SPATIAL EXPRESSIONS AND CASE IN SOUTH ASIAN LANGUAGES

Dissertation zur Erlangung des akademischen Grades
des Doktors der Philosophie (Dr. phil.)
an der Universität Konstanz
Geisteswissenschaftliche Sektion
Fachbereich Sprachwissenschaft

vorgelegt von

Tafseer Ahmed Khan

1. Referentin: Prof. Dr. Miriam Butt, Universität Konstanz
2. Referentin: Dr. Ashwini Deo, Yale University
3. Referent: Prof. Dr. Frans Plank, Universität Konstanz

Tag der mündlichen Prüfung: 24.08.2009

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor, Miriam Butt, whose understanding, encouraging and personal guidance have provided a good basis for the present dissertation. I am thankful for the detailed review, constructive criticism and excellent advice during this research.

I owe my most sincere gratitude to the *Sonderforschungsbereich 471* which provided the funding during my PhD work. I am thankful to Aditi Lahiri, Frans Plank and Irene Wolke for this regard.

I warmly thank Ashwini Deo, Scott Grimm, Ingrid Kaufmann, Tikaram Poudel and Gilliam Ramchand for their constructive comments and advice during the work.

I am deeply grateful to my informants who always answer my queries during the whole period of research. As it is a long list, I may forget to list some important names. So firstly my sincere thanks to any worthy name that was not mentioned here. I owe my sincere gratitude to Jameel Ahmed (Balochi), Nayyara Atif (Punjabi), Muhammad Bilal (Pashto), Sajid Iqbal (Saraiki), Ameeruddin Mehar (Sindhi) Tikaram Poudel (Nepali, Manipuri), Sarju Poudel (Manipuri), Ghulam Raza (Sariaki) and Anila Varghese (Malayalam) for the huge amount of linguistic data they provided for this work.

I am thankful to Susanne Trissler for editing the final manuscript of this dissertation. However, any remaining errors are mine.

I am also thankful to Nadeem Ahmed, Inam ullah, Zeeshan Amjad, Naila Ata, Imran ul Haq Arfi, Farhan Faiz, Rehman Gul, Atif Gulzar, Zahid Hussain, Bushra Jawaid, Amir Kamran, Akbar Khan, P. Madhavan, Tara Mohanan. Tahira Naseem, Adnan Rauf, Sarfraz Raza, Mutee-ur-Rehman, Lateef Sagar, Hassan Sajjad, Farooq Umar, Sohail Yousuf and Imran Zaheer for their responses.

I have great regard for Ghulam Ali Allana, Nabi Baksh Baloch, Sarmad Hussain, Mohammad Nadeem, Kamran Nishat, Mohammad Islam Nishtar, Jameel Pal, Nasir Touheed, Anjuman-e-taraqqi-e-Urdu (Karachi), Baluchistan University (Quetta), Centre of Research in Urdu Language Processing (Lahore), National Language Authority (Islamabad), Radbound University (Nijmegen), University of Bergen, University of

Karachi, Universitaet Stuttgart and Urdu Dictionary Board (Karachi) for their help in conducting the research and providing opportunities for good discussion.

I wish to express my warm and sincere thanks to my fellows Tina Boegel, Melanie Hochstaetter, Jaouad Mousser, Melanie Seiss, Sebastian Sulger and Alexandros Tantos who helped me and facilitated my work and the preparation of the dissertation.

Last but not the least, I owe loving thanks and gratitude to my family who have always supported me.

Zusammenfassung

Diese Dissertation untersucht die Beziehung zwischen Kasus und Raumausdrücken in südasiatischen Sprachen. Die Dissertation gibt einen Überblick über südasiatische Sprachen hinsichtlich des synchronischen Gebrauchs von Kasus, identifiziert interessante und/oder ungewöhnliche Muster und gibt eine lexikalisch-semantische Erklärung für diese Muster an, basierend auf einem unterspezifizierten, Merkmal-basierten Modell für Raumausdrücke sowie einer Identifizierung von Metaphern zur Erweiterung auf diejenigen Ausdrücke, die sich nicht auf den Raum beziehen.

Die Daten für diese Dissertation wurden erhoben durch eine Untersuchung von Kasusmarkierern in zehn südasiatischen Sprachen. Südasien wird als ein zusammenhängendes „linguistisches Areal“ oder *Sprachbund* (Emeneau 1956) betrachtet. Dies führt zu einer Konvergenz von linguistischen Merkmalen durch Sprachkontakt, der über Jahrhunderte hinweg besteht. Aus diesem Grund wurden sechs indoarische (Haryani, Nepali, Punjabi, Saraiki, Sindhi und Urdu/Hindi) und zwei iranische (Balochi und Pashto) Sprachen sowie eine dravidische (Malayalam) und eine tibetobirmanische (Manipuri) Sprache für die Studie ausgewählt. Die Untersuchung bestätigt die Beobachtung, dass gemeinsame räumliche Merkmale in Sprachen unterschiedlicher Familien vorliegen.

Bei der Untersuchung der Verwendung von Kasusmarkierern liegt das Hauptaugenmerk auf synchronischen Fragen. Die Untersuchung widmet sich jedoch auch der Herkunft der Kasusmarkierer. Es stellt sich heraus, dass viele der modernen Hauptkasusmarkierer (z.B. die Ergativ- und Akkusativ-Markierer) ihren Ursprung in Raumausdrücken haben. So haben manche dieser Formen, z.B. das Ergativ-/Akkusativ-*nae* im Haryani und das Akkusativ-/Dativ-*ko* in Urdu/Hindi, in den jeweiligen heute gebräuchlichen Sprachen auch eine räumliche Verwendung. Diese Beobachtung zeigt die tiefe Verbindung zwischen Kasus und Raumausdrücken.

Jene Verbindung wird von zwei Gesichtspunkten genauer betrachtet. Zum einen versucht die Dissertation, die Muster in der Verwendung und der Synchronie unter den untersuchten Sprachen zu verstehen, zunächst nur innerhalb der Domäne der

Raumausdrücke. Zum anderen identifiziert die Dissertation neue Verbklassen im Südasiatischen. Diese Identifizierung von Verbklassen stellt einen neuen Beitrag auf dem Gebiet der lexikalischen Semantik dar. Durch das Verstehen der Muster in der Kasusmarkierung innerhalb dieser Verbklassen untersucht die Dissertation außerdem, wie die Markierer der Raumausdrücke auch in eine Domäne der nicht raumbezogenen Kasusmarkierung ausgedehnt werden können.

Betrachtet man ausschließlich die raumbezogene Domäne, so wird gezeigt, dass die theoretischen Systeme, die in früheren Arbeiten zu Raummarkierern entwickelt wurden, wie z.B. Ostler (1979), Jackendoff (1990) und Kracht (2002), nicht dazu ausreichend sind, alle Probleme der feinen Unterscheidungen und Polysemien unter den raumbezogenen Kasusmarkierern, die in südasiatischen Sprachen vorhanden sind, zu erklären. Zu einer kohärenten und systematischen Erklärung dieser Probleme bringt diese Dissertation daher ein alternatives, unterspezifiziertes Modell ein.

Beim Problem der Erweiterung der Raummarkierer in eine nicht raumbezogene Ebene erörtert die Dissertation die Polysemien, die zwischen raumbezogenen und nicht raumbezogenen Kasusmarkierern bestehen. Manche der Formen, die als Raummarkierer benutzt werden, markieren außerdem den Instrumental, den Adressierten und nicht kanonische zweite Argumente. Die Untersuchung von Konstruktionen, die nicht kanonisch markierte zweite Argumente enthalten, führt zur Etablierung von sechs Verbklassen, die den meisten der genannten Sprachen zuteil sind. Diese Methode der Verbklassifizierung ist im Einklang mit Levin (1993), wo die Annahme formuliert wird, dass Verben mit einer syntaktisch identischen Struktur eine zusammenhängende semantische Klasse bilden. Die Bestimmung von Verbklassen mit einer systematischen Verwendung von nicht kanonischen zweiten Argumenten oder Non Canonical Second Argument (NCSA) führt uns zum Versuch, die semantischen Faktoren hinter der Wahl der Kasusmarkierer von NCSA zu verstehen.

Die Dissertation legt folgende Haupterkenntnisse zu südasiatischen Sprachen dar. Manche südasiatischen Sprachen unterscheiden zwischen statischen und dynamischen Quellen, z.B. die Ablative *dekhi* vs. *baaTa* im Nepali. Die Domäne des Raumes stellt Metaphern für diejenigen Domänen bereit, die sich nicht auf den Raum beziehen. Daraus

resultiert die Benutzung von räumlichen Markierern in nicht räumlichen Domänen. Es kann mehrere Metaphern in der räumlichen Domäne für eine einzige nicht räumliche Verwendung geben. Verschiedene Sprachen können unterschiedliche Metaphern für dieselbe Verwendung wählen, z.B. kann der Adressierte des Verbs ‘fragen’ als ein (mit dem Dativ markierter) Empfänger oder als eine mögliche (mit dem Ablativ markierte) Quelle betrachtet werden.

Die Wahl des Kasusmarkierers an einem zweiten Argument hängt nicht allein von seinen semantischen Merkmalen ab. Die semantischen Merkmale des gesamten Gliedsatzes können die Wahl des Markierers beeinflussen. Die identifizierten NCSA-Verbklassen wurden durch keine der vorangegangenen Theorien indiziert. Nicht kanonische, z.B. dativische, Markierung eines Subjekts der erfahrenden Person ist ein bekanntes Phänomen (Verma & Mohanan 1990), wohingegen die nicht kanonische Markierung eines zweiten Arguments in einer Konstruktion mit einem Subjekt der erfahrenden Person vorher noch nicht systematisch untersucht worden ist.

Somit zeigen sowohl die diachrone Entwicklung moderner Kasusmarkierung als auch synchrone Polysemien raumbezogener und nicht raumbezogener Markierer die tiefe Verbindung zwischen Raumausdrücken und Kasus. Die nicht raumbezogene Domäne entleiht ihre Metaphern der räumlichen Domäne und Raummarkierer können hierdurch auf eine sehr reguläre Weise um nicht raumbezogene Verwendungen erweitert werden. Vorherige Modelle der Raumverwendung in Verbindung mit NCSAs wurden als unzureichend analysiert, um die Daten aus südasiatischen Sprachen zu erklären. Diese Dissertation schlägt daher ein alternatives, Merkmal-basiertes Modell für die räumliche Domäne vor und erweitert unser Wissen über diejenigen semantischen Faktoren, die für die nicht räumliche Domäne eine Rolle spielen.

Abstract

This dissertation investigates the relation of case and spatial expressions in South Asian languages. The dissertation surveys South Asian languages with respect to synchronic case usage, and identifies interesting and/or unusual patterns and proposes a lexical semantic explanation for the patterns in terms of an underspecified feature-based model for spatial relations and an identification of metaphors for the extension into the non-spatial domain.

The data for this dissertation were obtained by a survey of case marker usages in ten South Asian languages. South Asia is considered as a "linguistic area" or *Sprachbund* (Emeneau 1956) leading to the convergence of linguistic features due to language contact persisting over centuries. Because of this reason, six Indo-Aryan (Haryani, Nepali, Punjabi, Saraiki, Sindhi and Urdu/Hindi), two Iranian (Balochi and Pashto), one Dravidian (Malayalam) and one Tibeto-Burman (Manipuri) language were selected for the study. The survey confirms the observation that there are areal features that are found in common among the languages of different families.

The study of case marker usages mainly focuses on synchronic issues. However, it also investigates the origin of case markers. It is found that many modern core case markers (e.g., ergative and accusative markers) originate from spatial terms. Some of these forms e.g., Haryani ergative/accusative *nae* and Urdu/Hindi accusative/dative *ko* have spatial usages in today's language as well. This observation shows the deep connection between case and spatial expressions.

This connection is explored in more detail from two perspectives. For one, the dissertation tries to understand the patterns of usage and multiple senses across the surveyed languages just within the domain of spatial relations. For another, the dissertation identifies new verb classes within South Asian. This identification of verb classes represents a novel contribution to the field of lexical semantics. In trying to understand the pattern of case marking on these verb classes, the dissertation further investigates how spatial markers can be extended into a non-spatial case marking domain.

With respect to just the spatial domain, it is shown that the theoretical frameworks developed in earlier studies on spatial markers such as Ostler (1979), Jackendoff (1990) and Kracht (2002) cannot explain all the problems of fine-grained differences and polysemy of spatial case markers found in South Asian languages. For a coherent and systematic explanation of these problems, this dissertation therefore proposes an alternative underspecified feature-based model.

With respect to the issue of how spatial marking can extend into a non-spatial domain, the dissertation explores the non-spatial usages of spatial case markers. Some of the forms that are used as spatial markers also mark instrument, addressee and non-canonical second arguments. The study of the constructions with non-canonically marked second arguments provides six verb classes that are common in most of these languages. This method of verb classification is parallel to Levin (1993), who claims that the verbs sharing the same syntactical structure form a coherent semantic class. The establishment of verb classes with a systematic use of non-canonical second argument (NCSA) leads us to try to understand the semantic factors behind the choice of the case marker for NCSAs.

The dissertation provides the following major findings for the South Asian languages. Some South Asian languages distinguish between static and dynamic sources, e.g., the Nepali ablatives *dekhi* vs. *baaTa*. The spatial domain supplies metaphors to the non-spatial domains. This is the reason for the use of spatial markers in non-spatial domains. There can be more than one spatial metaphor for a non-spatial usage. Different languages may select different metaphors for the same usage, e.g., the addressee of the verb 'ask' can be considered as a (dative marked) recipient or a (ablative marked) potential source.

The choice of the case marker on an argument does not solely depend on its semantic features. The semantic features of the whole clause may influence on the choice of the case marker. The identified NCSA verb classes were not predicted by any earlier theory. Non-canonical, e.g., dative, marker for the experiencer subject is a known phenomenon (Verma & Mohanan 1990), but the non-canonical marking on the second argument of an experiencer subject construction was not systematically studied earlier.

Hence both the diachronic development of modern case markers and synchronic non-spatial usage of spatial markers show the deep relation between spatial expression and

case. The non-spatial domain borrows its metaphors from the spatial domain and spatial markers can therefore be extended to mark non-spatial usages in a very regular manner. Previous models of spatial usage and NCSAs were shown to be inadequate for accounting for the South Asian data. This dissertation therefore proposes an alternative feature-based model for the spatial domain and extends our understanding of the semantic factors involved with respect to the non-spatial domain.

Abbreviations

ABL Ablative
ACC Accusative
ALL Allative
ANT Anterior
BEN Benefactive
CAUS Causative
COM Comitative
COP Copula
DAT Dative
DEIC Deictic
ERG Ergative
EVD evidential
F Feminine
GEN Genitive
IMP Imperative
IMPF Imperfective
INF Infinitive
INST Instrument
LOC Locative
M Masculine
NCSA Non Canonical Second Argument
NOM Nominative
NPST Non-past
OBL Oblique
PERL Perlative (‘through’)
PART Partitive
PST Past
PERF Perfect
PL Plural
PRES Present,
PROG Progressive
REAL Realis
SOL Solicitive
SG Singular

Chapter 1

Introduction

1.1. Survey of South Asian Languages

This dissertation discusses the relationship of spatial expressions and case in South Asian languages. South Asia is considered as a "linguistic area" or *Sprachbund*. Language contact over the centuries has resulted in a convergence of linguistic features among languages of different language families. This dissertation analyzes the data from ten languages of South Asia that belong to the Indo-Aryan, Iranian, Dravidian and Tibeto-Burman families. The languages studied are: Balochi (Iranian), Haryani (Indo-Aryan), Malayalam (Dravidian), Manipuri (Tibeto-Burman), Nepali (Indo-Aryan), Pashto (Iranian), Punjabi (Indo-Aryan), Saraiki (Indo-Aryan), Sindhi (Indo-Aryan) and Urdu/Hindi (Indo-Aryan).

1.2. Case and Case Markers

Case is a system of marking dependent nouns for the type of relationship they bear with their heads (Blake 2001). There are different ways to label case. Sanskrit is an Old Indo-Aryan language and is the predecessor of Indo-Aryan languages. The declinations of the Sanskrit word *deva-* 'god' are given in Table 1.1.

Table 1.1: Case inflections in Sanskrit (cf. Blake 2001)

Case	**Noun + Case Marker**
Nominative	*deva-s*
Accusative	*dev-am*
Dative	*deva-aya*
Instrument	*dev-ena*
Ablative	*deva-at*
Locative	*deva-sya*
Genitive	*dev-e*

It is important to note that the Sanskrit grammatical tradition does not use the case labels: nominative, accusative and dative, etc. These names come from the grammatical tradition of Western European languages. There is another important point regarding case marking. In Table 1.1, all the cases are morphological inflections. In other languages, the case maker can be an inflection, clitic or pre/postposition (Butt & King 2005). For example, Urdu/Hindi uses clitics and postpositions for case marking. Table 1.2 gives the case markers used along with Urdu/Hindi *diyotaa*[1] 'god'.

Table 1.2: Case markers of Urdu/Hindi

Case	**Noun + Case Marker**
Nominative	*diyotaa* (unmarked)
Ergative	*diyotaa ne* (clitic)
Accusative	*diyotaa ko* (clitic)
Dative	*diyotaa ko* (clitic)
Instrument	*diyotaa se* (clitic)
Ablative	*diyotaa se* (clitic)
Locative	*diyotaa par* (clitic), ……, *diyotta tak* (clitic), *diyotaa ke paas* (postposition)
Genitive	*diyotaa kaa* (clitic) *diyotaa kii* (clitic) *diyotaa ke* (clitic)

There is more than one way of classifying cases into different groups. A distinguishing feature is that some cases mark mandatory arguments like subject, object and (mandatory) oblique arguments. The examples of these cases are nominative, accusative and dative, while the other cases do not (canonically) appear on this kind of argument.

[1] In transcription, *a*, *i* and *u* are used for short vowels and *aa*, *ii* and *uu* are used for the long ones. *ai* is used for open mid front unrounded vowel and *oa* are for open mid back rounded vowel. Capital letters are used for retroflex consonants except capital *S* which is used for voiceless palatal fricative. Capital *N* used after a vowel shows nasalization. Small *c* is used for voiceless alveolar affricate. The consonants followed by *1,* e.g., *b1,* are implosive sounds. However, *n1* is used for retroflex nasal.

However, there can be another classification scheme for the cases. Locative and ablative mark the spatial relation with the head, while the other cases primarily mark a non-spatial relationship with the head. Thus there is a spatial vs. non-spatial distinction in the inventory of cases. This classification is of more interest for me because I want to investigate the use of the forms of spatial case in the non-spatial domain.

The study of spatial case introduces a problem of nomenclature. There are two spatial cases listed in Table 1.1 and 1.2. However, we find that Urdu/Hindi has more than one case marker corresponding to the locative case. Are different labels required for each of these markers? This problem is not unique to Urdu/Hindi. Many other languages, e.g., Hungarian and Avar, have a lot of spatial cases. The spatial case pattern in Hungarian is presented in Table 1.3 (Creissels 2008).

Table 1.3: Spatial cases in Hungarian

	IN	**ON**	**AT**
Location	*-ban/ben* (inessive)	*(o/e/ö)n* (superessive)	*-nál/nél* (adessive)
Destination	*-ba/be* (illative)	*-ra/re* (sublative)	*-hoz/hez/höz* (allative)
Source	*-ból/bó‘l* (elative)	*-ról/ró‘l* (delative)	*-tól/tó‘l* (ablative)

Table 1.3 introduces different labels for spatial case. These labels are not a standard, and different authors use different naming schemes. The nomenclature of spatial terms is discussed in Chapter 2. The point here is that the analysis of spatial case is not as simple as listing two spatial case markers (i.e., locative and ablative marker).

1.3. Origin of Case Markers

Table 1.2 presents the case markers of Urdu/Hindi that is a New Indo-Aryan (NIA) language. Table 1.1 presents the case markers of Sanskrit that is an Old Indo-Aryan (OIA) language and predecessor of Urdu/Hindi. The comparison of these two tables shows that the case markers of both languages are different. The case markers of the predecessor language are lost, and the successor language developed new case markers. The study about the origin of new case markers can tell us about the relation

of space and case. A lot of work is available on the origin of South Asian especially Indo-Aryan case markers (Beames 1872, Kellog 1893). In Chapter 2, the origin of non-spatial markers is especially explored to find any link between case and space.

1.4. Polysemy of Spatial Case Markers

The study of the South Asian data shows two types of pattern related to polysemous spatial markers. The other sense(s) of a polysemous spatial marker may be spatial or non-spatial. The study of both of these patterns provides interesting results that are presented in this dissertation (Chapter 3 and 4).

1.4.1. Same Form having Different Spatial Usages

The South Asian data shows that a single form may be used to mark more than one spatial usage. Urdu/Hindi ablative marker *se* marks the source of motion as well as the path.

(1) ahmad karaacii=se aa-yaa
Ahmad Karachi=ABL come-PERF.M.SG
'Ahmad came from Karachi.' <Urdu/Hindi>

(2) hamid baG=se guzr-aa
Hamid garden=through pass-PERF.M.SG
'Hamid passed through the garden.' <Urdu/Hindi>

Ostler (1979), Jackendoff (1990) and Kracht (2002) proposed models of spatial markers. Can these models explain all the polysemy patterns and other issues related to South Asian languages? Chapter 2 focuses on this question.

1.4.2. Spatial Markers marking Non-Spatial Usages

Usage of the same form for more than one spatial usage is only a part of the problem. A more interesting observation is the use of a spatial marker for non-spatial usages. Table 1.2 shows that the Urdu/Hindi ablative marker *se* is also used as the instrument marker. Similarly, Pashto uses the same form for marking the location and instrument. Can these observations be dismissed as coincidence, i.e., homophony?

The answer is in the negative because there are more examples of spatial markers acting in a non-spatial domain. A careful survey of examples from South Asian languages reveals that there are classes of verbs whose second (mandatory) argument

is marked by spatial markers. Two examples of spatial marking on such arguments are presented below.

(3) maiN **saaNp=se** Dar-taa huuN
1SG snake.M.SG=ABL fear-IMPF.M.SG be.1SG
'I fear snakes.' <Urdu/Hindi>

(4) jamiil=ko **zaahid=par** bharosaa hai
Jameel.M.SG=DAT Zahid.M.SG=LOC-on trust.M.SG be.SG
'Jameel trusts Zahid.' <Urdu/Hindi>

The claim that there is a relationship between spatial expressions and case markers is not new (see Butt 2006). In the localist approach, core case marking is explained in terms of local constructs (Anderson 1971). Ostler (1979) developed a linking theory for Sanskrit that uses spatial features. Case markers in Ostler's system contain spatial features like "source" and "goal".

Beside the localist approach, there are theories about argument realization that may explain why a non-canonical marker appears on the secondary mandatory argument. Hence, an investigation is required to find whether argument realization theories and other semantic models can explain all the usages of spatial forms in non-spatial domains. This is done in Chapter 3.

