M.SG
th-ā	*ki*		
be-PST.M.SG	that		

koi daivī śakti merī madad kar rahī hai.

"It seemes to him that some divine power was helping him."

The hypothesis that *jān paṛnā* does not express a general knowledge, such as *jānnā* "know", is also supported by the fact that it displays a different aspectual distribution as compared to its cognate simple verb. The aspectual analysis of *jān paṛnā* is reported in Table 49 and it reveals that unlike *jānnā* which almost never occurs in the perfective aspect (

Table **48**), *jān paṛnā* appears in the perfective 76 times over 217. The frequency relatively low in comparison to the imperfective occurrences tells us that, like *jānnā*, *jān paṛnā* prefers the imperfective aspect, however its semantics is not restricted to the expression of a cognitive state as it frequently allows a perfective characterization and consequently a telic reading (as in 348 above).

Table 49: Complement types and aspectual distribution of *jān paṛnā.*

Expericencer	Complement type	IPRF	PRF	PRGR	FUT	INCH	Tot
Present	Finite compl clause	17	19	1	1	0	38
	Predicative clause	13	34	1	1	1	50
Absent	Predicative clause	51	8	0	0	0	60
	Finite compl clause	52	15	0	1	1	69
Total		**133**	**76**	**3**	**3**	**2**	**217**

Another verb etymologically related to *jannā* and *jān paṛnā* is the complex predicate *jñāt honā* "know, understand, realize" composed by the adjective *jñāt* "known" (derived from the past participle *jñāta* of the Sanskrit verb *jñā-* "know") and the light verb *honā* "be". Unlike *jān paṛnā* which may appear in a single argument construction, this verb consistently appears with a dative Experiencer. The Stimulus can be differently encoded: most frequently, it is expressed with a complement clause, as in 350 or 352), but it also rarely occurs with a NP Stimulus in the nominative, as in 351. This verb occurs rarely in the corpus (only 60 occurrences), however, it is interesting because it seems to have a very specific aspectual characterization, as it almost only appears in the perfective. Table 50 shows the aspectual distribution of this verb across the constructions in which it occurs.

Table 50: Complement types and aspectual distribution of *jñāt jonā.*

Complement type	**PRF**	**PRGR**	**PRS**	**FUT**	**Total**
Finite complement clause	46	3	6	2	57
Nominative NP	1	0	0	1	2
Total	**47**	**3**	**6**	**3**	**59**

As it is clear from a quick comparison of this table with

Table 48, which shows the aspectual characterization of *jannā*, *jñāt honā* displays a distribution that is complementary to that of *jannā*. They both express knowledge, but *jannā* prompts an atelic reading which construes knowing as an unbounded mental state, while *jñāt honā* seems to point toward a telic interpretation and depict knowledge as a mental achievement in which the Experiencer becomes aware of some piece of information.

350. उसे देखकर मुझे अपने दोनों भाइयों की, पिता की याद आयी और तब मुझे ज्ञात हुआ कि यही एक ज्ञानी पुरुष है।

use dekh=kar mujhe apne donoṁ bhāiyoṁ=kī, pitā=kī yād āyī aur tab
Seeing him, I remembered my two brothers, my father and

mujhe	*jñāt*	*hu-ā*	*ki*	*yah=ī*	*ek*	*jñānī*
1SG.DAT	known	be.PRF-M.SG	that	this.NOM=EMPH	one	wise

puruṣ	*hai.*
man(M).SG.NOM	be.3SG.PRS

"Seeing him, I remembered my two brothers and my father and I realized that he was a wise man."

351. तुम्हें यह शीघ्र ही ज्ञात हो जायेगा।

tumheṁ	*yah*	*shīghr*	*hī*	*jñāt*	*ho*	*ja-ye-g-ā*
2PL.DAT	this	soon	EMPH	known	be	go-FUT-M.SG

"You will soon understand/realize this."

352. अब मुझे ज्ञात हुआ कि वह कौनसा रहस्य था जिसने तुझे मेरी दृष्टि में इतना सुन्दर, इतना चित्ताकर्षक बना दिया था।

ab	*mujhe*	*jñāt*	*hu-ā*	*ki*	*vah*	*kauns-ā*
now	1SG.DAT	known	be.PRF-M.SG	that	3SG.NOM	which-M.SG
rahasy		*th-ā*				
secret(M.SG.NOM)		be.PST-M.SG				

jis ne tujhe merī dṛṣṭi meṁ itnā sundar, itnā cittākarṣak banā diyā thā.
that made you so beautiful, so captivating in my eyes.
"Now I realized what was the secret that made you so beautiful, so captivating in my eyes."

Besides the verbs related to the Sanskrit root *jñā-*, Hindi displays other verbs expressing knowledge. In particular, two complex predicates are quite frequent: one is formed by the adjective *mālūm* which is a Persian-Arabic borrowing and means "known" and the other is formed by the Indo-Aryan noun *patā* "information". These two nominal hosts differ in the way they construe knowledge, and this is shown by the light verbs they appear with and the distribution of their aspectual forms. As the data in Table 51 show *mālūm* strongly prefers the light verb *honā* which mainly occurs with the imperfective aspect but also appears 46 times (over 179 times) with the perfective. The aspectual characterization contributes to construe the event differently. The fact that *mālūm honā* mostly occurs with the imperfective suggests that this verb is basically atelic. When it occurs in the perfective, the verb does no longer profile an unbounded state, as the perfective aspect implies that the event is profiled in its globality with a focus on the ending point, hence the verb here acquires an achievement reading such as "realize, understand". See the contrasting examples showing an imperfective (353) and a perfective (354) form of *mālūm honā*.

353. मैं तुमसे कितना मुहब्बत करता हुँ, यह शायद तुम्हें मालूम न हो।

maiṁ	*tum=se*	*kitnā*	*muhabbat*	*kar-t-ā*	*hūṁ,*
1SG.NOM	2PL=INS	how_much	love(M.SG.NOM)	do-IPRF-M.SG	be.1SG.PRS
yah	*śāyad*	*tumheṁ*	*mālūm na*	*ho.*	

this.NOM maybe 2PL.DAT know not be.3SG.SBJV
"Maybe you don't know how much I love you."

354. मैंने आश्चर्य से कहा, 'आपको उसका नाम कैसे मालूम हुआ?

maiṁ=ne	*āścary=se*		*kah-ā,*	*āp=ko*
1SG.NOM	surprise(M.SG.OBL)=INS		said-PRF.M.SG	2SG.HON
us=kā	*nām*	*kaise*	*mālūm*	*hu-ā?"*
3SG.OBL=GEN	name(M.SG.NOM)	how	know	be.PRF-M.SG

"I said in surprise, 'How did you find out his name?'"

The basically stative semantics of this verb is shown also by the fact that it most frequently occurs with what I have referred to as the habitual imperfective form of the verb *honā* "be" in Table 51. This form is realized by the imperfective participle of the verb *honā* followed by the same verb in its auxiliary function and it is used in Hindi when the speaker wants to refer to events or situations that are always and generically true (McGregor 1986, Kachru 2006, Montaut 2004). Consider for example the use of the general present of *honā* in 355(b) as opposed to the normal present in 355(a) (adapted from McGregor 1986: 18). The same form of the verb may also be used as an auxiliary adding an iterative reading (as in 356, adapted from Montaut 2004: 122). As is evident from these examples, the imperfective participle of the verb *honā* may add various semantic nuances, but in general it highlights the stative durative character of the verb "be", as it is used to express situations that are considered to be always true.

355. *(a) yah* *gār-ī* *lāl* *hai*
this.NOM car(F)-SG.NOM red be.3SG.PRS
"This car is red."

(b) gār-iyāṁ *mahaṁg-ī* *ho-t-ī* *haiṁ*
car(F)-SG.NOM expensive-F be.IPRF-F be.3SG.PRS
"Cars are expensive."

356. *jab kabhī* *maiṁ* *rasoī=ke andar* *dekh-t-ā*
whenever 1SG.NOM kitchen(F.SG.OBL)=inside look-IPRF.M.SG
th-ā *vah* *pī* *rah-ī* *ho-t-ī* *th-ī*
be.PST-M.SG 3SG.NOM drink PRGR-F be-IPRF-F.SG be.PST-F.SG
"Whenever I looked inside the kitchen she used to be drinking."

357. तुम्हारे चरित्र की विचित्रता मुझे बहुत भली मालूम होती थी।

tumhāre	*caritr=kī*	*vicitratā*	*mujhe*
2PL.GEN	character(M.SG.OBL)=GEN	strangeness(F)SG.NOM	1SG.DAT

bahut bhalī mālūm ho-t-ī *th-ī.*
very well known be-IPRF-F.SG be.PST-F.SG

"I knew very well (meaning I was very well aware of) the strangeness of your character."

Table 51: Aspectual distributions across the light verbs occurring with *mālūm.*

	FUT	HAB IPRF	IPRF	INCH	PRF	PRGR	SUBJ	Tot
honā	7	68	45	3	46	5	5	179
ho jānā	4	0	5	0	9	0	2	20
rahnā	0	0	0	0	0	0	1	1
Total	**11**	**68**	**50**	**3**	**58**	**5**	**7**	**200**

The complex predicate *mālūm honā* most frequently appears with a dative Experiencer. However, it also often occurs in a single argument construction in which only the Stimulus is encoded, and which is generally interpreted with an impersonal reading (as in 359 below). Notably, the Stimulus NP is frequently linked to a predicative element, in a construction that resembles the predicative construction that I have already discussed in relation to the complex predicate *jān paṛnā* (see 345, 346 and 347). Interestingly, as it happens with *jān paṛnā*, *mālūm honā* mostly features a noun or an adjective as secondary predication, but rarely also appears with a participle (as in 358). In 358, the verb seems to encode a situation in which the experiencer (Birbal Singh) experiences some perception that is triggered by an external stimulus, and it clearly does not encode a cognitive situation. As I discussed above in this section, this is interesting because the participle predicative construction tends to be specifically used by verbs of direct perceptions due to its capacity to encode simultaneity between two events. When it is used with cognition verbs, this construction adds its specific semantics and contributes to construing the event as lying somewhere in between perceptions and cognitions. In other words, in this sentence a verb that is typically associated with cognition acquires the meaning of a verb of perception because it occurs is a construction that is specifically associated with perceptions, thus showing that the construction may contribute its own a specific semantics.

358. बीरबल सिंह को आज उनके चेहरों पर एक नयी स्फूर्ति, एक नया उत्साह, एक नया गर्व झलकता हुआ मालूम होता था।

bīrbal siṁh=ko	*āj*	*un=ke*	*cehr-oṁ=par*	*ek*	*nay-ī*
Birbal Singh=DAT	today	3PL.OBL=GEN	face(M)-PL.OBL=on	one	new-F
sphūrti,	*ek*	*nay-ā*	*utsāh,*	*ek*	*nay-ā*
vitality(F.SG.NOM)	one	new-M	enthusiasm(M.SG.NOM)	one	new-M

garv jhalak-t-ā hu-ā mālūm ho-t-ā th-ā .
pride(M.SG.NOM) reflect-IPRF-M.SG be.PRF.M.SG known be-IPRF-M.SG be.PST-M.SG
"Birbal Singh could see a new vitality, a new enthusiasm, a new pride reflected on their faces today."

Even though it clearly prefers the light verb *honā*, the adjective *mālūm* can also appear with other light verbs which contribute a different semantic construal of the event and highlight different components of meaning. In sentence 359, *mālūm* is followed by the light verb *rahnā* "stay" which highlights even more a stative reading. While in sentence 360, the verb occurs with the V-V complex predicate *ho jānā* formed by the main verb *honā* "be" and the light verb *jānā* "go" and means "become". This V-V sequence implies a certain degree of dynamicity that is contributed by the light verb *jānā* "go" and points toward an achievement reading as shown in 360.

359. "यह हिंदुस्तान है, यूरोप नहीं है।""इंसान का स्वभाव सारी दुनिया में एक-सा है।"

"मगर यह भी मालूम रहे कि हर एक कौम में एक ऐसी चीज है जिसे उसकी आत्मा कह सकते हैं।"

yah hiṁdustān hai, yūropanahīṁ hai. iṁsān=kā svabhāv sārī duniyā=meṁ ek-sā hai.
'This is India, not Europe.' 'Human nature is the same all over the world.'

magar yah	*bhī*	*mālūm*	*rah-e*	*ki*	*har*	*ek*	
but this.NOM	also	known	stay-SBJV.3SG	that	every	one	
kaum=meṁ		*ek*	*ais-ī*	*cīj*	*hai*		*jise*
community(F.SG.OBL)=in		one	such-F	thing(F.SG.NOM)	be.3SG.PRS		REL.ACC
us=kī	*ātmā*			*kah*	*sak-t-e*	*haiṁ*.	
3SG.DAT=GEN	soul(F.SG.NOM)			say	can-IPRF-M.PL	be.3PL.PRS	

'But it should also be known that every community has something which can be called its soul.'"

360. अगर उसे मालूम हो जाए कि उसके रूपये तत्काल मिल सकते हैं, तो वह शांत हो जाएगी।

agar	*use*	*mālūm ho*	*jā-e*	*ki*	*us=ke*		
If	3SD.DAT	known be	go-SBJV.3SG	that	3SG.OBL=GEN		
rūpye		*tatkāl*	*mil*	*sak-t-e*	*haiṁ,*		*to*

rupees(M.PL.NOM) immediately find can-IPRF-M.PL 3PL.PRS then
vah *śāṁt* *ho* *jāegī.*
3SG.NOM calm be go-3SG-FUT-F
"If she finds out that she can get her money immediately, she will be calm."

Table 52 presents data for the complex predicates formed by the noun *patā* "information, clue". As is evident from the table, *patā* differs from *mālūm* in many respects. First of all, it allows a far wider variation on the light verb, in particular it frequently occurs with the verb *honā*, but shows a comparatively similar frequency with the verb *calnā* "walk, move" often appearing in the V-V complex predicate *cal jānā* "lit. walk go" contributing a dynamic reading. When it appears with *honā* "be", as in 361, this verb construes knowledge as a state, as is also supported by the correlation of the light verb *honā* with the imperfective aspect (Table 52). In contrast, *calnā* and *cal jānā* contribute a dynamic interpretation of the event, yet they do not show a striking aspectual characterization as *honā*. In particular, *patā calnā* seems to prefer the perfective aspect which indicates that is basically used as a telic verb (as in 362), but it may also appear in the imperfective (as in 363).

361. तुम्हें पता है सुरजीत की यह दूसरी या तीसरी शादी है?

tumheṁ *pat-ā* *hai* *surjīt=kī* *yah*
2PL.DAT information(M)-SG.NOM be.3SG.PRS Surjeet=GEN this.NOM
dūsr-ī *yā* *tīsr-ī* *śād-ī* *hai?*
second-F or third-F marriage(F)-SG.NOM be.3SG.PRS.
"Do you know whether this is Surjeet's second or third marriage?"

362. मुझे आज अच्छी तरह पता चल गया है कि जिस घर में मैं रहता हूँ, वह मेरा घर नहीं है और जिसे मैं अपनी पत्नी समझता हूँ, वह मेरी पत्नी नहीं है।

mujhe *āj* *acch-ī* *tarah* *pat-ā* *cal*
1SG.DAT today good-F way information(M)-SG.NOM walk
gay-ā *hai* *ki*
go.PRF-M.SG be.3SG.PRS that
jis ghar=meṁ maiṁ rah-t-ā hūṁ, vaha merā ghar nahīṁ hai aur jise maiṁ apnī patnī samajh-t-ā hūṁ, vah merī patnī nahīṁ hai.
"I have realized very well today that the house in which I live is not my house and the one whom I consider my wife is not my wife."

363. न जाने मुझमें ऐसी कौनसी बात है, जिससे दूसरों को फौरन पता चल जाता है कि मैं कौन हूँ, या क्या हूँ।

na jāne mujh=meṁ aisī kaunsī bāt hai,				*jis=se*		*dūsroṁ=ko*	
Don't know what it is about me				REL.OBL=INS		other-PL.OBL=DAT	
phauran	*pat-ā*			*cal*	*jā-t-ā*	*hai*	*ki*
immediately	information(M)-SG.NOM			walk	go-IPRF-M	be.3SG.PRS	that
maiṁ	*kaun*	*hūṁ,*	*yā*	*kyā*	*hūṁ.*		
1SG.NOM	who	be.1SG.PRS	or	what	be.1SG.PRS		

"Don't know what it is about me that let's others know at once who or what I am. Lit. Don't know what thing there is in me from which others know at once who or what I am. "

The use of the light verb *calnā* in an experiential complex predicate expressing knowledge is interesting for many reasons. First, the use of *calnā* as a light verb is quite rare, and usually experiential complex predicates involve other light verbs to construe the event as dynamic, mainly *ānā* "come" or *lagnā* "adhere". Second, when used as a main verb, *calnā* "walk, move" is an atelic durative verb, so it is quite unusual for it to appear in a complex predicate expressing knowledge, which is typically construed either as a state (atelic and durative, but no dynamic) or as an achievement (dynamic and telic). Notably, even in the imperfective, this verb seems to allow a telic reading while retaining the dynamic semantics contributed by the verb *calnā*. Consider sentence 363, in which the verb *calnā* occurring in the V-V complex predicate *cal jānā* seems to be associated with a telic iterative reading: the speaker is wondering what is it in his behavior or personality that allows others to form an immediate understanding of the speaker (telic reading), and he implies that this happens every time and it is not a one-time event (iterative reading). Notably, the verb shows a high range of alternation between many light verbs which seem to show a similar meaning. Consider for example sentence 364 in which the noun *patā* occurs with the light verb *lagnā* "adhere, be attached" with apparently no semantic difference from the use of *patā calnā* in sentence 362.

Interestingly, *patā* also appears with the light verbs *karnā* "do" and *lagānā* "attach (tr.)" which contribute to construing the situation in a completely different way. Recall from the introduction to this section, that knowledge is typically conceptualized either as a state or as an achievement. The complex predicates *patā karnā* and *patā lagānā*, in contrast, construe the event as a controlled activity. Their meaning is not "know" and they rather mean "find out (after an investigation), investigate". The agentive semantic component is contributed by the light verbs which typically encode activities and not states and achievements. Unlike the complex predicates expressing knowledge discussed above, these two complex predicates consistently occur with an

ergative/nominative Experiencer, supporting the hypothesis according that these two verbs display an agentive reading that is not present in the other verbs referring to knowing.

364. आज मैं तुझसे कह देती हुँ कि अगर इस तरह की बात फिर हुई और मुझे पता लगा, तो तीनों में से एक भी जीते न रहेंगे।

āj *maiṁ* *tujh=se* *kah* *de-t-ī* *hūṁ* *ki*
today 1SG.NOM 2SG.OBL=INS say give-IPRF-F.SG be.1SG.PRS that
agar *is* *tarah=kī* *bāt* *phir* *huī* *aur*
if this.OBL way=GEN fact(F.SG.NOM) again be.PRF-F.SG and
mujhe *patā* *lag-ā,*
1SG.DAT information(M.SG.NOM) attach-PRF.M.SG
to tīnoṁ=meṁ=se ek bhī jīte na raheṁge.
Then none of the three will survive.
"Today I tell you that if something like this happens again and I find out, none of the three will survive."

365. देवीदीन ने तत्परता से कहा, "तुमने जिस दिन मुझसे कहा था, उसी दिन से मैं इन बातों का पता लगा रहा हुँ।

devīdīn=ne tatpartā=se kahā,
Devidin said quickly
"tum=ne *jis* *din* *mujh=se* *kah-ā* *th-ā,*
2PL=ERG REL.OBL day 1SG.OBL=INS say-PRF.M.SG be.PST-M.SG
us=ī *din=se* *maiṁ* *in* *bāt-oṁ=kā*
that.OBL=EMPH day=INS 1.SG.NOM this.PL.OBL fact(F)-PL.OBL=GEN
patā *lagā* *rah-ā* *hūṁ."*
information(M.SG.NOM) attach-CAUS PRGR-M.SG be.1.SG.PRS
"Devidin said quickly, 'I have been investigating these things since the day you told me'".

Table 52: Light verbs and the aspectual distribution occurring with *patā.*

Light verb	CVB	Fut	Imp	Iprf	Prf	Tot
calnā	0	12	0	20	47	**79**
honā	0	1	0	82	0	**83**
lagnā	0	6	0	6	5	**17**
milnā	0	0	0	2	2	**4**

karnā	1	3	1	2	0	**7**
lagānā	0	1	2	7	2	**12**
pānā	0	0	0	0	1	**1**
Total	**1**	**23**	**3**	**117**	**57**	**204**

8.6. The acquisition of knowledge: Learn/Understand

Acquisition of knowledge implies a change of a state in the Experiencer's mind that results in the acquisition of some piece of information. The acquisition of knowledge always implies a mental change of state, hence it is either construed as an achievement (i.e. as the sudden grasp of a concept or of a state of affairs) or as an accomplishment (i.e. as the understanding of a concept or a situation by examining it). Since verbs expressing this cognition always imply a telic reading, they frequently occur in the perfective aspect, but they also allow the imperfective. I have already discussed the relation between the aspectual characterization of a verb and its interpretation in the previous section (8.5.1), where I addressed verbs of knowing, such as *mālūm honā,* which construe knowledge as a state in the imperfective aspect and as an achievement in the perfective. In this section, I will focus on the verb *samajhnā,* which is frequently associated with the meaning "understand, perceive", and on two constructions used to encode understanding which employ the etymologically related noun *samajh* "understanding" and the verb *ānā* "come". Some dictionaries of Hindi (McGregor 1994) report "understand" as the primary meaning of *samajhnā*. However, the aspectual distribution and the argument structure preferences of this verb seem to suggest that it is primarily used to express states rather than achievements. The etymology of the verb can be traced back to the Sanskrit *sam-ā-jñā-*, meaning "know or understand thoroughly, become acquainted with, learn, perceive, acknowledge" (Monier-Williams 2008), composed by the upasarga *sam*, which generally expresses the completeness of the action expressed by the verb, and the root *jñā-* "know" or "understand" (see also Butt and Deo 2013), which as mentioned in the previous section is also etymologically related to the verbs *jānnā*, *jān paṛnā* and *jñāt honā*.

The verb *samajhnā* is used with different meanings in Hindi: it can be used to express cognitive achievements such as "understand", "perceive" or "grasp", but also states or atelic situations such as "think something about something", "deem" or "consider". These different but related meanings of the verb are mirrored by differences in its argument structure. Indeed, this verb shows well how argument structure constructions may contribute various specific semantics to the construal of the situation and how different argument structures may add different semantic components to the meaning of the verb. That different meanings correspond to different constructions is something that has already been observed in the previous literature (Goddard 2007). Hare, McRae and Elman (2003,

2004), for example, showed that there is a probabilistic relationship between a verb's sense and its "subcategorization preferences" (i.e. the argument structure constructions they appear in) in corpora. They analyzed how the semantics of a verb drives its syntactic distribution and showed that argument structure constructions are conditioned by the meaning of the verb and that of the construction. For example, they analyzed the English verb *find* and they found out that it tends to occur in the transitive frame with a direct object when it means "locate", while it occurs with a complement clause when it means "realize". In a similar way, the different meanings expressed by *samajhnā* correlate with different argument structure constructions.

Table 53 shows the constructions occurring with the verb *samajhnā*, together with their aspectual distribution. When used in a transitive construction with an NP Stimulus, *samajhnā* usually means "understand" as in 366, while when it occurs with a predicative second argument construction it means "think sth about sth, consider sth as, deem" as in 367 and 368. Lastly, when occurring with finite complement clauses, it may mean both "understand, comprehend" and "consider, think" as in 369, in which the verb seems to express a thought or an opinion that is derived from the understanding of an external situation.

366. मैंने तुम्हारा मतलब नहीं समझा।

maiṁ=ne	*tumhāre*	*matlab*	*nahīṁ*	*samjh-ā*
1SG=ERG	2PL.GEN	meaning(M.PL.NOM)	not	understand-PRF.M.SG

"I didn't understand what you mean."

367. आप इसे उचित समझते हैं?

āp	*is-e*	*ucit*	*samajh-t-e*	*haiṁ?*
2SG.HON.NOM	this-ACC	appropriate	understand-IPRF-M.PL	be.PRS.3PL

"Do you think this is appropriate?"

368. उन लोगों के सामने जो उन्हें इतना गलत समझ रहे थे सच बोलना कितना निराशाजनक होता।

un	*logoṁ=ke sāmne*	*jo*	*unheṁ*	*itnā*	*galat*
that.PL.OBL	people-PL.OBL=in_front_of	REFL	3PL.DAT	so_much	wrong
samajh	*rah-e*	*th-e*	*sac*	*bolnā*	*kitnā*
understand	PRGR-M.PL	be.PST-M.PL	truth	tell-INF	how_much
nirāśājanak	*ho-t-ā.*				
disappointing	be-IPRF-M.SG				

"How disappointing it would have been to tell the truth to people who were misunderstanding them so much."

369. मैं समझता था कि तुम लखनऊ में ही हो।

maiṁ	*samajh-t-ā*	*th-ā*	*ki*	*tum*
1SG.NOM	understand-IPRF-M.SG	be.PST-M.SG	that	2SG.NOM
lakhnau=meṁ	*hī*	*ho*		
Lucknow=in	EMPH	be.PRS.2SG		

"I thought that you were still in Lucknow."