1.5. Main Questions and Dissertation Plan

The main questions of the dissertations are:

Question 1: Did many of the case markers originate from spatial nouns?

Question 2: Can a model be proposed that explains the different spatial usages of the same form?

Question 3: Why is a core spatial marker used for non-spatial usages?

The plan of the dissertation is as follows. Chapter 2 provides a survey of case markers in ten South Asian languages. It identifies some interesting patterns in the data that are discussed in the following chapters. It also investigates the origin of some core non-spatial case markers, and the presence of spatial and non-spatial cognate forms in sister languages. This investigation provides the answer to the first question.

Chapter 3 discusses the spatial models proposed by Kracht (2002), Jackendoff (1990) and Ostler (1979) in relation to the South Asian data found in Chapter 2. It points out merits and shortcomings of these models and proposes a new feature-based model for spatial relations that provides a single underspecified lexical entry for different usages of the same marker.

Chapter 4 discusses the non-spatial usages of spatial markers. It identifies and presents classes of verbs with non-canonically marked second arguments. To explain the patterns of marking in these verb classes, theories of argument realization are discussed. Finally, some semantic features are identified that can model non-spatial usages like instrument, addressee and originally spatial marking on the second (mandatory) argument. The patterns of case marking on verbs with non-canonical second arguments are shown to be not random, but the result of a systematic metaphorical extension of spatial concepts along with the systematic and contrastive encoding of semantic features such as animacy, dynamicity and intended vs. achieved goal, etc.

Chapter 5 then concludes the dissertation.

Chapter 2

Survey of South Asian Case Markers

2.1. Introduction

This chapter represents the empirical heart of the dissertation. It presents data with respect to case marking from these ten South Asian languages: Balochi, Haryani, Malayalam, Manipuri, Nepali, Pashto, Punjabi, Sindhi, Saraiki and Urdu/Hindi. There are several interesting observations and patterns which emerge from the data. These are briefly discussed in section 2.5 of this chapter and some, but not all of the issues are analyzed as part of this dissertation. That is, while many interesting issues arise with respect to the data presented here, the dissertation will concentrate on the following two issues: 1) accounting for crosslinguistic patterns and polysemy in the spatial domain; 2) attempting to understand the spread of originally spatial markers to marking non-spatial arguments.

This chapter is organized as follows. Section 2.2 of this chapter first goes through various terminology used in this dissertation, just to establish what is meant by which label. In particular, names of thematic roles, case markers and spatial markers are presented. Section 2.3 goes through the empirical data. The case markers, their distribution and their semantic usages are detailed by means of examples. The synchronic data is followed by a look at what is known (or has been postulated) about the historical origin of these case markers (section 2.4). Sections 2.5 then presents some observations with respect to the data presented in section 2.3. One can find many interesting observations from the data related to the case marking. However, as the main topic of this dissertation is related to space, the observations related to spatial markers or spatial origins of non-spatial markers are highlighted.

2.2. Establishing Terminology

2.2.1. Morphosyntactic Status of Case Markers

The prototypical use of case markers is to encode the relation of the argument with the head (Blake 2001, Butt 2006). Since classical languages like Latin and Sanskrit have case markers that are morphological inflections, it is often assumed that case markers must always be morphological inflections. For example, Kachru (2006) admits only three cases for Urdu/Hindi: direct, oblique and vocative. This is because these are three possibilities encoded by morphological inflection in Urdu/Hindi. For example, the Urdu word *laRka* 'boy' is assumed to be in the direct case. When the same word is inflected with *-e*, it can express a vocative or be used before a postposition (oblique case).

However, not all authors agree with this view of Urdu/Hindi case. Masica (1991) discusses three layers of case markers in South Asian languages. These layers are equivalent to inflection affixes, clitics and postpositions proposed by Butt & King (2005). However, Butt & King (who were working on Urdu/Hindi) do not consider postpositions as case markers. I use the idea of three layers of case markers from Masica with the labels provided by Butt & King. The reason for considering postpositions as case markers comes from the empirical data presented in 2.3. For example, the Punjabi instrument marker *naal* is a postposition. As it marks the instrument, it must be considered as a case marker.

Butt & King list tests for distinguishing inflection, clitic and postposition. Inflectional affixes do not scope over co-ordination. Take the example of the Urdu/Hidi locative inflection *-e* that appears only with words ending in *-aa*. The example (1) has the unmarked (nominative) form of two city names (section 2.2.2 introduces the label "nominative" in some detail).

(1) laaRkaanaa aur kalkattaa baRe Seher haiN
Larkana.NOM and Kolkata.NOM big city.M.PL be.PRES.PL
'Larkana and Kolkata are big cities.' <Urdu/Hindi>

In example (2), these cities are used as the goal. The city names in example (2) have the locative inflection *-e*. As the inflection is required to mark all the elements joined by the co-ordination, the locative inflection appears on both of these nouns.

(2) vo laaRkaane aur kalkatte gaye
3PL Larkana.LOC AND Kolkata.LOC go.PERF.M.PL
'They went to Larkana and Kolkata.' <Urdu/Hindi>

(3) *vo [laaRkaan aur kalkatt]-e gaye
3PL Larkana.LOC AND Kolkata.LOC go.PERF.M.PL
'They went to Larkana and Kolkata.' <Urdu/Hindi>

Example (3) is not valid as the inflectional affix is required on each stem and it does not have scope over the co-ordination.

The second type of case marker is a clitic and does not behave in the same way. Clitics do have scope over a co-ordination. See the example of the Urdu/Hindi ablative clitic *se* in (4).

(4) vo [laaRkaane aur kalkatte]=se aa-ye
3PL Larkana.OBL AND Kolkata.OBL=ABL come-PERF.M.PL
'They came from Larkana and Kolkata.' <Urdu/Hindi>

There is another test to distinguish between inflectional affixes and clitics. Urdu/Hindi contains a focus clitic *hi*. This focus clitic cannot split the affix from its stem, but it can appear between a stem/word and a case clitic. See the example of the Urdu/Hindi clitic *hi* with the affix *-e* and the clitic *se*.

(5) vo [laaRkaane=hi / *laaRkan-hi-e] gayaa
3PL Larkana.LOC=FOC / Larkana-FOC-LOC go.PERF.M.SG
'He went only to Larkana.' <Urdu/Hindi>

(6) vo laaRkaane=hi=se gayaa
3PL Larkana.OBL=FOC=ABL go.PERF.M.SG
'He went only from Larkana.' <Urdu/Hindi>

In example (5) it is not possible to put the focus clitic *hi* between the stem and the affix *-e*, while in example (6) it is possible to put the focus clitic between the word and the clitic *se*. Hence, we can distinguish between inflectional affixes and clitics using these two tests.

It is important to note that the Urdu/Hindi locative affix *-e* is homophonous with the oblique affix *-e*. It is why *laaRkaane* is used both as oblique (6) and as locative (5).

The third type of case marker is a postposition. The postposition is preceded by the genitive marker.

(7) billi mez=ke niice hai
cat.F.SG.NOM table.F.SG.OBL=GEN.M.SG.OBL under be.PRES.SG
'The cat is under the table.' <Urdu/Hindi>

The above example has the Urdu/Hindi postpostion *niice*. The genitive marker *ke* joins the noun and the postposition.

Section 2.3 mainly concentrates on different semantic usages of South Asian case markers. But before moving on to that discussion, it is necessary to introduce some relevant vocabulary. Hence, in section 2.2.2 we focus on thematic roles. These thematic roles are used to classify different semantic usages of case markers in the survey presented in section 2.3.

After the introduction of thematic roles, the next sections 2.2.3–2.2.4 give different labels used for the case markers. I discuss the relation of case markers and thematic roles in these sections as well.

2.2.2. Thematic Role Labels

Thematic role labels are generally used to describe the semantic relationship between an argument and its predicate. There is a small number of thematic roles that are widely used. These roles describe different types of semantic arguments used in a clause. See the following examples.

(8) John broke the window with the hammer.

(9) Michael killed Fred with the gun.

The example (8) is about the predicate *break* and the example (9) is about the predicate *kill*. We find many common semantic attributes in different arguments of these two sentences. *John* and *Michael* are semantically similar as both are performing an action. The *window* and *Fred* are semantically similar as both are affected by the action. And the *hammer* and the *gun* are similar as both are used to perform the action. Due to these similarities, both members of these three pairs have the same thematic roles. *John* and *Michael* are "agent", *window* and *Fred* are "patient" and *hammer* and *gun* are "instrument".

Historically the concept of thematic roles dates back to Sanskrit grammarian Panini who lived around the 4th century BC. In more recent times, the concept of thematic roles is discussed by many authors. See Gruber (1965, 1976), Fillmore (1968), Jackendoff (1972, 1987, 1990) and Givon (1990) among others. It is important to note that all authors do not use the term "thematic role" for this concept. The concept is called "deep semantic case" by Fillmore (1968), "semantic role" by Givon (1990) and "thematic relation" by Gruber (1976). The term "thematic role" was used in work on Government-Binding (Chomsky 1981).

Similarly, there is no full agreement on the complete number of thematic roles or on the names of different thematic roles. Fillmore (1968) used the term "objective" for thematic roles that are labeled as "patient" and "theme" by the other authors. The participant causing the action is usually labeled as "agent" by many authors, but Van Valin (1990) uses two thematic roles: "agent" and "effector".

In this section, we present the thematic roles that are used by most of the authors and for which a general consensus exists. These thematic role labels are used in sections 2.2.3 and 2.3 to describe different semantic usages of case markers pretheoretically (i.e., as convenient labels).

2.2.2.1. Agent

The agent is the instigator of an action. In the following example, *John* is the agent.

(10) John opened the door.

2.2.2.2. Causee

The causee is the intermediate agent who performs an action and is controlled/instigated by the (primary) agent. In the following example, *John* is the causee.

(11) Michael made John open the door.

2.2.2.3. Theme

The theme undergoes change of location or state by an action. In the following example, *door* is the theme.

(12) Michael opened the door.

2.2.2.4. Patient

The patient is the entity affected by an action. In the following example, *tree* is the patient.

(13) Michael burned the tree.

2.2.2.5. Recipient

The recipient receives something in an action. In the following example, *John* is the recipient.

(14) Michael gave a book to John.

2.2.2.6. Experiencer

The experiencer has cognition or perception of an action. In the following examples, *John* is the experiencer.

(15) John fears snakes.

(16) The snake frightens John.

2.2.2.7. Benificiary

The benificiary is the person for whom the action is performed. In the following example, *John* is the beneficiary.

(17) Michael bought a book for John.

2.2.2.8. Instrument

The instrument is the inanimate entity that is used for performing the action. In the following example, *knife* is the instrument.

(18) He cut the apple with the knife.

2.2.2.9. Goal

The goal is the end point of the motion. In the following example, *Frankfurt* is the goal.

(19) He went to Frankfurt.

2.2.2.10. Source

The source is the origin or starting point of the motion. In the following example, *Frankfurt* is the source.

(20) He came from Frankfurt.

2.2.2.11. Path

The thematic role path is not a commonly used role, but we need it to describe the South Asian data. It is the place through which a moving body passes during its motion. In the following example, *garden* is the path.

(21) He passed through the garden.

2.2.2.12. Purpose

In the following example, *(to) meet Michael* is the purpose of the action performed.

(22) John went to Frankfurt to meet Michael.

2.2.2.13. Manner

In the following example, *in circles* is the manner of the action.

(23) He walks in circles.

2.2.2.14. Summary

The thematic roles listed above are the labels needed to describe and pretheoretically classify the different usages of the South Asian case markers surveyed. However, the data includes a number of further usages that do not fall within the set of thematic role labels that are usually assumed. We will therefore introduce further labels as necessary, but note that the thematic roles listed above make a very good starting point for our analysis of the South Asian data. Before moving on to that discussion, however, in the next section, we briefly introduce the labels for case markers, along with assumptions about them, that are used in this dissertation.

2.2.3. Non Spatial Case Markers

As the relationship between thematic role labels and the use of case markers is not one to one, this section presents the labels assumed for core cases. In addition, I list all the thematic roles that are usually related to a particular case marker.

2.2.3.1. Nominative

Many languages, e.g., Sanskrit, have a separate inflectional affix for the nominative form. For example, the Sanskrit root word *deva* 'god' has the nominative form *deva-s.* Here the affix *-s* is used as the nominative case marker. In contrast, the modern South Asian languages generally express subjects in a bare stem form without any overt case marking (either of affix, clitic or postposition) after it. This bare or unmarked form is generally called the nominative (e.g., Mohanan 1994, Butt & King 2005). This definition is somewhat different from other definitions of nominative found in the literature on other language areas, but it makes sense in the South Asian context.

There is another issue related to the nominative. Languages have been classified as nominative-accusative or ergative-absolutive on the basis of case marking of subjects and objects (e.g., Dixon 1979, 1994). English (with pronouns as subject and object) is an example of a nominative-accusative language. The subjects of both transitive and intransitive verbs of such languages are marked by the nominative case, while the object is marked by the accusative case.

(24) I arrived.

(25) He arrived.

(26) He saw me.

(27) I saw him.

The first and third person pronouns are in nominative form, i.e., *I* and *he* respectively, in the subject position. When used as an object, the accusative forms *me* and *him* are used, respectively. Hence in English, the subjects of both transitive and intransitive verbs have the same (nominative) form.

Ergative-absolutive languages behave differently. The subject of a transitive verb is marked by the ergative marker, while the subject of an intransitive verb and the object is marked by the same marker (called absolutive marker). Basque is an ergative-absolutive language (Ignacio 2003).

(28) Gizon-a dator
man.ABS-DET is.coming
'The man is coming.'

(29) gizon-ak zakurr-a ikusi du
man-ERG dog.ABS-DET see AUX
'The man has seen the dog.'

The above examples show that the same marker *-a* marks both the subject of the intransitive verb and the object.

The surveyed South Asian languages do not fit in any of these patterns completely. For example, the subject of an intransitive verb in Urdu/Hindi can be unmarked, or it can be marked by the ergative marker (for some verbs in a dialect). Similarly, the subject of the transitive verbs can be either unmarked or marked by the ergative marker. Similarly, the object can be unmarked or marked by the accusative marker. Hence we have all the following patterns.

(30) Unmarked Subject

(31) Ergative Subject

(32) Unmarked Subject Accusative Object

(33) Ergative Subject Unmarked Object

(34) Ergative Subject Accusative Object

(35) Unmarked Subject Unmarked Object

The question arises whether we label the unmarked subject of (32) as nominative marked (as in nominative-accusative languages) and the unmarked object of (33) as absolutive marked (as in ergative-absolutive languages). Or should all the unmarked forms be labeled similarly? Following many authors, e.g., Mohanan (1994) and Butt (1995), I consider all the unmarked forms as the (unmarked) nominative. In the following example,

the uninflected noun *daraxt* ‘tree’ is used as the subject.

(36) daraxt kaT-aa
tree.M.SG.NOM (get) cut-PERF.M.SG
‘The tree got cut.’ <Urdu/Hindi>

The same form of the noun can be used as object too.

(37) laRkaa daraxt kaaT rahaa hai
Boy.M.SG.NOM tree.M.SG.NOM cut PROG.M.SG be.PRES.SG
‘The boy is cutting the tree.’ <Urdu/Hindi>

Hence in South Asian languages, nominative form means the absence of any case marker or oblique form.

2.2.3.2. Ergative

In the examples (36) and (37), the subjects of both Urdu/Hindi sentences are in the unmarked nominative form. However, the patterns shown in (31), (33) and (34) illustrate that in many South-Asian languages the subject could also be marked with the ergative case marker.

The South Asian languages studied here in fact show split ergativity. Usually, the ergative marker is found with the subject of transitive verbs with past or perfect tense/aspect. The subject in other constructions does not have the ergative marker. It is the reason why the subject of (37) is in nominative case, while the subject of (38) is marked by the ergative marker.

(38) laRke=ne daraxt kaaT-aa
Boy.M.SG.OBL=ERG tree.M.SG.NOM cut-PERF.M.SG
‘The boy cut the tree.’ <Urdu/Hindi>

It is important to note that the noun *laRke* ‘boy.OBL’ is the oblique form of the unmarked root *laRkaa*, when it is used before the ergative case marker. In many South Asian languages, the noun changes to an oblique form if it is followed by a case marker.

In addition to the ergative and split-ergative pattern described above, some South Asian languages also allow for other uses of the ergative. Urdu/Hindi can use it to express *want* modality (Butt & King 2005), Nepali and Manipuri use it for a contrast between stage- and individual-level predication (Butt & Poudel 2007, Poudel 2008a).

2.2.3.3. Accusative

As illustrated by the patterns (31)–(34), the accusative marker optionally appears on the object of many South Asian languages and generally marks themes and patients. Reasons for the appearance of the accusative as in (39) have to do with factors such as specificity or animacy, for example (Comrie 1981).

(39) puliis=ne laRke=ko maar-aa
Poilce.F.SG.NOM boy.M.SG.OBL=ACC beat-PERF.M.SG
'The police beat the boy.' <Urdu/Hindi>

2.2.3.4. Dative

The dative case marker can be used to mark the recipient or both recipient and beneficiary. In Nepali, the same form *laai* is used to mark both the recipient and the beneficiary.

(40) mai=le gaai=laai ghaaNs de
1SG=ERG cow=DAT fodder give.PST.1SG
'I gave the cow fodder.' <Nepali>

(41) us=le raam=laai ek-Taa kitaab kin-di-yo
3SG=ERG Ram=BEN one-CLF book buy-give-PERF.M.SG
'He bought a book for Ram.' <Nepali>

However, the dative markers of all the surveyed languages do not all mark both recipient and beneficiary usages. The Urdu/Hindi dative marker *ko* is used to mark the recipient only. The beneficiary usage in Urdu/Hindi is marked by the case marker *liye*.

There is another interesting point related to the dative marker in South Asian languages. In Urdu/Hindi and Nepali (along with Sindhi, Saraiki and Punjabi), the same form is used to mark the patient/theme and the recipient.

(42) raam=le bhaai=laai piT-yo
Ram.M.SG=ERG brother.M.SG=ACC beat-PERF.M.SG
'Ram beat the brother.' <Nepali>

The patient in the above sentence is marked by the marker *laai.* The same form is used as the dative marker in (40) and (41). This raises the question of whether we should label the *laai* used in (42) as the dative marker too. Similarly, the same form is used to mark patient/theme (as in (39)) and the recipient in Urdu/Hindi.

(43) maiN=ne laRke=ko kitaab dii
1SG.OBL=ERG boy.M.SG.OBL=DAT book.F.SG.NOM cut-PERF.F.SG
'I gave the book to the boy.' <Urdu/Hindi>

There is a debate about the question whether the label "dative" can be used with patient/theme arguments marked by *ko*. Mahajan (1990) and Davison (1998) say that the optional *ko* marking the patient/theme is the dative marker too. The accusative marker does not exist in Urdu/Hindi.

On the other hand, Mohanan (1994) and Butt (1995) recommend using two different labels accusative and dative for Urdu/Hindi case marker *ko*. They argue that the optional marking on the patient/theme (object) depends on the animacy and definiteness/specificity. This phenomenon is found crosslinguistically, and such optional markers on the object are traditionally called accusative. Hence, it is better to use accusative *ko* or *laai* etc. for these arguments.

2.2.3.5. Benefactive

As described above, the beneficiary role in many languages is marked by the dative marker. For example, the classical languages Latin and Sanskrit use the same (dative) form to mark both the recipient and the beneficiary. The Nepali dative marker behaves similarly. However, in many South Asian languages, the marker for the beneficiary usage is different from the marker of the recipient, i.e., the dative marker.

Urdu/Hindi uses the postposition *liye* to mark the beneficiary.

(44) maiN=ne laRke=ke liye kitaab xariid-ii
1SG=ERG boy.OBL=GEN BEN book.F.SG.NOM buy-PERF.F.SG
'I bought the book for the boy.' <Urdu/Hindi>

In the above example, the postposition *liye* is labeled as BEN, i.e., the benefactive marker. The term "benefactive marker" is not in common use, but we do find it in the literature as Heine & Kuteva (2002), Trask (1997) and Moravcsik (2003) have used the term. Hence, if the beneficiary role is marked by any form that is different from the dative marker, such marker is termed as the beneficative marker in the context of this dissertation.

2.2.3.6. Instrument

The instrument marker marks the instrument by means of which an action is performed. The form used as the instrument marker usually has other semantic usages too. In Nepali, for example, the same form is used as instrument marker as well as ergative marker.

(45) us=le camsaa=le bhaat khaa-yo
3SG=ERG spoon=INST meal eat-PST.M.SG
'He ate the meal with a spoon.' <Nepali>

2.2.3.7. Comitative

Like the benefactive marker, the comitative marker too is not found in the grammars of classical languages like Sanskrit and Latin. Many languages have a polysemous marker for instrument and accompaniment usages. But we also find examples in which the language has two different forms as comitative (marking accompaniment) and instrument (marking instrument) markers. Many South Asian languages, especially Dravidian languages, have a distinct marker for the accompaniment.

Masica (1991) presents an inventory of case markers in South Asian languages. He includes comitative markers in the list of case marker labels relevant for South Asian languages. Nepali, an Indo-Aryan language, also has different forms to mark accompaniment and instrument roles. The former is marked by the comitative marker *sange*, while the latter is marked by the instrument marker *le*.

(46) u ma=sanga bazaar ga-yo
3SG 1SG=COM market go-PST.M.SG
'He went to the market with me.' <Nepali>

2.2.3.8. Summary

The list of case marker labels above includes core case markers such as nominative or accusative, but also non-core case markers such as comitative or instrument. These labels will be used in section 2.3 to describe different South Asian case markers.[1]

What the case markers presented above have in common is that these are all referring to non-spatial concepts. However, there are also case markers which mark spatial

[1] Vocative case is not considered because it is only a form used to address the noun. It does not have a relation to the head.

concepts. These are discussed in the next section. Before moving on to these pure spatial markers, a point is required to be noted. If we compare the list of thematic roles given in 2.2.1 and the list of case markers presented above, we notice that all the semantic roles do not have a corresponding case marker. For example, there is no purpose or manner case marker in the above list. The reason for this omission is that none of the languages studied has a distinct case marker that is used to mark only manner or purpose. These thematic roles are usually marked by the instrument marker or by some other means. This is illustrated by the discussion in the next section.

2.2.4. Spatial Markers

The case systems of classical languages like Sanskrit and Latin have two cases for spatial relations. These cases are locative and ablative. The ablative marker marks the source of motion and the locative marker marks the static location.

However, these two cases are unable to express all the spatial relations found in the languages. That is why more terms are introduced to express other semantic varieties of spatial relations. For details see Creissels (2008) who describes different terms used for spatial case.

There are four important spatial markers that are widely discussed. The case labels for these cases are: ablative (source), perlative (path), allative (goal) and locative (static location). See the example of these relations in the following.