Table 53: Aspectual characterization *samajhnā.*

Construction	CP	IPRF	INCH	PRF	Tot
Finite complemental clause	1	41	0	16	**56**
Passive (predicative subject)	0	8	0	0	**8**
Predicative second argument	7	71	3	6	**90**
Transitive	0	39	1	7	**47**
Total	**8**	**159**	**4**	**29**	**201**

Interestingly, sometimes *samajhnā* occurs with a nominative Experiencer in the corpus even when the verb stands in the perfective aspect, as in 370. This happens rarely, only 12 times over 229 perfective occurrences that I manually scrutinized. According to some grammars (Milanetti 2008: 125), the lack of the ergative marking typically occurs when the verb appears without a direct object, and this would imply that the verb occurs in an intransitive construction. However this does not seem to be the case, as I found instances in the corpus in which the verb occurs with a direct object but with a nominative Experiencer (as in 371 and 372) and cases in which the verb occurs without a direct object but with an ergative Experiencer (as in 373).

370. वह समझी , शायद आज वह कंगन के रूपए दे देंगी।

vah	*samjh-ī,*	*śāyad*	*āj*	*vah*	*kaṁgan=ke*
3SG.NOM	understand-PRF.F	maybe	today	3SG.NOM	bracelet(M.SG.OBL)=GEN
rūpae	*de*	*d-eṁ-g-ī.*			
rupees(M.PL.NOM)	give	give-3SG-FUT-F			

"She thought, maybe today she will give the money for the bracelet."

371. वृद्ध पुरुष ने उसकी ओर बिना ताके ही उत्तर दिया- "मैं तुम्हारी बात नहीं समझा"।

vr̥ddh puruṣ ne uskī or binā tāke hī uttar diyā-

The old man replied without even looking at him:

“maiṁ	*tumhār-ī*	*bāt*	*nahīṁ*	*samjh-ā”*
1SG.NOM	2PL.GEN-F	words(F.PL.NOM)	not	understand-PRF.M

“The old man replied without even looking at him –“I did not understand what you said”.

372. तू इतनी मोटी-सी बात भी नहीं समझी !

tū	*itn-ī*	*moṭ-ī=s-ī*	*bāt*	*bhī*	*nahīṁ*
2SG.NOM	as_much-F	stupid-F=like-F	words(F.PL.NOM)	even	not

samjh-ī!

understand-PRF.F

“You didn’t even understand such a simple thing!”

373. मिर्जा जी ने हंस कर कहा - बस-बस, तूने खूब समझा।

mirjā ji ne haṁskar kahā -	*bas-bas,*	*tū=ne*	*khūb*	*samjh-ā.*
Mirza ji laughed and said	stop-stop	2SG=ERG	well	understand-PRF.M

“Mirza ji laughed and said – That’s it, you have understood well.”

Interestingly, in all cases where the verb appears with a nominative Experiencer despite being in the perfective aspect, *samajhnā* conveys the meaning of “understanding”. As I will argue below, a similar behavior also characterizes the verb *bhūlnā* “forget” and seems to indicate that these experiential types are typically construed as lacking an agentive Experiencer.

Besides the verb *samajhnā*, two constructions are used in Hindi to express the understanding of external elements or states of affairs. Both constructions employ the noun *samajh* “understanding, mental grasp, intelligence” (McGregor 1994) and the light verb *ānā* “come” which contributes the semantics of change-of-state that is typical of this experiential type. These two constructions construe the event differently and accordingly encode the participants in different ways. The first construction, which is also the most frequent one, construes the understanding as an abstract mental part of the Experiencer. The Experiencer is in the genitive, and it is conceptualized as the possessor of the understanding (*samajh*). The noun *samajh* is followed by the locative postposition *meṁ* “in” thus suggesting that the understanding is conceptualized as the location toward which the Stimulus is directed. The Stimulus may be encoded either by a nominative NP (as in 374) or by a complement clause (as in 375).

374. इस गीत का एक शब्द भी मेरी समझ में न आया।

is	*gītā=kā*	*ek*	*śabd*	*bhī*	*merī*	*samajh=meṁ*
this.OBL	song=GEN	one	word.M.SG.NOM	even	my	understanding=in

na *āy-ā*
not come-IPRF.M.SG
"I didn't understand a word of this song."

375. मेरी समझ में ही नहीं आता, वह किस बात से खुश और किस बात से नाराज होते हैं।

merī	*samajh=meṁ*	*hī*	*nahīṁ*	*ā-t-ā,*
1SG.GEN	understanding=in	EMPH	not	come-IPRF-M.SG

vah kis bāt se khuś aur kis bāt se nārāj hote haiṁ.
What makes him happy and what makes him angry.
"I don't understand what makes him happy and what makes him angry."

In the second construction the Experiencer is marked with the dative and the noun *samajh* is in the nominative and agrees with the verb *ānā* "come". Once again, the Stimulus may be expressed either by a genitive NP attached to the noun *samajh* (376) or by a complement clause (377). While in the locative construction the understanding is construed as the part of the Experiencer toward which the Stimulus is directed, hence it is the Target of the event, in the dative construction it is construed as the Source that moves toward the Experiencer.

376. मुझे इस लडक़ी के दिमाग़ की कुछ समझ नहीं आती।

mujhe	*is*	*laṛkī=ke*	*dimāg=kī*	*kuch*
1SG.DAT	this.OBL	girl-F.SG.OBL=GEN	mind(M.SG.OBL)=GEN	INDF.PRN

samajh *nahīṁ* *atī.*
understanding.F.SG.NOM not come-IPRF-F.SG
"I don't understand the mind of this girl."

377. उन्हें समझ नहीं आता कि उन्होंने अपनी झुकी हुई आंखों को उठाकर अच्छा किया है या बुरा किया है।

unheṁ	*samajh*	*nahīṁ*	*ā-t-ā*	*ki*
3PL.DAT	understanding(F.SG.NOM)	not	come-IPRF-M.SG	that

unhoṁ=ne apnī jhukī huī āṁkhoṁ=ko uṭhākar acchā kiyā hai yā burā kiyā hai.
they have done good or bad by raising their downcast eyes.

"They do not know whether they have done good or bad by raising their downcast eyes."

Both constructions, however, construe an event in which the Experiencer is not agentive. These patterns resemble two constructions that are used to express bodily sensations in Hindi. As I discussed in section 6.1, feelings localized on a specific area of the body, such as pain and itch, are typically expressed by two constructions. The first is the dative construction. The second is a locative construction in which the Expertum stands in the nominative and the part of the body where the sensation is localized stands in the locative case (followed by the postposition *meṁ* "in"). The Experiencer is marked with the genitive and is construed as an attributive possessor of the body part. The fact that the expression of understanding in Hindi patterns with the expression of bodily sensations is indicative of how understanding is construed in the language and seems to suggest that the Experiencer of this cognition is typically conceived as devoid of volitionality and control. This aligns with the fact that the verb *samajhnā* may occur without the ergative marking when it expresses understanding, as the ergative is generally associated to high agentivity.

8.7. Remember and forget

In this section, I will discuss the verbs used in Hindi to express memory. There are several studies on how memory is conceptualized cross-linguistically (see, for example, Chafe 1973, Van Valin and Wilkins 1993, Schwartz 1999, Tao 2001, 2003). However, as Amberber (2007) points out, only a few investigate the linguistic encoding of this experiential subtype. In previous literature, remember and forget are not always grouped together. For example, D'Andrade (1995: 158-160) treats remembering as a perception and groups verbs expressing this experiential type with verbs expressing sight, hearing, etc. At the same time, D'Andrade classifies forgetting as a cognitive event, thus associating remembering and forgetting to two different experiential subdomains. D'Andrade's observations make sense, as remembering is generally conceptualized as the retrieval of a piece of information in the experiencer's mind, possibly triggered by the perception of a stimulus. In contrast, forgetting is not linked to external stimuli and typically encodes either the absence of information or the experiencer's inability to access it. However, as Amberber (2007) points out remembering is an internal state, or change of state, that refers to some mental content just as forgetting is. Thus, following Amberber (2007, and Baddeley 1999), I will consider memory as both the capacity to encode, store and retrieve information and the inability to do s. Hence, in this section, I will investigate how Hindi encodes both concepts related to remembering and concepts related to forgetting. However, I want to emphasize that remembering and forgetting also show important differences. In particular, they differ significantly in their semantic properties, especially with respect

to the degree of agentivity and volitionality that the Experiencer may exhibit. As I will discuss in more detail, the Experiencer of remembering may show some degree of volitionality, while the Experiencer of forgetting is typically construed as inadvertent and unintentional.

8.7.1. Remember: complex predicates with *yād*

Wierzbicka (2007) argues that memory is a culture-specific concept and not a universal one. In particular, she notes that many languages do not express remembering with specific lexical items. This is the case of some Australian languages that do not exhibit a specific verb to encode remembering. For instance, Pitjantjatjara, an aboriginal language, uses the expression *pinangku kulini*, which literally means "think with (one's) ears", to encode the concept that in English is expressed by the verb *remember* (see also Goddard 1996, 2003). Speakers of the Australian language Dalabon, instead, use expressions such as "cause to have in mind" in order to encode experiential situations referring to memory (Evans 2007). Wierzbicka's assumption on the non-universality of memory is very controversial and many scholars have rejected it. For instance, relying on philosophical and psychological observations, Sutton (2007) argues in favor of the universality of the concept of memory. What is certain is that languages around the world exhibit great variation in how they lexicalize memory: some do not use specific lexical items, while others display a detailed pattern of lexicalization. Furthermore, languages may also employ constructional methods to distinguish between different construals of memory.

While analyzing the expression of remembering in English, Van Valin and Wilkins (1993: 511) argue that memory can be differently construed and that these different construals correspond to three actionality classes singled out by Vendler (1976). These different conceptualizations are exemplified in 378, 379 and 380. In sentence 378, *remember* indicates an active recalling of some piece of information previously stored in the mind of the Experiencer. In sentence 379, it expresses an achievement in which the content of the memory is activated in the Experiencer's mind without them controlling the event. While in sentence 380 it indicates a mental state in which a cognitive content is stored in the mind of the Experiencer.

378. *Sarah was trying to remember the address of her friend's house.*
379. *While speaking with Sarah, John remembered his talk with Mark.*
380. *I still remember the first time we met.*

As sentences 378, 379 and 380 show, English does not use lexical means to distinguish among different construals and employs the verb *remember* for all of them. As I will illustrate, Hindi exhibits a high degree of specificity with respect to the expression of this experiential type. This highly

specific lexicalization derives from the fact that in Hindi memory is expressed by a complex predicate and, as I discussed in section 3.2.5.2.1, complex predication allows alternations on the light verb that generally result in changes in the interpretation of the event. These different lexical items also pattern with different constructions. As I will show below in this section, two main constructions are used in Hindi to express remembering: the transitive construction and the dative construction.

The complex predicates expressing remembering in Hindi are all formed by the feminine noun *yād* "memory" and differ according to the verb used as light verb. Table 54 shows the light verbs that occur with *yād* in the corpus and their relative frequencies over the occurrences that I manually scrutinized.

Table 54: Light verbs occurring with the nominal host *yād* in the corpus and their frequency.

Light verb		**Tot**
Inransitive	*ānā* "come"	**101**
	rahnā "stay"	**13**
	honā "be"	**26**
	ho jānā/ānā "become"	**10**
		150
Transitive	*karnā* "do"	**44**
	rakhnā "keep"	**11**
		55
Tot		**205**

As I have already discussed for complex predicates referring to thought (8.4.2), the light verb conveys information on how the event is conceptualized, and the alternation of light verbs encodes differences in the lexical aspect. Additionally, the light verb also conveys information on the semantic properties of the Experiencer. In sentence 381, the complex predicate *yād karnā* (*karnā* "do") implies that the situation is construed as an ongoing controlled activity and the Experiencer is construed as agentive. This is supported also by the fact that the verb occurs in the imperative, thus implying that the Experiencer is conceived as endowed with a certain degree of control on the mental activity. In 382, the verb *yād honā* (*honā* "be") construes the experience as a cognitive state in which some piece of information is stored in the mind of the Experiencer. In this case, the Experiencer is encoded with the dative and they are not conceived as an agentive participant volitionally recalling something in their own mind (see also Butt 2006b, Butt and Ahmed 2010).

381. मेरी मुहब्बत याद करो।

merī	*muhabbat=kī*	*yād*	*kar-o*
my	love(F.SG.OBL)= GEN	memory(F.SG.NOM)	do-IMP

"Remember my love."

382. मरी माता को तुम्हारी सुरत याद है।

merī	*mātā= ko*	*tumhār-ī*	*sūrat*	*yād*	*hai*
1SG.GEN-F	mother=DAT	2PL.GEN-F	face	memory(F.SG.NOM)	be.3SG.PRS

"My mother remembers your face."

Hindi thus lexically distinguishes two different ways of conceptualizing memory, either as an activity or as a state. This distinction is present in many languages, for example in Polish (Wierzbicka 2007), but is also absent in many others, for example in English and Italian. What is remarkable in Hindi, however, is that this language exhibits an extremely specialized lexicalization structure (Table 54). In sentences 383, 384 and 385, memory is expressed by three other complex predicates, all formed by the nominal host *yād*. Each light verb adds a specific semantic component and results in a different construal of the event. In particular, the verb *rakhnā* (383) is similar to the light verb *karnā* as it occurs with agentive Experiencers, while the light verbs *ānā* (384) and *rahnā* (385) resemble *honā* as they imply a non-agentive reading. However, the verb *yād rakhnā* differs from *karnā* as it conveys a durative reading that is implied in the semantics of the light verb, which means "put" but also "keep". While the verb *yād ānā* contrasts with *yād honā* as *ānā* "come" contributes an inchoative semantics that construes the event as a change of state in which some memory is suddenly activated in the Experiencer's mind. *Yād rahnā*, instead, emphasizes a durative interpretation of the event that is supplied by the semantics of *rahnā*, which means "stay".

383. हमें हर क्षण इसे याद करना चाहिए।

hameṁ	*har*	*kṣaṇ*	*ise*	*yād*	*rakh-nā*
1PL.DAT	every	moment	this.ACC	memory(F.SG.NOM)	keep-INF
cāhiye					
should					

"We should remember this every moment."

384. सहसा उसे गोबर की याद आयी।

sahsā	*use*	*gobar=kī*	*yād*	*ā-ī*
Suddenly	3SG.DAT	gobar=GEN	memory(F.SG.NOM)	come-PRF.F.SG

"Suddenly he remembered Gobar."

385. यह तमाचा बच्चे को जन्म-भर याद रहेगा।

yah	*tamācā*	*bacce=ko*	*janm-bhar*
this.NOM	slap(M)-SG.NOM	child(M)-SG.OBL=DAT	life-full
yād	*rah-e-g-ā.*		
memory(F.SG.NOM)	stay-3SG-FUT-M.SG		

"The child will remember this slap for the rest of his life."

Table 55 shows the distribution of aspectual markers across the light verbs occurring with *yād*. The aspectual distribution supports the hypothesis that different light verbs correspond to different ways of construing the event. Indeed, the data shown in the table clearly point toward some correlations. First, the light verb *honā* "be" only occurs with the imperfective aspect. This is not surprising considering that *honā* points toward a stative reading. Second, the perfective aspect correlates with the verb *ānā* "come", as more than 70% of the occurrences of the perfective forms are with the complex predicate *yād ānā*. Notably, even though the perfective aspect privileges the verb *ānā*, this preference is not bidirectional: e.g. the perfective correlates with *ānā*, but *ānā* does not correlate with the perfective aspect. Indeed, the data indicates that this light verb actually shows a much more complex semantic characterization. In particular, it occurs frequently also in the imperfective, with almost a 50/50 distribution between the perfective and the imperfective aspect. Interestingly, despite the achievement reading implied by *ānā* "come", this light verb may also encode durative events and construe memory as a state, when occurring with specific aspectual forms. Compare sentence 386 with sentence 384 above. In 384, the complex predicate *yād ānā* construes remembering as an achievement: this reading is suggested the perfective aspect of the verb. In 386, the same verb stands in the imperfective and seems to construe memory as an atelic event. This reading is further suggested by the adverb *ab=hī=tak*, lit. "till now, still", which adds a durative characterization to the event.

386. पिताजी के कन्धों पर चढ़कर उचकने की याद मुझे अभी तक आती है।

pitā-jī=ke	*kandhoṁ=par*	*caṛh=kar*	*ucak-n-e=kī*	
father-HON=GEN	shoulder(M)-PL.OBL=on	climb=CP	be_lifted-INF.OBL=GEN	
yād	*mujhe*	*ab=hī=tak*	*ā-t-ī*	*hai.*
memory(F.SG.NOM)	1SG.DAT	now=EMPH=till	come-IPRF-F.SG	be.PRS.3SG

"I still remember climbing up on my father's shoulders."

Table 55: Distribution of aspectual markers across light verbs occurring with *yād.*

Light verb		**CP**	**IMPF**	**PRF**	**Tot**
Intransitive	*ānā*	1	53	47	**101**
	rahnā	0	8	5	**13**
	honā	0	26	0	**26**
	ho jānā/ānā	0	1	7	**8**
					148
Transitive	*karnā*	12	28	4	**44**
	rakhnā	0	10	1	**11**
					55
Tot		**13**	**126**	**64**	**203**

The V-V complex predicates *ho jānā* and *ho ānā* mentioned in Table 55 are formed by the verb *honā* "be" followed by either the vector verb *jānā* "go" or *ānā* "come". These complex predicates occur only in the perfective aspect in the corpus (387) except for one occurrence (given in 388). As I have already mentioned in section 3.2.5.1, it is generally acknowledged in the previous literature that one of the main functions of vector verbs in V-V constructions is to convey completion of the event expressed by the main verb (Hook 1974, Chakraborty 1992, Butt 1995) and for this reason some scholars have identified it as a marker of perfectivity (Hook 1991). The verbs *yād ho jānā* and *yād ho ānā* thus imply an achievement reading that is missing when the verb *honā* "be" occurs alone. As a consequence, the V-V sequence shows a different aspectual characterization in comparison to *yād honā* and mostly occurs in the perfective aspect (Table 55).

387. यह कहते हुए मुझे सहसा नीलिमा के नाम लिखे उसके पत्रों की याद हो आयी, और मेरा मन एक विचित्र उत्सुकता से भर गया।

yah	*kah-t-e*	*hu-e*	*mujhe*	*sahsā*
this.NOM	say-IPRF-M.OBL	be.PRF-M.OBL	1SG.DAT	suddently
nīlimā=ke	*nām*	*likh-e*	*us=ke*	
nilima=GEN	name(M.SG.OBL)	write-PRF.M.OBL	that.OBL=GEN	
patr-oṁ=kī	*yād*	*ho*	*ā-yī,*	
letter(M)-SG.PL=GEN	memory(F.SG.NOM)	be	come-PRF.F	

aur merā man ek vicitra utsuktā=se bhar gayā.

"Saying this, I suddenly remembered the letters he had written to Neelima, and my mind was filled with a strange curiosity."

388. उसके सिर के बाल काफ़ी उड़ गये थे, जिससे उसे देखते ही सिलविक्रिन के विज्ञापन की याद हो आती थी।

us=ke sir=ke bāl kāphī uṛ gaye the, jis=se use dekhte hī

The hair on his head had fallen out a lot, which

silavikrin=ke	*vijñāpan=kī*	*yād*
silvikrin=GEN	advertisement(M.SG.OBL)=GEN	memory(F.SG.NOM)

ho	*ā-t-ī*	*th-ī.*
be	come-IPRF-F	be.PST-F

"The hair on his head had fallen out a lot, which reminded him of a Silvikrin advertisement."

Interestingly, the vector verb *ānā* "come" appears more frequently than the vector verb *jānā* "go" with the complex predicate *yād honā* (notwithstanding their general low frequencies, which make any generalization problematic). In particular, *yād ho jānā* occurs only once in the corpus, while *yād ho ānā* occurs 15 times. This is interesting because the V-V *ho jānā* is far more productive than the V-V *ho ānā* in the grammar. In the Literary corpus, for examplee, *ho jānā* occurs 4477 times while *ho ānā* only 51 times. In the language, the frequency of *ho jānā* is so high that this V-V sequence today has been lexicalized with the meaning "become" (as in 389).

389. वहाँ दो बार आई हुँ, और जब-जब गई हुँ, बीमार हो गई हुँ।

vahāṁ	*do*	*bār*	*ā-ī*	*hūṁ,*	*aur*	*jab-jab*
there	two	time	come-PRF.F.SG	be.1SG.PRS	and	when-when

ga-ī	*hūṁ,*	*bīmār*	*ho*	*ga-ī*	*hūṁ.*
go.PRF-F.SG	be.1SG.PRS	sick	be	go.PRF-F.SG	be.1SG.PRS

"I've been there twice, and every time I got sick (lit. I have become sick)."

The difference in the absolute frequencies of these V-V complex predicates suggest that the higher frequency of *yād ho ānā* in comparison with *yād ho jānā* is significative when expressing remembering, despite the general low frequency of these complex predicates in the corpus. As I mentioned before, both *jānā* and *ānā* imply dynamicity and add an achievement reading to the situation that results is a correlation with the perfective aspect. Butt and Geuder (2001) investigate how the vector verb contributes to the semantics of the whole V-V complex predication in Hindi. They argue that it is limitative to consider vector verbs as mere perfectivity markers as suggested in

previous literature. In particular, they point out that the connection between perfectivity and vector verbs is not as strong as it is generally believed, and they argue that the vector verb also highlights different semantic components that are generally connected to the meaning that the verb has when it is used as a main verb. Under this light, the fact that *ānā* is more frequent that *jānā* is significative. These two verbs contribute a different meaning. They both supply a semantic characterization that implies directionality, but *jānā* "go" implies a movement directed outward, while *ānā* implies that the movement is inward. Events of remembering imply that some piece of information is stored or is suddenly activated inside the mind of the Experience and this probably favors the choice of the V-V sequence *ho ānā*.

Table 56 shows the constructions used to express memory in the corpus. When the noun *yād* occurs in complex predicates with intransitive light verbs (such as *ānā* "come", *honā* "be", *ho ānā/ jānā* "be come / be go" and *rahnā* "stay"), the Experiencer is consistently marked with the dative, while when it appears with transitive light verbs (*karnā* "do" and *rakhnā* "keep") the Experiencer is encoded with the nominative/ergative. Interestingly, light verbs with dative Experiencers are far more frequent, thus suggesting that Hindi prefers to construe remembering as a state or as an involuntarily change of state rather than as a controlled activity. In both transitive and intransitive complex predicates, the Stimulus may be expressed by an NP or by a complement clause.

Table 56: Constructions occurring with complex predicates with *yād* (1= Experiencer; 2=Stimulus).

Experiencer	**Stimulus**	**Form**	**Tot**
Dative subject	Complemental clause	1-dat *yād* LV.subj[*yād*] + compl. clause	31
	Nominative NP	1-dat 2-nom *yād* LV.subj[2]	65
	Genitive NP	1-dat 2-gen *yād* LV.subj[*yād*]	56
Canonical subject	Complemental clause	1-nom *yād* LV.subj[1] + compl. clause	5
	Genitive NP	1-nom 2-gen *yād* LV.subj[*yād*]	13
	Direct object	1-nom 2-nom *yād* LV.subj[1]	34
	Passive	(1-instr) 2-nom *yād* LV.pass.subj[2]	1
Tot			**205**

When it is expressed by an NP, the complex predicate may show either internal or external agreement. In the case of transitive complex predicates, the Stimulus is either encoded as a direct object (external agreement, as in 383), or it is in the genitive (internal agreement, as in 381). Similarly, when the

complex predicate is intransitive, the Stimulus NP may be either in the nominative (external agreement, as in 385), or in the genitive (internal agreement, as in 382). I have already discussed this ambiguous treatment of the nominal host in complex predication in section 4.2.3.1, in which I specifically addressed the case of complex predicates expressing memory. In that section, I concluded that the genitive marking on the Stimulus should not be interpretated as bearing any semantic difference. Consider , for instance, sentence 390 and 391. These two sentences refer to very similar situations: in both cases the event is construed as a controlled activity, as it is suggested by the transitive light verb occurring in the imperative form. Moreover, in both cases the Stimulus refers to a human participant, which is the speaker, expressed by the first-person singular pronoun *maiṁ*; while the Experiencer is the interlocutor. As it is clear, these two examples are identical both with respect to the semantic properties of the event and with respect to the properties of the participants. Moreover, both experiential constructions are used in very similar contexts, where the speakers invite the interlocutor to remember that they will be available to help whenever needed. Despite all these similarities, in the first example (390) the Stimulus is encoded as a direct object, marked because of differential object marking, while in the second example (391) the Stimulus is in the genitive. These two sentences show that the genitive marking on the Stimulus is not determined by semantic reasons. Instead, the genitive case derives from the fact that the noun-verb sequence in 391 does not fully functions as a complex predicate, and the nominal host is also treated as an argument. Consequently, the genitive case is selected because the Stimulus is an argument of the noun *yād*, rather than an argument of the whole complex predicate. And for this reason, it is marked with the case typically used for marking arguments of nouns in Hindi.

390. कोई जरूरत हो, तो मुझे याद कीजिए।

koī	*zarūrat*	*ho,*	*to*	*mujhe*	*yād*
INDF.ADJ	need(F.SG.NOM)	be.SBJV.3SG	then	1SG.ACC	memory(F.SG.NOM)
kījie					
do.IMPER					

"If you need anything, remember me."

391. जिस वक्त तुम्हें कभी जरुरत हो, मेरी याद करना।

jis	*vakt*	*tumheṁ*	*kabhī*	*jarurat*	*ho,*
REL.ADJ	time	2PL.DAT	sometimes	need(F.SG.NOM)	be.SBJV.3SG
mer-ī	*yād*		*kar-nā.*		
1SG.GEN-F	memory(F.SG.NOM)		do-INF		

"Remember me whenever you need."

8.7.2. Forget

In Hindi forgetting is expressed by the verb *bhūlnā*. In sentence 392, the verb occurs in the present and both the Experiencer and the Stimulus are in the nominative.