(47) The book is **on** the table. (Locative)

(48) He went **to** Karachi. (Allative)

(49) He went **from** Karachi. (Ablative)

(50) He passed **through** the garden. (Perlative)

The above four kinds of spatial markers again provide a good starting point, but they do not cover all the varieties of spatial markers. We need more terms for the spatial relations that are not covered by these four types. The English preposition *towards*, for example, is neither allative nor perlative. The term “approximative” is used for the marker of this semantic usage.

(51) The moved towards the station. (Approximative)

In many languages, we need to have a fine-grained distinction between locative, ablative or other spatial markers. The locative expresses that the object is present at a static location. The object can be 'in', 'on', 'above', 'beside' or 'near' the location. Similarly, the source of motion can be 'in', 'on' or 'at' a location. Hence, we need two dimensions (a) source/goal/path/location and (b) the relative orientation of theme/figure with respect to the location/ground to express spatial relations of many languages. There is no agreement on the names of these bidimensional relations. For example, Creissels (2008) presented a spatial case system for Hungarian and Avar. The system uses different labels for the similar concept in these two languages.

Table 2.1: Spatial cases in Hungarian (Creissels 2008)

	IN	**ON**	**AT**
Location	Inessive	Superessive	Adessive
Destination	Illative	Sublative	Allative
Source	Elative	Delative	Ablative

Table 2.2: Spatial cases in Avar (Creissels 2008)

	ON	**AT**	**IN1**	**IN2**
Location	Superessive	Apudessive	Interessive	Inessive
Destination	Superlative	Apudlative	Interlative	Illative
Source	Superelative	Apudelative	Interelative	Inelative
Path	Supertranslative	Apudtranslative	Intertranslative	Intranslative

The above tables show that there is no agreement between the naming convention of spatial cases in Hungarian and Avar. In Hungarian, the relation [Source, ON] is called delative, while it is called apudelative in Avar. For the discussion in section 2.3 in particular and the dissertation as a whole, we need to use some convention to label

different spatial case markers. The naming conventions shown in Table 2.1 and 2.2 are neither an agreed upon standard nor easy to remember. For this reason, we stick to the four (relatively) widely used labels, i.e., ablative, perlative, allative and locative. To express the relative orientation or any other fine grained difference, the labels like locative-in, locative-on and ablative-in, etc. are used.

Hence, if a language has a single marker for 'in', 'on' and 'at' version of the (static) locations, then that marker is termed as locative. The Nepali marker *ma* is used to mark the locations 'in' and 'on'. This marker is termed as a locative marker. But, if a language has different markers for 'in', 'on' and 'at', then the terms locative-in, locative-on, locative-near, etc. are used for these varieties of locations. Urdu/Hindi has two different markers *par* and *meN* for 'on' and 'in', respectively. These markers are labeled as locative-on (*par*) and locative-in (*meN*) markers.

Similarly, the moving object can have a differing orientation with respect to the source. Sindhi has three markers *maaN* ('from in'), *taaN* ('from on') and *khaaN* ('from') for the source. According to the convention adopted here, *maaN* is termed as ablative-in and *taaN* is termed as ablative-on. The third marker *khaaN* that represents the semantically less specific usage is termed as ablative marker. Similarly, we can use hyphenization for other varieties of fine-grained differences.

Having established the terminology that will be used in this dissertation with respect to thematic roles and case markers, we now move on to the empirical heart of the dissertation, namely, a survey of the synchronic use of case marking in ten different South Asian languages.

2.3. Survey of Case Markers in South Asian Languages

South Asia is home to a lot of languages belonging to different language families. A survey of the semantic usages of case markers in ten South Asian languages is conducted. Six of these languages belong to the Indo-Aryan family, two belong to the Iranian family, one belongs to the Dravidian family and one belongs to the Tibeto-Burman family.

I used grammar books and some textual data, e.g., news papers, novels, etc., to find example sentences for the analysis. However, the survey is conducted primarily by

consulting one or more native speakers as informants. These informants are mentioned in the Acknowledgements. The informants are given a set of about 60 sentences. These sentences are listed in Appendix A. The sentences were in Urdu/Hindi (written in Roman script), as all of the informants were fluent in Urdu/Hindi. About all of the informants could be contacted through internet, so I send some other related sentences to them for translation. This helped me to clarify some ambiguities and find replies of additional questions/problems that arose during the analysis.

The data gathered during the survey is detailed below.

2.3.1. Urdu/Hindi

Urdu and Hindi are different styles of the same language, which is mainly spoken in Pakistan and India. The syntactical structure of both languages is almost identical, but they differ in script, vocabulary and to some extent phonology and derivational morphology.

Urdu is the national language of Pakistan. According to Grimes & Grimes (2000), there are more than 60 million speakers of Urdu in Pakistan, India and other countries. It is the second or third language of most Pakistanis, for whom it is not the mother tongue. Urdu is written in a modified Arabic script and its vocabulary has borrowed massively from Arabic and Persian.

Hindi is mainly spoken in India. There are more than 360 million first language speakers of Hindi. Nearly 50% of the Indian population use Hindi as a second language (Grimes & Grimes 2000). Hindi is written in the Devanagri script and its vocabulary is mainly of Sanskrit origin.

Urdu/Hindi is an Indo-Aryan language. Most of the data presented in the following examples are from the author, who is a native speaker of Urdu and belongs to Karachi. However, as mentioned in the Acknowledgements, some judgments are asked/verified from other speakers.

2.3.1.1. ne

The clitic *ne* is used as the ergative marker. In Urdu, the ergative marker appears with transitive verbs in perfective form. The example (52) has the verb in perfective form. The

subject of this sentence is marked by the ergative marker. On the other hand, example (53) has the verb in imperfective form. The subject of this sentence is nominative.

(52) laRke=ne daraxt kaaT-aa
Boy.M.SG.OBL=ERG tree.M.SG.NOM cut-PERF.M.SG
'The boy cut the tree.'

(53) laRkaa daraxt kaaT-taa hai
Boy.M.SG.NOM tree.M.SG.NOM cut-IMPF.M.SG be.PRES
'The boy cuts the tree.'

The *ne* marker is also used to mark intention. The following alternation is not accepted by all speakers of Urdu/Hindi, but many speakers do accept these alternations.

(54) a. vo ciix-aa
3SG.NOM scream-PERF.M.SG
'He screamed (despite himself).'

b. us=ne ciix-aa
3SG=ERG scream-PERF.M.SG
'Ram screamed (on purpose).'

(55) a. naadyaa=ko zu jaa-naa hai
Nadya.F.SG=DAT zoo.M.SG go-INF.M.SG be.PRES.SG
'Nadya has/wants to go to the zoo.'

b. naadyaa=ne zu jaa-naa hai
Nadya.F.SG=ERG zoo.M.SG go-INF.M.SG be.PRES.SG
'Nadya wants to go to the zoo.'

The alternation in (54) occurs with a small set of intransitive verbs. The action in (54b) with the ergative subject is considered as more intentional than the action of (54a) that has a nominative subject (Butt 1995). There is a similar difference between the sentences in (55). Kaifi (cf. Naqvi 1998), Butt & King (1991) and Bashir (1999) show that the ergative marker in (55b) expresses the intention of the subject to perform the action. The dative marked subject in (55a) in contrast either has a neutral reading or it shows obligation on the part of the subject.

2.3.1.2. ko

The clitic *ko* is used as accusative and dative marker. It is used to mark some other semantic usages too. See Ahmed (2006) for a detailed discussion on different semantic usages of Urdu/Hindi *ko*.

The accusative *ko* marks the specific and animate object. The relation between specificity and *ko* marking is a wide subject that requires extensive research (cf. Mohanan 1994 for discussion and further references). Here only some examples are presented to show different aspects of the problem. The examples need some explanation that goes beyond animate/unanimate or specific/unspecific distinction.

(56) maiN=ne aik film dekh-ii
1SG=ERG one film.F.SG see-PERF.F.SG
'I saw a film.' (inanimate, unspecific without *ko*)

(57) maiN=ne aik laRkaa dekh-aa
1SG=ERG one boy.M.SG see-PERF.M.SG
'I saw a boy.' (animate, unspecific without *ko*)

(58) maiN=ne bilaal=ko dekh-aa
1SG=ERG Bilal.M.SG=ACC see-PERF.M.SG
'I saw Bilal.' (animate, specific with *ko*)

(59) maiN=ne vo film dekh-ii
1SG=ERG that film.F.SG see-PERF.F.SG
'I saw that film.' (inanimate, specific without *ko*)

(60) maiN=ne us film=ko dekh-aa
1SG=ERG that film.F.SG=ACC see-PERF.M.SG
'I saw that film.' (inanimate, specific with *ko*)

The marker *ko* is also used as dative marker. It marks recipient, experiencer and purpose.

(61) maiN=ne bilaal=ko kitaab dii
1SG.OBL=ERG Bilal.M.SG=DAT book.F.SG.NOM give-PERF.F.SG
'I gave the book to Bilal.' (recipient)

(62) umair=ko iinaam mil-aa
Omair.M.SG=DAT prize.M.SG get-PERF.M.SG
'Omair got the prize.' (recipient as subject)

(63) umair=ko buxaar hai
Omair.M.SG=DAT fever.M.SG be.PRES.SG
'Omair has fever.' (experiencer subject)

(64) umair sair=ko gayaa
Omair.M.SG walk.M.SG=DAT go.PERF.M.SG
'Omair went for a walk.' (purpose)

The marker *ko* is used to mark point of time too. In old Urdu, it is used to mark the goal and possessor, as shown in (66) and (67).

(65) cor raat=ko aa-yaa
thief.M.SG night.F.SG=LOC come-PERF.M.SG
'The thief came at night.'

(66) baap hindustaan=ko cal-aa
father.M.SG India.M.SG=ACC walk-PERF.M.SG
'Father went towards India.' (Dehalvi 1804)

(67) jin=ko caSm-e-biina hai
who=DAT eye-EZ-seeing be.PRES.SG
'Who have vision.' (Online Urdu Dictionary)

2.3.1.3. liye

Urdu/Hindi has different case markers to mark recipient and beneficiary. The benefactive postposition *liye* is used to mark beneficiary and purpose.

(68) maiN=ne laRke=ke liye kitaab xariid-ii
1SG=ERG boy.OBL=GEN BEN book.F.SG.NOM buy-PERF.F.SG
'I bought a/the book for a/the boy.'

(69) umair sair=ke liye gayaa
Omair.M.SG walk.M.SG=DAT BEN go.PERF.M.SG
'Omair went for a walk.'

The purpose usage in the above sentence is the same as the purpose usage of *ko* given in 2.2.1.2. The two markers can alternate freely for this usage.

The above two examples having the case marker *liye* show that it is a postposition (as it follows the genitive marker *ke*). The same form *liye,* however, can be used without the genitive marker in the following usage.

(70) vo is liye aa-yaa thaa ke …
3SG this BEN come-PERF.M.SG be.PST.M.SG CONJ
'He came for the purpose that …'

2.3.1.4. saat^{h}

The postposition *saat^{h}* is used as the comitative marker in Urdu/Hindi. See the following example.

(71) maiN us=ke saath baazaar gayaa
1SG 3SG=GEN COM market.F.SG go.PERF.M.SG
'I went to the market with him.'

2.3.1.5. se

The clitic *se* is used as the ablative and perlative marker.

(72) vo karaacii=se aa-yaa
3SG Karachi=ABL come-PERF.M.SG
'He came from Karachi.'

(73) vo baaG=se guzr-aa
3SG garden=ABL pass-PERF.M.SG
'He passed through the garden.'

The same form is used to mark the instrument, causee, demoted agent of the passive and manner.

(74) us=ne caabii=se darvaazaa khol-aa
3SG.OBL=ERG key=INST door.M.SG open-PERF.M.SG
'He opened the door with the key.' (instrument)

(75) us=ne nokar=se darvaazaa khulvaa-yaa
3SG.OBL=ERG servant=ABL/INST door.M.SG open.CAUS-PERF
'He made the servant to open the door.' (causee)

(76) darvaazaa us=se khol-aa gayaa
door.M.SG 3SG.OBL=ABL/INST open-PERF.M.SG PASS.M.SG
'He opened the door with the key.' (demoted agent)

(77) vo tezii=se bhaag-aa
3SG fastness=INST run-PERF.M.SG
'He ran fast.' (manner)

It is difficult to assign either the instrument or ablative label to these different usages of *se*. To distinguish between the instrument and ablative faces of the marker *se*, marking of

the same usages in the other surveyed languages is considered. The data presented in the following sections show that manner is usually marked by the instrument marker in sister languages of Urdu/Hindi. Hence manner is considered to be marked by the instrument case marker. Similarly, the causee and demoted agent are usually marked by the ablative marker in most of the sister languages. On the other hand, the marker for these usages is usually simply labeled as instrument in the literature (Mohanan 1994, Butt 1995, Bhatt 2003, Ramchand 2008). Hence, in this dissertation both the case labels are used for these usages.

Section 2.3.1.2 showed that the object in Urdu/Hindi can be marked by the accusative marker *ko*. However, our survey shows that there are many verbs whose second-argument is always marked by case markers other than *ko*. Similarly, the marking on the addressee shows interesting patterns too. This and data related to addressee marking shows interesting patterns and so this set of data is also presented for Urdu/Hindi and other languages.

In Urdu/Hindi, the second argument of the verbs like 'fear' and 'fight', and the addressee of the communication verbs 'say' and 'ask' are marked by the marker *se*.

(78) vo mujh=se laR-aa
3SG 1SG.OBL=COM/ABL/INST fight-PERF.M.SG
'He fought with me.'

(79) vo saaNp=se Dar-taa hai
3SG snake=COM/ABL/INST fear-IMPF.M.SG be.PRES.SG
'He fears snakes.'

(80) us=ne mujh=se kah-aa
3SG.OBL=ERG 1SG.OBL=COM/ABL/INST say-PERF.M.SG
'He said to me.'

(81) us=ne mujh=se puuch-aa
3SG.OBL=ERG 1SG.OBL=COM/ABL/INST ask-PERF.M.SG
'He asked me.'

Urdu has the marker *saat*^{h} that is used to mark accompaniment. The marker *se* is not used to mark the along-with usage shown in (71), but it marks the arguments that are involved in the action. In many South Asian languages, these arguments (e.g., the second argument of the verb 'love' and the addressee of the verb 'say') are marked with the comitative

marker. It is the reason why one of the glosses of *se* is comitative .

2.3.1.6. par

The clitic *par* is used as the locative-on marker. The same form is also used to mark the second argument of some verbs, e.g., 'trust'. See the following examples.

(82) kitaab mez=par hai
book table=LOC_on be.PRES.SG
'The book is on the table.'

(83) maiN=ne mez=par kitaab rakh-ii
1SG=ERG table=LOC_on book.F.SG put-PERF.F.SG
'I put the book on the table.'

(84) maiN=ne us=par bharosaa ki-yaa
1SG=ERG 3SG=LOC_on trust.M.SG do-PERF.M.SG
'I trusted him.'

The same form is used to mark the possessor.

(85) us=par bohot qarz hai
3SG.OBL=LOC_on much loan be.PRES.SG
'He has too much loan.'

2.3.1.7. meN

The clitic *meN* is used as the locative-in marker.

(86) bilaal ghar=meN hai
Bilal house=LOC_in be.PRES.SG
'Bilal is in the room.'

(87) bilaal=ne kitaab baste=meN rakh-ii
Bilal=ERG book.F.SG bag=LOC_in put-PERF.F.SG
'Bilal put the book in the bag.'

The same form is used on the subject to show affectedness and possession too.

(88) ghar=meN aag lag gayii
house=LOC_in fire.F.SG stick go.PERF.F.SG
'The house caught fire.' (affected location)

(89) bilaal=meN himmat hai
Bilal=LOC_in courage be.PRES.SG
'Bilal has courage.' (possessor)

2.3.1.8. -e

The inflection *-e* is used with nouns ending on *-aa* to mark a goal location.

(90) ye aik Daakxaanaa hai
this one post_office.NOM be.PRES.SG
'This is a post office.'

(91) bilaal Daakxaane gayaa
Bilal post_office.OBL go.PERF.M.SG
'Bilal went to the post office.'

Compare the examples (91) and (92). As the stem *bazaar* does not end with an *-aa*, the inflection *-e* is not used.

(92) bilaal baazaar gayaa
Bilal market.OBL go.PERF.M.SG
'Bilal went to the market.'

2.3.1.9. tak

The clitic *tak* is used as an endpoint marker, i.e., allative. It is used when the endpoint is not obligatorily included in the path. In example (91), the traveler entered the post office. But, in the following example the traveler may or may not enter the post office.

(93) bilaal Daakxaane=tak gayaa
Bilal post_office.OBL=ALL go.PERF.M.SG
'Bilal went to the post office.'

The clitic *tak* is also used as a discourse marker. See the following example.

(94) maiN=ne us=tak=ko bulaa-yaa hai
1SG-ERG 3SG.OBL=ALL=ACC invite-PERF.M.SG be.PRES.SG
'I have invited even him.' (p.c. CRULP team, Lahore)

2.3.1.10. paas

The postposition *paas* is used for locative-near usage. The same form is also used to mark the possessor.

(95) jhiil=ke paas aik pahaaR hai
lake=GEN LOC_near one mountain be.PRES.SG
'There is a mountain near the lake.' (*paas* marking the location)

(96) bilaal=ke paas aik qalam hai
Bilal=GEN LOC_near one pen be.PRES.SG
'Bilal has a pen.' (*paas* marking the possessor)

2.3.1.11. kaa/kii/ke

The clitic *k-* is used as the genitive marker. The marker inflects based on the gender and number of the head (that follows it). See the following examples.

(97) bilaal=kii kitaab
Bilal.M.SG=GEN.F.SG.NOM book.F.SG.NOM
'Bilal's book.'

(98) bilaal=kaa qalam
Bilal.M.SG=GEN.M.SG.NOM pen.F.SG.NOM
'Bilal's pen.'

The same form is used to mark the subject of a perfect participle.

(99) ye film bilaal=kii dekh-ii hu-ii hai
this film.F.SG Bilal.M=GEN.F.SG see-PERF.F.SG be-PERF.F.SG be.PRES
'This film is seen by Bilal.'

2.3.1.12. davaaraa

Urdu and Hindi use the same case markers for all of the above usages. But there is a difference in the case marking of the demoted agent in the passive construction. In Urdu, the demoted agent is marked by the clitic *se*. Hindi instead uses the postposition *davaaraa*.

(100) raamdavaaraa pustak paRh-ii ga-yii
Ram.M.SG by book.F.Sg read-PERF.F.SG go-PERF.F.SG
'The book was read by Ram.'

The same form is used to mark the means in Hindi too.

(101) aap es_em_es davaaraa bhii vout kar sakte haiN
2.HON SMS by too vote do can.HAB.PL be.PL
'You can vote by SMS too.'

The above data shows different semantic usages of Urdu/Hindi case markers. We find dative-accusative(-locative) usages for the marker *ko* that is common in other Indo-Aryan languages as well. However, the marker *se* presents an interesting example. This ablative-perlative marker has instrument and (marginally) comitative usages too. Hence, there are

many usages that in other languages are marked by four different markers, namely, comitative, instrument, ablative and perlative markers. This will be discussed in some detail in Chapter 3.

2.3.2. Haryani

Haryani, Haryanvii or Bangru is an Indo-Aryan language spoken in Haryana, India. Haryani region is near to the regions where Urdu/Hindi and Punjabi are spoken. The cognates of many Haryani case markers are present in Modern and Old Urdu/Hindi. The data presented below is taken from Singh (1970).

2.3.2.1. nae

The case marker *nae* is used as the ergative case marker.

(102) baalkaaN nae toRye hoN ge
child.PL ERG break-PERF be.SUBJN FUT.PL
'The children might have broken (these).' (Singh 1970:69)

Unlike in the sister languages, the same form is used as the dative and accusative marker.

(103) oh aapn1ii niiNd nae kah rhyaa thaa
3SG self.GEN sleep ACC say PROG be.PST
'He was referring to his sleep.' (Singh 1970:166)

(104) yaah bi raam_pyaarii nae e de diye
this too Ram_Pyari DAT EMP give give.PERF
'Give it too to Ram Pyari.' (Singh 1970:164)

The same form is used as a locative marker and for temporal usage.

(105) aoh aage nae caaly peR-yaa
3SG forward LOC_towards move fall-PERF
'He moved forward.' (Singh 1970:174)

(106) saaNj^{h} nae
morning LOC
'at morning' (Singh 1970:180)

2.3.2.2. kae

The case marker *kae* is used as the accusative and instrument marker.

(107) kute kae maary
dog ACC/DAT hit-PERF
'hit the dog' (Singh 1970:69)

(108) kute nae DanDe kae maary
dog ACC stick INST hit-PERF
'strike the dog with a stick' (Singh 1970:69)

2.3.2.3. khaatary

The case marker *khaatary* is used to express the purpose role.

(109) laakRii kaatan khaatary
wood cut.INF for
'for cutting wood' (Singh 1970:184)

We do not find any positive or negative example in Singh's book that shows whether the same form is used as a benefactive marker or not. However, its cognate *xaatir* is used as a benefactive marker in Urdu.

2.3.2.4. dhorae

The case marker *dhorae* is used as a locative-near marker.

(110) mere dhorae e us=kaa ghar sae
1SG.GEN LOC_near PART 3SG=GEN house is
'Her house is near me. /She lives nearby.' (Singh 1970:164)

The same form is used to mark the goal, possessor and intermediate agent/causee.

(111) aoh jaaTn1ii dhorae gayaa
3SG Jat's_wife ALL went
'He approached the Jat wife.' (goal) (Singh 1970: 186)

(112) raam dhorae ghoRaa sae
Ram LOC_near horse is
'Ram has a horse.' (Possessor)

(113) bann1i=ne aapn1i chohri dhorae puchwaya
Banni=ERG self.GEN girl through ask.CUAS.PERF
'Banni (grocer's wife) asked him through her daughter.' (causee)
(Singh 1970:162)

2.3.2.5. kaan1hi

The case marker *kaan1hi* is used as a locative-at marker.

(114) phuNglaa kaan1hi
outer-end LOC
'at outer end' (Singh 1970:184)

(115) ao bhaajya bhaamaan kaan1hi
3SG run Brahmin ALL
'He ran to/towards Brahmin.' (Singh 1970:162)

2.3.2.6. taaNhi

The case marker *taaNhi* is used to mark recipient and purpose. Hence we use the label dative in the following examples.

(116) bann;i nae … sao rapie ….. bahman taaNhi diwaa diye
grocer-wife ERG 100 rupees Brahmin DAT give-CAUS give.PERF
'Grocer's wife offered 100 rupees to the Brahmin.' (Singh 1970:162)

(117) raam sut kaatan; taaNhii ruii kharide sae
Ram yarn weave FOR cotton buy is
'Ram buys cotton to make yarn.' (Singh 1970:131)

The same form is used to mark the locative and temporal usages too.

(118) raajaa taaNhi … paoNcya
king ALL reached
'reached the king.' (Singh 1970: 166)

(119) taRke taaNhi
morning ALL
'till the morning' (Singh 1970: 190)

2.3.2.7. taeN

The case marker *taeN* is used to mark the source and second argument of a speech verb.