392. महाराज, आपने मेरे ऊपर जो एहसान किए, उनका मैं हमेशा कृतज्ञ रहुँगा। क्षत्रिय कभी एहसान नहीं भूलता।

mahārāj, āpne mere ūpar jo ehsān kie, unkā maiṁ hameśā kṛtajña rahūṁgā.
Your Majesty, I will always be grateful for the favor you have done to me.

kṣatriya	*kabhī*	*ehsān*	*nahīṁ*	*bhūl-t-ā.*
kshatrya(M.SG.NOM)	sometimes	favor(M.SG.NOM)	not	forget-IPRF-M.SG

"A kshatrya never forgets a favor."

This verb exhibits an interesting behavior that makes its syntactic status quite unclear. This ambiguity was already known by previous scholars of Hindi. For example, the Oxford Dictionary by McGregor (1994) classifies this verb as both transitive ("forget") and intransitive ("be forgotten") and reports the following meanings: "1. to be forgotten; 2. to go astray, to err; 3. to be deceived; 4. to forget; 5. to omit, to overlook". Notably, as I will discuss, the analysis of the occurrences that I sourced from the corpus suggests that the basic meaning of this verb is "forget" and not "be forgotten".

One of the most peculiar properties of *bhūlnā* is that this is a two-arguments verb that shows some syntactic properties of transitivity, for example it may be passivized (as in 393) and it allows differential marking on the second argument (as in 394), yet it rejects an ergative marking on the subject when it occurs in the perfective aspect (as in 395 and 396). Example 395 is particularly indicative of the ambiguous syntactic characterization of this verb, as *bhulnā* occurs in the perfective aspect and the Experiencer is in the nominative, not the ergative, and at the same time the Stimulus *bāt* "fact" is marked with the accusative *ko*, typically reserved to direct objects in transitive constructions. This is interesting because in the grammar there are no transitive verbs that do not alternate with an ergative marking on the first participant.

393. हां, एक बात तो भूली ही जा रही थी।

hāṁ,	*ek*	*bāt*	*to*	*bhūl-ī*	*hī*	*jā*
1SG.NOM	one	matter(F.SG.NOM)	then	forget-PRF.F.SG	EMPH	go
rah-ī	*th-ī.*					
PRG-F.SG	be.PST-F					

"Yes, one thing was being forgotten."

394. तुम तो वहाँ जाकर मुझे भूल गईं नैना, एक पत्र भी न लिखा।

tum *to* *vahāṁ* *jā=kar* *mujhe* *bhūl* *gaīṁ*
2SG.NOM then there go=CP 1SG.DAT forget go.PRF-F.PL
nainā, ek patr bhī na likhā.
Naina, you didn't even write a letter.
"You have forgotten me after going there Naina, you didn't even write a letter."

395. वह अब तक भी नौ साल पहले की उस बात को नहीं भूली कि मैं उससे कहकर भी उस शाम उसके यहां नहीं गया था।

vah *ab=tak* *bhī* *nau* *sāl* *pahle=kī* *us*
3SG.NOM now=till EMPH nine year before=GEN that.OBL
bāt=ko *nahīṁ* *bhūl-ī* *ki*
matter(F.SG.OBL)=ACC not forget-PRF.F that
maiṁ usse kahkar bhī us śām uske yahāṁ nahīṁ gayā thā.
I didn't go to her place that evening, despite telling her I would.
"She still hasn't forgotten that incident from nine years ago, when I didn't go to her place that evening, despite telling her I would."

396. मगर धनिया अभी तक गोबर के कठोर शब्द न भूली थी।

magar *dhaniyā* *ab=hī=tak* *gobar=ke* *kaṭhor* *śabd* *na*
but dhaniya.NOM now=EMPH=till gobar=GEN harsh word(M.PL.NOM) not
bhūl-ī *th-ī*
forget-PRF.F be.PST-F.SG
"But Dhaniya had not forgotten Gobar's harsh words yet."

That the transitive status of this verb is unstable is also shown by the fact that it mainly occurs in a V-V complex predicate in which the bare root *bhūl* is followed by the vector verb *jānā* "go". This tendency is evident from the corpus: as the data in Table 57 show, *bhūlnā* occurs more than 70% of the times with *jānā*.

Table 57: Vector verbs occurring in V-V sequences with *bhūlnā* and their aspectual distribution.

Vector	Prf	Imperfective			Imp	Fut	CP	Subj	Tot
		Cont	Hab	Prg					

jānā	78	0	25	1	13	13	0	14	144
no vect.	7	1	18	2	3	11	5	4	52
paṛnā	3	0	0	0	0	0	0	0	3
baiṭhnā	1	0	0	0	0	0	0	0	1
		1	**43**	**3**					
Tot	**89**	**47**			**16**	**24**	**5**	**18**	**200**

An example of this V-V complex predicate is given in 397, in which the verb followed by *jānā* occurs in the perfective aspect. The Experiencer does not show an ergative marking, it is in the nominative, and it agrees with the verb.

397. मैं फिर तुम्हारा नाम भूल गया।

maiṁ	*phir*	*tumhār-ā*	*nām*	*bhūl*	*ga-yā.*
1SG.NOM	again	2PL.GEN-M	name(M.SG.NOM)	forget	go.PRF-M.SG

"I forgot your name again."

In Hindi, V-V sequences in which the main verb and the light verb do not show the same characterization with respect to transitivity are quite rare. These cases have been widely discussed in the previous literature, with a particular focus on sequences in which a transitive main verb is followed by an intransitive vector verb and have been defined atypical (see Montaut 2004 and for a quantitative study on this see Drocco and Tiwary 2020a/b). The fact that the verb *bhūlnā* occurs 70% of the times with the intransitive vector verb *jānā*, thus, is indicative of its ambiguous status regarding transitivity, and suggests that even though this verb may show some properties of syntactic transitivity (DOM and passivization), it basically is intransitive.

The specific semantic properties of this experiential type might shed light on the ambiguous syntactic behavior exhibited by *bhūlnā,* as forgetting is almost always conceptualized as non-volitional and the Experiencer is almost never in control of the situation. Prototypically, forgetting implies that some piece of information that was once present in the mind of the Experiencer is no longer accessible or retrievable by them. Verbs expressing forgetting contrast with verb expressing remembering, as the latter can be both volitional and non-volitional, while the first typically refer to events that happen without the deliberate intention of the Experiencer. In sum, *bhūlnā* is typically connected to a non-agentive argument. This strong connection between the Hindi verb *bhūlnā* and lack of intentionality is proved also by the etymologically related noun *bhūl*, which means "forgetfulness, oversight, omission" but also "mistake, slip, error" in Hindi (398; McGregor 1994). Additionally, the language displays a fixed expression consisting of the conjunctive participle in *kar*

of the verb *bhūlnā* that has been lexicalized with the meaning "by mistake" (399). Sentences 398 and 399 indicate that this verb is deeply connected to inadvertency.

398. उससे कुछ आशा करना मेरी भूल थी।

us=se	*kuch*	*āsh-ā*	*kar-n-ā*	*mer-ī*
3SG.OBL=INST	some	hope(M)-SG.NOM	do-INF	1SG.GEN-F
bhūl		*th-ī*		
forgetfulness/mistake.FSG.NOM		be.PST-F		

"It was my mistake to expect anything from him."

399. अब मैं भूलकर भी गहनों का नाम न लूंगी।

ab	*maiṁ*	*bhūl=kar*	*bhī*	*gahn-oṁ=kā*	*nām*
Now	1SG.NOM	forget=CP	EMPH	jewelry-OBL.PL=GEN	name(M.SG.NOM)
na	*l-ūṁ-g-ī.*				
not	take-1SG-FUT-F				

"Now I will not mention jewelry even by mistake."

It is reasonable to hypothesize that the lack of an ergative marking is explained by the semantic properties of the event. The Experiencer is completely devoid of control and intentionality and thus resists an ergative marking which is typically connected to agentivity in Hindi (see section 3.2.2.1).

Another peculiarity characterizing this verb is that it also appears with a dative construction in the corpus. In sentence 400 and 401, the Experiencer stands in the dative while the Stimulus is in the nominative. One might object that in these sentences we are dealing with mere direct objects that refer to human participants and consequently are marked with the accusative because of differential object marking (as in sentence 394 above). However, that in these sentences we are dealing with dative subjects of the type generally found with experiential verbs is suggested by the semantic properties of the other participants in the sentence. In both 400 and 401 the dative argument is human, while the nominative argument is abstract and there is no evidence to interpret it as a metaphorical non-human Experiencer.

400. पर कुछ न बोलता था, जैसे उसे वह सब कायदे-कानून भूल गए।

par kuch na bol-t-ā th-ā, jaise.
But he didn't say anything, as if

use	*vah*	*sab*	*kāyade-kānūn*	*bhūl*	*ga-e*
3SG.DAT	that.NOM	all	rules-regulation(M.PL.NOM)	forget	go.PRF-M.PL

"But he didn't say anything, as if he had forgotten all those rules and regulations."

401. काले खां की याद उसे एक क्षण के लिए भी न भूलती और किसी गुप्त शक्ति की भांति उसे शांति और बल देती थी।

kāle khāṁ=kī	*yād*	*use*	*ek*	*kṣaṇ ke lie*	*bhī*	*na*
kale khan=GEN	memory(F.SG.NOM)	3SG.DAT	one	moment=for	also	not
bhūl-t-ī	*aur kisī gupt śakti kī bhāṁti use śāṁti aur bal detī thī.*					
forget-IPRF-F.SG	and like some secret power it gave him peace and strength.					

"He did not forget the memory of Kale Khan even for a moment and like some secret power it gave him peace and strength."

Notably, that *bhūlnā* may occur also in the dative construction has already been noted in previous studies (Montaut 2016) and Drocco (personal communication) pointed out that there seems to be a rising tendency to use *bhūlnā* in a dative construction in spoken Hindi, in particular in varieties that are spoken around the Panjabi area. These cases are very rare in the corpus, they only occur 7 times over the 200 sentences that I manually scrutinized. However, it should be considered that the corpus consists of literary texts written in Standard Hindi and that standard varieties typically resist neo-substandard variants. As I will discuss in chapter 9.1.4, this verb is clear evidence of a rampant semantically-based use of case in the language and of the connection between ergativity and the semantic property of agentivity.

In sum, the ambiguous syntactic status of the verb *bhūlnā* seems to be explained by the semantic properties of the event and in particular by the fact that the main argument of this verb is conceived to be a prototypical Experiencer, hence it is maximally distant from an Agent from a semantic point of view. On this line, we can explain all the peculiar behaviors found with this verb:

a) The lack of an ergative marking when the verb appears in the perfective aspect.
b) The high frequency of *bhūlnā* with the vector verb *jānā*, which typically is found with intransitive verbs.
c) The occurrence of this verb with the dative construction.

8.8. Discussion

The constructional analysis of verbs of cognitions showed that this semantic class tends to occur with many constructions that were found to be productive with the experiential classes investigated in the previous chapters. In particular, according to the semantic properties of the event they express, verbs of cognition pattern both with verbs of bodily sensation and with verbs of perception with respect to their argument structures. Cognitions that show a low degree of agentivity by the Experiencer

generally pattern with bodily sensations, as they are typically expressed by the dative construction or the locative construction (as for verbs expressing the acquisition of knowledge or understanding). In contrast, cognitions characterized by a high degree of control and volitionality pattern with agentive (visual) perceptions as they are typically expressed by the oblique Stimulus construction, which construes an agentive Experiencer, volitionally starting the event and directing their attention toward some mental content.

Verbs of cognition also allow a construction that is not found with the other semantic classes I analyzed in the previous chapters, that is the genitive Experiencer construction. As I discussed in section 8.4.2, this construction mainly occurs with complex predicates referring to thinking. In particular they refer to experiential states, in which the activity of thinking is profiled from the endpoint and the thought is construed as a thing possessed by the Experiencer. In other words, the genitive implies that the Experiencer in the genitive case is associated to a higher degree of agentivity and the construction is used to express events that can be construed as the result of a dynamic process in which the Experiencer is agentive.

We can conclude that the constructional analysis has proved itself very helpful as it has shown that the several different constructions do express different semantics and thy accordingly correlate with specific semantic classes. The only construction that is not strongly associated to a specific semantics and thus to a specific class of verbs is the transitive construction, which as I have already discussed in the previous chapters is semantically neutral. Interestingly, the analysis of some cognitions revealed the peculiar behavior of some verbs that are quite ambiguous with respect to their transitivity status: i.e. *samajhnā* "understand" and *bhūlnā* "forget". The semantics of these two verbs gives us valuable insights on the reason why they show this behavior and shed light on the productivity of the transitive construction in the language. I will further discuss on this in the next chapter. In the following sections, I would like to further delve into two aspects of verbs of cognition and draw some conclusions in this respect. The first is the realization of complement clauses with this semantic class, the second is the productivity of light verbs alternation in this class of verbs that differentiate it from the other two experiential classes analyzed in the previous chapters.

8.8.1. The semantic contribution of complement types

Verbs of cognition differ from verbs of perception and of bodily sensation as they show a much higher frequency of Stimuli realized via complement clauses rather than NPs. The most frequent complementation type is the finite clause, which occurs with all types of verbs of cognition, and is thus the most productive both with respect to type frequency and token frequency. This is obviously related to the semantics of verbs of cognition, which refer to mental states or activities that typically

involve a propositional mental content. This construction is used to indicate that a propositional thought is activated the Experiencer's mind, and it corresponds to the *that*-construction in English (Wierzbicka 1988, Goddard 2003). As I discussed in section 7.5, verbs referring to perception may allow a finite complement type as well. When perception verbs occur with finite clauses, they tend to express the indirect perception of some external information and tend to border with verbs of cognition, in the sense that they typically function as evidentiality markers.

Notably, verbs of cognition differ from verbs of perception as they also occur with non-finite complement clauses in the corpus, while I did not find any trace of non-finite complementation with verbs of perception. In this construction the Stimulus is realized by the infinitive form of the verb. As I briefly discussed in section 8.4.1, the meaning of the non-finite construction differs from that of the finite complement clause, as it does not represent a thought as it is in the Experiencer's mind, and it rather indicates the intention of the Experiencer to do something. This non-finite construction in Hindi corresponds to the English *to*-construction (Wierzbicka 1988, Goddard 2003). The semantics of the non-finite clause thus explains why this construction does not occur with verbs of perception in Hindi, as perception verbs never refer to events in which the Experiencer shows the intention to do something, they rather refer to the acquisition of information through the perception of external elements or states of affairs. In other words, the semantics of the non-finite construction does not align with the semantics of perception verbs.

That the choice of complement clauses may be semantically motivated has been suggested by many scholars (Bolinger 1968: 127; Wierzbicka 1988; Dixon 1984: 594; Goddard 2007). I have already discussed this while addressing the semantic difference between the predicative participle construction and the finite complement clause occurring with verbs of perception (section 7.5). In that occasion, referring to Dik and Hengeveld (1991)'s work, I argued that the predicative participle construction is typically used to express the direct perception of states of affairs, while the finite complement construction is used to express indirect perception. The semantics of the participle predicative construction is linked to its ability to represent simultaneity between two events. Given this distinctive semantics, this construction rarely occurs with verbs of cognition. This is not surprising, as not all verbs are equally likely to appear in a given construction, due to its semantic contribution (Perek 2015: 78). However, even verbs with different meanings can still occur with a given construction if there is a way of integrating the semantics of the verb into the constructional semantics. In such instances, the construction contributes its meaning and may change the interpretation of the verb. In other words, the meaning of an argument structure construction contributes a semantic content that is fused with the meaning of the verb. The use of the predicative participle clause with this verb class clearly exemplifies how complement constructions contribute

their semantics. As I argued in section 8.5.1, when cognition verbs are used with a predicative participle clause, the construction supplies its distinctive meaning, and the semantics of the whole sentence is somewhere in between perception and cognition (see example 358 above in section 8.5.1).

Another case exemplifying the semantically-motivated choice of complement clauses is proposed by Wierzbicka (1988: 71-72) who, while studying complementation with verbs of remembering, singles out three constructions that in English may occur with the verb *remember*. The *that*-complement clause is used when the Experiencer remembers that s/he did something but does not have a mental picture of him/herself doing it. *Remember* + *V-ing* is used when the Experiencer has a mental picture of the event s/he is remembering. While *remember* + infinitive differs from the previous complement types as it exhibits the semantic components of wanting and/or future. On a similar line, Goddard (2007) analyzes the complementation types of the verb *forget* in English. Relying on a Natural Semantic Metalanguage approach, he distinguishes between three types of clauses that can occur with this verb and argues that each type is specialized for the expression of a specific semantics. These types are the *to*-complement clause, as in *I forgot to lock the door*, the *that*-complement clause, as in *I forgot that the door was locked*, and the *wh*-complement clause, as in *I forgot where I put the key*. In particular, Goddard shows that the *to*-complement involves the semantic components of WANTING and DOING, while the *that*-complement construction involves KNOWING.

Hindi exhibits a similar constructional distinction. When occurring with verbs of remembering and forgetting, finite complement clauses usually encode a piece of information that is remembered/forgotten by the Experiencer. Sentences 402 and 403 are two examples of finite complementation expressing memory. In 402, the complement clause is not introduced by any particle, and it is merely juxtaposed to the experiential construction in order to express an information that is suddenly retrieved in the mind of the Experiencer. In sentence 403, the complement clause expressing the content of the cognition is introduced by the particle *ki*. Sentence 404 is an example of the same finite complement clause with the verb *bhūlnā* "forget". In this case the construction is used to indicate that the Experiencer forgot a piece of information that s/he previously knew.

402. सहसा उसे याद आई, थैली में आठ सौ रूपये थे।

sahsā	*use*	*yād*	*ā-yī,*
suddenly	3SG.DAT	memory(F.SG.NOM)	come-PRF.M.SG

thailī=meṁ āṭh sau rūpye the

there were eight hundred rupees in the bag.

"Suddenly he remembered, there were eight hundred rupees in the bag."

403. तेरा पिता मुझसे नाराज है, मेरी अनुनयविनय कुछ नहीं सुनता, इसलिए याद रख कि तेरे सिवा मेरा अब कोई नहीं है।

terā pitā mujh=se nārāj hai, merī anunayavinay kuch nahīṁ suntā, islie
Your father is angry with me, doesn't listen to my pleas, so

yād	*rakh*	*ki*	*tere*	*sivā*	*mer-ā*	*ab*	*koī*
memory(F.SG.NOM)	keep	that	2SG.GEN	except	1SG.GEN-M	now	INDF.PRN

nahīṁ hai.
not be.PRS.3SG
"Your father is angry with me, doesn't listen to my pleas, so remember that I have no one except you now."

404. मैं यह भूल ही गयी थी कि वे लोग तुम्हारे घनिष्ठ मित्र हैं।

maiṁ	*yah*	*bhūl*	*hī*	*ga-yī*	*ki*	*ve*
1SG.NOM	this.NOM	forget	EMPH	go.PRF-F.SG	that	that.NOM.PL

log	*tumhār-e*	*ghaniṣṭh*	*mitr*
people(M.PL-NOM)	1PL.GEN-M.PL	intimate	friend(M.PL-NOM)

haiṁ.
be.3PL.PRS
"I had forgotten that those people are close friends of yours."

Non-finite complement clauses, instead, are used to express different meanings related to actions and events rather than to pieces of information. In sentence 405, a non-finite complement clause occurs with the oblique infinitive form of the verb followed by the genitive postposition *kī* and expresses the memory of an event of which the Experiencer has a mental picture, and not a piece of information.

405. पिताजी के कन्धों पर चढ़कर उचकने की याद मुझे अभी तक आती है।

pitā-jī=ke	*kandhoṁ=par*	*caṛh=kar*	*ucak-n-e=kī*
father-HON=GEN	shoulder(M)-PL.OBL=on	climb=CP	be_lifted-INF.OBL=GEN

yād	*mujhe*	*ab=hī=tak*	*ā-t-ī*	*hai.*
memory(F.SG.NOM)	1SG.DAT	now=EMPH=till	come-IPRF-F.SG	be.PRS.3SG

"I (can) still remember climbing up on my father's shoulders."

8.8.2. The distribution of the light verbs in complex predicates

As I discussed diffusely in this chapter, besides information on the semantic properties of the Experiencer the alternation of the light verb also contributes lexical aspect properties. The analysis of complex predicates referring to cognition shed light on the high productivity that these alternations show within this semantic class. Some verbs referring to cognitions may alternate among a wide range of light verbs, each contributing its own semantic nuance. This is particularly true, for example, for the complex predicates formed by the noun *patā* "news, information" and for the complex predicates formed by the noun *yād* "memory". In this respect, verbs of cognition contrast with verbs belonging to other semantic classes. Consider for example the complex predicates referring to hunger and thirst, formed by the nominal hosts *bhūkh* "hunger" and *pyās* "thirst". As I discussed in chapter 6, these predicates allow an alternation between the light verb *honā* and the light verb *lagnā* which results in different characterizations of the event, a stative one in the case of *honā* and an achievement one in the case of *lagnā*; but they do not allow any further alternation. This difference between complex predicates referring to cognitions and sensations is presumably due to the fact that cognitions are far more complex than bodily sensations with respect to their semantic properties. Bodily sensations typically refer to states that always involve non-agentive Experiencers; the only deviation from this prototype is the case in which a bodily sensation is construed as a change of state, rather than a state. Cognitions, in contrast, can exhibit different semantic properties both in relation to the participants and in relation to the event. For example, a cognition as remembering may allow both an agentive and a non-agentive Experiencer, and may be construed as an ongoing activity, as an achievement or as state.

The difference in the alternations allowed by different nominal hosts is a consequence of the fact that light verbs contribute a specific semantics that must be in accordance with the semantics of the noun. Hence, not all light verbs are allowed with all nominal hosts, and there seems to be some restrictions imposed by the noun. Thus, if a noun refers to some experiential types that do not allow an agentive Experiencer, light verbs contributing agentive readings are ruled out. Hence, for instance, the light verb *karnā* "do "is never allowed with complex predicates expressing hunger or thirst because its agentive semantics does not align with the lack of control and volitionality that characterize these bodily sensations. Another example of how the semantics of the noun conditions the selection of the light verb is the alternation between *honā* "be" and *lagnā* "attach" that contributes an achievement reading. For complex predicates referring to sensations, the light verb *honā* "be" alternates with the light verb *lagnā* "attach" (as in 406 vs 407), for complex predicate referring to cognitions, the achievement reading is typically contributed by the light verb *ānā* (as in 408 vs 409).

406. आत्मा को प्रेम की प्यास है ।

ātmā=ko	*prem=kī*	*pyās*	*hai.*
soul(F.SG.NOM)=DAT	love(M.SG.OBL)=GEN	thirst(F.SG.NOM)	be.3SG.PRS

The soul is thirsty for love.

407. बन्दी उठ खड़ी हुई, 'हमें तो प्यास लग रही है।

bandī uṭh khad-ī hu-ī,	*'hameṁ*	*to*	*pyās*	*lag*	*rah-ī*	*hai'.*
the prisoner stood up	1PL.DAT	EMPH	thirst(F.SG.NOM)	attach	PRG-F	be.3SG.PRS

"Then the prisoner stood up, 'We are feeling thirsty'."

408. तुम्हारा क्या विचार है, यह मैं नहीं जानती।

tumhār-ā	*kyā*	*vicār*	*hai,*	*yah maiṁ nahīṁ jāntī*
2PL.GEN-M	what	though(M.SG.NOM)	be.3SG.PRS	I don't know this

"I don't know what you think."

409. तुम्हारे बारे में जब विचार आया तो मुझे बड़ा दुःख हुआ।

tumhāre	*bare=meṁ*	*jab*	*vicār*	*ā-yā*
2PL.GEN	about=in	when	though(M.SG.NOM)	came-PRF.M

to mujhe baṛ-ā duḥkh hu-ā.

I felt very sad.

"When I thought about you, I felt very sad."

Table 58 illustrates the functional distribution of the verbs *lagnā* and *ānā* over the complex predicates belonging to the semantic classes discussed in this dissertation. From the distribution is evident that the choice between *lagnā* and *ānā* seems to be associated to the type of involvement of the Experiencer in the event. In particular, the verb *lagnā* is used in complex predicates that refer to experiential types that are typically low in the intellectual hierarchy, i.e. sensations and perceptions. These experiential types typically involve the Experiencer physically. The verb *ānā*, in contrast, is used to refer to events that are higher in the intellectual hierarchy and that involve the Experiencer mentally. In particular, the use of *lagnā* is sporadic for cognitions, as it only occurs with the complex predicate *patā lagnā* "find out". The same verbs, on other hand is systematically used for the expression of bodily sensations. For instance, cold, hot, hunger, and thirst are frequently encoded with the verb *lagnā*. The same verb is also frequently used to express non-agentive perceptions, in particular when referring to sense modalities that do not border with the domain of cognitions, such as touch, taste and smell.

Table 58: Distribution of the light verbs *lagnā* and *ānā* among experiential complex predicates.

Bodily sensations					
	Temperature		Saturation		Pain
lagnā	✓		✓		✓
ānā	×		×		×
Perceptions					
	Sight	Hearing	Touch	Taste	Smell
lagnā	✓	✓	✓	✓	✓
ānā	×	×	×	×	✓
Cognitions					
	Thought	Know	Remember	Understand	
lagnā	×	✓	×	×	
ānā	✓	×	✓	✓	

Notably, the distribution is not clear-cut. For example, complex predicates expressing knowing formed by the nominal host *patā* allows the light verb *lagnā* (410). Similarly, perceptions referring to smell may be expressed both by a complex predicate using the verb *lagnā* (412) and a complex predicate using the light verb *ānā* (411).

410. आज मैं तुझसे कह देती हुँ कि अगर इस तरह की बात फिर हुई और मुझे पता लगा, तो तीनों में से एक भी जीते न रहेंगे।

āj *maiṁ* *tujh=se* *kah* *de-t-ī* *hūṁ* *ki*
today 1SG.NOM 2SG.OBL=INS say give-IPRF-F.SG be.1SG.PRS that
agar *is* *tarah=kī* *bāt* *phir* *huī* *aur*
if this.OBL way=GEN fact(F.SG.NOM) again be.PRF-F.SG and
mujhe *patā* *lag-ā,*
1SG.DAT information(M.SG.NOM) attach-PRF.M.SG
to tīnoṁ=meṁ=se ek bhī jīte na raheṁge.
Then none of the three will survive.
“Today I tell you that if something like this happens again and I find out, none of the three will survive.”

411. सरकारी अधिकारियों ने हिन्दू मंदिरों का नियंत्रण अपने हाथों में ले रखा है, क्योंकि उन्हें इसमें पैसों की गंध लगती है।

sarkārī adhikār-iyoṁ=ne hindū maṁdir-oṁ=kā niyaṁtraṇ apne hāth-oṁ=meṁ le rakhā hai,
Government officials have taken control of Hindu temples in their hands,

kyoṁki	*unheṁ*	*is=meṁ*	*pais-oṁ=kī*	*gaṁdh*
because	3PL.DAT	this.OBL=in	money-(M)PL.OBL=GEN	smell(F.SG.NOM)

lag-t-ī	*hai*
adhere-IPRF-F.SG	be.3SG.PRS

"Government officials have taken control of Hindu temples in their hands because they smell money in it." (from hiTenTen21)

412. मुझे पायसम की खुशबू आ रही है।

mujhe	*pāyasam=kī*	*khuśbū*	*ā*	*rah-ī*	*hai.*
1SG.DAT	payasam=GEN	scent(F.SG.NOM)	come	PRG-F	be.PRS.3SG

"I smell payasam." (from hiTenTen21)

Despite these idiosyncrasies, however, the type frequency of these two verbs clearly points toward a semantically-based distribution. It is reasonable to conjecture that this distribution is closely linked to the meaning that the two verbs have when used as main verbs. As I discussed in section 6.5, when used as a simple verb *lagnā* means "be attached, adhere" and implies a physical contact that the verb *ānā* "come" does not exhibit. The use of the verb *lagnā* thus indicates a higher involvement of the Experiencer, which is conceived as more affected by the event.

9. Analysis and discussions

9.1. Constructions productivity and semantic coherence

In this work I used a cognitive-functional framework, and I relied on insights from the usage-based constructional approach that I outlined in section 4.2.2. The constructional analysis of the Hindi expression of experience that I carried out in the previous chapters allows me to draw some conclusions that I will expose here. In the course of my discussion, I was able to establish clear evidence of the interplay between semantics and syntax in Hindi, in particular I showed that most constructions used for the encoding of the experiential subdomains bear their own semantics and co-occur with specific classes of verbs. In chapters 6, 7 and 8, I provided qualitative analysis and some quantitative data of the constructions used in Hindi to encode bodily sensations, perceptions and cognitions. On some occasions, I also compared the frequencies with which a given construction occurs with specific grammatical features, such as aspect for verbs of cognitions. In the following sections, I will systematically describe the main constructions I dealt with in this work from a semasiological point of view, with respect to their productivity, semantic coherence and functional distribution in the domain of experience.

As discussed in chapter 1.1, experiential situations are exceptionally variable from the semantic point of view, both in relation to the properties of the participants and in relation to the way the event is construed. For example, bodily sensations prototypically refer to states or changes of state profiling a single participant that never shows agentivity and is affected by the event. Perceptions, on the other hand, profile two participants and allowing a certain degree of control and awareness on the part of the Experiencer. To further complicate matters, significant variability also exists within the same subdomain, as in the case of cognitions, which encompass situations like forgetting which are typically construed as lacking control and volitionality, as well as activities like reflecting which always imply agentivity by the Experiencer over the mental process. Additionally, the same experiential situation may display high variability with respect to the way it is construed: knowledge, for instance is typically viewed as a state but may also be construed as an achievement. The extremely variable semantic nature of the experiential domain makes the study of experiential constructions extremely useful for the investigation of the syntactic-semantic interplay in a language and to evaluate the extent to which a given morphosyntactic choice is driven by semantic features. As I discussed in chapter 4.3, Hindi tends to use constructions in an iconic way, meaning that it tends to associate each construction to a given semantics and to use it accordingly across the verbal lexicon. The consequence is that certain classes of verbs reject certain constructions because their meaning do not align with the semantics of the verbs. This hypothesis drove my decision to frame this work within the theoretical

approach of cognitive Construction Grammar. As discussed in chapter 4, this approach is based on the assumption that constructions have a specific semantics and exhibit a distributional preference for particular verbs with meanings closely related to that of the construction (Goldberg et al. 2004; Perek and Lemmens, 2010, Perek 2015). If a verb shares semantic similarities with the constructional meaning, it is eligible for use in that construction.

In the previous chapters, I have analyzed the argument structures used to encode the different classes of verbs belonging to the cognitive functional domain of experience. In particular, in chapter 6, I addressed verbs of bodily sensation, in chapter 7, I dealt with verbs of perception and in chapter 8, I analyzed verbs of cognition. Table 59 shows the main constructions analyzed in the previous chapters and their distribution among bodily sensation, perception and cognition. In the following pages, I will discuss the productivity and semantic consistency of the constructions shown in the table.

Table 59: Main constructions analyzed in this study and their distribution across the experiential semantic classes. (The brackets indicate that the construction is lexically specified and not productive inside that experiential subdomain.)

Constructions	**Sensations**	**Perceptions**	**Cognitions**
Transitive construction		✓	✓
Oblique Stimulus construction		✓	✓
Dative construction	✓	✓	✓
Genitive construction			(✓)
Locative construction	(✓)		(✓)
Copular construction	(✓)		

As Barðdal (1999, 2008) and Goldberg (1995) point out, constructions are used in productive ways in languages, but they are not equally productive. A construction is productive when it occurs with new verbs, and it is frequent in use (Barðdal 1999, Barðdal 2008). In section 4.2.2, I distinguished two types of frequency: type frequency and token frequency. Type frequency is defined as the number of verbs occurring in a certain construction, while token frequency is the number of times a verb is used in the same construction. In order to establish the productivity of a construction we need to look at its type frequency. The higher the type frequency of a construction is the more productive and the less semantic coherent it will be. As a consequence, in Construction Grammar, productivity is conceived as a function of type frequency and semantic coherence, as there is an inverse correlation between the productivity (and frequency) of a construction and the degree of semantic consistency

holding among the verbs that instantiate it. The more a construction is extended to different classes of verbs, the less semantically based it is (Bybee and Thompson 1997, Barðdal 1999, 2008, Fedriani 2012, Luraghi 2020a). From the interaction between type frequency and semantic coherence three types of semantic productivity may emerge. These three types are represented in Figure 32 (taken from Fedriani 2012: 192, after Barðdal 2008). In the first type, the construction is associated to a vast set of verbs and thus shows low *lexical strength* (according to Bybee (1985)'s terminology). This results in a lower semantic specificity of the construction and a wider range of verb classes are allowed to occur with it. This construction type is associated to the area around point A in the productivity continuum represented in Figure 32. In the second type of productivity, the construction is associated to a smaller set of verbs, which are semantically highly coherent, but it is stable and even productive, as it can still be extended to other verbs which align to that specific semantics. This construction type is associated to the area around point B in Figure 32. Both low and high semantic specificity thus may result in productive constructions, even if characterized by different types of productivity. In contrast, when a construction occurs with a small set of verbs, which is not semantically consistent, it is associated to a low degree of productivity as there is no reason for the construction to be extended to other classes. This construction type is associated to the area around point C in Figure 32.

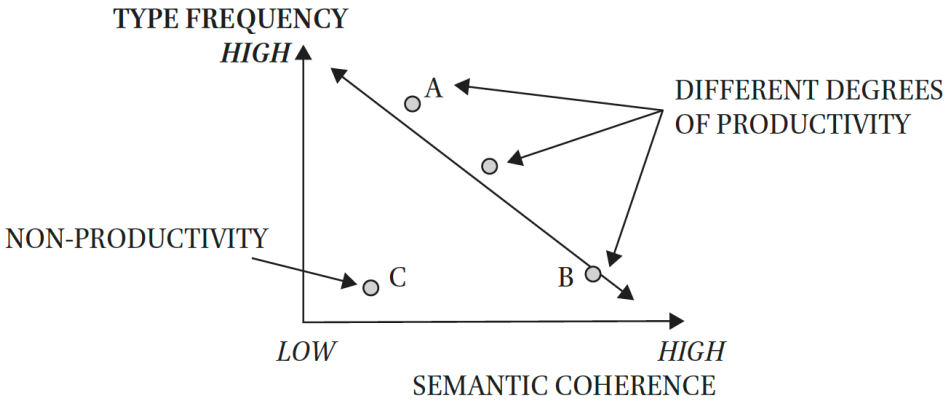

Figure 32: The productivity continuum (from Fedriani 2012: 192, after Barðdal 2008: 172).

By assessing the number of different verbs belonging to the experiential subdomains which occur with a given construction one can determine the type frequency of the construction. While by analyzing the semantic properties of the events encoded by the verbs occurring with a given construction one can establish its semantic coherence. In the following section, I will try to establish the semantic coherence and the productivity of the constructions in Table 59. I will mostly deal with the dative construction and the transitive construction, and their interaction within the grammar. The dative construction is the only one that can be used to encode all three experiential subdomains that I

analyzed in this work. In this construction the Stimulus or the Expertum are encoded in the nominative, while the Experiencer is marked with the dative. Experiential events are among the most distant from the semantic transitivity prototype. A prototypical transitive event is dynamic and concrete, characterized by an intentionally acting Agent and by a Patient that is directly affected in a perceptually salient way (Hopper and Thompson 1980, Kittilä 2002: 190, Haspelmath 2015, Malchukov 2015, Tsunoda 2015). A prototypical experiential situation typically lacks these properties: the prototypical Experiencer does not control the event and is not volitional, and the Stimulus is not affected by the event, it rather triggers it. As a consequence, the prototypical Experiencer marked with the dative is associated to the semantic role of the Goal/Recipient, and it is construed as a human participant involved in the event but that does not volitionally take part in it and that receives the consequences of the situation. As I discussed in the previous chapters, in experiential constructions the Experiencer is prototypically animate, most often humans, while the Stimulus may refer to a wider range of participant types. As a consequence, the Experiencer is typically in a higher position than the Stimulus in the salience hierarchy and it is the preferred candidate for the assignment of the subject relation, hence it is marked with the dative based on its semantic properties, and it is assigned the subject status based on its participant properties (section 2.2.1). The second most frequent construction among those shown in Table 59 is the transitive construction, which encodes the Experiencer in the nominative/ergative and the Stimulus in the nominative/accusative. While the dative construction reflects the semantic properties of the Experiencer and of the event in its globality, the transitive construction masks its semantic content, as this morphosyntactic coding is not driven by semantic properties but by default syntactic mechanisms triggered by grammatical relations.

9.1.1. The transitive construction

In the Hindi expression of experience, the transitive construction is quite productive, as it can occur with a wide set of semantic classes, and it can be associated to point A in the productivity continuum in Figure 32. As I discussed in previous chapters, this construction can be used for the expression of agentive and non-agentive perceptions and for the expression of cognitions. However, its semantic scope seems to be limited in Hindi, as it is not allowed for the expression of events that are characterized by the semantic properties of prototypical experiential situations, such as bodily sensations and many subtypes of uncontrolled mental change of states. Additionally, highly agentive situations in which a the Experiencer directs his attention toward a Stimulus construed as a target in Hindi may be encoded by the transitive construction, but also allow (and sometimes favor) a semantically based construction, in which the second argument is in the oblique case (the oblique Stimulus construction, see section 9.1.5.4 below). This suggests that both deviations from the

transitive prototype, i.e. lack of Agent properties and lack of Patient properties, lead to the rejection of the transitive construction in Hindi. I will address this latter type of deviation from the transitive prototype in section 9.1.5.4 below, while discussing the oblique Stimulus construction, and here I will focus on the deviation from the prototype characterized by the loss of the properties of the Agent, which favors the use of the dative construction.

As I mentioned in chapter 4.3, in this work I followed Haspelmath (2015) and Croft (2022), and I defined the transitive construction as the construction used to encode events that are prototypically transitive from a semantic point of view. The extension and the high productivity of the transitive pattern is a typological quite common phenomenon (Dixon 1994: 147, Haspelmath 2015, Malchukov 2015), and many languages around the world tend to associate this construction to non-prototypically transitive events on the base of syntactic principles (see the discussion on transitivity prominence in section 4.3). The reason to use the same pattern for the encoding of different semantic roles lies in the necessity to distinguish the two core arguments in a sentence from the peripheral ones. In this way, the participants that are associated to the two arguments of a transitive construction are unequivocally interpreted as the most prominent in the event as they are given the two highest syntactic relations, i.e. subject and object. This extension is a clear case of markedness (Malchukov 2015), in which syntactic constraints prevail on semantic ones. As I have shown in the previous chapters, it frequently happens also in Hindi that the transitive construction is extended according to this markedness strategy. This happens for example in the case of verbs of perception that occur with a transitive construction even when encoding non-agentive events. The fact that the transitive construction may express both agentive and non-agentive situations was already noticed by other scholars before. Masica (1991: 350) for example pointed out that the contrast between the transitive and the dative construction in South Asian languages "is not volitional / nonvolitional but rather neutral (unmarked) / nonvolitional (marked). It is not that Direct constructions [i.e. transitive constructions] are necessarily volitional, but rather that they are unmarked as to volitionality and thus may be (where the contrast is present may even tend to be) volitional, whereas the Dative construction is definitely nonvolitional." In sum, the transitive construction is often not associated to a specific semantics and its semantic coherence is low.

However, it should also be noted that the constructional analysis outlined in the previous chapters showed that this syntactically based extension of the transitive pattern is often restricted in Hindi by semantic constraints and there are clear cut-off points at which Hindi does not allow the transitive construction. For example, Hindi does not use the transitive construction to encode bodily sensations. Moreover, some verbs in Hindi show syntactic exceptions that seem to be caused by semantic properties. In section 8.6, I discussed the case of the verbs *samajhnā* "understand" and *bhūlnā*

"forget", which may occur in a perfective transitive construction without an ergative marking on the Experiencer. Recall from chapter 4.3, that in Hindi the transitive construction displays different instantiations based on aspectual properties (split ergativity) and on the participant properties of the Patient (differential object marking). In my definition of Hindi transitive construction, the alternation between the nominative and the ergative case on the subject is an intrinsic characteristic as it always occurs with verbs referring to prototypical transitive events, such as *break*. Yet, as discussed in section 4.3, the ergative case marking in Hindi is also typically associated to agentive roles (Mahajan 1990, Butt, Butt and King 2002, de Hoop and Narasimhan 2005, Butt and Ahmed 2010), and thus may also be subject to semantically based constraints. This semantic use of the ergative is typical of unergative verbs of bodily emission which may alternate between an ergative and a nominative marking of the argument. In this case, the ergative implies agentivity, while the nominative is semantically unmarked. As a consequence, verbs that are not agentive may tend to reject such a case marking due to semantic incompatibility. Since *samajhnā* "understand" and *bhūlnā* "forget" depict situations in which the Experiencer is not endowed with control and volitionality, they reject the ergative marking. As I discussed in section 8.7.2, this is particularly true for *bhūlnā*, which never occurs with an ergative Experiencer, while *samajhnā* may occur both with an ergative and a nominative Experiencer. This results in sentences such as 413, in which the verb *bhūlnā* occurs in a perfective aspect, but the first argument is in the nominative and agrees with the verb, while the Stimulus is marked with the accusative *ko* that is typically reserved to transitive constructions. [32]

413. वह अब तक भी नौ साल पहले की उस बात को नहीं भूली कि मैं उससे कहकर भी उस शाम उसके यहां नहीं गया था।

vah	*ab=tak*	*bhī*	*nau*	*sāl*	*pahle=kī*	*us*
3SG.NOM	now=till	EMPH	nine	year	before=GEN	that.OBL
bāt=ko		*nahīṁ*	*bhūl-ī*		*ki*	
matter(F.SG.OBL)=ACC		not	forget-PRF.F		that	

maiṁ usse kahkar bhī us śām uske yahāṁ nahīṁ gayā thā.

I didn't go to her place that evening, despite telling her I would.

[32] Recall from section 8.7.2 that the verb *bhūlnā* most frequently appears in a verb-verb complex predicate with the vector verb *jānā* "go", which being intransitive makes the whole construction intransitive. This might have influenced the marking of the first argument even in cases in which the verb is not followed by *jānā*. However, this is not problematic for my interpretation, as combinations of transitive verbs followed by intransitive vector verbs are extremely rare in Hindi and have been defined as atypical by some scholars (see Drocco and Tiwari 2020a, 2020b). The fact that the verb *bhūlnā* is most frequently followed by the verb *jānā* is further evidence that its status is ambiguous with respect to transitivity.

"She still hasn't forgotten that incident from nine years ago, when I didn't go to her place that evening, despite telling her I would."

In sum, the Hindi transitive construction is productive and can be associated to point A in the continuum in Figure 32, as its productivity derives from the fact that the verbs it occurs with are not characterized by a high semantic consistency. However, the transitive construction seems to be in decline since, as I will discuss below in this chapter, semantic constrains are getting stronger in the language. This is supported by the rejection of the ergative marking by verbs implying lack of control and volitionally, which seems to be a consequence of the fact that the ergative postposition tend to be associated to the semantic property of agentivity. The verbs *samajhnā* "understand" and *bhūlnā* "forget" thus can be considered in the middle of two different tendencies in the language. On the one hand, the transitive pattern is extended by default on the basis of syntactic strategies to the expression of two arguments simple verbs which profile at least a human participant, and as a consequence this construction is attracted to verbs such as *samajhnā* "understand" and *bhūlnā* "forget". On the other hand, the ergative case marking tends to be associated to a high degree of agentivity on the part of the Experiencer and is thus rejected based on semantic constraints.

9.1.2. The dative construction

The dative construction can be placed between point A and B in the productivity continuum. It is coherent from a semantic perspective, but at the same time its semantic scope is wide. The construction is generically associated to a participant that does not control the event and consequently it may occur with verbs encoding many situation types: experiential events, happenstances (as in example 414) and non-agentive events such as "meet unintentionally" or "found". The dative construction seems to be the most productive for the expression of the experiential classes investigated in this study: it is the only one used for the expression of every experiential subdomain, and thus shows the highest type frequency. This construction may encode stative and non-stative events; however, when expressing non-stative events, it refers to telic situations that cannot be interpreted as activities.

414. *us=ko* *kyā* *huā?*
3SG.OBL=DAT what be.PRF-M.SG
"What happened to him?"

Previous literature (Ahmed 2006, 2009, Butt and Ahmed 2010, Reinöhl 2016) has extensively discussed the original allative function of the postposition *ko*. This origin is supported both by

diachronic and by synchronic data. In particular, as already discussed in section 3.2.2.1, the dative *ko* comes from a spatial marker that is the original locative form of the Sanskrit noun *kakṣā* "armpit" (*kakṣe*). In Old Hindi, the postposition *ko* still retained the function of a spatial marker with an allative meaning (consider for example sentence 415 and 416 from Hindi, adapted from Ahmed 2009: 71). In modern Hindi, the postposition *ko* may still optionally be used to mark the allative function (see example 417), even if most of the times spatial Goals are expressed by the simple oblique form of the noun.

415. *us* *simt=ko* *cal-ā*
that.OBL direction=towards move-PRF.M.SG
"(He) moved to that direction." (Dehalvi 1804, Old Urdu/Hindi)

416. *ab* *vatan=ko* *jā-t.ā* *hūṁ*
now home-country=to go-IPRF-M.SG be.1SG.PRS
"Now (I) go to (my) country." (Dehalvi 1804, Old Urdu/Hindi)

417. *maiṁ* *dīllī(=ko)* *jat-ī* *hūṁ*
1SG.NOM delhi=DAT go-IPRF-F be.1SG.PRS
"I am going to Delhi."

As Ahmed (2009) argues, data from other Indo-Aryan and Indo-Iranian languages support the original spatial meaning of this postposition. For instance, Bengali displays a locative marking *kāche* that literally means "near" and that is etymologically connected to the same locative form of the Sanskrit noun *kakṣā*. The Iranian language Pashto marks locative functions with the postposition *kī,* which according to Hewson and Bubenik (2006) derives from the Avestan word *kaṣa* "armpit", a cognate of the Sanskrit *kakṣā*. And the same is true for the postposition *khe* present in some dialects of Pashto. The extension of an originally allative marker for the expression of Recipients and Experiencers is quite common typologically. Luraghi (2014: 112-113), for example, points out that Experiencers may be encoded as spatial roles, and particularly frequently as locations or as directions. On a similar line, Haspelmath (2003, 2004) and Narrog (2010) note that in many languages the same element that is used to mark directionality is also used to mark Recipients and Experiencers. Consider for example the use of the English preposition *to* in sentences 418-420 and of the French preposition *à* in sentences 421-423, both expressing a spatial Goal (418 and 421), a Recipient (419 and 422) and a Beneficiary (420 and 423).

418. *He went to the supermarket*.

419. *She gave the keys to her husband.*

420. *It seemed impossible to her husband ,*

421. *Il est allé au supermarché*

422. *Elle a donné les clés à son mari.*

423. *Cela semblait impossible à son mari.*

Based on a cross-linguistic comparison of the different meanings used by the dative case in the languages of the world, Haspelmath (2003) draws the conceptual space in Figure 33, which depicts the semantic network of the functions typically encoded by dative markers.

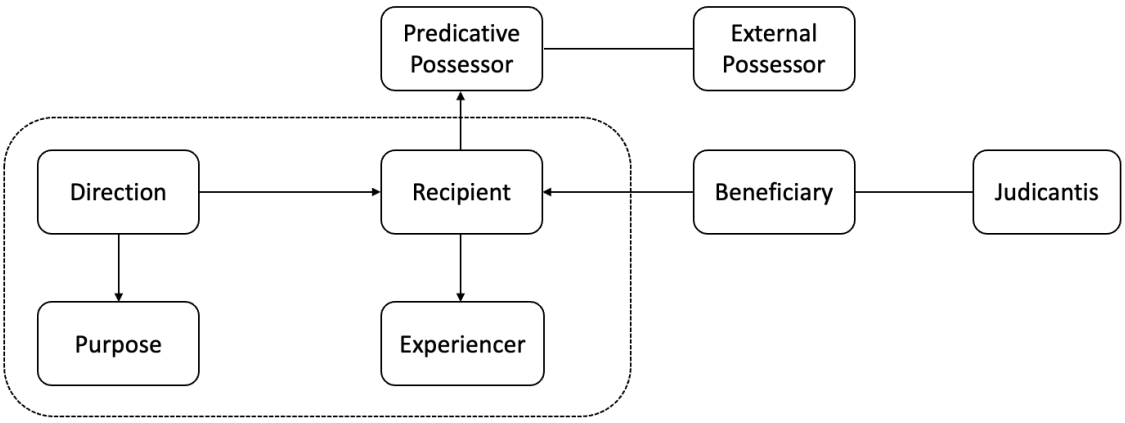

Figure 33: Semantic map of the postposition *ko* in Hindi, over Haspelmath's conceptual space of typical dative functions (adapted from Haspelmath 2003: 234).

As it is shown in the figure, there is a typological tendency to use the same marker that encodes directionality to also encode Recipients and Experiencers. The dotted line in the figure shows the semantic functions covered by the Hindi postposition *ko*. In particular, *ko* is today used to encode Purpose (as in 424), Recipient, and Experiencer.

424. तुम्हारी बातें सुनने को तैयार हुँ।

tumhār-ī	*bāt-eṁ*	*sun-n-e=ko*	*taiyār*	*hūṁ.*
2PL.GEN-F	word(F)-PL.NOM	listen-INF-OBL=DAT	ready	be.1SG.PRS

"I am ready to listen to what you have to say."

Ahmed (2009) analyses non spatial usages of spatial markers in South Asian languages and proposes a model according to which different usages of the same marker are accounted for by a single unspecified lexical entry and the use of spatial markers to encode non spatial arguments is the result of a systematic metaphorical extension of spatial concepts. Such metaphorical extensions are quite

common typologically and have been discussed in detail by many scholars within the cognitivist approach (see Taylor 1993, Tyler and Evans 2003, Fedriani 2012, Luraghi and Narrog 2014). They have been explained with the notion of embodiment (2.1), according to which complex and abstract meanings tend to be metaphorically mapped onto more concrete and physical ones. For example, the use of the postposition *ko* to express Purpose in Hindi can be explained via the conceptual metaphor PURPOSES ARE DESTINATIONS (Lakoff and Johnson 1980). A similar metaphor operated diachronically and resulted in the use of the Sanskrit spatial marker *kakṣe* (> Hindi *ko*) for the expression of Recipients in modern Indo-Aryan languages. The metaphor is RECIPIENTS ARE DESTINATIONS and, as the typological observations discussed above show, it is overwhelmingly popular among the languages of the world (see Blansitt 1988, Rice and Kabata 2007, Luraghi 2014). The use of the same postposition *ko* to express both Recipients and Experiencers can be accounted for in terms of another metaphor according to which EXPERIENCERS ARE RECIPIENTS OF FEELINGS AND SENSATIONS and FEELINGS AND SENSATIONS ARE THINGS (Luraghi 2014: 112). As the constructional analysis in the previous chapter shows, every experiential situation can be conceptualized through this metaphor in Hindi, since the prototypical Experiencer type is similar to a prototypical Recipient: they are both human and non-volitional, and both receive the consequences of by the event.