(120) yeh mire pihar taeN aye saeN
this.PL 1P.GEN parents ABL come are
'These are from my parents.' (Singh 1970: 216)

(121) raajaa taeN bolyii
king.M.Sg to say.perf.F.Sg
'(She) said to the king.' (Singh 1970: 216)

2.3.2.8. saathy

The case marker *saathy* is used as a comitative marker. The same form is used to mark some verbs like 'meet' etc.

(122) apn1e dhan;ii kii saathy
self.GEN husband GEN COM
'with her husband' (Singh 1970: 206)

(123) raja nae us=kii saathy byaah kery liyaa
kingERG 3SG=GEN COM marriage do take.PERF
'The king married her.'

The survey of Haryani case markers provides some interesting observations. Many of these case markers have cognates in Punjabi, Modern Urdu/Hindi and Old Urdu/Hindi. The Haryani ergative marker *nae* has cognates in Urdu/Hindi and Punjabi. However, in Haryani, it is used as dative, accusative and locative marker too. This is a rare example of polysemy, but we find in section 2.4 that it sheds light on the origin (and grammaticalization) of case markers.

2.3.3. Punjabi

Punjabi is an Indo-Aryan language spoken in Punjab, Pakistan and Punjab, India. Punjabi is usually written in the Gurumukhi script. In Pakistan, Punjabi is usually written in a modified Persio-arabic script that is sometime called Shahmukhi.

The survey for Punjabi was conducted in Lahore, Pakistan. As Lahore is a cosmopolitan city, the families of the informants belong to different areas of (Pakistani) Northern Punjab. Beside the informants, some newspapers, novels and short stories, e.g., Yad (1997) and Ahmed (2005), are consulted too.

2.3.3.1. ne

The case marker *ne* is used as the ergative marker. The subject of the transitive verb with perfect aspect is marked by the marker *ne*.

(124) haamid=ne kitaab paRh-ii
3SG=ERG book.F.SG read-PERF.M.SG
'Hamid read the book.

2.3.3.2. nuuN

The clitic *nuuN* is used as the accusative marker. It marks the specific and animate object.

(125) jamiil=ne haamid=nuuN vekhii-yaa
Jameel=ERG Hamid=ACC see-PERF.M.SG
'Jameel saw Hamid.'

The same form is used to mark the recipient (dative marker) and the addressee of communication verbs.

(126) jamiil=ne haamid=nuuN kitaab ditii
Jameel=ERG Hamid=DAT book.F.SG give.PERF.F.SG
'Jameel gave Hamid the book.' (recipient)

(127) jamiil=ne haamid=nuuN aakh-yaa
Jameel=ERG Hamid=DAT/ACC say-PERF.F.SG
'Jameel said to Hamid.' (addressee)

(128) jamiil=ne haamid=nuuN/koloN savaal puc-yaa
Jameel=ERG Hamid=DAT-ACC/ABL question ask-PERF
'Jameel asked Hamid a question.' (addressee)

2.3.3.3. laii

The postposition/clitic *laii* is used as a benefactive marker.

(129) haamid=ne mere laii kitaab xariid-ii
Hamid=ERG 1SG.GEN BEN book.F.SG buy-PERF.F.SG
'Hamid bought a book for me.'

Section 2.2 discussed definitions and differences between a clitic and a postposition in the South Asian context. A postposition is preceeded by a genitive marker, while a clitic occurs without the genitive marker. The behavior of the marker *laii* (and many other Punjabi, Saraiki and Sindhi markers) is a hybrid of these two types of case markers. The case maker *laai* needs mandatory genitive form of the pronoun before it. But for nouns, the preceding genitive marker is optional, i.e., both of the following is allowed.

(130) [viir laii] / [viir=de laii]
brother BEN brother=GEN BEN
'for the brother'

2.3.3.4. -oN

The inflection *-oN* is used as an ablative marker.

(131) maiN g^har-oN aa-yaa
3P.M.SG house.M.Sg-abl come-PERF.M.SG
'I came from the house.'

2.3.3.5. toN

The clitic/postposition *toN* is used as ablative and perlative marker.

(132) main g^har toN aa-yaa
3P.M.Sg house.M.Sg ABL come-PERF.M.SG
'I came from the house.'

(133) haamid baaG toN nangi-aa
Hamid garden PER pass.PERF.M.SG
'Hamid passed through the garden.'

The same form is used to mark the intermediate agent/causee and the second argument of some verbs like 'fear'.

(134) o=ne mazduraaN toN ghar ban-vaa-yaa
3SG=ERG laborer.M.PL ABL house.M.SG make-CAUS-PERF.M.SG
'He caused the laborers to make the house.' (causee)

(135) haamid=nuuN saaNp toN Dar lag-yaa
Hamid=DAT snake.M.SG ABL fear.noun stick-PERF.M.SG
'Hamid feared a/the snake.' (second argument)

(136) haamid=nuuN mere toN Dar lag-yaa
Hamid=DAT iSG.GEN ABL fear.noun stick-PERF.M.SG
'Hamid feared me.' (second argument)

The marker *toN* needs a mandatory genitive marked pronoun in (136), but no genitive marker is present before *toN* in (135) when it is marking a noun.

The marker *toN* is not the only ablative/perlative marker in Punjabi. Other case markers, e.g., *naaloN*, *valoN*, *koloN* and *vicoN,* are used as ablative/perlative markers. These markers are used for specific semantic usages, while *toN* is used for all the usages.

For example, *vicoN* is used to show passing through the inside of a location. Similarly, the marker *naaloN* is used for comparisons.

(137) haamid baaG vicoN/toN nangi-aa
Hamid garden PER/PER pass.PERF.M.SG
'Hamid passed through the garden.'

(138) o haamid naaloN/toN lambaa ae
3SG Hamid ABL/ABL tall be.PRES.SG
'He is taller than Hamid.'

These ablative/perlative markers consist of more than one element. The markers *toN*, *vicoN* and *naaloN* are formed as:

toN = *te* (locative-on) + -*oN*

naaloN = *naal* (locative-near) + -*oN*

vicoN = *vic* (locative-in) + -*oN*

2.3.3.6. naal

The case marker *naal* is used as comitative and instrument marker.

(139) maiN o=de naal baazaar gayaa
1SG 3SG=GEN COM market.M.SG go.PERF.M.SG
'I went to the market with him.'

(140) o=ne caabi naal buuhaa k^{h}ol-iaa
3SG=ERG key.F.SG INST door.M.SG open-PERF.M.SG
'He opened the door with the key.'

The same form is used to mark location, manner and the second argument of some verbs like 'talk' and 'fight', etc.

(141) maiN kampiutar=de naal kitaab rak^{h}-ii
1SG computer=GEN LOC_beside book.F.SG put.PERF.F.SG
'I put the book beside/near the computer.'

(142) maiN o=de naal gal kiit-ii
1SG 3SG=GEN COM talk do.PERF.F.SG
'I talked with him.'

(143) o cetii=naal b^{h}ag-yaa
3SG fast=COM/INST run-PERF.M.SG
'He runs fast.'

2.3.3.7. vic

The case marker *vic* is used as the locative-in marker.

(144) haamid ghar vic ae
Hamid bouse LOC_in be.PRES.SG
'Hamid is in the house.'

2.3.3.8. te

The case marker *te* is used as the locative-on marker.

(145) kitaab mez te ae
book.F.SG table.F.SG LOC_on be.PRES.SG
'The book is on the table.'

The same form is used to mark the second argument of some verbs like 'trust' etc.

(146) zaahid=ne jamiil te bharosaa kitaa
Zahid=ERG jameel LOC_on trust do.PERF.M.SG
'Zahid trusted Jameel.'

2.3.3.9. kol

The case marker *kol* is used as the locative-near marker.

(147) mere ghar=de kol aik baazaar ae
1SG.GEN house=GEN LOC_near one market be.PRES.SG
'There is a market near my house.'

The same form is used to mark the possessor too.

(148) mere kol aik kitaab ae
1SG.GEN LOC_near one book be.PRES.SG
'I have a book.'

2.3.3.10. taaiiN

The case marker *taaiiN* is used as an allative marker.

(149) ae saRak karaacii toN laahor taaiiN jaandii ae
this road Karachi ABL Lahore ALL go.IMPF be.PRES
'This road goes from Karachi to Lahore.'

2.3.3.11. daa/dii/de

The clitics *daa*/*dii*/*de* are used as the genitive marker. It inflects according to the number and gender of the head following it.

(150) laahor=dii saRak
Lahore=GEN.F.SG road.F.SG
'Lahore's road'

The above data about Punjabi case marking shows some interesting observations. It has the example of polysemous use of the same form as locative-comitative-instrument marker. And as compared to a single Urdu/Hindi ablative marker *se,* we find many ablative markers in Punjabi. Many of the markers show hybrid properties of postposition and clitic, e.g., *laii* and *naal*.

2.3.4. Saraiki

Saraiki is an Indo-Aryan language spoken in Southern Punjab, Pakistan. It is written in a modified Arabic script. The data is collected from informants from Multan who are aware of other dialects of Saraiki too.

2.3.4.1. ne

The ergative case marker *ne* marks the subject of transitive verbs with perfective form. It is only used with the noun. The pronouns have an ergative form (as in (152)) that is different from the nominative form (as in (153)).

(151) alii=ne khaan1aa khaad-aa
Ali=ERG meal.M.SG eat-PERF.M.SG
'Ali ate the meal.' (ergative *ne* with transitive verb)

(152) oN khaan1aa khaad-aa
3SG.ERG meal.M.SG eat-PERF.M.SG
'He/She ate the meal.' (ergative pronoun with transitive verb)

(153) o aa-yaa
3SG.NOM come-PERF.M.SG
'He came.' (nominative pronoun with intransitive verb)

2.3.4.2. kuuN

The case marker *kuuN* is used as accusative and dative marker in Saraiki.

(154) yaasir=ne qalam=kuuN taroR-aa
Yasir=ERG pen=ACC break-PERF.M.SG
'Yasir broke the pen.'

(155) alii=ne gohar=kuuN qalam di-yaa
Ali=ERG Gohar=DAT pen.M.Sg give-PERF.M.SG
'Ali gave the pen to Gohar.'

The same form is used to mark the addressees of the communication verb 'say'.

(156) oN alii=kuuN aakh-aa
3SG Ali=DAT/ACC say-PERF.M.SG
'He said to Ali.'

2.3.4.3. kiite

The case marker *kiite* is used to mark the beneficiary.

(157) alii=ne nidaa kiite anghotii xariid-ii
Ali=ERG Nida BEN ring.F.Sg buy-PERF.F.SG
'Ali bought a ring for Nida.'

2.3.4.4. naal

The case marker *naal* is used as locative-beside, comitative and instrument marker.

(158) maiN Teliifuun naal kitaab rakh-ii
1SG telephone LOC_beside book.F.SG put.PERF.F.SG
'I put the book beside/near the telephone.'

(159) maiN alii naal baazaar gayaa
1SG Ali COM market.M.SG go.PERF.M.SG
'I went to the market with Ali.'

(160) oN caabi naal jandaraa khol-aa
3SG.ERG key.F.SG INST door.M.SG open-PERF.M.SG
'He opened the door with the key.'

The same form is used to mark manner (162) and second argument of some verbs like 'talk', 'love' and 'fight', etc. (161).

(161) alii(=ne) nidaa(=de) naal gal kiit-ii
Ali=ERG Nida=GEN COM talk do.PERF.F.SG
'Ali talked with Nida.'

(162) oN jaldii=naal xat likh-aa
3SG hurry=COM/INST letter.M.G write-PERF.M.SG
'He wrote the letter in a hurry.'

2.3.4.5. vic

The case marker *vic* is used as the locative-in marker.

(163) haamid ghar vic he
Hamid bouse LOC_in be.PRES.SG
'Hamid is in the house.'

2.3.4.6. te

The case marker *te* is used as the locative-on marker.

(164) kitaab mez=te he
book.F.SG table.F.SG=LOC_on be.PRES.SG
'The book is on the table.'

The same form is used to mark the second argument of some verbs like 'trust' etc.

(165) gohar=ne alii=te aitbaar kiitaa
Gohar=ERG Ali=LOC_on trust do.PERF.M.SG
'Gohar trusted Ali.'

2.3.4.7. -eN

The case marker *-eN* is used as a locative marker. The following examples show that it is used as a variant of locative-on.

(166) nidaa=de sar-eN dopaTTa he
Nida=GEN head-LOC scarf be.PRES
'Nida has a scarf on her head.'

(167) unde haat-eN hik mundarii pa-vaa-o
3SG.GEN hand-LOC one ring put_on-CAUS-IMP
'Put on the ring on his hand.'

(168) to safr-eN meDe naal aa
2SG journey-LOC 1SG.GEN COM come
'Come on the journey with me.'

2.3.4.8. kol

The case marker *kol* is used as the locative-near marker.

(169) meDe ghar kol hik bazaar he
1SG.GEN house LOC_near one market be.PRES.SG
'There is a market near my house.'

The same form is used to mark the possessor too.

(170) meDe kol hik kitaab he
1SG.GEN LOC_near one book be.PRES.SG
'I have a book.'

2.3.4.9. kan

The case marker *kan* too is used to mark the location 'near' and the possessor.

(171) meDe ghar kan hik bazaar he
1SG.GEN house LOC_near one market be.PRES.SG
'There is a market near my house.'

(172) meDe kan hik kitaab he
1SG.GEN LOC_near one book be.PRES.SG
'I have a book.'

2.3.4.10. taaiiN

The case marker *taaiiN* is used as an allative marker.

(173) ai saRak karaacii tuuN laahor taaiiN veN-dii he
this road Karachi ABL Lahore ALL go-IMPF be.PRES
'This road goes from Karachi to Lahore.'

2.3.4.11. toR

The case marker *toR* is an alternative of the marker *taaiiN*. It is used as an allative marker.

(174) ai saRak karaacii tuuN laahor toR veN-dii he
this road Karachi ABL Lahore ALL go-IMPF be.PRES
'This road goes from Karachi to Lahore.'

2.3.4.12. -uuN

The inflection *-uuN* is used to mark the source of motion.

(175) o multaan-uuN aay-aa
3SG Multan-ABL come-PERF.M.SG
'He came from Multan.'

2.3.4.13. tuuN and other ablatives

The case marker *tuuN* is used as the primary ablative/perlative marker in Saraiki. There are also other ablative markers, e.g., *kinuuN*, *vicuuN, koluuN,* etc., that mark the different semantic usages. However, *tuuN* marks (alternates freely with) all these semantic usages. The source of motion is marked by the marker *tuuN*.

(176) o multaan tuuN aay-aa
3SG Multan ABL come-PERF.M.SG
'He came from Multan.'

No other marker can be used to mark the above primary ablative usage. However, there are other usages that are marked by any specialized ablative marker or *tuuN*.

The perlative, second arguments of verbs like 'fear', the addressee of the verb 'ask' and comparison are marked by the ablative markers.

(177) gohar galii vicuuN/tuuN langh-yaa
Gohar.M.SG street ABL/ABL pass-PERF.M.SG
'Gohar passed through the street.'

(178) nidaa naang kinuuN/tuuN Dar-dii he
Nida.F.SG snake ABL/ABL fear-PERF.F.SG be.PRES
'Nida fears snakes.'

(179) oN alii=tuuN/kinuuN savaal puc-aa
3SG Ali=ABL/ABL question ask-PERF.M.SG
'He asked Ali a question.

(180) maiN salmaan kinuuN/koluuN/tuuN coTaa haaN
1SG Salman ABL/ABL/ABL small be.PRES.1SG
'I am younger than Salman.'

These ablative markers are formed by appending the inflection *-uuN* after various locative markers like *te* (locative-on) and *kan* (locative-near).

The survey of Saraiki case markers shows that many of Punjabi and Saraiki case markers are similar. The Saraiki locative marker *-eN* has some unique spatial relation that is not found in the other South Asian languages surveyed. The language has more than one makers for some usages that can alternate freely.

2.3.5. Sindhi

Sindhi is an Indo-Aryan language spoken mainly in Sindh, Pakistan and some parts of India. In Pakistan, it is written in a modified Arabic script. Some characters (especially retroflex and aspirated characters) in this script are different from the characters in the script used for Urdu, Punjabi and Saraiki.

The following data come from informants and printed texts. The principal informant is from Mirpur Khas. A literature book (Sindhi 1999), some newspapers and two grammar books (Bulchand 1901, Allana 1999) were also used to collect data.

2.3.5.1. Oblique

In Sindhi, there is no distinct ergative maker. If a noun is followed by a case marker, then the noun will be in oblique form. The same form of the noun is used as the ergative form. The subjects of transitive verbs with perfect morphology are in the oblique/ergative form.

(181) cokro huna=khe d1is-e tho
boy.NOM 3SG.F.OBL=ACC see-IMPF.M.SG be.PRES
'The boy saw her.'

(182) huna cokre=khe d1is-yo
3SG.F.OBL boy.OBL=ACC see-PERF.M.SG
'She saw the boy.'

(183) cokre huna=khe d1is-yo
boy.OBL 3SG.F.OBL=ACC see-PERF.M.SG
'The boy saw her.'

In (181), the nominative form of the noun (*cokro*) is used. The oblique form of the third person pronoun *huna* is used before the case marker in (181) and (183) as well as an agent in (182). Similarly, the oblique form of the noun *cokre* is used before the case marker in (182) and as agent in (183). The pronoun *huna* in (182) and the noun *cokre* in (183) are the ergative marked subjects. Moreover, the verb does not agree with the feminine subject of (183).

2.3.5.2. khe

The clitic *k^{h}e* is used as an accusative marker to mark the specific or animate second argument. The animate and specific objects in (182)–(183) are marked by the accusative marker. The same form is used as a dative marker to mark the recipient and experiencer.

(184) huna cokre=khe kitaab d1in1o
3SG.F.OBL boy.M.SG.OBL=DAT book.M.SG.NOM give.PERF.M.SG
'She gave the book to the boy.'

(185) muuN=khe tapo aahe
1SG.OBL=DAT fever be.PRES.SG
'I have fever.' (Bulchand 1901:145)

The same form is used to mark possession and the addressee of the communication verb 'say/tell'.

(186) muuN=khe hikRo kitaab aahe
1SG.OBL=DAT one book be.PRES.SG
'I have a book.' (Bulchand 1901:145)

(187) maan1uu-a=khe b1a TangoN aahin
Man-OBL=DAT two leg.PL be.PRES.PL
'A man has two legs.' (Bulchand 1901:145)

(188) muuN to=khe ciyo
1SG.OBL 2SG=ACC/DAT tell-PERF.M.SG
'I told you.'

2.3.5.3. laae

The case marker *laae* is used as the benefactive marker. It marks beneficiary and purpose.

(189) huna cokre laae kitaab
3SG.F.OBL boyM.SG.OBL BEN book.M.SG.NOM
xariid-o
buy-PERF.M.SG
'She bought the book for the boy.'

(190) muuN=khe paRh-an;aa laae hiko kitabo ghurje
1SG.OBL=DAT read-INF BEN one book want
'I want a book to read.'

The marker *laae* acts as clitic when followed by a noun. But when it is preceded by a pronoun, the genitive marking is mandatory.

(191) asaan=je laae
1PL.OBL=GEN.OBL BEN
'for us'

2.3.5.4. meN

The clitic *meN* is used as a locative-in marker.

(192) Paan1ii piyaali=meN aahe
Water.NOM bowl.OBL=LOC_in be.PRES
'There is water in the bowl.'

2.3.5.5. te

The clitic *te* is used as a locative-in marker. The same form is used to mark the second argument of some verbs like 'trust' and 'blame', etc.

(193) kitaab mez=te aahe
book.NOM table.OBL=LOC_on be.PRES
'The book is on the table.'

(194) hun cokre=te aitbaar ki-o
3SG.OBL boy.OBL=LOC_on trust d-oPERF.M.SG
'He trusted the boy.'

2.3.5.6. taaiiN

The clitic *taaiiN* is used as an allative marker.

(195) hun karaacii=khaaN hedaraabaad=taaiiN saRak thaa-ii
3SG.OBL Karachi=ABL Hyderabad=ALL road build-PERF
'He made a road from Karachi to Hyderabad.'

2.3.5.7. saaN

The clitic *saaN* is used as comitative and instrument marker.

(196) maaN cokre=saaN baazaar vayo
1SG.NOM boy.OBL=COM market go.PERF.M.SG
'I went to the market with the boy.'

(197) darzii keNcii-a=saaN kapRo katre tho
tailor scissor-OBL=INST cloth cut.IMPF PRES.M.SG
'A tailor cuts cloth with a pair of scissors.'

The same form is used to mark the second argument of some verbs like 'meet' and 'love'.

(198) maaN cokre=saaN mil-yo
1SG.NOM boy.OBL=COM meet.PERF.M.SG
'I met with the boy.'

2.3.5.8. vaT

The clitic *vaT* is used as locative-near marker. The same form is used to mark the possessor too.

(199) mez =vaT kursii aahe
table=LOC_near chair be.PRES.SG
'The chair is near the table.'

(200) muuN=vat hikRo kitaab aahe
1SG.OBL=LOC-near one book be.PRES.SG
'I have a book.'

2.3.5.9. -aaN

The inflection -*aaN* is used as an ablative marker.

(201) ho ghar-aaN aa-yo
3SG.M.NOM house.M.SG-ABL come-PERF.M.SG
'He came from the house.'

The marker -*aaN* is used with few words. It can be used with *ghar* 'house', but it cannot be used with the word *kaalij* 'college'.

(202) ho kaalij*-aaN aayo
3.SG.M.NOM college-ABL come.PERF.M.SG
'He came from the college.'

2.3.5.10. khaaN

The clitic *khaaN* is used as an ablative marker.

(203) ho ghar/kaalij=k^{h}aaN aayo
3SG.M.NOM house/college=ABL come.PERF.M.SG
'He came from the house/college.'

The same form is used to mark the causee, the addressee of the communication verb 'ask' and the second argument of some verbs like 'fear'.

(204) huna mazdoraaN=khaaN gharo joR-aa-yo
3SG.M.OBL laborer.M.PL=ABL house.M.SG make-CAUS-PERF.S.SG
'She caused the laborers to build the house.'

(205) muuN cokre=khaaN savaal puc-yo
1SG.OBL boy.OBL=ABL question.NOM ask-PERF.M.SG
'I asked the boy a question.'

(206) maan1uu-a=khaaN na dij1u
Man.PL-OBL=ABL NOT fear
'Do not fear the people.'

2.3.5.11. taaN

The clitic *taaN* is used as the ablative-on marker.

(207) muuN mez=taaN kitaab khaN-yo
1SG.M table.SG=ABL_on book.M.SG take-PERF.M.SG
'I took the book from the table.'

The same form is used in non-spatial contexts too.

(208) Gizaai jinsan=taaN Teksan=jo xaatmoN
food item.PL=ABL_on tax=GEN end
'removal of tax from food items'

2.3.5.12. maaN

The clitic *maaN* is used as ablative-in and perlative-in marker.