9.1.3. The rise of the dative construction

Barðdal (1999, 2008) remarks that besides type frequency, the other parameter used to establish the productivity of a construction is its occurrence with new verbs: either new verbs that are introduced in the vocabulary through borrowings or verbs already existing in the lexicon that begin to be used with a different meaning. As discussed in 3.2.4.1, dative subject constructions are rather recent in Indo-Aryan languages and Old Indo-Aryan languages lack non-canonical subject constructions (Hock 1991, Butt and Deo 2013, Montaut 2016). Sanskrit was a nominative-accusative language with a strong tendency to select the transitive frame to encode most semantic verb classes, including experiential verbs and other non-semantically transitive verbs. Previous scholars linked the spread of non-nominative subjects in Indo-Aryan languages to the split ergative alignment of these languages, arguing that the ergative construction opened the way to the possibility of marking the first argument of a verb with a non-nominative case, favoring the rise of differential subject markings (see on this Verbeke, Kulikov and Williams 2015). However even if the two phases at the extremes of this diachronic development are clear, i.e from a nominative-accusative language to an ergative-absolutive language with many non-canonical subjects, the reconstruction of the intermediate phases poses some challenges. In the following sections I will try to shed light on this development.

As Montaut remarks (2016), non-nominative subjects increased in the grammar between the 14th and 16th century, also as a consequence of the long contact between North Indian languages and Persian that occurred during the Moghul empire. This long-lasting contact with Persian caused a massive borrowing that restructured the verbal system in New Indo-Aryan and that boosted a shift in Indo-Aryan alignment. Montaut points out that in this period the dative subject construction was the rising one for the expression of experiences. At that time, this construction occurred mainly with the verb *milnā* "meet" and was used in the context of popular devotion to express the experience of the mystical fusion of the devotee with the deity. For example, Montaut (2013, 2016) points out that in the works of Kabir, a mystic poet of the 15th century, the dative construction is mostly used with the verb *milnā*, and with only a few other verbs, such as *bhānā* "please, like" (see section 9.1.4 below). As Barðdal (2009) notes, given that new verbs are generally attracted to the most frequent construction in the grammar, a rapid change in the lexicon of the language may increase the spread of a construction "precisely because the bulk of new verbs will be attracted by the high type frequency constructions, thereby lowering the proportional type frequency of the low type frequency constructions, increasing the chances of them becoming extinct" (2009: 142). The lexical restructuring of the predicate system started by the contact with Persian attracted the dative construction, causing its extension to more and more predicates (for a detailed discussion on this, see Montaut 2016). Dative subject constructions were thus boosted by the restructuring of the verbal lexicon that occurred during the Moghul empire. But why should the new class of verbs be attracted by the emerging dative construction and not by the already stable and productive transitive construction? I think that the reason lies in the fact that during the Moghul empire the language was already undergoing a restructuring of its case system that developed into a semantically based use of case-marking. Old Indo-Aryan was characterized by a highly inflecting nominal morphology, while modern Hindi displays a simpler nominal morphology integrated by postpositions. As Butt (2006b) and Butt and Ahmed (2010) argue, one of the main factors governing this re-development of the case system in NIA languages is the expression of systematic semantic contrasts (Butt and Ahmed 2010: 2). Hence, the period marked by the restructuring of the verbal lexicon saw the development of a semantically based use of case-marking. This resulted in semantically constrained selections of constructions allowed with the new experiential complex predicates. Thus, the dative construction became the construction on the rise as it was the most semantically coherent with the experiential class. As a consequence of the contact with Persian, the new dative construction, unlike the old transitive construction, ended up being linked to a set of verbs with a quite consistent semantic characterization. This resulted in the strengthening of the construction, in a sort of positive reinforcing cycle that ultimately led to a higher degree of productivity. Hence, the dative construction displays a

high type frequency, but it is still semantically consistent: dative subjects are allowed with many verb classes (verbs of bodily sensation, perception, cognition, emotion, happenstance, etc.) all showing a non-agentive human participant.

The role that the restructuring of the verbal lexicon had in the spread of the dative construction in modern Hindi explains why the massive presence of non-nominative subjects correlates with noun-verb complex predicates. As is well known, only a handful of simple verbs require a dative construction in Hindi (see the list below (I-V)), while hundreds of experiential complex predicates require a dative subject.

I. *milnā* "get, meet".
II. *lagnā* "adhere" when used with the meaning "think".
III. *ānā* "come" when used with the meaning "know" (see section 8.67.4.1).
IV. *dikhnā* "see" (see section 7.4.1).
V. *sūjhnā* "think, occur to someone's mind" (see section 8.4.1).

Following Montaut and Barðdal, we can hypothesize that the dative pattern originally associated to a very specific semantics attracted all the new verbs whose semantics aligned with that of the dative construction and became very productive, developing a strong connection with complex predicates that still exists in contemporary Hindi. Additionally, since the massive lexical borrowing favored high type frequency constructions, it also boosted emerging phenomena in the language, such as the semantically based use of case-marking discussed above. This is resulted in a shift towards a more semantically based use of constructions in the grammar.

9.1.4. Two competing constructions in the language

In this section, I address the increasing productivity of the dative construction in modern Hindi, and I show that the dative construction is in competition with the transitive one as it is slowly extending its domain to verbs that previously belonged to the domain of the transitive constructions. This is shown by the fact that the same verbs that reject the ergative marking on the Experiencer also seem to attract the dative marking. This behavior is attested for the verb *bhūlnā* "forget" in my corpus and is mentioned in the previous literature also with respect to other verbs, such as *samajhnā* "understand" (Montaut 2013: 105). See the contrasting examples for the verb *bhūlnā* in 425 and 426 below. In sentence 425 the verb is used with a nominative Experiencer, while in sentence 426 it occurs with a dative Experiencer and a nominative Stimulus.

425. तुम मुझे भूल जाओगी, किन्तु मैं तुम्हें न भूलूंगा।

tum	*mujhe*	*bhūl*	*ja-o-g-ī,*	*kintu*	*maiṁ*
2SG.NOM	1SG.ACC	forget	go-2PL-FUT- F	but	1SG.NOM
tumheṁ	*na*	*bhūl-ūṁ-g-ā*			
2SG.ACC	not	forget-1SG- FUT-M			

"You will forget me, but I won't forget you."

426. मुझे अब सब दर्द भूल गया।

mujhe	*ab*	*sab*	*dard*	*bhūl*	*ga-yā*
1SG.DAT	now	all	pain(M.SG.NOM)	forget	go.PRF-M.SG

"Now I forgot all the pain."

Both *samajhnā* "understand" and *bhūlnā* "forget" are prototypical experiential verbs, in which the Experiencer is not agentive and does not control the event. The occurrence of these two verbs with the dative construction indicates that this construction has grown so productive for experiential verbs that it does not only extend to new verbs but also to verbs that already exist in the verbal lexicon and that are now being recategorized as dative subjects verbs due to their semantics. Remarkably, this extension occurs even at the expense of the default transitive pattern and suggests that the semantic reasons are becoming stronger that the syntactic ones in the language. Another case exemplifying this tendency is represented by the complex predicates expressing non-agentive visual perception *dikhāī denā* "see" and auditory perception *sunāī denā* "hear". As I discussed in section 7.3.3, the verb *denā* "give" is a ditransitive verb, occurring with an Agent, a Theme and a Recipient, respectively encoded as subject, direct object and indirect object. As any other (di)transitive verb in Hindi, *denā* shows two alignments depending on aspect: split ergativity occurs not only when the verb is used as a simple verb with its basic meaning "give", but also when it is used as a light verb in other complex predicates. However, the complex predicates *dikhāī denā* and *sunāī denā* never allow an ergative marking in their argument structure and consistently occur in a dative construction. These two verbs were recategorized as complex predicates requiring a dative subject, due to their semantics, as they only express events in which the Experiencer does not control or start the perception intentionally but perceives it by chance. In sum, the dative and the transitive construction seem to be in conflict with one another within the grammar, as they represent the two encoding strategies in a language: markedness and iconicity. The study of the semantic classes analyzed in this dissertation suggests that the iconicity strategy seems to be prevailing within Hindi grammar.

This tendency was already rampant in early New Indo-Aryan languages (see section 9.2.1 below in this chapter). For example, the verb *milnā* "meet" originally occurred with a canonical nominative subject in Sanskrit and with a second argument encoded in various cases (accusative,

genitive, instrumental) but not dative. However, as early as in NIA, *milnā* may occur with an oblique subject construction (see the contrastive examples in 427, and 428 adapted from Montaut 2013: 97, 112). The same happens in Modern Hindi, where the verb *milnā* may occur both with a dative construction and with an oblique second argument construction. As discussed in section 4.2.3, the choice is driven, as it is typical in the language, by the semantic properties of the event.

427.	*premīṁ=kauṁ*	*premīṁ*	*milai,*	*tab*	*sab*	*biṣ*
	Lover=DAT	lover	meet/find.PRS.3SG	then	all	poison
	amrt hoi					
	nectar be.PRS.3SG					
	"When the lover finds the lover, all poison becomes nectar."					

428.	*tan*	*mati=maiṁ*	*mili*	*gayā*
	body.M.SG	earth=LOC	mix/meet.CP	go.PFV.M.SG
	"The body got mixed with/into the earth."			

The diachronic evolution briefly outlined in this section depicts a complex picture in which a massive borrowing triggered by contact with Persian resulted in a restructuring of the verbal lexicon. This interacted with an emerging semantically constrained use of postpositions, leading to a shift from a highly nominative-accusative language (OIA), characterized by a high transitivity prominence and a preference of markedness over iconicity, to a language low in transitivity prominence that favors a semantically based use of morphosyntax. Hence, the property that enabled the transitive construction to become so productive in OIA, that is its semantic vagueness, is leading to its decline in the NIA language. This suggests that what makes a construction productive in one linguistic system does not yield the same results in others. Productivity is influenced by how the encoding strategies of markedness and iconicity interact in the language. Languages that privilege an iconic marking of semantic roles will promote constructions with high semantic coherence, thus favoring the productivity type associated to point B in Figure 32 and marginalizing the transitive construction.

9.1.5. Constructions with limited semantic scope

Beside the dative and the transitive constructions, other constructions are used in Hindi for the expression of experience. These constructions tend to be associated to a high semantic coherence. In particular, as it emerged from the constructional analysis of the previous chapters, four main constructions are used for the expression of experiences in Hindi which show a limited semantic scope. These are the genitive construction, the locative construction, the copular construction and the

oblique Stimulus construction. This section is dedicated to a final discussion of their semantics and their functional distribution over the experiential domain.

9.1.5.1. Genitive Experiencer construction

The genitive Experiencer construction is used with few complex predicates referring to cognitions. The semantics of the genitive-nominative construction is not very productive within the verb classes belonging to the experiential subdomain. This construction in Hindi is generally used to refer to situations such as inalienable possession (429), prototypical ownership (430) or meronymic relations (431) and is prototypically associated with stative situations depicting a relationship between two entities. In these situations, the participant that is marked with the genitive generally displays a higher status than the other participant. For example, when this construction is used to encode ownership, the genitive argument expresses a possessor that shows some control over the nominative possessee. When it expresses a meronymic relation the genitive argument encodes the whole, while the nominative argument refers to one of the whole's parts. This semantic characterization explains why the genitive construction is found only with a small set of experiential predicates, i.e. with verbs of cognitions that may refer to mental states and that show some degree of agentivity on the part of the Experiencer (such as *vicār honā* "think" or *irādā honā* "intend", see section 8.4.2). As a consequence, this construction is not allowed with bodily sensations and perceptions.

429. उनकी तीन लड़कियाँ थीं।

un=kī	*tīn*	*laṛk-iyāṁ*	*th-īṁ*
3PL.GEN-F	three	daughter(F)-PL.NOM	be.PST-F.PL

"He had three daughters."

430. उसके मौरूसी पांच बीघे खेत हैं।

us=ke	*maurūsī*	*pāṁc*	*bīghe*	*khet*	*haiṁ*
3SG.OLB=GEN	inherited	five	bighe	field(M.PL.NOM)	be.3PL.PRS

"He has an inherited field of five bighe (Indian unit of measure)."

431. उस घर के चार कमरे हैं।

us	*ghar= ke*	*cār*	*kamr-e*	*haiṁ*
that.OBL	house.OBL=GEN	four	room(M)-PL.NOM	be.3PL.PRS

"That house has four rooms." Lit. "Of that house, there are four rooms."

9.1.5.2. The locative construction

Another construction that shows a high degree of semantic coherence is the locative construction used for the expression of pain and itching. This construction is not used for the expression of other bodily sensations such as feelings of temperature and saturation, as it requires the experience to be localized on a specific area of the body (section 6.1). Despite its limited semantic scope, this construction is highly productive, as shown by the fact that it is the most frequent with verbs referring to localized sensations. This means that, even though it shows a low type frequency and is restricted from a semantic point of view, the locative construction is still productive as it is higher in token frequency in comparison to the other constructions expressing localized sensations. Remarkably, the same construction may also be used to encode understanding in Hindi (section 8.6). This is interesting because it shows that internal bodily sensations and internal mental achievements are construed in a similar way. In both cases, the Experiencer remains in the background while his/her part is construed as a container in which the Expertum is contained (sensations) or towards which it is directed (understanding).

9.1.5.3. The copular construction

The copular construction is limited to the expression of feelings of hunger and thirst in Hindi and cannot express other experiential types. Once again, I argue that the reason for the limited extension of this construction is connected to its high semantic coherence. The copular construction is used to encode transitory states and the possession of inalienable qualities, and requires only one argument, encoded in the nominative. The extension of this construction to experiential situations such has perceptions, thus, is not allowed for two main reasons: 1. perceptions are typically dynamic situations, not states; 2. perceptive situations involve two participants and perception verbs require an argument structure with two arguments. This latter point also explains why cognitions, which are often construed as states (for example knowing), are not expressed by this construction in Hindi. Cognitive situations always imply two participants and thus require two arguments argument structures. Additionally, recall from section 6.2 that this construction cannot encode many bodily sensations either and it is only used to express feelings of saturation. This is because the copular construction is specifically meant to express states (such as *to be tall*) rather than feelings of states. As I argued in section 6.3, this construction can metonymically be used to express feelings of saturation due to the fact that being hungry and feeling hungry always correlate, but it does not allow the expression of pain or temperature since these bodily sensation types do not necessarily correlate with specific bodily states (Verhoeven 2007: 43).

9.1.5.4. The oblique Stimulus construction

The oblique Stimulus construction is highly coherent from a semantic perspective, and it is associated with verbs that exhibit a high degree of agentivity and imply directionality. In this work, I used the term oblique Stimulus construction as an umbrella term covering many subtypes, which differ according to the postposition marking the Stimulus. The reason to consider these constructions as all belonging to the same macro-type lies in the fact that they deviate from the transitive prototype in a similar way. They all imply an intentional Experiencer who is in control of the situation, which alternates between a nominative and an ergative marking, and a Stimulus that is not physically affected by the event and that may be involved in various ways. The role of the Stimulus in the event and its conceptualization determines its case marking. The postpositions that may instantiate this construction are represented in Table 60. Notably, different postpositions also vary in how they contribute to the construal of the Stimulus: the three postpositions *meṁ* "in", *par* "on" and *kī or/taraf* "towards" all supply a spatial meaning and construe the Stimulus as a Target in the experiential event. In contrast, the postposition *se* expresses both a comitative and a delative meaning and it construes the Stimulus either as a Companion or as a Source.

Table 60: Postpositions allowed in the Oblique Stimulus construction and their semantics.

Postposition	Spatial relation	Meaning	Verb with which it occurs
meṁ	inessive	"in"	Cognition (*socnā*)
se	delative	"from"	Emotion (*prem karnā, viśvās karnā*)
par	superessive	"on"	Cognition, emotion (*socnā, dayā/gussā honā*)
kī or/taraf	directionality	"toward"	Perception (*dekhnā*, *tāknā*)

In this work, I did not specifically address the oblique Stimulus construction featuring the postposition *se* as it is typically associated to the experiential subdomain of emotions in Hindi. Outside the experiential domain, the postposition *se* on the second argument is typically restricted to verbs that express interactive situations such as "meet someone purposefully" or "marry someone" (see the verb *vivāh karnā* lit. "marriage do" in 432) and are thus associated with events in which a second participant collaborates with the first participant to bringing about the event. Verbs of experiences occurring with this construction all belong to the subdomain of emotions, such as *viśvās karnā* "trust", *prem karnā* "love", *pyār karnā* "love", etc. In sentence 432, the oblique Stimulus occurs with the verb *prem karnā* "love, lit. love do".

432. और जो मैं कहुँ कि मैं तुमसे प्रेम करता हुँ, तो तुम मुझसे विवाह करोगी?

aur *jo* *maiṁ* *kah-ūṁ* *ki* *maiṁ* *tum=se*
and if 1.SG.NOM say-SBJV.1SG that 1SG.NOM 2PL=INS
prem *kar-t-ā* *hūṁ,* *to* *tum* *mujh=se*
love(M.SG.NOM) do-IPRF-M.SG be.1SG.PRS then 2PL.NOM 1SG.OBL=INS
vivāh *kar-o-g-ī?*
marriage(M.SG.NOM) do-2PL-FUT-F
"If I tell you that I love you, will you marry me?"

The two most frequent postpositions marking the oblique Stimulus occurring with the verbs analyzed in this work are the superessive postposition *par* "on", which is very frequently used for the expression of controlled mental activities such as "think about, reflect on" (section 8.4.2), and the postposition *kī or/taraf,* which implies directionality and is mostly used with verbs of agentitve visual perceptions (section 7.3.1). These two constructions may be associated to the so-called English conative construction (Goldberg 1995, 2006, Perek 2015). The English conative construction consists of an argument encoded as a subject and a second argument realized through a prepositional phrase headed by the preposition *at*. In English, this construction typically occurs with originally transitive verbs and, semantically, it has been described as a "detransitivizing" construction as it supplies semantic features that contribute to construe the event as deviating from the transitive prototype (Hopper and Thompson 1980, Langacker 1991, Taylor 1995). According to Dixon (1991: 280), for example, the prepositional phrase in the conative construction typically implies "that [the object] lacks some of the salient properties associated with the syntactic relation 'object'". Previous literature (Dixon 1991, Levin 1993, Goldberg 1995, Schlesinger 1995, Broccias 2001, Perek 2015) lists several semantic components contributed by this construction that construe the event as deviating from semantic transitivity. Among these semantic components, missed contact and lack of affectedness on the Patient are prominent, as this construction generally depicts situations in which the Agent directs his/her action towards a second participant, but does not make contact with it, or situations in which the Agent is able to have contact with the second argument, but his/her action does not physically affect it. Some scholars also consider lack of intentionality as another relevant semantic aspect contributed by the conative construction (Dixon, 1991; Guerrero Medina, 2011; Van der Leek, 1996). However, this is not the case for Hindi, as the oblique Stimulus construction clearly always implies an agentive reading. As a matter of fact, this construction is even more closely associated with agentivity than the transitive construction, as the latter may be extended to encode events that do not have a first human participant in control of the situation and intentionally bringing it about. In contrast, the Hindi oblique Stimulus construction does not allow the expression of events lacking an agentive participant. Goldberg (1995: 63) proposes to consider directionality as the basic semantic

component of the English conative construction and thus she identifies the basic semantics of the construction in the meaning "X DIRECTS ACTION AT Y" (see on this also Pinker 1989). As Perek (2014) highlights, this analysis accounts for the fact that this construction can occur not only with contact verbs (such as *hit* or *kick*), but also with verbs of agentive perception (as in *She was looking at me*), verbs of sound emission (as in *She was yelling at me*) and verbs of facial expression (as in *She was smiling at me*). In sum, this construction occurs with verbs referring to an action that is directed by the Agent towards a Target, that generally does not imply some kind of physical contact and that is only allowed with verbs characterized by an "orientational component" (see on this also Perek and Lemmens 2010: 27). The Hindi oblique Stimulus construction featuring the postposition *kī or*/*kī taraf* and *par* seems to imply a similar semantics. This is supported also by the fact that this construction may be used to encode the same meanings that are expressed by the conative construction in English. Sentences 433 and 434 are examples of this construction occurring with verbs of sound emission.

433. काफी देर तक अमरीश आमिर पर चिल्लाते रहे।

kāphī	*der=tak*	*amrīś*	*āmir=par*	*cillā-t-e*
enough	period_of_time=till	Amrish.NOM	Amir=on	shout-IPRF-M.PLG
rah-e				
stay-PRF-M.PL				

"Amrish kept shouting at Aamir for a long time." (From HiTenTen21)

434. वह सड़क के दूसरी ओर खड़े एक दस बारह साल के लड़के की ओर चिल्लाया।

vah	*saḍak=ke*	*dūsrī*	*or*	*khaḍe*	*ek*
3SG.NOM	street=GEN	other-F	direction	standing-M.PL	a
das	*bārah sāl=ke*		*laḍk-e=kī or*	*cillā-yā.*	
ten	twelve year=GEN		boy(M)SG.OBL=towards	shout-PRF.M.SG	

"He shouted towards a ten/twelve-year-old boy standing on the other side of the road." (From HiTenTen21)

The oblique Stimulus construction is highly productive within the semantic classes that allow it, as it is shown by the fact that it is the most frequent, if not the only, construction used with certain verbs lexicalizing the agentivity of the Experiencer. As discussed in chapter 7.3.1, the verb *tāknā* "stare, observe", for example, most frequently occurs with the oblique Stimulus construction. The verb *vicār karnā* "lit. thought do" which highlights the high degree of agentivity and control of the Experiencer over the mental activity only allows an oblique Stimulus, most frequently featuring the postposition *par* "on" (see section 8.4.2). The preference for the oblique Stimulus construction among verbs

overtly expressing agentivity indicates that decline of the productivity of the transitive construction is evident not only with respect to the semantic classes associated with the loss of Agent's properties (lower chain in Malchukov's transitivity hierarchy) but also with respect to the classes featuring the loss of the Patien's properties (the upper chain).

Croft (1993, 2012, 2022) notes that one of the main characteristics of experiences is that the same event can be expressed with a difference regarding its force dynamic construal. In particular, Croft notes that every event can be schematically represented in terms of how participants act and are acted on one another. This interaction goes under the name of transmission of force from one participant, the initiator, and another participant, the endpoint (Croft 1991: 166-67). Experiential situations may be construed as involving two different causal relations: either with an Experiencer directing his/her attention to the Stimulus, or with the Stimulus altering the mental state of the Experiencer. The opposition between the oblique Stimulus construction and the dative construction mirrors these two opposing conceptualizations of experiential events. The dative construction construes the experiential event as something happening to the Experiencer, in which the Experiencer resembles a Goal and is the target of the event, while the Stimulus or the Expertum are the Source. The oblique Stimulus construction, on the other hand, construes the event as an experiential activity controlled and started by the Experiencer, which is the Source of the event and directs his/her attention toward the target Stimulus. In Figure 34, I propose a schematic summary of the different conceptualization and construal of the same event contributed by the oblique Stimulus construction and the dative Experiencer construction.

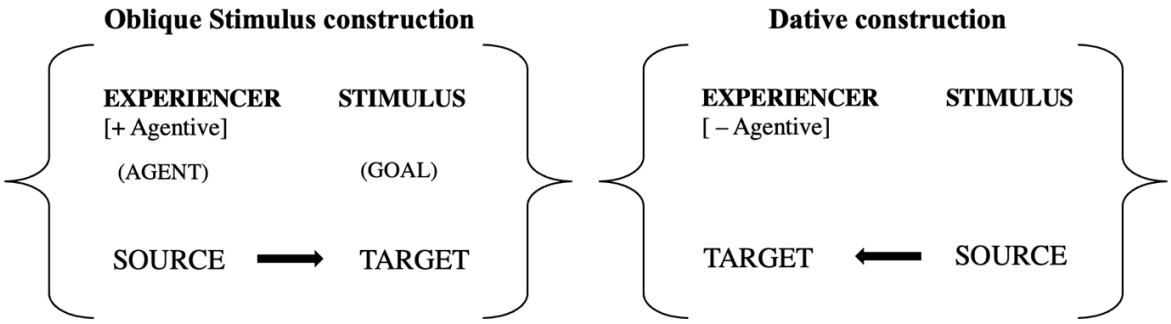

Figure 34: Schematic representation of the different construal contributed by the oblique Stimulus construction and the dative construction in the expression of visual perceptions in Hindi.

Note that, given the agentive interpretation of the oblique Stimulus construction, the alternation in this construction between a nominative and an ergative first argument may be one of the factors that is contributing to entrench the association between ergative case and agentivity in the grammar. Indeed, when the first argument is not agentive, for example with the expression of negative and

uncontrolled emotions such as fear, the first argument does not alternate between the ergative and the nominative, but always occurs in the nominative (as in 435 and 436).

435. पर मैं उससे ऐसा डरा कि सूरत तक न दिखायी।

par	*maiṁ*	*us=se*	*aisā*	*ḍar-ā*	*ki*
but	1SG.NOM	3SG.OBL=INS	so_much	fear-PRF.M.SG	that

sūrat	*tak*	*na*	*dikhā-yī.*
face.F.SG.NOM	till	not	show-PRF.F.SG

"But I feared him so much that I didn't even show my face."

436. वकील साहब ने कई मिनट चुप रहने के बाद कहा, 'मैं मौत से डरता नहीं, टीमल!

vakīl sāhab=ne	*koi*	*minaṭ*	*cup*	*rah-n-e=ke bād*	*kah-ā*
lawyer sir=ERG	some	minute	silent	stay-INF-OBL=after	say-PRF.M.SG

maiṁ	*maut=se*	*ḍar-t-ā*	*nahīṁ, tīmal!"*
1SG.NOM	death=INS	fear-IPRF.M.SG	not Timal

"The lawyer was silent for a minute and then said, 'I am not afraid of death, Timal'".

9.2. Transitive-intransitive (anti)causative alternation

Previous literature has established the cross-linguistic stability of two universal clause types that seem to exist in all languages of the world (Dixon 1994, Dixon and Aikhenvald 2000). These are an intransitive clause consisting of a single core argument, generally labelled S, and a transitive clause with two core arguments, generally labelled A and O. In most languages, a further argument may occur, which typically refers to a Beneficiary or Recipient and which is labelled E. This argument appears in extended versions of the two basic clause types, and results in an extended intransitive construction and an extended transitive construction. The structure of these four construction types is given in Table 61. Typologically, the two extended constructions tend to occur with specific classes of verbs that require a Recipient/Beneficiary-like argument. In particular, the extended transitive construction tends to occur with verbs of giving, showing and telling, while the extended intransitive construction, which is most interesting for the present work, typically occurs with verbs of seeing, hearing, liking and wanting (Dixon 1994: 122), hence with experiential verbs. In these extended intransitives, an argument is added to the (anticausative) intransitive form and is typically marked as the third argument of a ditransitive construction, i.e. as a Recipient (Dixon 1994: 122–4).

Table 61: The four basic construction types (from Dixon 1994).

Intransitive	S
Transitive	A + O
Extended intransitive	S + E
Extended transitive	A + O + E

In this section, I would like to summarize and draw some conclusions on the functions of the (anti)causative system within the expression of experiential events in Hindi, focusing on how this system operates with the semantic properties of experiential situations. As discussed in detail in the chapters dedicated to verbs of perception (7) and verbs of cognition (8), the (anti)causative alternation changes the argument structure of the experiential verbs and/or their semantic characterization, thus resulting in different construals of the same event. The anticausative form typically appears with a dative Experiencer, thus instantiating the extended intransitive clause type. Indeed, extended intransitives are quite frequent in Hindi, and their high productivity is obviously related to the high frequency of non-nominative subjects in the language. Interestingly, in Hindi, the argument extending the intransitive clause can also be marked differently from a Recipient and extended (anticausative) intransitives are not just limited to experiential verbs (as suggested by typological observation made by Dixon 1994), but they span across various semantic classes, including verbs typically associated with transitive events. As sentences 437 and 438 show, for instance, when occurring with prototypical transitive verbs, the anticausative form may appear with an argument marked with the instrumental postposition *se* referring to a participant unintentionally causing the event. The extended intransitive clause type is an expedient used by the language to re-introduce with a non-causal semantics the argument removed by the anticausative alternation. The instrumental construction in 438, for example, is used to construe the event as spontaneously happening and to construe the human participant not as an agent, but as a non-volitional causer. As discussed in the previous chapters, the dative construction with experiential verbs has a similar semantic purpose: it construes the experiential situation as spontaneous and adds an Experiencer, construing it as non-agentive. The Experiencer is typically marked with the dative, but other case-markings are allowed with specific semantic properties (such as the genitive, section 9.1.5.1).

437. *maiṁ =ne* *gilās* *toṛ-ā*
1SG=ERG glass(M.SG.NOM) break-PRF.M.SG
"I broke the glass."

438. *mujh=se* *gilās* *ṭūṭ* *gayā*
1SG.OBL=INS glass(M.SG.NOM) break go.PRF-M.SG
"I broke the glass inadvertently."

In sum, in a language that tends to use case markings iconically and that it is not high in transitivity prominence, the anticausative is used to encode events that are distant from the transitive prototype. It is used to construe the event as a state or as a spontaneously happening inchoative event in which the main participant is not a prototypical Agent. The anticausative intransitive construction is extended by an additional argument which is iconically marked, according to its semantic properties, and may take subjecthood properties.[33] The anticausative alternation thus allows the language to display pairs of cognate verbs for most experiential situations, excluding bodily sensations as they never show agentive Experiencers. See for example: *yād karnā* vs *yād honā* "remember" agentive vs non-agentive, *dekhnā* vs *dikhnā* "see" agentive vs non-agentive, *socnā* vs *sūjhnā* "think" agentive vs non-agentive, *vicār karnā* vs *vicār honā* "think" agentive vs non-agentive, and so on.

9.2.1. Some remarks on the origin of the extended (dative) intransitive construction

In section 9.1.3, I tried to depict the diachronic reasons that led to the spread of the dative construction in NIA languages. In this section, I would like to dive into the origin of the extended dative intransitive construction that spread during the contact with the Persian language. In section 9.1.3, I mentioned that there are many reasons that make it difficult to trace the origin of non-canonical constructions in Hindi. In particular, the main problem lies in the fact that MIA and Early NIA were transitional phases shifting from an inflecting morphology to postpositions, so that the case-marking of semantic roles is often difficult to interpret in MIA and Early NIA texts (Butt 2006b, Butt and Ahmed 2010, Montaut 2013). Montaut (2013: 111), for example, notes that "in Kabir [15th century], non-agents of two place predicates are more frequent in the oblique or locative, a case which has then become extremely syncretic, to such an extent that we find it for animate objects, experiencers, locatives, ablatives, and transitive agents in pre-ergative sentences". However, there are some roots whose evolution from Sanskrit is continuously attested and whose analysis might give us some interesting insights. One of these roots is the Sanskrit verb *ruc-* "shine". As Deshpande (1991) notes (see also Cardona 1991, Montaut 2013, Butt and Deo 2013), the evolution of a verb like *ruc-* in later stages of Sanskrit could be evidence of the process that led to the formation of Experiencers in the dative case in Middle Indo-

[33] Shibatani and Pardeshi (2001) discuss the basic intransitivity of dative constructions in Hindi and advance the hypothesis that these are not variants of double-subject patterns (as previous analyses of non-canonical constructions in other languages assume), as according to them only one noun phrase is a lexically selected argument, while the other is sanctioned by a clausal predicate.

Aryan languages and later in New Indo-Aryan. The development of the syntactic construction occurring with this root unfolds through two main phases. In an initial phase, the root *ruc-* expresses a spontaneous event of appearance and occurred with a single nominative argument, that is the thing that shines. This single argument construction most probably allowed the addition of a Recipient/Beneficiary in the dative case. However, this dative element was an adjunct within the event, and it was not conceived as an argument. The increasingly frequent occurrence of this verb with a Recipient/Beneficiary adjunct led to a reanalysis of its semantics, which shifted to the meaning "attract, draw interest". Consequently, the dative element was reanalyzed as an Experiencer argument. Since in this new configuration the Experiencer is the most salient argument, it acquires subjecthood syntactic properties (Lehmann 2002, Comrie 1981). Hence, the verb is recategorized as a non-nominative subject verb. The same root still exists in the Hindi verb *rucnā* "be pleasant, be desirable, be tasty" (McGregor 1994) and in its derivations, such as the noun *ruci* "light, beauty, interest, inclination, desire" and the adjective *rocan* "luminous, beautiful, desirable, charming". It is worth noting that only the derivatives, not the verb itself, have retained meanings related to the semantics of light in Hindi, while the verb, which today is used very rarely, has lost its original semantics of "shining" and only means "like, be interested in" (as in 439).

439. मैं जब तक उनका भोजन न बनाऊँ, उन्हें कोई चीज रुचती ही न थी?

maiṁ	*jab=tak*	*un=kā*	*bhojan*	*na*
1SG.NOM	when=till	3PL.OBL=GEN	food(M.SG.NOM)	not
banā-ūṁ	*unheṁ*	*koī*	*cīj*	*ruc-t-ī*
cook-1SG.SBJV	3PL.DAT	INDF.PRN	thing.F.SG.NOM	be_tasty-IPRF-F.SG
hī	*na*	*th-ī?*		
EMPH	not	be.PST-F.SG		

"Until I prepare their meal, didn't they like anything?"

The verb *bhānā* "please", from the Sanskrit root *bhā* "shine", was characterized by the same semantic shift and the same constructional change. Montaut (2013: 101) notes that in the 14th century this verb was clearly reanalyzed as bearing the meaning "please" and recategorized as a dative subject verb, in which an original beneficiary adjunct has become an Experiencer argument. Both the verb *rucnā* and the verb *bhānā* in modern Hindi have almost disappeared. The analysis of these roots supports the hypothesis of the metaphorical extension from Recipient/Beneficiary to Experiencer that I discussed in section 9.1.3 above in this chapter.

10. Conclusion

Table 62 shows the distribution of the main constructions analyzed in this study across the verbs belonging to the experiential subdomains of bodily sensations, perceptions, and cognitions discussed in the previous chapters. The domain of cognition shows the widest constructional variation: this is not surprising given that it also shows the widest semantic variation. However, as is evident from the table, constructions tend to be lexically specified and individual verbs rarely show alternation between two constructions.

Table 62: Distribution of the main constructions analyzed in this study over the experiential subdomains. (The brackets indicate that the construction exhibits a low frequency in the corpus.)

Meaning	**Hindi verb**	**Constructions**					
		Dat	**Tran**	**Obl Stim**	**Gen**	**Loc**	**Copular**
BODILY SENSATIONS							
Be hungry	*bhūkh honā*	✓					
	bhūkh lagnā	✓					
	bhūkhā honā						✓
Be thirsty	*pyās honā*	✓					
	pyās lagnā	✓					
	pyāsā honā						✓
Be hot	*garmī honā*	✓					
	garmī lagnā	✓					
Be cold	*ṭhaṇḍ honā*	✓					
	ṭhaṇḍ lagnā	✓					
Be in pain	*dard honā*	✓				✓	
	dard karnā		(✓)				
Perceptions							
Appear/see	*dikhnā*	✓					
	dikhāī denā	✓					
	dikhāī paṛnā	✓					
	najar ānā	✓					
See/look at	*dekhnā*		✓	(✓)			
Look at/stare	*tāknā*		(✓)	✓			

Be heard	*sunāī denā*	✓					
Be heard	*sunāī paṛnā*	✓					
Hear/listen	*sunnā*		✓				
Touch	*chūnā*		✓				
	sparś karnā		✓				
Taste	*cakhnā*		✓				
	svād lenā		✓				
	svād milnā	✓					
Smell	*khuśbū ānā*	✓					
Touch/smell/taste	*lagnā*	✓					
COGNITIONS							
Think	*socnā*		✓				
	sūjhnā	✓					
Understand	*samajhnā*		(✓)				
Know	*jānnā*		✓				
Forget	*bhūlnā*	✓	(✓)				
Think	*vicār karnā*			✓			
	vicār honā	(✓)			✓		
	vicār ānā	✓					
	khayāl karnā			✓			
	khayāl honā	(✓)			✓		
	khayāl ānā	✓					
Understand	*samajh ānā*	✓				(✓)	
Know	*jān paṛnā*	✓					
	jñāt honā	✓					
	malūm honā	✓					
	patā	✓	✓				
Remember	*yād*	✓	✓				

The focus on argument structure construction in this dissertation did not allow me to touch upon a number of other issues and to adequately discuss other phenomena that remained in the background.

For example, I limited my analysis to the constructional investigation of three experiential subdomains, leaving aside the expression of emotions and volitions in Hindi. The analysis of these two subdomains may provide further insights on the semantic-syntactic interplay in the language.

Moreover, one of the most important issues I couldn't delve into is the role of the verbal aspect and its relation to lexical aspect in many experiential verbs. In particular, while addressing the aspectual characterization of some verbs, I mainly focused on the macro distinction between perfective and imperfective and left in the background the discussion of more specific aspectual characterizations. For example, I did not focus on a striking correlation that seems to exist between the progressive aspect and verbs that are generally classified as states in the literature, such as "be hungry" in 440 and "see" in 441. This is interesting since the progressive aspect is generally associated to durative dynamic events, and therefore shouldn't lend itself to the expression of states. I believe that an analysis of the use of the progressive aspect might give us interesting insights on the semantics of this verb form in Hindi.

440. बन्दी उठ खड़ी हुई, 'हमें तो प्यास लग रही है।

bandī uṭh khad-ī hu-ī, 'hameṁ to pyās lag rah-ī hai'.
the prisoner stood up 1PL.DAT EMPH thirst(F.SG.NOM) attach PRG-F be.3SG.PRS
"Then the prisoner stood up, 'We are feeling thirsty'."

441. उसका शरीर एक लम्बी चौड़ी चादर से का हुआ था, जिससे उसका मुंह भी छिप गया था केवल दो आंखें दिखाई दे रही थीं।

uskā śarīr ek lambī cauṛī cādar se kā huā thā, jisse uskā muṁh bhī chip gayā thā
His body was covered with a long wide sheet, which hid his face
(use) keval do āṁkh-eṁ dikhāī de rah-ī th-īṁ,
3SG.DAT only two eyes-F.PL.NOM seeing(F) give PRG-F.PL be.PST-F.PL
"His body was covered with a long wide sheet, which hid his face, he could see only two eyes/ only two eyes were visible."

Turning back to the analysis of arguments structures, more insights on the interplay between semantics and syntax in Hindi could be gained from typological observations, in particular from a comparison between SAE languages and South Asian languages. One of the most distinctive features of SAE languages is that experiential events are very frequently expressed by the same construction used for the expression of possessive meanings (Haspelmath 1998). For example, in English, the *have*-construction used to encode prototypical possessive relations, such as ownerships or kinship, is

also used to express experiences, such as *I have a splitting headache* or *I had an* idea. Fedriani (2012: 108) argues that the extension of the have possessive construction for the expression of experiential events is favored by the syntactic rules of the European languages. In particular, while discussing the spread of the *habeo*-construction in Latin (and, later, in Romance languages), she points out that the *have*-construction "models the possessive situation by accommodating it to the syntax of two-participant events with a nominative subject and an accusative object, thus conforming to the general activity schema" (Fedriani 2012: 113, see on this also Heine 1997). In contrast, modern Indo-Aryan languages lack an *have*-verb and rarely use possessive constructions for the expression of experiential events. Further studies might shed light on the relation between the function of possessive constructions and the encoding strategies preferences in these two linguistic areas. In particular, it is reasonable to conjecture that the lack of a *have*-verb, the pervasive use of dative constructions and the fact that possessive constructions are not systematically extended to the expression of experiential events in Hindi are all connected and can be accounted for once we consider the high iconicity of the language.

References

Abbi, Anvita. 1990. Experiential Constructions and the 'Subjecthood' of the Experiencer NPS in South Asian languages. In Verma M and Mohanan K., *Experiencer Subjects in South Asian Languages*, pp. 253-269. Stanford: CSLI.

Abbi, Anvita, Gopalakrishnan, Devi. 1991. Semantics of explicator compound verbs in south Asian languages. *Language Sciences*, vol. 13/2: pp. 161-180.

Ackerman, Farrell and Philip LeSourd. 1996. Toward a Lexical Representation of Phrasal Predicates. In A. Alsina, J. Bresnan and P. Sells (eds.) *Complex Predicates*. Stanford: CSLI Publications.

Ahmed, Tafseer, 2006. Spatial, temporal and strucutural usages of Urdu *ko*. In Miriam Butt and Tracy Holloway King (eds.) *Proceedings of the LFG06 Conference*, University of Konstanz. CSLI Publications.

Ahmed, Tafseer. 2009. *Spatial expressions and case in South Asian languages*, PhD dissertation, University of Konstanz.

Ahmed, Tafseer, Butt, Miriam. 2011. Discovering Semantic Classes for Urdu N-V Complex Predicates, in *International Conference on Computational Semantics (IWCS 2011)*.

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford Linguistics.

Aikhenvald, Alexandra Y., Dixon, Robert M. W. and Onishi, Masayuki (eds.) 2001. *Non-canonical Marking of Subjects and Objects.* Amsterdam: John Benjamins Publishing Company.

Aissen, Judith, 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21: pp. 435–483.

Aldai, Gontzal and Søren Wichmann. 2018. Statistical observations on hierarchies of transitivity. *Folia Linguistica* 52(2): pp. 249–281.

Alsina, Alex. 1996. *The Role of Argument Structure in Grammar: Evidence from Romance*. Stanford, CA:CSLI Publications.

Amberber, Mengistu. 2007. Introduction. The Language of Memory. In: Mengistu Amberber (ed.) *The language of memory in crosslinguistic perspective*, pp. 1–12. Amsterdam: John Benjamins.

Amritavalli, R. 2004. Experiencer datives in Kannada. In Bhaskararo Peri and Subbarao Karumuri V. (eds.) *Non-nominative subjects (Vol 1)*, pp. 1-24. Amsterdam/Philadelphia: John Benjamins Publishing.

Athanasiadou, A. and Tabakowska E. (eds.) 2010. *Speaking of Emotions. Conceptualisation and Expression*, Berlin: Mouton de Gruyter.

Barcelona, A. 1989. Análisis contrastivo del léxico figurado de la ira en inglés y en español. *Actas del VI Congreso Nacional de Lingüística Aplicada* (AESLA), Santender: Universidad de Cantabria, Spain: pp. 141-148.

Barðdal, Johanna. 1999. Case in Icelandic, A Construction Grammar Approach, *TijdSchrift voor Skandinavistiek*, vol. 20, n 2, pp. 66-99.

Barðdal, Jóhanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins.

Barðdal, Jóhanna. 2009. "The Development of Case in Germanic". In Johanna Barðdal and Shobhana Chelliah (eds.) *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, pp. 123-159 Amsterdam: Benjamins.

Barðdal, Jóhanna. 2011. The rise of Dative Substitution in the history of Icelandic: A diachronic construction grammar account. *Lingua* (121), pp. 60-79.

Barlow, Michael and Suzanne Kemmer (eds.). 2000. *Usage-based models of language*. Stanford: Center for the Study of Language and Information.

Bashir, Elena. 1993. Causal Chains and Compound Verbs. In Manindra Verma (ed.), *Complex Predicates in South Asian Languages*, pp. 220-230. Delhi: Motilal Banarsidass.

Bashir, Elena. 1999. The Urdu and Hindi Ergative Postposition ne: Its changing role in the grammar. In *The Yearbook of South Asian Languages and Linguistics*, pp. 11-36. Berlin, New York: Mouton De Gruyter.

Beames, John. 1970 [1875]. *A Comparative Grammar of the Modern Aryan Languages of India—to wit Hindi, Punjabi, Sindhi, Gujarati, Marathi, Oriya and Bengali (Vol. II. The Noun and Pronoun*). London: Trübner.

Bertinetto, Pier Marco and Denis Delfitto. 2000. Aspect vs. Actionality why they should be kept apart. In Dahl Oesten (ed.) *Tense and Aspect in the Languages of Europe*, pp. 189–226. Berlin: Mouton de Gruyter.

Bhaskararao, Peri and Subbarao, Karumi, Venkata Nageswara Rao (eds). 2004. *Non-nominative Subjects*, John Benjamins Publishing Company.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23, pp. 757–807.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

Bickel, Balthasar. 2004. The syntax of experiencers in the Himalayas. In Bhaskararo Peri and Subbarao Karumuri V. (eds) *Non-nominative subjects (Vol 1)*, pp. 77-112. Amsterdam / Philadelphia: John Benjamins Publishing.

Blansitt, Edward. 1978. Stimulus as a semantic role, in W. Abraham (ed.), *Valence, semantic case and grammatical relations,* pp. 311–325. Amsterdam: John Benjamins

Blansitt Jr., Edward L. 1988. Datives and allatives. In Michael Hammond, Edith Moravcsik and Jessica Wirth (eds*), Studies in Syntactic Typology*, pp. 173–191. Amsterdam / Philadelphia: Benajamins.

Bolinger, Dwight L. 1968. Entailment and the meaning of structures. *Glossa* 2.2: pp. 119-127.

Bossong, Georg. 1985. *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bossong, G. 1998. Le Marquage de l'Expérient dans les Langues d'Europe, in Feuillet J. (ed.) *Actance et Valence dans les Langues d'Europe*, pp. 259-294. Berlin: Mouton de Gruyter.

Broccias, Cristiano. 2001. Allative and ablative at-constructions. In M. Andronis, C. Ball, H. Elston, and S. Neuvel (eds.), CLS 37: *The main session. Papers from the 37th meeting of the Chicago Linguistic Society* (Vol. 1), pp. 67–82. Chicago: Chicago Linguistic Society.

Bugenhagen, Robert D. 2001. Emotions and the nature of persons in Mbula. In *Emotions in crosslinguistic perspective,* J. Harkins and A. Wierzbicka (eds.), pp. 69-114. Berlin, New York: de Gruyter.

Butt, Miriam, 1993. Object specificity and agreement in Hindi/Urdu. *Papers from the 29th Regional Meeting of the Chicago Linguistic Society* (1), pp. 80–103.

Butt, Miriam, 1995. *The Structure of Complex Predicates in Urdu.* Stanford: CSLI Publications.

Butt, Miriam, 2001. Case, Agreement, Pronoun Incorporation and Pro-Drop in South Asian Languages.

Butt, Miriam. 2002. The morpheme that wouldn't go away, talk presented at the *Workshop on Pertinacity*, Shloss Freudental.

Butt, Miriam, 2006a. *Theories of Case.* Cambridge: Cambridge University Press.

Butt, Miriam. 2006b. The Dative-Ergative connection. In O. Bonami, P. Hofherr (eds.) *Empirical Issues in Syntax and Semantics (6)*, pp. 69-92. Paris: CSSP.

Butt, Miriam. 2014. Control vs. complex predication: Identifying non-fnite complements. *Natural Language and Linguistic Theory* 32, pp. 165–190.

Butt, Miriam, King, Tracy H. 1991. Semantic case in Urdu. In Dobrin, L., L. Nichols, and R.M. Rodriguez (eds.), *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*: pp. 31–45.

Butt, Miriam, Tracy Holloway King. 1997. Null Elements in Discourse Structure. In (K.V. Subbarao) *Papers from the NULLS Seminar*. Delhi: Moti Lal Banarsi Das.

Butt, Miriam and Geuder, Wilhelm. 2001. On the (semi)lexical status of light verbs. In N. Corver and H. van Riemsdijk, (eds.), *Semi-lexical Categories: On the content of function words and the function of content words*, pp. 323–370. Berlin, Boston: Mouton de Gruyter.

Butt, Miriam and King Tracy H. 2002. Case systems: Beyond structural distinctions. In Brandner, Ellen and Heike Zinsmeister (eds.), *New Perspectives on Case Theory*, pp. 53–87. Stanford: CSLI Publications.

Butt, Miriam and King, Tracy H. 2004. The status of case. In V. Dayal and A. Mahajan (eds.), *Clause Structure in South Asian Languages,* pp. 53–198. Berlin: SpringerVerlag.

Butt, Miriam and Jafar Rizvi. 2010. Tense and Aspect in Urdu. In Patricia Cabredo-Hofherr and Brenda Laca (eds.) *Layers of Aspect*. Stanford, CA: CSLI Publications.

Butt, Miriam and Ahmed, Tafseer. 2010. The redevelopment of Indo-Aryan case systems from a lexical semantic perspective. *Morphology, 21:* pp. 545–572.

Butt, Miriam and Gillian Ramchand. 2003. Complex Aspectual Structure in Hindi/Urdu. In Nomi Ertishik-Shir and Tova Rappaport (eds.) *The Syntax of Aspect*. Oxford: Oxford University Press.

Butt, Miriam and Tracy Holloway King. 2006 . Restriction for morphological valency alternations: The Urdu causative. In Butt, Miriam, et al. (eds.) *Intelligent Linguistic Architectures : Variations on Themes by Ronald M. Kaplan. Stanford*, pp. 235-258. California: CSLI Publications.

Butt, Miriam, Scott Grimm and Tafseer Ahmed. 2006. Dative Subjects. Paper presented at the NWO/DFG *Workshop on Optimal Sentence Processing*. Nijmegen, June 2006.

Butt, Miriam and Ashwini Deo. 2013. A Historical Perspective of Dative Subjects in Indo-Aryan. Paper presented at the *2013 Lexical Functional Grammar Conference*, Debrecen, July 2013.

Butt, Miriam, Carnesale Lucrezia and Ahmed Tafseer. 2023. Experiencers vs. Agents in Urdu/Hindi Nominalized Verbs of Perception. In *Proceedings of the 28th Annual Lexical-Functional Grammar Conference (LFG-23)*, Vol. 28. University of Rochester, 22-24 July 2023.

Bybee, Joan. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: Benjamins.

Bybee, Joan and Perkins, Revere, Pagliuca, William. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: The University of Chicago Press.

Bybee, Joan and Sandra Thompson. 1997. Three frequency effects in syntax. *Berkeley Linguistic Society* 23: pp. 378–388.

Bybee, Joan and Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In Joan Bybee and Paul Hopper (eds*.), Frequency and the emergence of linguistic structure*, pp. 1–24. Amsterdam: Benjamins.

Caracchi Pinuccia. 1992. *Grammatica hindi*. Torino: Magnanelli Edizioni.

Cardona, George. 1991. A Note on 'Dative Agents in Sanskrit'. In Verma M.K and Mohanan K.P. (eds.) *Experiencer Subjects in South asian languages*, pp. 141-145. Stanford: CSLI Publications.

Carnesale, Lucrezia. 2022. Predicative Possessive Constructions in Hindi, *Bhasha*, 1(1): pp. 133-172.

Chakraborty, Jayshree. 1992. Perfectivity and the resultative state in Hind. In *South Asian Language Review*, Volume II (1), pp. 55-67.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects and the Theory of Syntax*. Cambridge: MIT Press.

Chung, Sandra and Timberlake, Alan. 1985. Tense, Aspect and Mood. In Timothy Shopen (ed.) *Language typology and syntactic description*, pp. 241–258. Cambridge: Cambridge University Press.

Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1978. Ergativity. In Lehmann W.P. (ed.) *Syntactic typology. Studies in the phenomenology of language*, pp. 329-393. Sussex: The Harvester Press.

Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*: *Syntax and Morphology* (2 ed.). Oxford: Blackwell.

Comrie, Bernard. 1989. *Language universals and linguistic typology*. Chicago, IL: University of Chicago Press.