(209) ho ii Seher=maaN aa-yo
3SG.NOM this city=ABL_in come-PERF.M.SG
'He came from this city.'

(210) kapRaa peTii-a=maaN b1aahar kaDh
Cloth.PL box-OBL=ABL_in outside take-out
'Take out the clothes from the box.'

(211) ho ii Seher=maaN guzar-yo
3SG.NOM this city=PERL_in pass-PERF.M.SG
'He passed through this city.'

The same form is used in non-spatial contexts too.

(212) uhaaN unhan maN1han=maaN aahiiyo …….
2PL.OBL those man.PL=ABL-in be.PRES.PL
'You are among those people ……'

(213) uhaaN taariix=maaN sabaq na siikh-yo
2PL.Obl history=ABL_in lesson not learn-PERF.M.SG
'You did not learn your lesson from history.'

The ablative clitics in Sindhi are composed of a location and the ablative inflections.

taaN = *te* (locative-on) + -*aaN*

maaN = *meN* (locative-in) + -*aaN*

Similarly, *khaaN* has the inflection -*aaN* at the end and the dative marker *khe* originated from a Sanskrit locative (explained in section 2.4.1). Punjabi and Siraiki have compound ablative markers too. But there is a difference between those markers and Sindhi ablative markers. Sindhi ablative markers cannot alternate freely. One cannot use the marker *khaaN* in place of *taaN* in any example of 2.2.5.11, while Punjabi *taaN* and Saraiki *tuuN* can be used in place of other compound ablative/perlative markers. The only exception in Sindhi is the source of (actual) motion. A city or country can be considered as a point from which the motion starts and hence the marker *khaaN* is used. In reality, the city or country is not a point and the motion actually starts from its inside. Hence, the marker *maaN* is also used to mark the source city or country.

2.3.6. Nepali

Nepali is an Indo-Aryan language spoken mainly in Nepal. It is also spoken in Bhutan and India. There are more than 17 million speakers of Nepali (Gordon & Grimes 2005). It is written in the Devanagri script. The following data is collected from an informant belonging to Nepal.

2.3.6.1. le

The case marker *le* is used as an ergative marker. It marks the subject of transitive verbs in the past tense.

(214) Raam=le kitaab paRh-yo
Ram=ERG book Read-PST.3SG.M
'Ram read a book.'

It marks also the subject of a few intransitive verbs, e.g., *khok-* 'cough', *mut-* 'piss', *hag-* 'pass stool', *nuhaau-*'bathe', *pasinaa kaaDh-* 'sweat', *thuk-* 'spit', etc. (Butt & Poudel 2007). Most of these verbs allow ergative marking in Urdu/Hindi too.

(215) mai=le khok-e
1SG=ERG cough-PST.1SG
'I coughed.'

The same form marks the instrument subject and the instrument (adjunct).

(216) kalam=le lekh-yo
pen=ERG write-PST.3.SG.M
'The pen wrote.' (instrument subject)

(217) us=le camsaa=le bhaat khaa-yo
3SG=ERG spoon=INST meal eat-PST.M.SG
'He ate the meal with spoon.' (instrument)

The Nepali ergative marker also marks the subject of transitive verbs in non-past tense, if an individual level predicate is used (Poudel & Butt 2007).

(218) raam-le angreji jaan-da-cha
Ram-ERG English know-IMPERF-NPST.3.SG.M
'Ram knows English.'

The same form marks the subject of a perfective participle clause.

(219) yo film mai-le her-eko ho
this film 1SG-ERG see-PPART be.NPST.SG
'This film is seen by me.'

2.3.6.2. laai

The case marker *laai* is used as dative and accusative. The accusative marker marks specific and animate objects.

(220) raam=le bhaai=laai piT-yo
Ram.M.SG=ERG brother.M.SG=ACC beat-PERF.M.SG
'Ram beat the brother.'

The same form is used as the dative marker. It marks the recipient, experiencer and possessor.

(221) mai=le gaai=laai ghaaN s de
1SG=ERG cow=DAT fodder give.PST.1SG
'I gave the cow fodder.' (recipient)

(222) ma=laai khoki laag-yo
1SG=DAT cough attach-PST.3SG.M
'I had a cough.' (experiencer)

(223) raam=laai dherai rin cha
Ram=DAT much loan be.NPST
'Ram has much loan.' (possessor)

The same form is used to mark the beneficiary, purpose and addressee of the verb 'say'.

(224) us=le raam=laai ek-Taa kitaab kin-di-yo
3SG=ERG Ram=BEN one-CLF book buy-give-PERF.M.SG
'He bought a book for Ram.' (beneficairy)

(225) u phurmaas=laai das rupiyaaN PaauN-cha
3SG expense=DAT ten rupee.PL get-NPT.3.SG.M
'He gets 10 rupees for expenses.' (purpose)

(226) mai=le us=laai bhan-eN
1SG=ERG 3SG=ACC say-PERF
'I said to him.' (addressee)

2.3.6.3. baaTa

The Nepali marker *baaTa* is used as the ablative/perlative marker.

(227) u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=ALL ran-PST
'He ran from Delhi to Kathmandu.'

The same form is used to mark the causee and second argument of some verbs like 'resign'.

(228) mai=le khetaalaa=baaTa ghar ban-aa-eN
1SG=ERG laborer.PL=ABL house (be) make-CAUS-PST
'I caused the laborers to build the house.'

(229) us=le jaagir=baaTa raajinaamaa di-yo
3SG=ERG job=ABL resignation give-PST.3SG.M
'He resigned from the job.'

2.3.6.4. dekhi

The case marker *dekhi* is used as an ablative marker. It marks the starting point of a static path.

(230) u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=LOC-to ran-PST
'He ran from Delhi to Kathmandu.'

The same form is used to mark the second arguments of some verbs like 'fear'.

(231) u sarpa=dek^{h}i DarauuN-cha
3SG snake=ABL fear-NPST
'He fears snakes.'

2.3.6.5. sanga

The case marker *sanga* is used as a comitative marker.

(232) u ma=sanga bazaar ga-yo
3SG 1SG=COM market go-PST
'He went to the market with me.'

The same form is used to mark possessor, the addressee of the communication verb 'ask' and second arguments of verbs like 'marry' and 'love'.

(233) ma=sanga ek kitaab cha
1SG=COM one book PRES.SG
'I have a book.' (possessor)

(234) us=le ma=sanga sodh-yo
3SG=ERG 1SG=COM ask-PST.3SG.M
'He asked me.' (addressee)

(235) raam=le siita=sanga bıhaa gar-yo
Ram=ERG Sita=COM marriage do-PST.3SG
'Ram married Sita.' (second argument)

2.3.6.6. maa

The case marker *maa* is used as a locative marker. It can be used for locative-in and locative-on usages.

(236) kitaab mez=maa cha
book table=LOC be.NPST
'The book is on the table.'

The same form is used to mark the second argument of some verbs like 'trust' and 'attack'.

(237) ma=laai us=maa viswaas cha
1SG=DAT 3SG=LOC trust be-NPST
'I trusted him.'

The same form is used to mark the inanimate subject of the following sentence.

(238) rukh=maa aago laag-yo
tree=LOC fire stick-PST.3SG
'The tree caught fire.'

2.3.6.7. samma

The case marker *samma* is used as the allative marker.

(239) u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=LOC-to ran-PST
'He ran from Delhi to Kathmandu.'

2.3.6.8. ko/kaa/kii

The clitics *ko*/*kaa*/*kii* are used as the genitive marker. The genitive marker inflects according to the head following it.

(240) siitaa=ko bhaii
Sita.F.SG=GEN.M.SG brother.M.SG
'Sita's brother'

The Nepali data shows that semantic usages of case marking in Nepali are somewhat different from the other surveyed South Asian languages. The ergative marker of Nepali is used with individual level predicates too. It is also used with a perfect participle that usually takes genitive marked subjects in the sister languages. The Nepali ablatives encode static vs. dynamic distinction. The Nepali dative marker *laai* covers both recipient and beneficiary usages that are not found in the other surveyed Modern Indo-Aryan languages. These languages usually have a cognate of *laai* as the benefactive marker that is different than the dative marker for the recipient.

2.3.7. Pashto

Pashto is an Iranian language spoken in Pakistan and Afghanistan. It is written in a modified Arabic script. Pashto has a complex system of case marking having prepositions, postpositions and circumpositions. Different dialects have some differences in the usage of the adpositions. We present the data obtained from Khyber Pakhtoonkhwa (former NWFP), Pakistan.

In the description of Pashto and Balochi case markers we use the terms preposition and postposition for both the clitics and the postpositions. For these languages, we use these terms to show whether the case marker follows or precedes the noun.

2.3.7.1. ta

The postposition *ta* is used as an allative marker.

(241) haGa ma sara bazaar ta laaR
3SG 1SG COM market ALL went
'He went to the market with me.'

The same form is used to mark the second argument of the speech verb 'say/tell'.

(242) jamiil haamid ta uuvii
Jameel Hamid ALL said
'Jameel told Hamid.'

The same form is used to mark the recipient or experiencer subject of certain predicates.

(243) haamid ta xazaanaa milao Swa
Hamid ALL treasure get was
'Hamid got a treasure.'

(244) haamid ta Tuuxay lagiidalay day
Hamid ALL cough suffer is
'Hamid has a cough.'

It is important to note that examples given by Tegey & Robson (1996), who worked on Pashto in Afghanistan, have *ta* as the marker for recipient.

(245) za daa kitaab xapal waror ta warkawam …..
1SG this book my brother ALL/DAT gave
'I am giving this book to my brother …...' (Tegey & Robson 1996:211)

But the survey conducted in Pakistan shows that *la* is used as a recipient marker, as given

in the next subsection.

2.3.7.2. la

The postposition *la* is used as the dative and benefactive marker to mark recipient, beneficiary and purpose.

(246) ma ahmad la kitaab warkoo
1SG Ahmad DAT book gave
'I gave a book to Ahmad.' (recipient)

(247) jamiil haamid la io kitaab waaxast
Jameel Hamid BEN one book bought
'Jameel bought a book for Hamid.' (beneficiary)

(248) jamiil haamid sara milaaviiduu la raGle
Jameel Hamid COM meet BEN came
'Jameel came to meet Hamid.' (purpose)

The same form is used to mark the experiencer subject.

(249) Jamiil la xob ne warzi
Jameel DAT sleep NEG do
'Jameel cannot sleep.'

(250) Jamiil la pa haamid kaar raale
Jameel DAT LOC Hamid anger come
'Jameel got angry with Hamid.'

2.3.7.3. pa

The preposition *pa* is used as the locative-at and locative-on marker.

(251) kitaab pe mez de
book LOC_on table is
'The book is on the table.'

(252) haamid kitaab pa mez kekhod
Hamid book LOC_on table put
'Hamid put the book on the table.'

The same form is used as the perlative marker.

(253) haamid pa baG tiir sho
Hamid PERL garden pass was
'Hamid passed through the garden.'

The same form is used to mark the instrument, manner and second argument of some verbs like 'trust'.

(254) asad paray pa chaara prekaR
Asad rope INST knife cut
'Asad cut the rope with a knife.' (instrument)

(255) jamiil pa teeza manDa kRa
Jameel LOC/INST fast run did
'Jameel ran fast.' (manner)

(256) jamiil pa haamid baawar ooko
Jameel LOC Hamid trust did
'Jameel trusted Hamid.' (second argument)

2.3.7.4. ke

The circumposition *pa-ke* is used as a locative-in marker.

(257) pa kase ke obe Sta
LOC bowl LOC-in water is
'There is water in the bowl.'

In some dialects, the preposition *pa* can be dropped. These dialects drop the preposition part of the circumposition, and have the postposition only.

(258) kase ke obe Sta
bowl LOC-in water is
'There is water in the bowl.'

2.3.7.5. na

The postposition *na* is used with the genitive preposition *da* to mark the source of motion.

(259) haGha [(da) hedaraabaad na] raaGhlay
3SG GEN Hyederabad ABL came
'He came from Hyderabad.'

The same form is used to mark the second argument of some verbs like 'fear' and the addressee of the verb 'ask'.

(260) haamid da maar na yariigii
Hamid GEN snake ABL fear
'Hamid fears snakes.' (second argument)

(261) haamid da jamiil na tapuus ookRo
Hamid GEN Jameel ABL ask did
'Hamid asked Jameel.' (addressee)

2.3.7.6. sara

The postposition *sara* is used as the comitative marker.

(262) haGa maa sara bazaar ta laaR
3SG 1SG COM market ALL went
'He went to the market with me.'

The same form is used to mark the second argument of some verbs like 'love'.

(263) da haamid jamiil sara miina daa
GEN Hamid Jameel COM love is
'Hamid loves Jameel.'

The same form is used as a locative-near marker. It is used to mark the possessor.

(264) zama kor send sara day
1SG.GEN house sea COM is
'My house is near the sea.' (location)

(265) maa sara io kitaab day
1SG COM one book is
'I have a book.' (possessor)

2.3.7.7. da

The preposition *da* is used as the genitive marker.

(266) da jamiil kor
GEN Jameel house
'Jameel's house'

The Pashto data provide some important observations. Like Punjabi and Saraiki, the form used as Pashto comitative marker also has locative-beside usage. Pashto has experiencer/recipients subjects that are marked either by an allative or by a dative marker. It has circumpositions and the preposition part of the circumposition can be dropped in some dialects. The instrument marker of Pashto also acts as the locative marker. This polysemy needs explanation, and I return to this issue in Chapter 3.

2.3.8. Balochi

Balochi is an Iranian language spoken in Pakistan and in some parts of Iran and Afghanistan. In Pakistan, it is spoken in Baluchistan and in some other areas, especially in Karachi (located in Sindh). Different dialects of Balochi differ from each other with regard to the choice of postposition and preposition. The same form is used as preposition in one dialect and as postposition in another dialect. Almost all the data presented here was obtained from a Balochi informant in Karachi. A language-learning book (Baloch 1981) was also consulted, but the informant's choices of the case marker positions (pre- or post-) were given preference.

Balochi is written in a modified Arabic script that is almost identical to the Urdu script.

2.3.8.1. -a

The inflection *-a* is used as ergative, accusative and dative marker. In the example (268), the inflection *-a* is used with both the agent and the recipient.

(267) tafseer-a gilaas poorosh thaa
Tafseer-ERG glass break was
'Tafseer broke the glass.'

(268) jamiil-a tafsiir-a kitaab daat
Jameel-ERG Tafseer-DAT book gave
'Jameel gave the book to Tafseer.'

The same form is used to mark the subject of some intransitive verbs, e.g., 'laugh'. The list of verbs is similar to the unergative verbs of Urdu/Hindi and Nepali that allow/have ergative marker.

(269) jamiil-a kandith
Jameel-ERG laughed
'Jameel laughed.'

The same form is used to mark the goal and addressee of the verb 'say'.

(270) haa baazaar-a Sotha
3SG market-ALL went
'He went to the market.' (Goal)

(271) jamiil-a tafsiir-a gushta
Jameel-ERG Tafseer-ACC/DAT said
'Jameel said to Tafseer.' (Addressee)

2.3.8.2. raa

The postposition *raa* is used as an accusative marker in some dialects of Balochi.

(272) kucik-aa ham-aa jinik-araa dist
dog-OBL that.very girl-ACC see.PAST
'The dog saw this girl.' (Farrel 1995:221 cf. Korn 2009)

In the Sarawani dialect of Balochi, the postpoition *araa* is used as both the ergative and the dative marker.

(273) tafsiir-araa jamiil-araa kitaab daat
Tafseer-ERG Jameel-DAT book gave
'Tafseer gave the book to Jameel.'

2.3.8.3. cii/Saa

The postposition *ce/cii/Saa/Se* is used as the ablative marker.

(274) haa hedaraabaad-a ce hathka
3SG Hyderabad-OBL ABL came
'He came from Hyderabad.'

The same form is used to mark the second argument of some verbs like 'fear' and the addressee of the verb 'ask'.

(275) tafsiir leraan cii tursi
Tafseer wave.PL ABL fear
'Tafseer fears waves.' (second argument)

(276) tafsiir-a jamiil-a cii justh kothaa
Tafseer-ERG Jameel-OBL ABL question did
'Tafseer asked Jameel.' (addressee)

2.3.8.4. sara

The postposition *sara* is used as the locative-on marker.

(277) jamiil kohain sara rawaga hain
Jameel mountain LOC_on climb was
'Jameel climbed on the mountain.'

The same form is used to mark the second argument of some verbs like 'trust' and manner.

(278) jamiil-a tafsiir-e sara barossaa hain
Jameel-DAT Tafseer-OBL LOC_on trust was
'Jameel trusted Tafseer.' (second argument)

(279) jamiil teziyeh sara tatka
Jameel fastness LOC_on ran
'Jameel ran fast.' (manner)

2.3.8.5. guun

The postposition *guun* is used as the comitative and instrument marker.

(280) haa mane guun baazaar-e shotha
3SG 1SG.GEN.OBL COM market-ALL went
'He went to market with me.'

(281) tafsiir-a caabii-e guun darwa pac kotha
Tafseer-ERG key-OBL INST door open did
'Tafseer opened the door with the key.'

The same form is used to mark the second argument of some verbs, e.g. 'love' and 'marry', etc.

(282) aah jamiil-a guun piyaar kaan
3SG.F Jameel-OBL COM love do
'She loves Jameel.'

The case marking data of Balochi shows that it has the same form for comitative-instrument marking. It also has ergative marking in some dialects, but most of its dialects do not have a separate ergative marker. The Sarwani dialect has the same form *-araa* as both ergative and dative marker. Hence, it is the second language (along with Haryani) found in the survey that has the same form as both ergative and dative marker.

2.3.9. Manipuri

Manipuri or Meitei is a Tibeto-Burman language spoken mostly in Manipur, India. It has more than one million speakers (Gordon & Grimes 2005). It is written in the Eastern Nagri script. The data collected in this survey are from a resident of Manipur.

2.3.9.1. naa

The case marker *naa* is used as an ergative marker. It marks the subject of a past/perfective transitive clause.

(283) raam-naa gilaas thugaai-re
Ram-ERG glass break-ANT.REAL
'Ram broke the glass.'

The same form is used to mark the subject of individual level predicates in a past/perfective transitive clause (Poudel 2008a).

(284) carulata-naa inglis ha-i
Carulata-ERG English know-REAL
'Carulata knows English.'

The same form is used to mark the instrument, manner and the subject of a comparative clause (Poudel 2008a).

(285) raam-naa so-naa thong hang-i
Ram-ERG key-INST door open-REAL
'Ram opened the door with a/the key.' (instrument)

(286) mohan kan-naa cel-le
Mohan hard-INST run-ANT.REAL
'Mohan run fast.' (manner)

(287) manaunupaa-naa mayaamba-dagi hennaa waang-i
younger brother-ERG /INST elder brother-from Comp tall-REAL
'The younger brother is taller than the elder brother.' (comparison)

In many languages, the subject of perfect participle is marked by the genitive. In Manipuri the ergative/instrument markers are used to mark the subject of perfect participles.

(288) asi film mohan-naa yeng-kh-re
this film Mohan-ERG see-EVD-ANT.REAL
'This film is seen by Mohan.'

2.3.9.2. damak

The case marker *damak* is used as a benefactive marker.

(289) mohan-naa raam-gi damak lairik amaa lau-bi-re
Mohan-ERG Ram-GEN BEN book one buy-BI-ANT.REAL
'Mohan bought a book for Ram.'

2.3.9.3. daa/taa

The case marker *da/ta* is used as a locative marker.

(290) laairik tebal mathak-taa lai
book table top-LOC COP.REAL
'The book is on the table.'

The same form is used as dative marker to mark the recipient.

(291) tomba-naa caoba-daa laairik amaa pi-i
Tomba-ERG Chaoba-DAT book one give-REAL
'Tomba gave Chaoba a book.'

The same form is used to mark the purpose, the addressee of the verbs 'say' and 'ask', the causee and the object of contact verbs.

(292) nang una-ba-gi damak-taa ai cen-lak-i ne
2SG see-NOM-GEN BEN-LOC 1SG run-DIST-REAL SOL
'I ran here just to see you.' (purpose)

(293) mohan-naa raam-da haa-i
Mohan-ERG Ram-LOC say-REAL
'Mohan said to Ram.' (addressee)

(294) saatra-naa oja-daa wahang hang-i
student-ERG teacher-LOC question ask-REAL
'The student asked question of the teacher.' (addressee)

(295) mohan-naa raam-daa yum saa-han-le
Mohan-ERG Ram-LOC house make-CAUS-ANT.REAL
'Mohan caused Ram to built the house.' (causee)

(296) tomba-naa caoba-daa phu-i
Tomba-ERG Chaoba-LOC beat-REAL
'Tomba beat Chaoba.' (object of contact verbs)

2.3.9.4. dagi

The case marker *dagi* is used as an ablative marker.

(297) maa hederaabaad-dagi laak-i
1SG Hyderabad-ABL come-REAL
'He came from Hyderabad.'

2.3.9.5. gaa

The case marker *gaa* is used as a comitative marker.

(298) maa ai-gaa bazaar cat-lu-i
3SG 1SG-COM market go-DEIC-REAL
'He went to market with me.'

The same form is used as conjunction, when used with all the nouns in a conjunction.

(299) raam-gaa mohan-gaa bazaar cat-le
Ram-COM Mohan-COM market go-ANT.REAL
'Ram and Mohan went to the market.'

The same form is used to mark the second argument of some verbs like 'marry'.

(300) maa-naa mohan-gaa luhong-i
3SG-ERG Mohan-COM marry-REAL
'She married Mohan.'

2.3.9.6. bu

It is an open question whether *bu* is a case marker. Subbarao, Hakacham & Devi (2007) and Chelliah (1990) gloss *bu* as accusative. On the other hand, it does not appear on all of the objects. Poudel (p.c.) does not gloss it as the accusative marker. According to him, there are other (mostly not well understood) semantic reasons for the *bu* marking.

The marker is used after the objects in the sentences that have unexpected events (Poudel p.c.).

(301) tomba=naa ma-paa-bu phu-i
Tomab=ERG 3-father-FOC beat-REAL
'Tomba beat his father.'

The same form is used after the second arguments of these verbs.

(302) maa-naa mohan-bu nungsi-i
3SG-naa Mohan-BU love-REAL
'She loves Mohan.'

(303) Raam-naa lin-bu ki-ja-i
Ram-ERG snake-BU fear-REFL-REAL
'Ram fears snakes.'

There are some interesting points about Manipuri case marking data. The language does not have a proper accusative marker. Objects either have unmarked nominative or are marked by locative-dative/ablative/comitative markers. The locative marker *da*/*ta* has a lot of semantic usages including recipient and object/goal of the contact verbs. Besides marking the agents, the ergative marker has many extra-agentive usages and the role of *bu* remains to be understood properly.

2.3.10. Malayalam

Malayalam is a Dravidian language spoken mainly in Kerala, India. There are more than 35 million speakers of Malayalam (Gordon & Grimes 2005). It is written in a Malayalam script that is derived from the Grantha script. Most of the data presented below is obtained from a Malayalam native speaker from India. A grammar book (Asher & Kumari 1997) was also consulted.