Consolaro, Alessandra. 2003. *Madre India e la parola. La lingua hindi nelle università*. Torino: Edizioni dell'Orso.

Colman, Andrew (ed.). 1994. *Companion encyclopedia of psychology*. Vol.1. London: Routledge.

Croft, William. 1985. Indirect object 'lowering.' In Mary Niepokuj et al. (ed) *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, pp. 39–51. Berkeley: Berkeley Linguistics Society.

Croft, William. 1991. *Syntactic Categories and Grammatical Relations. The cognitive organization of Information*. Chicago: The University of Chicago Press.

Croft, William. 1993. Case marking and the semantics of mental verbs. In James Pustejovsky (ed.), *Semantics and the Lexicon*, pp. 55–72. Boston and London: Kluver Academic Publishers.

Croft, William. 1998. Event Structure in Argument Linking. In Miriam Butt and Wilhelm Geuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, pp. 21–63. Stanford: CSLI Publications.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, William 2003. *Typology and Universals*. 2nd Edition. Cambridge: Cambridge University Press.

Croft, William. 2012. *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.

Croft, William 2022. *Morphosyntax*. Cambridge: Cambridge University Press.

Croft, William and D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

Dahl, Eystein. 2009. Some semantic and pragmatic aspects of object alternation in Early Vedic. The role of semantic, pragmatic and discourse factors in the development of case, in Jóhanna Barðdal and Shobhana L. Chelliah (eds.) *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, pp.23-55. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Dahl, Eystein. 2014. The morphosyntax of the Experiencer in Vedic. In Luraghi Silvia and Narrog Heiko (eds.) *Perspective on Semantic Roles*, pp 181-204. Amsterdam/Philadelphia: John Benjamins Publishing.

Dahl, Eystein. 2022. Alignment and alignment change in the Indo-European family and beyond. In Dahl Eystein (ed.) *Alignment and alignment change in the Indo-European family*, pp. 1-23. Oxford: Oxford University Press.

Dahl, Eystein and Fedriani, Chiara 2012. The Argument Structure of Experience: Experiential Construction in Early Vedic, Homeric Greek and Old Latin, *Transactions of the Philological Society* 110(3), pp. 342-362.

Dahl, Östen 1990. Standard Average European as an exotic language, in: Bechert, Johannes, Bernini, Giuliano and Buridant, Claude (eds.), *Toward a typology of European languages*, pp. 3-8. Berlin: Mouton de Gruyter.

Dasgupta, Probal. 2004. Some non-nominative subjects in Bangla. In Bhaskararo Peri and Subbarao Karumuri V. (eds) *Non-nominative subjects (Vol 1)*, pp. 130-153. Amsterdam / Philadelphia: John Benjamins Publishing.

Davison, Alice. 2005. Phrasal predicates: How n combines with V in Hindi/Urdu. In Tanmoy Bhattacharya (ed.) *Yearbook of South Asian Languages and Linguistics*, pp. 83–116. Amsterdam/NewYork: Mouton DeGruyter.

Davison, Alice. 1999. Ergativity: Functional and Formal Issues, in M. Darnell, E.A. Moravcsik, F. Newmeyer, M. Noonan, K. Wheatley (eds.) *Functionalism and Formalism in Linguistics*. *Volume I: General Papers*, pp. 177–208. Amsterdam: John Benjamins.

Deo, Ashwini. 2006. *Tense and Aspect in Indo-Aryan Languages: Variation and Diachrony*. Ph.D. thesis, Stanford University.

Dik, Simon Cornelis. 1978. *Functional Grammar*. Amsterdam: North-Holland Publishing Company.

Dik, Simon Cornelis. 1981. *Functional Grammar*. Dordrecht: Foris.

Dik, Simon Cornelis and Kees Hengeveld. 1991. The Hierarchical Structure of the Clause and the Typology of Perception-verb Complements. *Linguistics* 29: pp. 231–259.

Dixon, Robert M.W. 1979. Ergativity. In *Language*, 55: pp. 59–138.

Dixon, Robert M.W. 1984. The semantic basis of syntactic properties. *Berkeley Linguistics Society, Proceedings* 10: pp. 583-595.

Dixon, R.M.W. 1991. *A new approach to English grammar, on semantic principles*. Oxford: Oxford University Press.

Dixon, Robert M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Donohue, Mark and Wichmann Soren (eds) 2008. *The Typology of Semantic Alignment.* Oxford: Oxford University Press.

Dowty, David R. 1978. Lexically governed transformations as lexical rules in a Montague Grammar. *Linguistic Inquiry* 9, pp. 393–426.

Dowty, David R. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language*, 67(3), pp. 547–619. DOI: 10.1353/lan.1991.0021.

Drocco, Andrea. 2008. *L'ergatività in hindi. Studio diacronico del processo di diffusione della posposizione 'ne'*. Alessandria: Edizioni dell'Orso.

Drocco, Andrea. 2018. An attempt to understand the encoding of reduced transitivity in Hindi: the case of compound verbs with jānā. In *Trends in Hindi Linguistics,* pp. 265-284. Berlin, New York: Mouton de Gruyter.

Drocco, Andrea. 2019. In Search of the Roots of Khaṛī-bolī Hindī. *International Journal of Language and Linguistics Vol. 6, No. 2,* pp. 112-120. doi:10.30845/ijll.v6n2p15.

Drocco, Andrea. Tiwari, Neha. 2020a. Atypical Compound Verb Constructions in Hindi/Urdu. The case of transitivity Polar Verbs with Intransitive Vector Verbs. *Annali di Ca'Foscari. Serie Orientale*, 56, pp. 307-334.

Drocco, Andrea. Tiwari, Neha. 2020b. A Pragmatic Approach to Compound Verbs in Hindi/Urdu: The case of (Inter)subjectivity. *Linguistica e Filologia*, 40, pp. 157-195.

Ekman, Paul. 1992. *An argument for basic emotions*. Cognition and Emotion 6(3/4), pp. 169-200.

Ekman, Paul and Richard J. Davidson (eds). 1994. *The Nature of Emotion: Fundamental Questions*. Oxford: Oxford University Press.

Emanatian, Michele. 1995. Metaphor and the expression of emotion: the value of cross-cultural perspectives. *Metaphor and Symbolic Activity* 10, pp. 163–182.

Evans, Nicholas and David Wilkins. 2000. In the Mind's Ear: the Semantic Extension of Perception verbs in Australian Languages. *Language* 76(3), pp. 546–592.

Evans, Nicholas and David Wilkins. 1998. The knowing ear: An Australian test of universal claims about the semantic structure of sensory verbs and their extension into the domain of cognition. *Cologne, Institut fur Sprachwissenschaft: Arbeitspapier 32 (Neue Folge).*

Fedriani, Chiara. 2012. *Experiential Constructions in Latin*, Leiden and Boston: Brill.

Filip, H. 1996. Psychological predicates and the syntax-semantics interface. In Goldberg A. (ed.) *Conceptual Structure, Discourse, and Language*, pp. 131-147. San Diego: CSLI Publications.

Fillmore, Charles J. 1976. Frame semantics and the nature of language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, 280: pp. 20-32.

Fillmore, Charles J. 1977a. Scenes-and-frames semantics. In Antonio Zampolli (ed.), *Linguistic Structures Processing*. Amsterdam: North Holland Publishing.

Fillmore, Charles J. 1977b. The need for a frame semantics in linguistics. In Hans Karlgren (ed.), *Statistical Methods in Linguistics*. Stockholm: Scriptor.

Fillmore, Charles J. 1977c. The case for case reopened. In P. Cole and J. M. Sadock (eds), *Syntax and Semantics (VIII) Grammatical Relations*, pp. 59-81. New York: Academic.

Fillmore, Charles J. 1982. Frame Semantics. In *Linguistics in the Morning Calm: Selected Papers from SICOL.1981*, pp. 111-37. Seoul: Hanshin.

Fillmore, Charles J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, 6, pp. 222-254.

Fillmore, Charles J. 1986. Pragmatically controlled zero anaphora. In *Berkely Linguistic Society*, 12, pp. 95-107.

Fillmore, Charles J. 1988. The Mechanisms of "Construction Grammar" in *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Societ*y, pp. 35-55.

Fillmore, Charles J., Paul Kay and Mary Kay O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: the case of *let alone*. *Language* 64, pp. 501-38.

Fisher, Cynthia, Hall, Geoffrey, Rakowitz, Susan and Gleitman Lila. 1991. When it Is Better to Receive than to Give: Syntactic and Conceptual Constraints on Vocabulary Growth. In IRCS Report, pp. 91 – 41. Philadelphia: University of Pennsylvania.

Fillmore, Charles J. and Paul Kay. 1993. *Construction grammar coursebook*. Berkeley: University of California.

Fillmore, Charles J. and Collin F. Baker. 2001. Frame Semantics for Text Understanding. In *Proceedings of WordNet and Other Lexical Resources Workshop*, Pittsburgh. NAACL.

Foley, William A. and Robert D. Van Valin. 1984. *Functional Syntax and Universal Grammar.* Cambridge: Cambridge University Press.

Fortescue, Michael. 2001. Thoughts about thought, *Cognitive Linguistics* 12 (1), pp. 15-45.

Gibbs, Raymond, W. Jr. 1994. *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge: Cambridge University Press. doi:10.2277/0521419654.

Gibbs, Raymond W. Jr., Lima Paula Lenz Costa and Francozo Edson. 2004. Metaphor is grounded in embodied experience. *Journal of Pragmatics* 36, pp. 1189–1210.

Givón, Talmy. 1991. Isomorphism in the Grammatical Code: Cognitive and Biological Considerations. *Studies in Language*, 15-1, pp. 85-114.

Goddard, Cliff. 2003. Thinking Across Languages and Cultures: Six Dimensions of Variation. *Cognitive linguistics* 14 (2-3), pp. 109-140.

Goddard, Cliff. 2007. A 'lexicographic portrait' of forgetting. In Mengistu Amberber (ed.) *The language of memory in crosslinguistic perspective*, pp. 119-138. Amsterdam: John Benjamins.

Goddard, Cliff and Wierzbicka, Anna. 2007. NSM analyses of the semantics of physical qualities: Sweet, hot, hard, heavy, rough, sharp in cross-linguistic perspective. *Studies in Language* 34(1), pp. 675–800.

Goldberg, Adele E. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele. E. 1996. Words by Default: Optimizing Constraints and the Persian Complex Predicate. In *Proceedings of the Twenty-Second Annual Meeting of the Berkeley Linguistics Society.*

Goldberg, Adele. E. 1999. The emergence of the semantics of argument structure constructions. In B. MacWhinney (ed.), *The emergence of language*, pp. 197–212. Mahwah, NJ: Lawrence Erlbaum.

Goldberg, Adele E. 2003. Words by default: inheritance and the Persian complex predicate construction, in E. Francis and L. Michaelis (eds.), *Mismatch: Form-Function Incongruity and the Architecture of Grammar*, pp. 84–112. Stanford, CA: CSLI Publications.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Grierson, G. A. 1904. *Linguistic Survey of India. Vol. VI. Indo-Aryan Languages, Mediate Group*, Delhi: Low Price Publications.

Grierson, G. A. 1919. *Linguistic Survey of India. Vol. IX. Indo-Aryan Languages, Central Group*. Delhi: Low Price Publications.

Gries, S. 2003. Towards a Corpus based Identification of Prototypical Instances of Constructions. *Annual Review of Cognitive Linguistics*, I, pp. 1-27.

Gries S. T., Stefanowitsch A. 2006. *Corpora in cognitive linguistics : corpus-based approaches to syntax and lexis.* Berlin. New York : Mouton de Gruyter.

Guerrero Medina, P. 2011. An antipassive interpretation of the English "conative alternation": Semantic and discourse-pragmatic dimensions. In P. Guerrero Medina (ed.), *Morphosyntactic Alternations in English: Functional and Cognitive Perspectives*, pp. 182–203. London: Equinox.

Haig, Geoffrey. 2009. Non-canonical Subjects and Alignment Splits: Where's the Connection? Paper presented at the Workshop on Reconstructing Alignment Systems, Bergen, May 14–15, 2009.

Hare, M, McRae, K., and Elman, J. L. 2003. Sense and structure: meaning as a determinant of verb subcategorization preferences, *Journal of Memory and Language*, 48, pp. 281–303.

Hare, M, McRae, K., and Elman, J. L. 2004. Admitting that admitting verb sense into corpus analyses makes sense, *Language and Cognitive Processes*, 19, pp. 181–224.

Harkins, J. and Wierzbicka, Anna. (eds.) 2001. *Emotions in crosslinguistic perspective* [Cognitive Linguistics Research, 17]. Berlin / New York: de Gruyter.

Hartmann, Iren, Martin Haspelmath and Bradley Taylor. 2013. *Valency Patterns Leipzig: A database of valency information in 35 languages around the world*.

Harris Zellig. 1970. Distributional structures. In *Papers in structural and transformations Linguistics*, pp. 775-794.

Haspelmath, Martin. 1998. How young is Standard Average European? *Language Sciences* (209), pp. 271-287. 10.1016/S0388-0001(98)00004-7.

Haspelmath, Martin. 1999. External possession in a European areal perspective. In, Doris L. Payne and Immanuel Barshi (eds.) *External Possession*, pp. 109–35. Amsterdam / Philadelphia: John Benjamins.

Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. In *2. Halbband Language Typology and Language Universals 2.Teilband,* pp. 1492-1510. Berlin / Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110194265-044

Haspelmath, Martin 2003. The Geometry of Grammatical Meaning: Semantic Maps and Cross-Linguistic Comparison. In Michael Tomasello (ed.), *The New Psychology of Language*, Vol. II, pp. 211–242. Mahwah, NJ: Lawrence Erlbaum.

Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization, in Olga Fischer, Muriel Norde and Harry Perridon (eds.) *Up and down the Cline – The Nature of Grammaticalization*, pp. 17–44. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Haspelmath, M. 2015. Transitivity prominence. In Malchukov, Andrej L and Comrie, B. (eds.) *Valency classes in the world's languages*, pp. 131-147. Berlin: De Gruyter Mouton.

Haspelmath, Martin 2022. Valency and voice constructions. https://www.semanticscholar.org/paper/Valency-and-voice-constructions-Haspelmath/c9bc40ca5d43cb4811dbd1e3fcb07cd435abed7b

Haspelmath, Martin, König, Ekkehard 1995. *Converbs in Cross-Linguistic Perspective, Structure and Meaning of Adverbial Verb Forms - Adverbial Participles, Gerunds.* Berlin: Mouton de Gruyter Mouton. https://doi.org/10.1515/9783110884463.

Heine, Bernd. 1997. *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd. 2001. Ways of explaining possession. In Baron, I. et al (eds.), *Dimensions of Possession*, pp. 311-328. Amsterdam-Philadelphia: John Benjamins.

Hensel, Herbert. 1981. *Thermoreception and Temperature Regulation*. London: Academic Press.

Hettrich, Heinrich. 2014. Some remarks on the adverbal genitive in Ṛigvedic Sanskrit, in Jared Klein and Elizabeth Tucker (eds.) *Vedic and Sanskrit Historical Linguistics*, pp. 129–152. New Delhi: Motilal Banarshidas.

Hewson, John and Vit Bubenik. 2006. *From Case to Adposition: The Development of Configurational Syntax in Indo-European Languages.* Amsterdam/Philadelphia: John Benjamins Publications.

Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.

Hilpert, Martin. 2014. *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.

Hock, Hans, Henrich. 1990. Oblique subject in Sanskrit? In Verma, Manindra and Mohanan *Experiencer Subjects in South Asian Languages*. Stanford: CSLI Publications.

Hock, Hans, Henrich. 1991. Possessive Agents in Sanskrit? In Hock, Hans, Henrich (ed.) *Studies in Sanskrit Syntax*, pp. 55-70. Delhi: Motilal Banarsidass.

Hock, Hans, Henrich. 2016a. Oblique experiencers and oblique subjects. In Hock, Hans, Henrich and Elena Bashir (eds.) *The languages and linguistics of South Asia. A comprehensive guide*, pp. 544-549. Berlin/Boston: De Gruyter Mouton.

Hock, Henrich Hans. 2016b. Finite and nonfinite subordination. In Hock and Bashir (eds.), *The Languages and Linguistics of South Asia. A Comprehensive Guide,* pp. 567- 584. Amsterdam: John Benjamins

Hock, Henrich H., Bashir, Elena (eds.) 2016. *The Languages and Linguistics of South Asia. A Comprehensive Guide*. Amsterdam: John Benjamins.

Hook, Peter E. 1974. *The Compound Verb in Hindi*. Center for South and Southeast Asian Studies: The University of Michigan.

Hook, Peter E. 1991. The Emergence of Perfective Aspect in Indo-Aryan Languages. In E. Traugott and B. Heine (eds.), *Approaches to Grammaticalization, volume II: Types of grammatical markers*, pp. 59–89. Amsterdam: John Benjamins.

Hook, Peter E. 1993. Aspectogenesis and the Compound Verb in Indo-Aryan. In M. Verma (ed.), *Complex Predicates in South Asian Languages*, pp. 97–113. Delhi: Manohar Publishers & Distributors.

de Hoop, Helene and Narasimhan, Bhuvana. 2005. Differential case marking in Hindi. In Amberber M. – de Hoop H. (eds.) *Competition and Variation in Natural Languages: The Case for Case*, pp. 334-359. Amsterdam: Elsevier.

Hopper, Paul J., Thompson and Sandra. A. 1980. Transitivity in Grammar and Discourse. *Language*, *56*(2): pp. 251–299. https://doi.org/10.2307/413757.

Horie, Kaoru 1985. Lexico-syntactic analysis of verbs of cognition, conception and perception. *Sophia Linguistica* 18, pp. 9-48.

Huumo, Tuomas. 2010. Is perception a directional relationship? On directionality and its motivation in Finnish expressions of sensory perception. *Linguistics* 48(1), pp. 49–97. doi: 10.1515/ling.2010.002

Ibarretxe-Antuñano, Iraide. 1999. *Polysemy and metaphor in perception verbs: A crosslinguistic study.* PhD Dissertation. University of Edinburgh.

Ibarretxe-Antuñano, Iraide. 2002. "MIND-AS-BODY as a Cross-linguistic Conceptual Metaphor", *Miscelánea. A Journal of English and American Studies*. 25, pp. 93-119.

Izard, Carroll. 1977. *Human emotions*. New York: Plenum Press.

Jamison, Stephanie. 1976. Functional Ambiguity and Syntactic Change: The Sanskrit Accusative. In *Papers from the Parasession on Diachronic Syntax*, *12th Regional Meeting of the Chicago Linguistic Society*, pp 126–135.

Jayaseelan, K.A. 2004. The possessor-experiencer dative in Malayalam. In Bhaskararo Peri and Subbarao Karumuri V. (eds) *Non-nominative subjects (Vol 1)*, pp. 227-244. Amsterdam / Philadelphia: John Benjamins Publishing.

Johnson, M. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: The University of Chicago Press.

Johnson, Mark and Tim Rohrer. 2007. We are live creatures: Embodiment, American Pragmatism and the cognitive organism. In Jordan Zlatev, Tom Ziemke, Roz Frank and René Dirven, *Body, Language and Mind. Vol. 1: Embodiment*, pp. 17–54. Berlin: Mouton de Gruyter.

Johnson-Laird, Philip N. and Keith Oatley. 1989. The language of emotions: an analysis of a semantic field. *Cognition and Emotion* 3(2), pp. 81-123.

Johnson-Laird, Philip N. and Keith Oatley. 1992. Basic emotions, rationality, and folk theory. *Cognition and Emotion* 6(3/4), pp. 201-223.

Kachru, Yamuna. 1979. Pragmatics and Verb Serialization in Hindi-Urdu. *Studies in the Linguistic Sciences*, 9(2), pp. 157-69.

Kachru, Yamuna. 1980. *Aspects of Hindi Syntax*. Delhi: Manohar Publications.

Kachru, Yamuna. 1981. Transitivity and Volitionality in Hindi-Urdu. *Studies in the Linguistic Sciences*, 11(2), pp. 181-93.

Kachru, Yamuna. 1990. Experiencer and other oblique subjects in Hindi. In Manindra K. Verma, K.P. Mohanan, *Experiencer subjects in South Asian Languages*, pp. 59-77. Stanford, Center for the Study of Language (CSLI).

Kachru, Yamuna. 2006. *Hindi*. Amsterdam/Philadelphia: John Benjamin.

Kachru, Yamuna and Pandharipande, Rajeshwari. 1980. Toward a Typology of Compound Verbs in South Asian Languages. In *Studies in the Linguistic Sciences*, 10(1), pp. 113-24.

Kay, P. and Fillmore, Charles. J. 1999. Grammatical Constructions and Linguistic Generalizations: The What's X Doing Y? Construction. *Language*, 75(1), 1–33. https://doi.org/10.2307/417472

Karimi-Doostan, Gholamhossein. 1997. *Light Verb Constructions in Persian*. PhD thesis, University of Essex.

Keenan, Edward L. 1976. Towards a Universal Definition of 'Subject'. In C. Li (ed.) *Syntax and Semantics: Subject and Topic*, pp. 305-333. Academic Press, New York.

Keidan, Artemij. 2008. Predicative possessive constructions in Japanese and Korean. *Rivista degli Studi Orientali 81*: 339–367.

Kellogg, Samuel H. 1972 [1875]. *A Grammar of the Hindi Language*. London: Routledge & Kegan Paul Ltd.

Kemmer, Suzanne. 1993. *The Middle Voice*. Amsterdam: Benjamins.

Kibrik, Aleksandr. 1997. Beyond subject and object: Toward a comprehensive relational typology. *Linguistic Typology, 1(3)*, pp. 279-346.

King, Christopher R. 1994. *One Language, Two Scripts: The Hindi Movement in Nineteenth Century North India*. Delhi: Oxford University Press.

Kittilä. Seppo. 2002. Remarks on the basic transitive sentence. *Language Science* 24, pp. 107-130.

Kittila, Seppo, Katja Vasti and Jussi Ylikoski. 2011. Introduction to case, animacy and semantic roles. In Seppo Kittila, Katja Vasti and Jussi Ylikoski (eds.) *Case*, *Animacy and Semantic Roles*, pp. 1–26. Amsterdam/Philadelphia: John Benjamins.

Klaiman, M. H. 1980. Bengali dative subjects. *Lingua*, 51, pp. 275-295

Koptjevskaja-Tamm, Maria. 2011. "It's boiling hot!" On the structure of the linguistic temperature domain across languages. In Sarah Dessì Schmid, Ulrich Detges, Paul Gèvaudan, Wiltrud Mihatsch and Richard Waltereit (eds.) *Rahmen des Sprechens. Beitrage zur Valenztheorie, Varietatenlinguistik, Kognitiven und Historischen Semantik. Peter Koch zum 60*, pp. 393–410. Tübingen: Narr.

Koptjevskaja-Tamm, Maria (ed.) 2015. *The linguistics of temperature*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Koptjevskaja-Tamm, Maria. 2015. Introducing "The linguistic of temperature". In Koptjevskaja-Tamm, Maria. 2015. *The linguistics of temperature*, pp. 1-43. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kövecses, Z. 1986. *Metaphors of Anger, Pride and Love: A Lexical Approach to the Structure of Concepts*. Amsterdam / Philadelphia: John Benjamins.

Kövecses, Z. 1990. *Emotion Concepts*. New York: Springer-Verlag.

Kövecses, Z. 2000. *Metaphor and Emotion: Language, Culture, and Body in Human Feeling*. Cambridge: Cambridge University Press. doi:10.2277/0521641632

Kulikov, Leonid I. and Lavidas, Nikolaos. 2014. Introduction: Typology of labile verbs. *Linguistics* 52(4), pp. 871–877. (doi:10.1515/ling-2014-0010)

Kulkarni, Amba. 2011. Agreement in Hindi Conjunct Verbs. In *Proceedings of ICON-2011: 9th International Conference on Natural Language Processing.* India: Macmillan Publishers.

Lakoff, George. 1970. *Irregularity in syntax*. New York: Holt, Rinehart and Winston.

Lakoff, George. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: Chicago University Press.

Lakoff, George. 1993. The contemporary theory of metaphor. In Ortony, Andrew. (ed.). *Metaphor and Thought*, pp. 202-251. Cambridge: Cambridge U. P.

Lakoff, George abd Zoltán Kövecses. 1987. The cognitive model of anger inherent in American English. In Dorothy Holland and Naomi Quinn (eds.), *Cultural models in language and thought*, pp. 195–221. Cambridge: Cambridge University Press.

Lakoff, George and Mark Johnson. 1980. *Metaphors we live by*. Chicago: Chicago University Press.

Lakoff, George and Mark Johnson. 1999. *Philosophy in the flesh. The embodied mind and its challenge to western thought*. New York: Basic Books.

Lakshmi Bai, B. 2004. Acquisition of dative subject in Tamil. In Bhaskararo Peri and Subbarao Karumuri V. (eds) *Non-nominative subjects (Vol 2)*, pp. 245-268. Amsterdam / Philadelphia: John Benjamins Publishing.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar, vol. I: theoretical prerequisites.* Stanford.: Stanford University Press.

Langacker, Ronald W. 1991. *Foundations of cognitive grammar, vol. II: descriptive application.* Stanford.: Stanford University Press.

Langacker, Ronald. W. 1995. Possession and possessive constructions. In Taylor J.R., MacLaury R. E. (eds.) *Language and the Cognitive Construal of the World*, pp. 51-79. Berlin/New York: Mouton de Gruyter.

Langacker, R Ronald. W. 2001. Topic, subject, and possessor. In Simonsen, H. G. and Endresen, R. T. (eds.), *A Cognitive Approach to the Verb. Morphological and Constructional Perspectives*, pp.11-48. Berlin/New York: Mouton de Gruyter.

Langacker, Ronald. W. 2009. *Investigations in Cognitive Grammar*. Berlin/New York: Mouton de Gruyter.

Lazzeroni, Romano. 1990. La diatesi come categoria linguistica: studio sul medio indoeuropeo. *Studi e Saggi linguistici* 30, pp. 1–22.

Lehmann, Christian. 2002. *Possession in Yucatec Maya*, Second revised edition. In ASSidUE, 10, Seminar für Sprachwissenschaft der Universität, Erfurt.

Lehmann, Christian. 2015. *Thoughts on grammaticalization* (3rd edition) Berlin: Language Science Press.

Letuchiy, Alexander. 2009. Towards a typology of labile verbs: Lability vs. derivation. In Alexandre Arkhipov and Patience Epps (eds.), *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, pp. 223–244. Berlin / New York: Mouton de Gruyter.

Levin, Beth. 1993. *English verb classes and alternations*. Chicago, IL: The University of Chicago Press.

Levin, Beth and Malka Rappaport-Hovav. 2005. *Argument Realization*. Cambridge, MA: MIT Press.

Li, Charles and Thomson, Sandra. 1976. Subject and Topic. A New Typology of Language. In Li Charles (ed.), *Subject and Topic*, pp. 457-491. New York: Academic Press.

Liljegren Henrik and Haider Naseem. 2015. Facts, feelings and temperature expressions in the Hindu Kush. In Koptjevskaja-Tamm, Maria (ed.) 2015.*The linguistics of temperature*, pp. pp. 440-471. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Luraghi, Silvia. 2001. Some remarks on Instrument, Comitative, and Agent in Indo-European. In *Sprachtypologie und Universalienforschung* 54/4, pp. 385- 401.

Luraghi, Silvia. 2003. *On the Meaning of Prepositions and Cases. The expression of semantic roles in Ancient Greek*. Amsterdam: Benjamins.

Luraghi, Silvia. 2014. Plotting Diachronic Semantic Maps: the Role of Metaphor, in: Silvia Luraghi and Heiko Narrog (eds.) *Perspectives on Semantic Roles*, pp. 101–152. Amsterdam: John Benjamins.

Luraghi, Silvia. 2015. Asymmetries in Italian temperature terminology. In Koptjevskaja-Tamm, Maria. 2015. *The linguistics of temperature*, pp. pp. 333-353. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Luraghi, Silvia. 2020a. *Experiential Verbs in Homeric Greek*. Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004442528.

Luraghi, Silvia. 2020b. External possessor constructions in Indo-European. In Jóhanna Barðdal, Spike Gildea, Eugenio R. Lujan Martines (eds.) *Reconstructing Syntax*, pp. 162-196. Leiden: Brill.

Luraghi, Silvia and Alessandra Caviglia and Erica Pinelli. 2014. Accusative and genitive objects with verbs of perception and mental activity in Homeric Greek, Early Vedic, and Old Church Slavic. Silvia Luraghi, Alessandra Caviglia and Erica Pinelli. Paper presented at the Workshop "Case, Arguments And Clause Structure In Indo-European" Ghent, 5-6 May 2014.

Luraghi, Silvia and Narrog, Heiko. 2014. (eds.) *Perspective on Semantic Roles*. Amsterdam/Philadelphia: John Benjamins Publishing.

Luraghi, Silvia and Eleonora Sausa. 2015. Hate and Anger, Love and Desire: the Construal of Emotions in Homeric Greek. In Dag T.T. Haug (ed.) *Historical Linguistics* 2013, pp. 233–255. Amsterdam: John Benjamins.

Luraghi, Silvia and Eleonora Sausa. 2017. Pensare, sapere, ricordare: i verbi di attività mentale in greco omerico, in Felicia Logozzo and Paolo Poccetti (eds.) *Ancient Greek Linguistics: New approaches, insights, perspectives*, pp. 745–774. Berlin: Mouton De Gruyter.

Luraghi, Silvia and Eleonora Sausa. 2019. Aspects of aural perception in Homeric Greek, in Egle Mocciaro and William Short (eds.) *The Embodied Basis of Constructions in Greek and Latin,* pp. 149–175. Berlin: Mouton De Gruyter.

Luraghi Silvia and Narrog Heiko. 2014. *Perspective on semantic roles*, John Benjamins Publishing Company Amsterdam/Philadelphia

Malchukov, Andrei L. 2005. Case Pattern Splits, Verb Types and Construction Competition". In Amberber M. and de Hoop H. (eds.), *In Perspectives on Cognitive Science, Competition and Variation in Natural Languages*, pp: 73-117. Amsterdam: Elsevier.

Malchukov, Andrei. 2015. Valency classes and alternations: parameters of variation. In Malchukov A. and Comrie B. (eds.) *Valency Classes in World's Languages,* Vol. 1, pp. 74-130. De Gruyter Mouton, Berlin.

Malchukov Andrei. Comrie Bernie (eds.). 2015. *Valency Classes in World's Languages,* Vol. 1. Berlin: De Gruyter Mouton.

Manning, C. and Schütze, H. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge Massachusetts: The MIT Press.

Masica, Colin P. 1976. *Defining a Linguistic Area: South Asia*. Chicago and London: University of Chicago Press.

Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Masica, Colin P. 2001. The Definition and Significance of Linguistic Areas: Methods, Pitfalls, and Possibilities (with Special Reference to the Validity of South Asia as a Linguistic Area). In Rajendra Singh and Probal Dasgupta (eds.) *The Yearbook of South Asian Languages and Linguistics*, pp. 205-268. Berlin, New York: De Gruyter Mouton.

Masini, Francesca. 2016. *Grammatica delle costruzioni: un'introduzione*. Roma: Carocci Editore.

McEnery, A. and Wilson, A. 2001. *Corpus Linguistics.* Edinburgh: Edinburgh University Press.

McEnery, Tony, Xiao, Richard, Tono, Yukio. 2006. *Corpus-based Language Studies: An Advanced Resource Book*. London: Routledge.

McGregor, Ronald S. 1967. The Rise of Standard Hindi and Early Hindi Prose Fiction. In *Journal of the Royal Asiatic Society of Great Britain and Ireland, no. 3/4*, pp. 114–132.

McGregor, R. S. 1972. *Outline of Hindi Grammar. With Exercises*. Delhi: Oxford University Press.

McGregor, Ronald S. 1977. *Outline of Hindi Grammar, with exercises (2nd edition)*. Delhi: Oxford University Press.

McGregor Ronald S. 1994. *The Oxford Hindi-English Dictionary*. Oxford: Oxford University Press.

Michaelis, Susanne Maria and the APiCS Consortium. 2013. Experiencer constructions: headache. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath and Magnus Huber (eds.), *The Atlas of Pidgin and Creole Language Structures* pp. 260–263. Oxford: Oxford University Press. Moyse-Faurie, Claire.

Milanetti, Giorgio. 2002. La tradizione inventata: in qual modo una bella lingua indiana senza un nome preciso fu chiamata hindi e trasformata in power construction. In E. Basile e M. Torri (eds.) *Il subcontinente indiano verso il terzo millennio. Tensioni politiche, trasformazioni sociali ed economiche, mutamento culturale,* pp. 449–499. Milano: Franco Angeli.

Mistry, P.J. 2004. Subjecthood of non-nominatives in Gujarati. In Bhaskararo Peri and Subbarao Karumuri V. (eds.) *Non-nominative subjects (Vol 2),* pp. 1-32. Amsterdam / Philadelphia: John Benjamins Publishing.

Mohanan, Tara, 1994. *Argument Structure in Hindi*. Stanford, CA: CSLI Publications.

Molesworth, James Thomas, George Candy and Thomas Candy. 2005. *Molesworth's Marathi-English Dictionary*. Pune: Shubhada-Sarawat Prakashan.

Monier-Williams, Monier. 2008. *Sanskrit-English Dictionary*. Varanasi: Indica Books.

Montaut, Annie. 2004a. *A Grammar of Hindi*. LINCOM EUROPA.

Montaut, Annie. 2004b. Oblique main arguments as localizing predications in Hindi/Urdu. In Bhaskararao P., Subbarao K.V. (eds.), *Non nominative subjects*, pp, 178-210. John Benjamins Publishing Company.

Montaut, Annie. 2011. Basic Intransitivity: A Typologically Relevant Feature for Indo-Aryan. O.N. Koul (ed.). *Indo-Aryan Linguistics,* pp.29-46. Central Institute of Indian Languages.

Montaut, Annie. 2013. The rise of non-canonical subjects and semantic alignments in Hindi. In Seržant Ilja A., Kulikov Leonid (eds.) *The Diachronic Typology of Non-Canonical Subjects*, pp. 91-118. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Montaut, Annie. 2016. Noun-Verb Complex Predicates in Hindi and the rise of non-canonical subjects. In Pollet Samvelian, Lea Nash (eds.) *Approaches to Complex Predicates,* pp. 142-174. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Montaut, Annie. 2018. On the nature of the Hindi infinitive: History as an answer to its syntactic behavior? In: Sharma Ghanshyam and Rajesh Bhatt, *Trends in Hindi Linguistics*, pp. 115-146. Berlin / Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110610796-006

Namrata Paul, Abhijeet Satsangi and Gosh, Sanjukta. 2023. Cognitive Semantic Analysis of Fear Expressions in Hindi and Bangla, talk presented at the 37th South Asian Languages Analysis Roundtable (SALA-37), Department of Asian and North African Studies, Ca' Foscari University of Venice, October 4-7, 2023.

Napoli, Maria. 2006. *Aspect and Actionality in Homeric Greek. A contrastive Analysis*. Milano: Franco Angeli.

Narasimhan, Bhuvana. 1998. A lexical semantic explanation for `quirky' case marking in Hindi. *Studia Linguistica 52(1),* pp. 48-76.

Narrog, Heiko. 2010. A diachronic dimension in maps of case functions. In *Linguistic Discovery* 8 (1), pp. 233-254.

Narula, Shamsher S. 1976. *Hindi Language – A Scientific History*. Delhi: Oriental Publishers & Distributors.

Nespital, Helmut. 1996. Verbal Aspect in Indo Aryan and Dravidian languages. The relation of Simple verbs to Verbal expressions. *Berliner Indologische Studien 9/10*, pp. 247-258.

Næss, Ashild. 2007. *Prototypical Transitivity*. Amsterdam / Philadelphia: John Benjamins Publishing

Niemeier, Susanne. Dirven Ren. (eds.) 1997. *The language of emotions: conceptualization, expression and theoretical foundation*. Amsterdam / Philadelphia: John Benjamins Publishing.

Noonan, Michael. 1985. Complementation. In Shopen T. (ed.) *Language Typology and Syntactic Description*, vol. 2, pp- 41-140. Cambridge: Cambridge University Press.

Noonan, Michael. 2007. Complementation. In Shopen T. (ed.) *Language Typology and Syntactic Description,* pp. 52–150. Cambridge: Cambridge University Press.

Orsini, Francesca, 2010. *Before the Divide: Hindi and Urdu Literary Culture*. Hyderabad: Orient Blackswan.

Osmond, Meredith. 1997. The prepositions we use in the construal of emotion: Why do we say fed up with but sick and tired of? In Susanne Niemeier and Ren. Dirven (eds.), *The language of emotions: conceptualization, expression and theoretical foundation*, pp. 111–133. Amsterdam / Philadelphia: John Benjamins Publishing.

Onishi, Masayuki. 2001. Non-canonically marked subjects and objects: Parameters and properties. In Aikhenvald, Alexandra Y., Dixon, Robert M. W., Onishi, Masayuki (eds.) *Non-canonical*

Marking of Subjects and Objects, pp. 1-53. Amsterdam: John Benjamins Publishing Company.

Perek, Florent. 2015. *Argument Structure in Usage-Based Construction Grammar. Experimental and corpus-based perspectives*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Perlmutter, David M. 1978. Impersonal Passives and the Unaccusative Hypothesis. In *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*. UC Berkeley, pp. 157–189.

Pinker, Steven. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, Mass.: MIT Press.

Plank, Frans. 2003. Temperature talk: The basics. A talk presented at the Workshop on Lexical Typology at the ALT conference in Cagliari, September 2003.

Pokorny, Julius. 1959 [1887]. *Indogermanisches Etymologisches Wörterbuch*. Bern: Francke.

Rai, Amrit. 1984. *A house divided: the origin and development of Hindi/Hindavi*. Delhi: Oxford University Press.

Raineri, Sophie and Camille Debras. 2019. Corpora and Representativeness: Where to go from now? CogniTexts [online] 19.

Reh, Mechthild (ed.). 1998. *Experiens-Kodierung in afrikanischen Sprachen typologisch gesehen: Formen und ihre Motivierungen*. Hamburg: Universität Hamburg, Institut für Afrikanistik und Äthopistik.

Reh, Mechthild and Christiane Simon. 1998. Experiens-Konstruktionen in Mande-Sprachen. In Reh Mechthild (ed.) *Experiens-Kodierung in afrikanischen Sprachen typologisch gesehen: Formen und ihre Motivierungen,* pp. 41-88. Hamburg: Universität Hamburg, Institut für Afrikanistik und Äthopistik.

Reinöhl, Uta. 2016. *Grammaticalization and the Rise of Configurationality in Indo–Aryan*. Oxford: Oxford University Press.

Reznikova, Tatiana, Ekaterina Rakhilina and Anastasia Bonch-Osmolovskaya. 2012. Towards a typology of pain predicates. *Linguistics* 50(3). 421–465. DOI: 10.1515/ling-2012-0015.

Rice, Sally and Kabata, Kaori. 2007. Crosslinguistic grammaticalization patterns of the allative. *Linguistic Typology* 11, pp. 451–514. DOI: 10.1515/LINGTY.2007.031

Rubin, David. 2002. *The return of Sarasvati: four Hindi poets. New Delhi* : Oxford University Press.

Saksena, Anuradha. 1980. "The affected agent". *Language, 56*, pp. 812–826.

Saksena, Anuradha. 1980. Causative Relations in Hindi. In *General Linguistics Vol.20 (10)*, pp. 23. Pennsylvania State University Press.

Saksena, Anuradha. 1982. *Topics in the Analysis of Causatives, with an Account of Hindi Paradigms*. Los Angeles: University of California Press.

San Roque, Lila, Kobin H. Kendrick, Elisabeth Norcliffe and Asifa Majid. 2018. “Universal meaning extensions of perception verbs are grounded in interaction”. *Cognitive Linguistics* 29(3), pp. 371–406.

Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? In *Linguistic Typology* 6(2), pp. 199–217.

Schlesinger, I. M. 1995. On the semantics of the object. In B. Aarts and C.F. Meyer (eds.), *The verb in contemporary English: Theory and description*, pp. 54–74. Cambridge, MA: Cambridge University Press.

Schomer, Karine.*1983. Mahadevi Varma and the Chhayavad Age of Modern Hindi Poetry*. Berkley: University of California Press.

Shapiro, Michael C. 1987. Hindi lagnā: A Study in Semantic Change. *Journal of the American Oriental Society*, Vol. 107, No 3, pp. 401-408.

Shapiro, Michael C. 1989. *A Primer of Modern Standard Hindi*. Delhi: Motilal Banarsidass.

Shapiro, Micheal C. 2003. *Hindi*. In G. Cardona e D. Jain (eds.) *The Indo–Aryan Languages*, pp. 276–314. New York: Routledge.

Shibatani, Masayoshi. 1957. *A Linguistic study of Causative Construction*. University of Calitfornia dissertation, reproduced by Indiana University Linguistics Club, Bloomington, Indiana.

Shibatani, Masayoshi and Pardeshi, Prashant. 2001. Dative Subject Constructions in South Asian Languages, In *The Yearbook of South Asian Languages and Linguistics*, pp.311-348.

Sinclair, John. 1996. *EAGLES Preliminary recommendations on corpus typology*. Tech. rep. EAG–TCWG–CTYP/P. Pisa: Expert Advisory Group on Language Engineering Standards.

Singh, Madhu. 2016. Altered Realities, New Experiences: Bhisham Sahni, Nirmal Verma, and the “Nayi Kahani” Movement. In *Comparative Literature Studies* (Volume 53, Number 2), pp. 312-333.

Singh, Rajendra, Agnihotri, RamaKant. 1997. *Hindi Morphology. A Word-based Description*. Delhi: Motilal Banarsidass.

Sinha, Rai, Mahesh, Kumar. 2009. Mining Complex Predicates In Hindi Using A Parallel Hindi-English Corpus. In *Proceedings of the 2009 Workshop on Multiword Expressions, ACL-IJCNLP 2009*, pp. 40–46.

Slade, Benjamin. 2013. The diachrony of light and auxiliary verbs in Indo-Aryan. *Diachronica 34:4*, pp. 531:578.

Smith, Michael B.,1993. Cases as conceptual categories: evidence from German' in Richard A. Geiger and Brygida Rudzka-Ostyn (eds), *Concpetualization and Mental Processes in Language*, pp. 531–66. Berlin: Mouton de Gruyter.

Stassen, Leon. 2009. *Predicative Possession*, Oxford, Oxford University Press.

Stefanowitsch, Anatol.2020. *Corpus linguistics A guide to the methodology*. Berlin: Language Science Press. DOI: 10.5281/zenodo.3735822.

Stolz, Thomas. 2001. "Comitatives vs. instrumental vs. agents". In W. Bisang (ed.) *Aspects of Typology and Universals*, pp. 153-174. Berlin: Akademie Verlag.

Subbarao Karumuri V. 2012. *South Asian Languages: A Syntactic Typology*. Cambridge: Cambridge University Press.

Subbarao Karumuri V. and Bhaskararao Peri. 2004. Non-nominative subjects in Telugu. In Bhaskararo Peri and Subbarao Karumuri V. (eds) *Non-nominative subjects (Vol 2)*, pp. 161-196. Amsterdam / Philadelphia: John Benjamins Publishing.

Sulger, S. (2015) *Modeling nominal predications in Hindi/Urdu*, PhD Thesis, University of Konstanz.

Sutton, John. 2007. Language, memory, and concepts of memory: Semantic diversity and scientific psychology. In Mengistu Amberber (ed.) *The language of memory in crosslinguistic perspective*, pp. 41-66. Amsterdam: John Benjamins.

Sweetser, Eve. 1990. *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Talmy, Leonard. 2000. Force dynamics in language and cognition. Cognitive Science 12.49–100. Revised and expanded version published in Talmy Leonard (ed.), *Toward a Cognitive Semantics* (vol. I), pp. pp. 409–470. Cambridge, Mass.: MIT Press.

Tatevosov, Sergej. 2002. The parameter of actionality. In *Linguistic Typology* 6(3), pp. 317–401.

Taylor, John, R. 1995. *Linguistic categorization*. Oxford: Oxford University Press.

Taylor, John R. 1993. Prepositions: patterns of polysemization and strategies of disambiguation. In Zelinsky-Wibbelt (ed.), The Semantics of Prepositions: From Mental Processing to Natural Language Processing, pp. 151–175. Berlin: Mouton de Gruyter.

Tyler, Andrea and Vyvyan Evans, 2003. *The Semantics of English Prepositions*. Cambridge: Cambridge University Press.

Tsunoda, Tasaku. 1981. Split case-marking patterns in verb-types and tense / aspect / mood". *Linguistics* 19 (5/6), pp. 389–438.

Tsunoda, Tasaku. 1985. Remarks on Transitivity. *Journal of Linguistics* 21, pp. 385-396.

Tsunoda, Tasaku. 2015. The hierarchy of two-place predicates: its limitations and uses. In Malchukov A. Comrie B. (eds.), in *Valency Classes in World's Languages,* Vol. 2, pp. 1597-1625. De Gruyter Mouton, Berlin.

Turner L. Ralph. 1971. *A comparative dictionary of the Indo-Aryan languages*. London: Oxford University Press.

Usoniene, Aurelia. 1999. Perception verbs revisited. *Working Papers* 47 *Lund University, Dept. of Linguistics*, pp. 211-225.

Van der Leek, F. 1996. The English conative construction: A compositional account. In L. Dobrin, K. Singer, and L. McNair (eds.), *CLS 32: The main session. Papers from the 32th meeting of the Chicago Linguistic Society*, pp. 363–378. Chicago: Chicago Linguistic Society.

Van Olphen, Herman. 1975. Aspect, tense, and mood in the Hindi verb. *Indo-Iranian Journal*, Vol. 16, No. 4, pp. 284-301.

Van Valin, Robert D. 1993. A Synopsis of Role ald Reference Grammar. In Robert D. Van Valin (ed.), *Advances in Role and Reference Grammar*, pp, 1–164. Amsterdam: Benjamins.

Van Valin, Robert D. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

Van Valin, Robert D. and Randy LaPolla. 1997. *Syntax: structure, meaning, and function*. Cambridge: Cambridge University Press.

Van Valin, Robert D. 1991. Another Look at Icelandic Case Marking and Grammatical Relations. Natural Language and Linguistic Theory 9: 145–194.

Van Valin, R.D. and Wilkins, D. 1993. Predicting syntactic structure from semantic representations: Remember in English and its equivalents in Mparntwe Arrernte. In R. D. Van Valin (ed.), *Studies in Role and Reference* Grammar, pp. 499–534. Amsterdam: John Benjamins.

Vasiljevic, Zorana and Graham, Jim. 2020. Conceptualization of Anger in English and Japanese in 「文学部紀要」 文教大学文学部 (34－1号), pp. 1－46.

Verbeke, Saartje. 2013. *Alignment and Ergativity in New Indo-Aryan Languages.* Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110292671

Verhoeven, Elisabeth. 2007. *Experiential Constructions in Yucatc Maya. A typologically based analysis of a functional domain in a Mayan language*. Amsterdam / Philadelphia: Benjamins Publishing Company.

Verma, Manindra K. 1976. *The notion of subject in South Asian languages*, Madison: South Asian Studies.

Verma, Manindra K., Mohanan, Karuvannur P (eds.) 1990. *Experiencer subjects in South Asian languages.* Stanford: CSLI.

Verma, Mahindra K. and K.P. Mohanan (eds.) 1993. *Complex Predicates in South Asian Languages*. Delhi: Manohar.

Viberg, Åke 1984. The verbs of perception: a typological study. In B. Butterworth, B. Comrie and Ö. Dahl (eds.) *Explanations for language universals*, pp. 123-162. The Hague: Mouton.

Viberg, Åke. 2001. The verbs of perception. In Martin Haspelmath, Ekkehard Knig, Wulf Oesterreicher and Wolfgang Raible (eds.), *Language typology and language universals*. An international handbook, pp. 1294–1309. Berlin: Mouton.

Viberg, Åke. 2005. The lexical typological profile of Swedish mental verbs. *Languages in Contrast* 5(1), pp.121–157. DOI: 10.1075/lic.5.1.09vib

Viberg, Åke. 2006. Towards a lexical profile of the Swedish verb lexicon. In *Sprachtypologie und Universalienforschung* 59(1), pp. 103–129.

Viberg, Åke 2008. Swedish verbs of perception from a typological and contrastive perspective. In Maria de los Angeles Gomez Gonzalez, J. Lachlan Mackenzie and Elsa M. Gonzalez Alvarez (eds.), *Languages and cultures in contrast and comparison*, pp. 123–172. Amsterdam: Benjamins. DOI: 10.1075/pbns.175.09vib

Viberg, Åke. 2015. Sensation, perception and cognition. Swedish in a typological-contrastive perspective. *Functions of Language* 22:1, pp. 96–131. doi 10.1075/fol.22.1.05vib.

Wali Kashi 2004. Non-nominative subjects in Marathi. In Bhaskararo Peri and Subbarao Karumuri V. (eds.) *Non-nominative subjects (Vol 2)*, pp. 223-252. Amsterdam / Philadelphia: John Benjamins Publishing.

Wierzbicka, Anna. 1981. Case marking and human nature. In *Australian Journal of Linguistics, 1,* pp. 43–80.

Wierzbicka, Anna. 1995. Everyday conceptions of emotion: A semantic perspective. In J. A. Russell, J.-M. Fernández-Dols, A. S. R. Manstead, and J. C. Wellenkamp (eds.), *Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion*, pp. 17–47. Kluwer Academic/Plenum Publishers. https://doi.org/10.1007/978-94-015-8484-5_2

Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Amsterdam: John Benjamins Publishing.Zalizniak.

Wierzbicka, Anna. 1999. *Emotions across languages and cultures: diversity and universals*, Cambridge: Cambridge University Press.

Wierzbicka, Anna. 2007. Is “remember” a universal human concept? “Memory” and culture. In Mengistu Amberber (ed.) *The language of memory in crosslinguistic perspective*, pp. 13-40. Amsterdam: John Benjamins.

Woolford, Ellen. 1997. Four-Way Case Systems: Ergative, Nominative, Objective and Accusative. In *Natural Language and Linguistic Theory* Vol. 15, 1, pp. 181-227.

Yadava, Yogendra P. 2004. Non-nominative subjects in Maithili. In Bhaskararo Peri and Subbarao Karumuri V. (eds.) *Non-nominative subjects (Vol 2)*, pp. 253-263. Amsterdam / Philadelphia: John Benjamins Publishing.

Yu, Ning. 1995. Metaphorical Expressions of Anger and Happiness in English and Chinese, in *Metaphor and symbolic activity* 10(2), pp. 59-92.

Zanchi Chiara and Inglese Guglielmo. 2022. Ancient Greek valency patterns and alternations. *Paper presented at the 3rd International Colloquium on Ancient Greek Linguistics*, Universidad Autònoma de Madrid, Spain, 16.18 June 2022.

Zúñiga, F., and Kittilä, S. (2019). Changing Semantic Valency: Causatives, Applicatives, and Related Constructions. In *Grammatical Voice* (Cambridge Textbooks in Linguistics), pp. 12-81. Cambridge: Cambridge University Press. doi:10.1017/9781316671399.002