2.3.10.1. -e

The case marker -*e* is used as the accusative marker.

(304) pooliiskaaran kutti-ye aticcu
policeman child- ACC beat- PAST
'The policeman beat the child.' (Asher & Kumari 1997:59)

(305) avan puuccaye enn-e eelpiccu
3SG cat-ACC 1SG-ACC entrust-PST
'He entrusted the cat to me.' (Asher & Kumari 1997:108)

(306) raman-∂ pamb-ine peeDi aaN∂
Raman-DAT snake-ACC fear be-PRES
'Raman fears snakes.'

2.3.10.2. -*ə* /-**kk***ə*

The case marker -*ə* and -*kkə* are used as the dative marker.

(307) hanipha enikka ii pustakam tannu
Hanifa 1-DAT this hook give-PAST
'Hanifa gave me this hook.' (Asher & Kumari 1997:107)

The same form is occasionally used to mark the benefactive usage too (Asher & Kumari 1997:209). It is also used to mark purpose.

(308) avan oolikkə pooyi
3SG work-DAT go-PST
'He went for work.' (Asher & Kumari 1997:217)

The same form is used to mark the subjects that are experiencer, recipient or possessor.

(309) raman-ə cuma uNDə
Ram-DAT cough be.PRES
'Ram has a cough.' (experiencer)

(310) raman-ə nid^{h}i kiTTi
Ram-DAT treasure get-PST
'Ram got the treasure.' (recipient)

(311) Tiiccar-kkə panam uNDə
teacher-DAT money be-PRES
'The teacher has money.' (possessor) (Asher & Kumari 1997:63)

The same form is used to mark the goal.

(312) avan Delhi-kkə / Delhi-yil-ekkə pookayaaNDə
3SG Delhi-DAT/ Delhi-LOC-DAT go.PRES.be.PRES
'He is going to Delhi.'

In Malayalam, there are many examples when more than one case marker appears on the noun. In this example, the dative -*kkə* for marking the goal is preceded by the locative marker -*il*.

2.3.10.3. -veeNDi

The case marker -*veeNDi* is used as the benefactive marker.

(313) mohan raman-ə-veeNDi oru pustagam meeDiccu
Mohan Ram-DAT-BEN one book buy-PST
'Mohan bought a book for Ram.'

As the goal of (312) has a dative marker preceded by the locative marker, the beneficiary of (312) has the benefactive marker that is preceded by the dative marker. This phenomenon seems common in Malayalam, and more examples are available in the following text. However, this phenomenon is not directly related to the basic question of this dissertation, hence I do not investigate the details.

2.3.10.4. -kkoND∂

The case marker *-kkoND∂* is used as the instrument marker.

(314) raaman caabi-kkoND∂ kadag∂ turannu
Raman key-INST door open-PST
'Raman opened the door with the key.'

The same form is used to mark the causee in causative constructions.

(315) Mohan raman-e-kkoND∂ viiD∂ paNiyiccu
Mohan Raman-ACC-INST house make-CAUS-PST
'Mohan caused Raman to build the house.'

It can occasionally be used for comitative usage too (Asher & Kumari 1997:211).

2.3.10.5. -kuuDe

The case marker *-kuuDe* is used as the comitative marker.

(316) avan enDe-kuuDe canda-yil vannu
3SG 1SG-COM Market-LOC go-PST
'He went to the market with me.'

The same form is used as the perlative marker.

(317) janalil kuudDe kaRR∂ varunnunT∂
window.LOC PERL wind come-IMPF-PRES
'A breeze is coming through the window.'

2.3.10.6. -ooD∂

In many of Malayalam grammars, the case marker *-ooD∂* is termed as the sociative marker. As I have used the label comitative marker for the other languages and *-ooD∂* can be used for accompaniment ('along with') usage, we are labeling it as the comitative marker.

The marker is used to mark the addressee of ‘ask’ and ‘say’.

(318) Raaman mohan-ooD∂ coodicu
Raman Mohan-COM ask-PST
‘Raman asked Mohan.’

(319) mohan raaman-ooD∂ paRañña
Mohan Raman-COM say-PST
‘Mohan said to Raman.’

The same form is used to mark the second argument of some verbs like ‘fight’ and ‘love’.

(320) mohan raaman-ooD∂ aDi-kuuDi
Mohan Raman-COM fought
‘Mohan fought with Raman.’

(321) raaman-∂ siita-ooD∂ sneeham uND∂
Raman-DAT Sita-COM love be.PRES
‘Raman loves Sita.’

It is important to note that when the verb *sneehik* ‘love’ is used, the second argument is marked by the accusative marker.

(322) raaman siita-ye sneehik-unnu
Raman Sita-ACC love-PRES
‘Raman loves Sita.’

The case marker *kuuDe* is used to mark the comitative usage, but *-ooD∂* can also be used as the comitative marker that marks the accompaniment (Asher & Kumari 1997:196).

2.3.10.7. -il

The case marker *-il* is used as the locative marker.

(323) ngaan skuuL-il ammuvine kanTu
1SG school-LOC Ammu-ACC see-PST
‘I saw Ammu at school.’ (Asher & Kumari 1997:23)

(324) joosaph hoQalil taamasikkunnu
Joseph hotel-LOC stay-PRES
‘Joseph stays in a hotel.’ (Asher & Kumari 1997:63)

The same form is used to mark the second argument of some verbs like ‘trust’.

(325) raman-∂ mohan-il vishvaasam uNd∂
Ram-DAT Mohan-LOC belief BE-Pres
‘Ram trusted Mohan.’

The goal is marked by the dative marker that is preceded by the locative marker, but some dialects prefer to drop the dative marking.

(326) avan Delhi-yil pookayaaND∂
3SG Delhi-LOC go.PRES.be.PRES
'He is going to Delhi.'

2.3.10.8. -ninn∂

The case marker *-ninn∂* is used as the ablative marker.

(327) avan hederaabaad-il-ninn∂ vannu
3SG Hyderabad-LOC-ABL come-PST
'He came from Hyderabad.'

(328) aa rooD∂ Delhi-yil-ninn∂ bombe vare uND∂
this road Delhi-LOC-ABL Bombay LOC_till be-PRES
'This road goes from Delhi to Bombay.'

2.3.10.9. mutal

The case marker *mutal* is used as the ablative marker.

(329) aa rooD∂ Delhi mutal bombe vare uND∂
this road Delhi ABL Bombay LOC_till be-PRES
'This road goes from Delhi to Bombay.'

The marker *mutal* marks the static origin.[2] Nothing moves from the starting point of the road in the above sentence. It only indicates the origin of a line. The marker *mutal* does not mark the source of actual motion. Hence the following sentence is ungrammatical.

(330) *avan hederaabaad mutal vannu
3SG Hyderabad ABL come-PST
'He came from Hyderabad.'

2.3.10.10. puRat̲∂

The case marker *puRat̲∂* is used as a locative-on marker.

(331) avan pustagam meesha-puRat∂ vaccu
3SG book table-LOC_on put-PST
'He put the book on the table.'

[2] I am thankful to Tara Mohanan and P. Madhavan for discussion related to the marker *mutal*.

The study of case marking in Malayalam raises some interesting points. Malayalam is the only language in this survey that does not have ergative marking or an ergative construction. The language has two ablative markers that encode fine-grained distances. There are examples when a single usage, e.g., accompaniment, can be marked by more than one case marker.

2.3.10.11. Summary

The survey of ten South Asian languages has provided a lot of data about the use and distribution of case markers in South Asian languages. The patterns of polysemy, related case markers and patterns of usage across verb classes raise many issues for further analysis. As the main focus of this dissertation is on space and spatial markers, patterns related to usages of spatial expression are pulled together in section 2.5 and discussed briefly. The basic patterns of usage identified through the above survey are then discussed in more detail in Chapters 3 and 4. But before analyzing and arranging the data of the current section, we provide a brief look at the history and origin of the case markers as far as we have been able to trace them. In particular, the next section explores the spatial origin of some non-spatial markers.

2.4. Origin and cognates of South Asian case markers

In the introduction the question was posed whether the case markers used in the languages surveyed originate from spatial terms. A look at what is known about the history of these case markers shows that many case markers indeed originate from spatial terms. The diachronic evidence is supported by comparative synchronic data across sister languages. In particular, many of the South Asian case markers share a common origin. In some cases, a spatial marker in one language has a cognate in another language, but in this language the marker is used on non-spatial arguments. This section goes through a few examples that illustrate the common spatial origin of different non-spatial markers.

2.4.1. ko/kuuN/khe/kii

The dative/accusative marker in many Indo-Aryan languages begins with *k/k*^{h}. The case marking data presented above shows the dative/accusative marker of three languages, i.e., Urdu/Hindi *ko*, Saraiki *kuuN* and Sindhi *k*^{h}*e* begin with *k/k*^{h}. Similarly, Bengali *ke* and

Oriya *ku* fall into the same category. We can thus ask the question: Do all these markers share a common origin and what is that common origin?

Beames (1872) proposed that the origin of these case markers is Sanskrit *kakSa* 'armpit, side'. The locative of *kakSa* is *kakSe* which means 'in the armpit', 'at the side'. In the early stages of Modern Indo-Aryan languages, *kakSa* became *kakha*. Its accusative was *kaakham*. After a series of changes, it changed into the modern day case markers. We find the use of an intermediate form *kahun* in Chand Bardai's (1149– ca. 1200) work (Beames 1872). The marker is used to mark the recipient, purpose and object of the verb 'seek'. We find the use of this old form *kahun* in the work of Tulsi Das (1532–1623) too.

In most of the Modern Indo-Aryan languages, the accusative/dative markers starting with *k*/*k*^{h} have lost their spatial usages. But, we find some examples of spatial usages in synchronic and diachronic data. Bengali has a locative marker *kaache* that means 'near'. The marker is also derived from the same Sanskrit root *kakSe* 'in the armpit' or 'at the side'. In Old Urdu/Hindi, the accusative/dative *ko* is used to mark spatial relations too. See the following examples.

(332)	us	simt	ko	cal-aa
	that	direction	towards	move-PERF.M.SG
	'(He) moved to that direction.'		(Dehalvi 1804)	<Old Urdu/Hindi>

(333)	ab	vatan	ko	jaa-taa	huuN
	now	home-country	to	go-IMPF.M.SG	be.PRES.1SG
	'Now (I) go to (my) country.'			(Dehalvi 1804)	<Old Urdu/Hindi>

In addition, a non-standard dialect of Nepali uses *khaiN* for locative and speech object usages. (Poudel p.c.)

All of these examples are from Indo-Aryan languages. However, the *k/kh* marker is also found in Pashto, an Iranian language. The Pashto postposition *kii* is used as a locative-in marker. Hewson & Bubenik (2006) mentioned that a dialect of Pashto uses *khe* for this locative usage. They proposed that it is derived from the Avestan word *kaSa* 'armpit', which is a cognate of the Sanskrit *kakSa* from which the many dative/accusative markers of modern Indo-Aryan languages seem to be derived.

2.4.2. liye/laae/laaii

Urdu/Hindi, Sindhi and Punjabi use *liye*, *laae* and *laii,* respectively, as a benefactive marker to mark the beneficiary and purpose. In Nepali, *laai* is used to mark dative/accusative and the beneficiary. Similarly, *la* is used as a dative/accusative marker in Marathi. Kellog (1893) suggested that all these forms originate from the Sanskrit past participle *lagya* whose root is *lag* 'to be attached'. Hoernle (1880) suggested that these forms derive from Prakrit locative *laaiahuN*. This locative is itself derived from the Sanskrit locative *labdhe* 'for the benefit of'. Both of these theories have a point in common that these modern case markers originated from an old locative.

2.4.3. ne/nae/nuuN

The case marker *ne* is used as an ergative marker in Urdu/Hindi, Punjabi and Marathi. It is used as dative/accusative marker in Rajasthani and Gujarati. Hence, the same form is used as an ergative marker in some languages and as a dative/accuative marker in other languages. Butt (2005) proposed that these markers along with Punjabi dative/accusative *nuuN* came from the same origin. In Haryani, we find ergative, dative/accusative and locative usages of *nae*. The same use of the same form for both ergative and dative/accusative usage supports the claim that *ne/nuuN* markers in other languages too have a common origin.

There are many theories about the origin of these markers. Trumpp (1872) suggested that the ergative marker *ne* originates from the Sanskrit instrument inflection *-ina*. Beames (1872) and many other authors do not agree with this idea because Sanskrit instrument had been lost long before the emergence of the ergative marker in Modern Indo-Aryan languages, especially Urdu/Hindi.

Beames (1872) and Kellog (1893) instead proposed that these *n-* markers have the same origin as dative/benefactive *l-* markers discussed above (section 2.4.2). They argue that due to a phonological change 'l' became 'n'. The presence of the Nepali ergative marker *le* in addition to the ergative marker *ne* in Urdu/Hindi and many other languages strengthen this claim. According to Kellog (1893), Nepali *le*, Urdu/Hindi *ne* and other markers are derived from the Sanskrit root *lag* meaning 'to be attached'. Its active past

participle is *lagya* that became *laggio* in Prakrit, from which the various ergative and dative/accusative markers are supposed to be derived.

Butt (2005) instead proposed that these *n-* markers originated from Sanskrit locative *janiye* or *janiyaa* 'for the sake of'/'because of'. These two different meanings of *janiye* explain why similar forms are used as either ergative or dative/accusative markers in different languages. The 'for the sake of' reading is postulated to be responsible for the origin of dative markers. This is supported by the fact that in Modern Bengali, *jono* is used as a benefactive postposition, which also still simultaneously means 'because of'. This 'because of' reading of this locative is argued to be responsible for the origin of ergative markers. This hypothesis is supported by the data from Haryani, where the same form *nae* is used as both dative and ergative marker.

Both of these theories have a point in common that these modern case markers originated from an old locative. This provides an evidence for my claim that many core case markers of the studied Indo-Aryan languages originated from spatial terms.

2.4.4. dhorae/davaaraa

The postposition *dvaaraa* is used in Hindi, Nepali and some other Indo-Aryan languages. It is used to mark the demoted agent in the passive construction. Haryani has a marker *dhorae* that has locative usages. The same form is used to mark the causee too.

The origin of this *davaaraa* can be traced to the Urdu/Hindi word *davaar* that means 'door'.

(334) aap=ke sivaa kis=ke davaar=par jaoN
2PL=GEN except who=GEN door=LOC_on go.SUBJ.SG
'To whose door, except yours, I should go.' (Prem Chand) <Urdu/Hindi>

Hence, the postposition for the demoted agent also originated from a spatial term. This is another evidence of the relation between non-spatial case and space.

2.4.5. kane/kan1hi

There are locative markers in several of South Asian languages that begin with *kan-*. The examples are Haryani *kan1hi*, Saraiki *kan* and Old Urdu/Hindi *kane* (see the examples

presented above in 2.3.4.9 and 2.3.2.5). The postposition *kane* is used as a locative-near marker in Old Urdu. It is also used to mark the recipient and possessor.

(335) cacaa buzurgavaar kane gayaa
uncle honorable LOC go.PERF.M.SG
'Went to the uncle.' (Karbal Katha:106, cf. Narang 2007:240)

(336) tum koN soNp-aa haq kane
2.FAM ACC hand_over-PERF truth DAT
'Handed over you to God.' (Karbal Katha: 99, cf. Narang 2007:240)

According to Insha (1988:306), *kane* is a synonym of *paas* 'near'. The locative usages of *kane* can be found in the poetry of Mir Taqi Mir (1723–1810) too.

The marker originated from Sanskrit locative *karna* meaning 'at the ear' or 'side' (Kellogg 1893). The example (336) shows that *kane* in Old Urdu/Hindi has non-spatial usages too. We find a non-spatial cognate in Nepali as well. In Old Neapli, *kana* is used as the dative marker.

(337) Thakur-kana sabeTTo di
king-DAT flesh_of_goat give.PST
'Gave flesh of goat to the king.' (Poudel 2008b)

Hence, we find that the markers derived from the locative 'at the ear' or 'side' are used as dative markers.

2.4.6. taaiiN/taiiN

In Punjabi, Saraiki and Sindhi, the marker *taaiiN* is used as the allative marker. In Haryani, the marker *taaNhii* has locative and dative usages. In Old Urdu, the marker *taiiN* is used as the dative and locative marker.

(338) laayaa na t^{h}aa tuu aaj taiiN haat^{h} suue teG
take.PERF not be.PST 2SG today ALL hand towards sword
'You did not took your hand towards the sword, till now.' (Dard 1996:42)

(339) mujh taiiN is baat kii kiyaa xabar
1SG.OBL DAT this matter GEN what news
'To me, this matter is not known./ I do not know this matter.'(Insha 1808:44)

According to Kellog (1893), *taiiN* is derived from Sanskrit locative *sthane*. In Sankrit, *sthane* means 'place'.

2.4.7. Summary

The above discussion has demonstrated that many of the modern South Asian non-spatial markers have a spatial origin. More examples exist, for example the connection between instrument markers and locative-beside or perlative markers. The connection between originally spatial terms and non-spatial participant marking is the topic of Chapter 4.

2.5. Polysemous Case markers

The data presented in section 2.3 provide a lot of evidence for polysemous case markers. A single form can be used as the marker of more than one semantic usage. Indeed, a spatial marker cannot only be used for more than one spatial meaning, but it can be used for non-spatial markings as well. In this last section, we pull together the data presented in section 2.3 in a different way and show that the crosslinguistic patterns provide evidence for a classification of verb classes that has not been established as such previously in other work on South Asian languages. In particular, given Levin's (1993) methodology of establishing verb classes by investigating the morphosyntactic marking of the event participants as well as argument alternations, a very systematic crosslinguistic classification of South Asian verb classes emerges. These verb classes furthermore show an interesting, but not random variation in their case marking possibilities.

This distribution of case marking across verb classes and the connection to spatial meanings is presented in the next section and then discussed in more detail in Chapter 4. In addition, we point out some interesting patterns related to multiple usages/polysemy of the same form, which are then dealt with at length in Chapter 3.

2.5.1. Addressee of Communication Verbs

A close look at the crosslinguistic patterns in section 2.3 shows that the addressee of verbs like 'say' and 'ask' are marked by different markers in the different South Asian languages. However, the pattern is not random. The table below shows that the addressees are usually marked by dative, comitative or ablative markers. The semantic reasons for these markings are discussed in detail in Chapter 4.

Table 2.3: Case markers on addressee in different South Asian languages

Language	**Say to**		**Ask (a question)**	
	Marker	**Ref.**	**Marker**	**Ref.**
Punjabi	DAT/ACC	2.3.3.2	DAT/ACC, ABL	2.3.3.2
Saraiki	DAT/ACC	2.3.4.2	ABL	2.3.4.13
Nepali	DAT/ACC	2.3.6.2	COM	2.3.6.5
Manipuri	LOC/DAT	2.3.9.3	LOC/DAT	2.3.9.3
Pashto	ALL	2.3.7.1	ABL	2.3.7.5
Balochi	DAT/ACC	2.3.8.1	ABL	2.3.8.3
Sindhi	DAT/ACC	2.3.5.2	ABL	2.3.5.10
Malayalam	COM	2.3.10.6	COM	2.3.10.6
Urdu/Hindi	ABL/INST/COM	2.3.1.5	ABL/INST/COM	2.3.1.5

2.5.2. Non Canonical Second Argument marking

As already discussed in section 2.2.3.1, in most of the South Asian languages, the object is canonically marked either by the unmarked nominative or by the accusative marker based on semantic reasons. We find many verbs in the surveyed South Asian languages that have a second mandatory argument[3] that is not marked by a canonical object marker. These arguments can be marked by a locative, ablative, comitative or dative. It is interesting that the non-canonical second argument (NCSA) of the same verb is marked by the same non-canonical marker in different languages. The following table shows a verb classification based on patterns of non-canonical subject and second argument marking. Again, the patterns are interesting and not random.

[3] I will not discuss the issue whether these second (non-subject) mandatory arguments are syntactic objects or not. As I am working on the semantics, determining the exact grammatical relation of this argument is beyond the scope of this study. Are these arguments syntactical objects? I leave this question for further inquiry. I cannot give a definite reply to this question because I do not find an agreed upon objecthood test for Urdu.

However, some constructions hint that at least some of these arguments are not syntactic objects, or they do not behave like syntactic subjects. Compare the perfect participles for a transitive verb having an object and a verb with a non-canonical second argument.

(i) a. puuliis= nE gaaRii rok-ii
Police.F.SG=Erg vehicle.F.SG stop-PERF.F.SG
'The police stopped the vehicle.'

b. laRkii saanp=se Dar-ii
girl.F.SG snake.M.SG=ABL fear-PERF.F.SG
'The girl feared the snake.'

(ii) a. [puuliis=kii rok-ii huuii] gaaRii
Police.F.SG=GEN stop-PERF.F.SG be.PERF.F.SG vehicle.F.SG
'The vehicle stopped by the police.'

b. saaNp=se Dar-ii huuii laRkii
snake.M.SG=ABL fear-PERF.F.SG be.PERF.F.SG girl.F.SG
'The girl afraid of the snake.'

In (iia), the syntactic object of (ia) becomes the head noun. However, the non-canonically marked (*se* marked) argument of (ib) does not appear in this position in (iib). It shows that canonical objects and NCSA have some difference in syntactic behavior.

Table 2.4: Classes of NCSA verbs in South Asian languages

Class	Subject Marking	2nd Arg. Marking	Examples
I	NOM/ERG, DAT	ABL	fear
II	NOM/ERG	ABL	ask, beg
III	NOM/ERG	LOC-on/DAT	bless, govern, attack, blame
IV	NOM/ERG, DAT	LOC-on/DAT	trust, doubt, suspect, believe
V	NOM/ERG	COM/DAT	talk, meet, marry
VI	NOM/ERG, DAT	COM	love, hate

All the surveyed languages do not strictly follow this table. For example, the second argument of the verb 'fear' (a Class I verb) has canonical marking in Manipuri in place of the ablative marker. However, if one of the surveyed languages has non-canonical marking on the second argument of 'fear', then it is always the ablative marker. The semantic reasons for non-canonical marking and these classes are discussed in detail in Chapter 4.

2.5.3. Experiencer Subjects

The experiencer subject in South Asian languages has already been written about fairly extensively. For a good introduction to experiencer subjects in South Asian languages, see Verma & Mohanan (1990). This dissertation does not deal with experiencer subjects in any detail. However, some interesting data from Pashto can be highlighted at this point.

In most of the South Asian languages, the dative marker marks the experiencer subject.

(340) muuN=khe tapo aahe
1SG.OBL=DAT fever be.PRES.SG
'I have fever.' <Sindhi>

But, an experiencer subject can be marked by other markers as well. For example, the experiencer subjects are marked by the genitive marker in Bengali.

(341) amar tomake cai
1SG.GEN 2SG.ACC wants
'I need you.' (Klaiman 1980) <Bengali>

However, in Sindhi, Bengali and most of the other South Asian languages, only one marker is used to mark the experiencer subjects. Pashto experiencer subjects are different in this matter. In Pashto, the experiencer (and recipient) subjects can be marked by allative *ta*, dative *la* or genitive *da*.

(342) [haamid ta] Tuuxay lagiidalay day
Hamid ALL cough suffer is
'Hamid has a cough.' <Pashto>

(343) [jamiil la] pa haamid kaar raale
Jameel DAT LOC Hamid anger came
'Jameel got angry with Hamid.' <Pashto>

(344) [da haamid] jamiil sara miina daa
GEN Hamid Jameel COM love is
'Hamid loves Jameel.' <Pashto>

The semantic factors governing the choice of Pashto allative, vs. dative vs. genitive subjects need future investigation.

2.5.4. Possession

The Urdu/Hindi examples in section 2.3.1 show that different locative markers can be used to mark the possessor or express part-whole relationships. These relations are discussed in detail by Mohanan (1994). The following examples point out some patterns of other languages that are not found in Urdu/Hindi and hence not discussed in Mohanan's work.

Alienable possession is usually marked by the locative-near marker in most of the above-discussed languages, but it is marked by the comitative marker in Nepali and Pashto. Compare these examples from Balochi and Pashto.

(345) mani kirra yak kitabi ya
1SG.GEN LOC_near one book is
'I have a book.' (locative subject) <Balochi>

(346) maa sara io kitaab day
1SG COM one book is
'I have a book.' (comitative subject) <Pashto>

Similarly, the dative marker is used in Malayalam, Sindhi and Old Urdu/Hindi for possession and part-whole relationship.

(347) muuN=khe hikRo kitaab aahe
1SG.OBL=DAT one book be.PRES.SG
'I have a book.' <Sindhi>

(348) burhaanulmulk uus koN laRkaa athaa
Burhan-ul-Mulk 3SG.OBL DAT boy be.PAST
'He has a son, Burhan-ul-Mulk.' <Old Urdu> (Shauqi 1564, cf. Jalibi 1971:79)

It is beyond the scope of this dissertation to investigate the semantic reasons for these possessive usages, but from a first look at the data it seems that one could easily extend the explanations and discussions offered in Chapter 4 to explain these non-canonical markings as well.

2.5.5. Spatial markers for Instruments

In the next two sections, we turn to interesting patterns related to polysemous case markers. For example, the instrument marker in South Asian languages has a lot of other usages. The following table shows that instrument markers have three major patterns: ergative-instrument, comitative-instrument and perlative-instrument. Much has already been written about the ergative-instrument markers as part of discussions of ergativity (e.g., see various papers in Plank 1979) and this pattern is not focused on in this dissertation. Rather, Chapter 4 discusses the semantic reasons for comitative-instrument and perlative-instrument markers.

Table 2.5: Multiple semantic usages of instrument markers

Language	Other usages of instrument marker	Reference
Punjabi	LOC-beside, COM	2.3.6
Saraiki	LOC-beside, COM	2.4.4
Nepali	ERG	2.6.1
Manipuri	ERG	2.9.1
Pashto	LOC, PERL	2.7.7
Balochi	COM	2.8.5
Sindhi	COM	2.5.3
Malayalam	-	2.10.
Urdu/Hindi	ABL, PERL, COM	2.1.5

2.5.6. Polysemous Spatial markers

The data in section 2.3 also shows the polysemy of different spatial markers. As described earlier, there are four broad types of spatial relations, i.e., ablative ('from'), perlative ('through'), allative ('to') and locative. Similarly, there are many types of locatives, but locative-in and locative-on are usually the ones analyzed as case markers because these tend to be clitics rather than postpositions (case marker plus a genitive) in the South Asian languages. Usually, the markers for other varieties of locatives are postpositions.

The data shows the following sets of multiple usages of these four types of spatial markers: ablative-perlative (e.g., Urdu/Hindi *se*, Sindhi *maaN*), locative-allative (e.g., Urdu/Hindi *par*, Punjabi *te*) and locative-perlative (e.g., Pashto *pa*).

I discuss these patterns in the next chapter since they do not fall out from currently available models of spatial relations. These South Asian patterns, along with the fact that Nepali distinguishes static vs. dynamic ablatives (*dekhi*/*baaTa)* but other South Asian languages do not, lead us to propose an alternative semantic model, which can account for the South Asian spatial markers and their polysemous behavior illustrated in this chapter.

Chapter 3

Semantic Models for Spatial Markers

3.1. Introduction

The previous chapter presented a survey of case markers in ten South Asian languages. It presented different spatial and non-spatial semantic usages of the same form. This chapter focuses on different spatial usages of the same form, that is, it deals with patterns of polysemous spatial markers identified in the previous chapter. This chapter also provides an answer to the following question posed in Chapter 1.

Question: Can we provide a model that explains different spatial usages of the same form?

The task to find an answer to this question has several parts. First, the basic patterns must be identified. We need to identify different patterns of spatial usages and fine-grained differences between similar spatial markers. The significant patterns have already been briefly discussed in the previous chapter. After having identified the relevant pieces of the South Asian data, we need to investigate whether existing spatial models can explain and model the data. If the existing models fail to explain or model some part of the data, then we need a better model that can model all the aspects of the data. I show that the existing models cannot adequately model the South Asian data and propose an alternative model in section 3.5.

This chapter is structured as follows. It begins with a brief introduction to three possible models for an analysis of spatial markers in section 3.2. These models have been proposed by Kracht (2002), Ostler (1979) and Jackendoff (1990). Kracht's model is selected as a representative of pure spatial models. The reason for selecting the other two models is different. As mentioned in Chapter 1, a further important goal of this dissertation is to investigate the non-spatial usages of originally spatial markers. It is required to find how we can explain the different varieties of arguments in terms of

spatial expressions. It is the reason for the selection of the models proposed by Ostler (1979) and Jackendoff (1990). Both of these models represent a localist approach that models predicates and their arguments in terms of primarily spatial constructs. However, as their models do not confine themselves to spatial concepts, but attempt to provide a general analysis or linking theory which explains the correspondences between thematic roles and argument marking, it is possible that these localist approaches can appropriately explain the use of spatial markers for non-spatial usages. This issue is dealt with in Chapter 4. The current chapter first focuses on the spatial parts of these models.

After a brief introduction to the three existing models in section 3.2, the next section (section 3.3) presents the South Asian data in some detail. Section 3.4 attempts to model the patterns with each of the three models. The issues related to modeling different spatial usages of the spatial markers in these models are discussed and their merits and shortcomings with respect to the South Asian patterns are identified. As a result, a list of problems that need to be solved are identified, resulting in requirements for a better model.

Section 3.5 then presents a new model that solves the problems identified in section 3.2. It shows how the proposed model succeeds in fulfilling the requirements identified in the previous section. Analyses of individual case markers are presented in terms of lexical entries in section 3.6. These lexical entries together with the new model account for the patterns related to polysemous spatial markers found in the South Asian data.

3.2. Literature Review

This section presents a brief introduction of Ostler's (1979), Jackendoff's (1990) and Kracht's (2002) models or systems for spatial markers. In this brief introduction, only those parts of these systems are mentioned that are relevant for the discussion with regard to patterns of multiple semantic usages of the same form in the following sections.

3.2.1. Jackendoff's Model

Jackendoff (1990) based his model on ideas coming from the localist theory of Gruber (1965). Jackendoff's model has three basic conceptual functions: [$_{\text{Event}}$ GO], [$_{\text{State}}$ BE] and [$_{\text{Stay}}$ BE]. All types of spatial and non-spatial predicates can be modeled by these

constructs. Take the example of GO — all of the sentences in (1) are examples of GO in which something moves from one point to another. This GO event can be spatial or it may be an instance of a metaphorical movement. See the examples.

(1) a. The bird flew to the tree. (spatial motion)

b. The inheritance went to Philip. (possession)

c. The light changed to red. (ascription of property)

The first sentence (1a) is an example of a physical motion. The bird moved to the tree. In (1b), the inheritance abstractly moved to Philip. In (1c), the light moved to the red state. The last two sentences have metaphorical motions, but all of these can be modeled by the same construct. The "conceptual structures" (CS) for the sentences of (1) are given below.

(2) a. $[_{Event}\ [GO_{Spatial}\ ([_{Thing}$ bird], $[_{Path}$ TO $[_{Thing}$ tree]])]]

b. $[_{Event}\ [GO_{Poss}\ ([_{Thing}$ inheritance], $[_{Path}$ TO $[_{Thing}$ Philip]])]]

c. $[_{Event}\ [GO_{Ident}\ ([_{Thing}$ light], $[_{Path}$ TO [red]])]]

The difference between the three different types of events (spatial motion, possession and ascription of property) is shown by the subscript on the function GO. In all of the cases, a thing (*bird*, *inheritance* or *light*) traverses a path and reaches a destination marked by TO. As we focus on the spatial usages of case markers in this chapter, hence we discuss only $GO_{Spatial}$ (along with $BE_{Spatial}$) here. The examples of the other paradigms are presented and discussed in the next chapter.

The examples for spatial $[_{State}$ BE] and $[_{Stay}$ BE] are:

(3) a. The bird is in the tree. (State BE function)

b. Harry kept the bird in the cage. (Stay BE function)

The conceptual structures of the above sentences are:

(4) a. $[_{State}\ [BE\ (\ [_{Thing}$ bird] , $[_{Thing}$ tree])]]

b. $[_{Stay}\ [BE\ (\ [_{Thing}$ bird] , $[_{Thing}$ bird])]]

The subscripts "Event", "Thing", "Path" and "Place" are termed "conceptual categories".

These are the semantic parts of speech in Jackendoff's model. The important conceptual categories in the spatial domain are:

(5) a. [PLACE] → [$_{Place}$ PLACE-FUNCTION ([THING])]

b. [PATH] → [$_{Path}$ {TO/FROM/TOWARD/AWAY-FROM}
([{THING/PLACE}])]

c. [EVENT] → { [$_{Event}$ GO ([THING], [PATH])] /
[$_{Event}$ STAY ([THING], [PLACE])] }

d. [STATE] → { [$_{State}$ BE ([THING], [PLACE])] /
[$_{State}$ ORIENT ([THING], [PATH])] /
[$_{State}$ EXT ([THING], [PATH])] }

The category PLACE consists of the place-function and the category THING. The place-function defines a region that is the place of the thing. For example, [$_{Place}$ UNDER ([$_{Thing}$ table])] is the representation of the expression *under the table*. We can use the category PLACE in the category PATH too. See the following sentence and its conceptual structure.

(6) The cat came from under the table.

(7) [$_{Event}$ [GO ([$_{Thing}$ cat], [$_{Path}$ FROM [$_{Place}$ UNDER ([$_{Thing}$ table])]])]]

We have the examples of PATH in (2). For example, in (2a) the thing bird traverses a path. The endpoint (TO) of the path is the thing tree.

There are three functions of STATE: BE, ORIENT and EXT. We have seen the example of BE in (3) and (4). The examples of the others are as follows.

(8) a. The sign points towards New York.

b. The road goes from New York to San Francisco.

The sentence (8a) is an example of the function ORIENT. It shows the orientation of a thing towards a place. The sentence (8b) is an example of the function EXT(end). It shows the spatial spread or extension of the object. The conceptual structures of these sentences are:

(9) a. [$_{State}$ ORIENT ([$_{Thing}$ sign], [$_{Path}$ TOWARDS ([$_{Place}$ New York])])]

b. [$_{State}$ EXT ([$_{Thing}$ road],

[$_{Path}$ FROM ([$_{Place}$ New York]) TO ([$_{Place}$ San Francisco])])]

We know that a preposition can have more than one sense. We saw different spatial usages of the same form in Chapter 2. Similarly, the English preposition *under* has three spatial usages or three senses. See the following sentences.

(10) a. The cat is under the table

b. The cat ran under the table.

c. The mouse ran under the table into the hole.

The structures corresponding to these senses of the preposition *under* are:

(11) a. [$_{Place}$ UNDER ([$_{Thing}$])]

b. [$_{Path}$ TO ([$_{Place}$ UNDER ([$_{Thing}$])])]

c. [$_{Path}$ VIA ([$_{Place}$ UNDER ([$_{Thing}$])])]

Jackendoff's model presents a way to write an abbreviated lexical entry of the preposition that covers the conceptual structures of all of its senses. This is achieved by allowing for a notation for optionality. Hence, we can write the first two senses as the following single entry.

(12) [$_{Path}$ TO ([$_{Place}$ UNDER ([$_{Thing}$])])]

The underlined parts of the above structure are optional. If we omit these optional elements, it becomes equivalent to (11a).

Jackendoff's system can also represent the disjunction or ORing of two elements by putting those into curly brackets. For example, if we want to include the CS of (11c) in the unified lexical entry, the resultant lexical entry will be the following. This lexical entry is compatible with all the three lexical entries of *under* in (11a)–(11c).

(13) [$_{Path}$ {VIA / TO} ([$_{Place}$ UNDER ([$_{Thing}$])])]

Hence Jackendoff's system provides constructs to model different semantic usages of English prepositions. It provides a single lexical entry for the prepositions with multiple usages. These constructs can be used to model the spatial markers of other languages too.

3.2.2. Ostler's Model

Ostler's (1979) system does not primarily focus on the spatial markers, but it is about modeling case using spatial constructs. Like Jackendoff, Ostler also follows the localist approach towards case. Due to this reason, most of the ideas and constructs used in both the systems are similar. I included Ostler's model in our discussion because it has some merits (related to our problem) over Jackendoff's system. After giving a brief introduction to Ostler's system, I present a comparison of the two models.

Ostler followed the idea that all predicates can be modeled as the motion of a theme from an abstract location, i.e., a state, to another abstract location, i.e., another state. For example, the melting of ice is the motion of water from solid state to liquid state. We will return to the representation of these abstract movements in the next chapter.

In the current chapter, I present the examples of physical motion only. Just like Jackendoff's system, different predicates in Ostler's model also can be expressed by BE (for states) and GO (for events). The physical motion (in contrast to abstract motion) is notated by the subscript "posit" for "position". Hence, this chapter discusses BE_{posit} and GO_{posit} only.

Ostler gave four roles that appear in BE and GO constructs. The roles are "source", "theme", "path" and "goal". According to him, copular constructions can be modeled as a theme present at a location. An example for BE_{posit} is:

(14) The book is on the table.

BE_{posit}	Theme	(stationary theme)	book
	Goal	(location)	table

For the events that involve change, the theme does not reside in the location but moves to the goal. The goal is the final location of the theme. See an example of GO_{posit} that involves a physical motion.

(15) The train traveled to New York.

GO_{posit}	Theme	(moving theme)	train
	Goal	(end point)	New York

As the theme moves towards a goal, it needs an optional starting point. The starting point is termed "source".

(16) The train traveled from Detroit to Cincinnati.

GO_{posit}	Theme	train
	Source	Detroit
	Goal	Cincinnati

The last role is "path". All the four roles can be used in a single sentence.

(17) The train traveled across Mid West from Detroit to Cincinnati.

GO_{posit}	Theme	train
	Path	Mid West
	Source	Detroit
	Goal	Cincinnati

How can we determine the role of an argument in a sentence? Ostler said that the theme starts from a source, moves through a path and finally ends in a location that is the goal. Related to this scenario, Ostler gave two tests for determining the role of any argument present in a sentence.

(18) a) An entity x is + Source in "GO_y ...x..."

iff (i) x is distinct from the theme of "GO_y"
(ii) the theme is at x not later than when GO_y ...x...

b) An entity x is + Goal in "GO_y ...x..."

iff (i) x is distinct from the theme of "GO_y"
(ii) the theme is at x not earlier than when GO_y...x….

The definitions of roles in terms of these features are given in Table 3.1.

Table 3.1: Ostler's roles in terms of source and goal features

	Source	**Goal**
Source	+	–
Theme	–	–
Path	+	+
Goal	–	+

After having provided a brief description of Ostler's model, we can now compare it to Jackendoff's model. The first difference is that Jackendoff's system recognizes that a single form can mark more than one spatial usage. It allows for declaring parts of the conceptual structure of a spatial marker as optional. Moreover, it provides an abbreviated notation to write two different lexical entries as a single entry. Ostler does not mention such a notation or requirement for his model.

The other difference is that Jackendoff's model has the concept of both PLACE and PATH for modeling spatial expressions. Many of the spatial models, e.g., Kracht (2002), have a similar concept. We do not find such concepts in Ostler's system. The three roles source, path and goal are related to the (concept) PATH. The static location has the same features as the goal. In Jackendoff's system, we can differentiate between *into* and *onto*. Both of these prepositions have the PATH-function TO, but the PLACE-functions are IN and ON respectively. In contrast, both of these are [–source, +goal] in Ostler's model. There is no feature to distinguish the fine difference between these two types of goals.

Jackendoff's system models more types of predicates than Ostler's system. We have seen the examples of ORIENT and EXT(end) in (8) and (9). There is no parallel construct in Ostler's model. EXT is the static or state counter part of the (dynamic) event that involves PATH. In contrast, Ostler claimed that the role "source" only appears with GO, i.e., with a dynamic situation. He claims that

[+source] → [+dynamic].

The source, path and goal of Ostler's system are equivalent to FROM, VIA and TO, respectively, of Jackendoff's system. Jackendoff's system allows for other types of

PATH, e.g., TOWARDS and AWAY-FROM, etc. Ostler's two-feature system does not allow the modeling of more than three types of path (in addition to theme).

The above comparison of the two systems shows that Ostler's system offers less granularity than Jackendoff's. But, Ostler's system has better granularity in one domain. It decomposes different kinds of path by using the two binary features [±source] and [±goal]. In section 3.5, I argue that the idea of decomposing FROM (source), VIA (path) and TO (goal) into features will help us to model polysemy of South Asian spatial markers.

3.2.3. Kracht's Model

Kracht's (2002) system is concerned with the spatial usages only. He proposed that locative expressions have two layers: "configuration" and "mode". The configuration is the way in which several objects are positioned with respect to one another. Examples of configurations are IN, ON, AT, UNDER and BEHIND, etc. A box can be *in* the car, *on* the car, *under* the car or *behind* the car. The relative position of the box with respect to the car is called the configuration.

The second layer "mode" describes the way in which an object moves with respect to the named configuration. The examples of mode are:

- coinitial (object moves away from the location), e.g., English *from*
- transitional (object enters and leaves the location), e.g., English *through*
- cofinal (object reaches the location), e.g., English *to*
- static (object is at the location), e.g., English *at*

Hence, we can represent the English preposition *in* as follows in Kracht's system.

English *in* : configuration = IN, mode = static

The evidence for these layers is found in the morphology of spatial markers of some languages. For the Finnish spatial markers given in Table 3.2, we can identify the morphemes for configuration and mode. The morpheme *-s* is related to the configuration IN and the morpheme *-l* is related to the configuration ON. Similarly, the morpheme *-ta* is related to the coinitial mode.

Table 3.2: Kracht's analysis of Finnish spatial markers (Kracht 2002)

Config/Mode	Static	Cofinal	Coinitial
Null	-na	-ne	-ta
IN	-s-sa	-s-se	-s-ta
ON	-l-la	-l-le	-l-ta

Hence, Kracht's system has the concept of two layers of spatial expressions just like Jackendoff's system. However, it does not have the localist constructs.

In the following sections I apply all of these models to the South Asian case marking patterns with respect to spatial markers and identify the merits and shortcomings of each of the models. This then leads to the proposal of a more adequate semantic model for spatial markers in section 3.5.

3.3. Peculiarities of South Asian Spatial Markers

Two interesting themes can be identified from the compilation of data in Chapter 2. For one, section 2.5.6 summarized different usage patterns related to polysemous spatial markers. This section revisits the examples of these patterns, but organizes the data differently to bring out the patterns more clearly.

The second interesting theme is the existence of different types of ablative markers in South Asian languages. The fine-grained differences in the usages of these markers are also revisited and discussed in some detail below. The different types of ablatives (discussed in the next section) and the patterns of multiple spatial usages of the same marker constitute the main data to be accounted for in this chapter.

3.3.1. Fine-Grained Distinctions Between Ablatives

It has already been mentioned above that many of the surveyed South Asian languages have multiple ablative markers. There are two categories of these groups of multiple ablative markers. The first category consists of the ablative/perlative markers of Sindhi, Saraiki and Punjabi that are formed by appending the ablative inflection with a locative.

In Sindhi, the ablative marker *-aaN* is used as the ablative inflection (cf. 2.3.5.9). The three ablative clitics *khaaN*, *maaN* and *taaN* are composed of a locative morpheme followed by the ablative marker.

taaN (ablative-on) = *te* (locative-on) + *-aaN*

maaN (ablative-in) = *meN* (loc-in) + *-aaN*

khaaN (ablative) = *khe* (an old locative) + *-aaN*

As mentioned in 2.4.1, the dative/accusative marker *khe* originated from an old locative. It is the first part of the marker *khaaN* that is the main ablative marker. It is used when the motion starts from a (logical/abstract) point. The other two markers are used to show whether the source of motion lies on the location (*taaN*) or inside the location (*maaN*).

(19) a. ho ghar=k^{h}aaN aayo
3SG.M.NOM house=ABL come.PERF.M.SG
'He came from the house.'

b. muuN mez=taaN kitaab khan-o
1SG.M table.SG=ABL_on book.M.SG take-PERF.M.SG
'I took the book off the table.'

c. kapRaa peTii-a=maaN b1aahar kaDh
Cloth.PL box-OBL=ABL_in outside take-out
'Take the clothes out of the box.'

The usages of the compound ablative markers in the above examples are in accord with the semantics of their components. It is important to note that the Sindhi ablative *khaaN* cannot be used in the case of specialized usages of (19b) and (19c).

Punjabi and Saraiki have compound ablative markers as well. Like Sindhi, the Punjabi markers too are composed of a locative followed by the ablative marker *-oN*. See the following examples.

toN (ablative) = *te* (locative-on) + *-oN*

vicoN (ablative-in) = *vic* (loc-in) + *-oN*

The specialized marker *vicoN* is used when the source of motion lies inside the location.

The marker *toN* is not a specialized marker for ablative-on usage. It is a general-

purpose marker that can be used in place of the ablative/perlative-in marker *vicoN*.

(20) havaa pakhe vicon/toN aandi payi ae
air.F.SG fan.M.SG ABL_in/ABL come PROG.F.SG be.PRES
'The wind came from the fan.' <Punjabi>

(21) haamid baaG=vicoN/toN langiaa
Hamid garden=PER_in/PER pass.PERF.M.SG
'Hamid passed through the garden.' <Punjabi>

Hence, unlike the Sindhi ablative *khaaN*, Punjabi ablative *toN* is grammatical with the specialized usages of other markers too.

These Sindhi and Punjabi markers are examples of the first category of a group of multiple ablative markers. The second category of such markers is found in Nepali and Malayalam. There are two ablative markers in Nepali. The difference in the usage of these ablative markers is shown in the following examples.

(22) us=le dilli=dek^{h}i kathmandu=samma baaTo banaa-yo
3SG=ERG Delhi=ABL Kathmandu=LOC_to road make.PST
'He built a road from Delhi to Kathmandu.' <Nepali>

(23) u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=LOC_to ran-PST
'He ran from Delhi to Kathmandu.' <Nepali>

The ablative-static marker *dekhi* marks the starting point of a static path. Hence, the starting point of the road is marked by the marker *dekhi* in (22). The ablative-dynamic marker *baaTa* marks the source of the (actual) motion. Hence, the point from which the runner starts the motion is marked by the marker *baaTa* in (23). Neither of these two markers can be used in place of the other marker. Many of the surveyed languages have a single form for these ablative usages. For example, Urdu/Hindi *se* has these usages.

(24) us=ne dilli=se kathmandu=tak saRak banaa-yii
3SG=ERG Delhi=ABL Kathmandu=LOC_to road make-PERF
'He built a road from Delhi to Kathmandu.' <Urdu/Hindi>

(25) vo dilli=se kathmanDu=tak bhaag-aa
3SG Delhi=ABL Kathmandu=LOC_to ran-PERF
'He ran from Delhi to Kathmandu.' <Urdu/Hindi>

Most of the other surveyed languages, e.g., Sindhi and Punjabi, etc., also behave like this.

However, a parallel to the Nepali example is found in Malayalam. The Malayalam ablative marker *ninnə* can be used as both the ablative-dynamic and the ablative-static marker. Hence, both the source of a dynamic motion (as in (26)) and the starting point of a static path (as in (27)) can be marked by the marker *ninnə*.

In contrast, the ablative marker *mutal* can only be used to mark the starting point of a static path, as in (27). It cannot be used in (26), a context which needs an ablative-dynamic marker.

(26) aa rooDə [Delhi-yil-ninnə / Delhi mutal] bombe vare uNDə
this road Delhi-LOC-ABL / Delhi ABL Bombay LOC_till be-PRES
'This road goes from Delhi to Bombay.'

(27) avan [Delhi-yil-ninnə / *Delhi mutal] vannu
3SG Delhi-LOC-ABL / Delhi ABL come-PST
'He came from Delhi.'

Section 3.4 discusses whether the spatial semantic models presented above can model these two categories of groups of ablative markers.

3.3.2. Ablative and Perlative usages

As already mentioned, in many South Asian languages, the same form, e.g., Urdu/Hindi *se,* is used as the ablative and the perlative marker.

(28) maiN gʰar=se aa-yaa
3SG.M house.M.SG=ABL come-PERF.M.SG
'He came from the house.' <Urdu>

(29) haamid baaG=se guzr-aa
Hamid garden=through pass.PERF.M.SG
'Hamid passed through the garden.' <Urdu>

In Punjabi and Sindhi, the markers ending on *-oN* and *-aaN* respectively are used as the ablative and perlative markers. See the following examples.

(30) havaa pakhe vicoN aandi payi ae
air.F.SG fan.M.SG ABL_in come PROG.F.SG be.PRES
'The wind came from the fan.' <Punjabi>

(31) havaa baari vicoN aandi payi ae
air.F.SG fan.M.SG ABL_in come PROG.F.SG be.PRES
'The wind came through the windows.' <Punjabi>

(32) ho ghar=maaN bhaag-o
3SG home=ABL_in run-PERF.M.SG
'He ran (away) from the home.' <Sindhi>

(33) ho darvaaze=maaN bhaag-o
3SG door=PER_in run-PERF.M.SG
'He ran (away) through the door.' <Sindhi>

A single polysemous marker for ablative-perlative usages is found in many of the surveyed languages, hence it is neither an accident nor a case of homophony. Any spatial model should explain why this pattern occurs in many languages.

3.3.3. Locative and Goal Usages

In most of the surveyed languages, the same form is used to mark the location and the goal. For example, Urdu/Hindi *par* and Pashto *pa* are primarily used to mark static locations, but they are also used to mark goals.

(34) kitaab mez=par hai
book table=LOC_on be.PRES
'The book is on the table.' <Urdu/Hindi>

(35) maiN=ne mez=par kitaab rakh-ii
1SG=ERG table=LOC_on book.F.SG put-PERF.F.SG
'I put the book on the table.' <Urdu/Hindi>

(36) kitaab pa mez de
book LOC_on table is
'The book is on the table.' <Pashto>

(37) haamid kitaab pa mez kekhod
Hamid book LOC_on table put
'Hamid put the book on the table.' <Pashto>

The Urdu/Hindi locative-in *meN*, Punjabi locative-on *te* and many other locatives of South Asian languages behave similarly.

3.3.4. Locative and Perlative Usages

Pashto has a further interesting pattern. The Pashto marker *pa* is used as both the perlative and the locative marker. See the following examples.

(38) kitaab pe mez de
book LOC_on table is
'The book is on the table.'

(39) haamid pa baG tiir sho
Hamid PERL garden pass was
'Hamid passed through the garden.'

As shall be seen in the section 3.4, this and the patterns shown above pose interesting challenges to the currently available spatial models.

3.4. Modeling South Asian Spatial Markers

This section evaluates the three semantic models developed by Ostler, Jackendoff and Kracht, respectively, in light of the fine-grained ablative meanings and the patterns of polysemous markers illustrated above. We show that the models are not able to provide satisfactory analyses of the South Asian patterns. Therefore, an alternative model is proposed, which incorporates some of the successful components of the existing models, but improves on them as well.

3.4.1. Location and Goal

As already mentioned in 3.3.3, many locative markers in the surveyed South Asian languages are used to mark both the static location and the goal. Take the example of the Urdu/Hindi locative-on marker *par* that is primarily used to mark the static location, but is also used to mark goals. See the following examples.

(40) a. kitaab mez=par hai
book table=LOC_on be.Pres
'The book is on the table.' <Urdu/Hindi>

b. maiN=ne mez=par kitaab rakh-ii
1SG=Erg table=LOC_on book put-Perf
'I put the book on the table.' <Urdu/Hindi>

This issue is not specific to Urdu/Hindi *par*. The Punjabi locative-on marker *te*, the Pashto locative marker *pa* and locative markers of some other South Asian languages too use the same form to mark both of these usages. The same issue is present in many European languages as well. In English, the prepositions for static locations, e.g., *in* and

on, behave similarly. The location and goal usages of the preposition *on* are shown in (41) and (42).

(41) a. The book is on the table. (locative)
b. He put the book on the table. (goal)

(42) He jumped on the floor. (locative/goal)

The example in (42) has two meanings. Either the person jumped from a certain place, and landed on the floor. This is the goal usage. Or the person was on the floor and jumped up and down several times on the same location. This is the locative usage.

There is a debate whether we need two lexical entries for these two usages or not. Thomas (2003) and Gehrke (2007) claim that English, German and Dutch prepositions encoding location do not have directionality in their lexical meaning. The evidence of this claim comes from the following sentence.

(43) He swam in the lake. (locative/*goal)

As the verb *swim* is related to a manner of motion, and its lexical entry does not provide directionality or path, hence a goal reading is not allowed for (43). It is in contrast to the verb *jump* that can represent directional motion, hence allows for both location and goal interpretation of the preposition.

Similarly, Urdu/Hindi *par* marks a location. If the direction/path is provided by the verb or the context, then it can mark the goal too, as shown in the example (40b). Otherwise, it allows only the location reading. See the following example. The *par* marked argument cannot be treated as the goal with the verb *doR* 'run'.

(44) vo saRak=par doR-aa
3SG road=LOC_on run-PERF.M.SG
'He ran on the road.' / *'He ran to the road.' <Urdu/Hindi>

The locative marker *par* would have a single entry in Kracht's system too. We can represent the locative *par* as:

par : mode = static, configuration = ON

Similarly, we can represent Urdu/Hindi *par* by the following lexical entry using Jackendoff's model.

(45) [$_{Place}$ ON ([$_{Thing}$])]

The entry does not have a path. However, if a verb, e.g., *rakh* 'put' used in (40b), provides a TO path, then the *par* marked argument is considered as goal.

In Ostler's system, both the location and the goal are described as [–source, +goal]. Hence, there is no modeling problem for the lexical entry of Urdu *par*.

The Urdu/Hindi locative-in marker *meN*, Punjabi locative-on marker *te* and many other location markers of the surveyed languages behave similarly. The use of the same form for location and goal usages does not introduce any problem in any of the three studied spatial models/systems and thus can be modeled straight-forwardly.

3.4.2. Modeling Path and Place

The section 3.3.1 says that there are two categories of groups of multiple ablatives in the surveyed South Asian languages. The first category consists of the ablative markers that differ from each other on the basis of configuration of the source with respect to the location. Example (19) shows the usages of Sindhi ablatives *khaaN* (ablative-at or ablative), *maaN* (ablative-in) and *taaN* (ablative-on). The marker *taaN* is used when the motion starts from the inside of a location. The marker *maaN* is used when the motion starts from the top of a location. The marker *khaaN* is related to a point, or to a place/thing that the speaker abstracts as a point, e.g., 'city' etc.

The lexical entries of these markers in Kracht's system are the following.

(46) a. *taaN* : mode = coinitial, configuration = ON

b. *maaN* : mode = coinitial, configuration = IN

c. *khaaN* : mode = coinitial, configuration = AT

As all of these are ablative markers, the mode part is the same, i.e., coinitial. The difference in these three markers is of configuration. There is a distinct marker for the configuration ON, IN and AT.

Similarly, these markers have different types of PLACE in Jackendoff's system. The conceptual structures corresponding to these markers are given below. All the entries have the same type of PATH, i.e., FROM, and different types of PLACE.

(47) a. *khaaN* : [$_{\text{Path}}$ FROM ([$_{\text{Place}}$ house])]

b. *taaN* : [$_{\text{Path}}$ FROM ([$_{\text{Place}}$ ON ([$_{\text{Thing}}$ table])])]

c. *maaN* :[$_{\text{Path}}$ FROM ([$_{\text{Place}}$ IN ([$_{\text{Thing}}$ box])])]

Unlike Kracht's and Jackendoff's systems, Ostler's system fails to distinguish between these varieties of markers. Provided with the roles/features "source", "path" and "goal", Ostler's system has three different types of PATH (in terms of Jackendoff) or mode (in terms of Kracht), but it does not have the concept of different types of PLACE (in terms of Jackendoff) or configuration (in terms of Kracht). In Ostler's model, all the three Sindhi ablative markers are therefore represented as [+source, –goal]. The system cannot model the varieties of the place associated with this path.

Hence we conclude that any system for modeling South Asian spatial markers should have constructs to account for both the path and the place. A system like Ostler's model will fail to distinguish between different varieties of ablatives found in some of the languages.

3.4.3. Multiple lexical entries for the same form

The above two issues do not involve multiple lexical entries for the same form. However, there are many spatial markers that may have more than one lexical entry. It is due to the polysemous nature of these markers. A single form can be used as the marker for more than one usages. Section 3.3 provided a list of polysemy patterns of South Asian spatial markers.

The first pattern is related to ablative and perlative markers. Relevant examples were presented in section 3.3.2. In many languages, the same form, e.g., the Urdu/Hindi marker *se,* is used for both the ablative and the perlative usages. See the following examples.

(48) a. maiN g^{h}ar=se aa-yaa
3SG.M house.M.SG=ABL come-PERF.M.SG
'I came from the house.' <Urdu/Hindi>

b. haamid baaG=se guzra-aa
Hamid garden=PER pass.PERF.M.SG
'Hamid passed through the garden.' <Urdu/Hindi>

The existence of two usages of the marker *se* implies that we need two lexical entries for it. The multiple entries for the same form can be modeled by Ostler's system. In his system, Urdu/Hindi marker *se* would have the following lexical entries.

(49) Semantic entries of *se* in Ostler's system

se (ablative): [+source, –goal]

se (perlative): [+source, +goal]

There seems to be a potential problem in these entries. The two entries for the same surface form *se* have conflicting values of the feature [goal]. The ablative usage has the feature [–goal] and the perlative usage has the feature [+goal]. Some other South Asian languages (e.g., Sindhi, Punjabi and Nepali, etc.) also have the same form as both the ablative and the perlative marker, and hence introduce the same problem.

Under the assumption that a proliferation of lexical entries for one and the same form is to be avoided, the concept of underspecification can be used to solve this problem. If we consider the feature [goal] to be underspecified for the Urdu/Hindi marker *se*, both the lexical entries can be replaced by a single lexical entry. The underspecified entry of Urdu/Hindi *se* in Ostler's system could then be as follows.

(50) *se* (ablative/perlative) : [+source]

Hence, we have a single lexical entry for the marker having ablative and perlative usages. The next question is whether this is a specific property of the ablative-perlative marker, or whether we can model other polysemous markers in a similar way as well. For example, the Pashto marker *pa* is used as a locative and perlative marker. See the following examples.

(51) a. kitaab pe mez de
book LOC_on table is
'The book is on the table.'

b. haamid pa baG tiir sho
Hamid PER garden cross was
'Hamid passed through the garden.'

In Ostler's system, the lexical entries for the above usages of *pa* would be the following.

(52) Semantic entries for *pa* in Ostler's system

pa (locative): [–source, +goal]

pa (perlative): [+source, +goal]

The locative usage has the feature [–source], while the perlative usage has the feature [+source]. Again, we have a conflict in features. And again, we can decide to underspecify the feature [source]. Hence, Pashto marker *pa* would have the following single lexical entry in Ostler's system.

(53) *pa* (locative/perlative) : [+goal]

The lexical entry is compatible with both the locative and the perlative entry of *pa* given in (52). Hence, it does not seem to introduce any problem for Ostler's system.

Jackendoff's model too tries to present a single lexical entry for multiple usages of a marker. As mentioned in 3.2.1, the parts of a lexical entry can be optional or two elements can be disjunct (ORed) to have a single unified lexical entry. The lexical entries for Urdu/Hindi *se* and Pashto *pa* in Jackendoff's abbreviated notation are the following.

(54) $[_{Path}$ {FROM / VIA } ($[_{Thing/Place}$])] (Urdu/Hindi *se*)

(55) $[_{Path}$ VIA ($[_{Thing/Place}$])] (Pashto *pa*)

In (54), the $[_{Path}$ VIA] part of Pashto *pa* is optional. The locative reading of *pa* does not need it. The lexical entry of Urdu/Hindi *se* in (53) states that the PATH part is not optional. However, it says that either the FROM or VIA reading of *se* can be used.

Hence, Ostler's system as well as Jackendoff's allow for the writing of a single lexical entry each for the ablative-perlative and the locative-perlative polysemous markers. However, Jackendoff uses both disjunction and underspecification, while Ostler uses underspecification only. In contrast, Kracht introduced no means for underspecification or abbreviation/disjunction. In his model, both Urdu/Hindi *se* and Pashto *pa* necessarily have two lexical entries each.

So far, all the three patterns identified in 3.3 have been discussed. However, the issue of multiple ablatives in Nepali and Malayalam remains (section 3.3.1). The difference between the distinct ablatives markers lies in the nature of the path. For example, the difference between Nepali ablative-dynamic *baTaa* and ablative-static *dekhi* is shown in the following examples.

(56) a. us=le dilli=dekhi kathmandu=samma baaTo banaa-yo
3SG=ERG Delhi=ABL Kathmandu=LOC-to road make.PST
'He built a road from Delhi to Kathmandu.' <Nepali>

b. u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=LOC-to ran-PST
'He ran from Delhi to Kathmandu.' <Nepali>

The English paraphrases of both sentences have the same preposition *from.* Similarly, Urdu/Hindi and many other South Asian languages use a single form for both of these usages. However, the Nepali examples have two different markers for these usages. This situation poses a question. Should we consider that the Urdu/Hindi marker *se* too has two distinct semantic usages ablative-dynamic and ablative-static? Do Urdu/Hindi *se*, Sindhi *khaaN* and Punjabi *toN* have to be termed as dynamic as well as static ablative markers? As shown in the next section, the answer to these questions is affirmative. Hence, we need to model multiple semantic usages of these spatial markers too. Can Ostler's and Jackendoff's systems provide a single entry for these two usages of Urdu/Hindi *se*? We cannot answer this question now, because we do not know about the semantic features or constructs for ablative-static markers in both of the systems. The next section therefore discusses this and finds that the ablative-static marker cannot be modeled by any of the three available spatial models. Hence, we cannot find a single lexical entry for polysemous dynamic and static ablative markers in any of the existing spatial models. Here all the three systems fail, and we need a new model or extension/modification in the existing ones.

3.4.4. Fine grained differences in the usages

There are some spatial markers whose semantic usages cannot be modeled by either of the three, i.e., Jackendoff's, Ostler's or Kracht's models. Nepali has two ablative markers. The usage of these markers is shown in (56). The lexical entries for both of these markers are identical in the studied spatial models.

(57) Lexical entries for Nepali ablative *baaTa*/*dekhi* in different spatial models

a. Mode = coinitial, Configuration = null (Kracht)

b. [+source, -goal] (Ostler)

c. [$_{\text{Path}}$ FROM ([Thing/Place])] (Jackendoff)

Kracht's model is based on only two attributes: "mode" and "configuration". Hence it cannot model the fine-grained difference between the two markers. Ostler's model can have more features beside [source] and [goal]. As mentioned in 3.2.2, one such features is [dynamic] that represents the dynamic motion or event. The feature can be used to differentiate static and dynamic ablatives. But, Ostler claimed that the feature [source] is always related to the feature [dynamic], i.e., [+source] → [+dynamic]. However, the semantic usage (starting point of static path) of Nepali marker *dekhi* provides a counter example of this claim. The ablative markers of Urdu/Hindi and many other languages can mark both [+dynamic] and [–dynamic] usages, while Nepali has two different markers for these usages.

The PATH and PLACE constructs of Jackendoff's model cannot distinguish between static and dynamic ablative usages. There is a relevant concept in Jackendoff's model but it is related to the verb and not to the spatial marker. This is the function EXT(end), which has usages similar to the Nepali *dekhi*. See the following two English sentences and their conceptual structures.

(58) a. The train goes from New York to Washington DC.

b. The road goes from New York to Washington DC.

(59) a. [$_{\text{State}}$ GO ([$_{\text{Path}}$ FROM ([$_{\text{Place}}$ New York]) TO ([$_{\text{Place}}$ Washington DC])])]

b. [$_{\text{Event}}$ EXT ([$_{\text{Path}}$ FROM ([$_{\text{Place}}$ New York]) TO ([$_{\text{Place}}$ Washington DC])])]

The above two conceptual structures are similar. The only difference is that (58a) has the function [$_{State}$ GO] that is similar to the usage of the ablative marker *baaTa,* and (58b) has the function [$_{Event}$ EXT] that is similar to the usage of the ablative marker *dekhi*. Although Jackendoff's model contains this concept, we cannot model the difference of Nepali ablative markers in Jackendoff's model because case markers/adpositions do not have functions like [$_{State}$ GO] and [$_{Event}$ EXT]. These functions are associated with lexical entries of verbs, and cannot appear in the lexical entries of case markers.

Hence, none of the three studied spatial models of Kracht, Ostler and Jackendoff, respectively, can model the fine grained difference of Nepali ablative-dynamic *baaTa* and ablative-static *dekhi*. We need new or modified constructs/features to model these markers.

The Malyalam ablative markers *ninn∂* and *mutal* show a similar contrast as illustrated in the examples given in section 3.3.1. The marker *ninn∂* is used for both static and dynamic ablative usages, while the marker *mutal* is used for the static ablative only. These markers also call for some new/modified features or constructs.

3.4.5. Requirements for a better model

The above discussion has compared the merits and shortcomings of Jackendoff's, Ostler's and Kracht's systems to model spatial markers of the surveyed South Asian languages. In light of the above discussion, the following requirements for an ideal spatial model for South Asian markers are proposed.

Requirement 1: The model should be able to represent the fine-grained differences between different varieties of a particular spatial usage.

Requirement 2: The model should have a single lexical entry for a single form. The lexical entry should be compatible with all of the spatial usages of this form.

Requirement 1 is about fine-grained differences. We identified two issues related to this requirement. The compound ablative markers of Sindhi, Punjabi and Saraiki need a model that has constructs or features for place and path. Both Jackendoff's and Kracht's models fulfill this requirement. Ostler's model fails with respect to this requirement. The

second issue is about Nepali and Malayalam static/dynamic ablative markers. None of the three systems can successfully model this difference. Ostler's and Jackendoff's models do provide some relevant constructs/features, but they cannot be used to model these spatial markers.

The second requirement is about having a single lexical entry covering all the semantic usages of a spatial marker. Kracht's model has no provision for unifying two lexical entries. Jackendoff's model has the concept of disjunction and optionality that can be applied to different types of paths that are the main reason for the existence of different semantic usages of the same form. Ostler's system decomposes different paths into simpler features, and makes underspecification possible. It means different paths can more easily be unified by using Ostler's system.

A better system for modeling South Asian spatial markers should incorporate all the successful ideas that came out of the above discussion and should avoid all the areas which caused failure in modeling the South Asian data.

3.5. A New Model for Spatial Representation

The previous section presented different issues related to the modeling of South Asian spatial markers. It identified two major requirements for a better model. As none of the studied spatial systems can model all the identified issues and fulfill the requirements, a new model is introduced here. This model incorporates the merits of the studied spatial models, but avoids their disadvantages

For one, the proposed model allows for underspecified features. Ostler's model has features. The layers "mode" and "configuration" of Kracht's model also can be termed as features. Jackendoff's model has conceptual structures, but some features creep into these structures too. Hence the decision to use a feature based model is taken. The underspecification is used to unify different usages of the same form and obtain a single lexical entry for it.

The main points of this feature-based model are the following.

- Every spatial marker has three primary features: PLACE, PATH and DYN(amic).

- Each of the above features may have a set of features as the value.
- The feature PLACE may have following values.

  null, ON, AT, IN, BESIDE, …
- The feature PATH may have following features as values.

  S(ou)RC(e) : the theme leaves the place

  END : the theme enters the place
- The feature DYN(amic) shows whether an action/activity is performed or it is a static situation.
- The features SRC, END and DYN have a positive (+) or negative (–) specified value or they can be underspecified. The underspecified feature is represented by the feature name only.
- The underspecified PATH feature is shown by a parenthesis around the structure.

The template for a feature-based semantic model entry of the spatial marker is:

[

PLACE null / IN / ON / AT / BESIDE /….

(PATH [SRC +/–, END +/–]),

DYN +/–

]

The feature [PLACE] is similar to the conceptual category PLACE in Jackendoff's system or the layer "configuration" of Kracht's system. The value [null] of the feature [PLACE] is used as the default or neutral location. If a language does not distinguish between locative configurations [IN] and [ON], then the value [null] is used.

The feature [DYN] is relevant for Nepali and Malayalam static/dynamic ablative markers. [DYN +] represents the actual motion away from the source, while [DYN –] represents the static origin of a path. This features remains underspecified for the ablative markers of other languages.

The feature [PATH] is similar to the conceptual category PATH in Jackendoff's system or the layer "mode" of Kracht's system, but it is decomposed into simpler features following Ostler's idea. The features [SRC] and [END] are similar to the features/roles "source" and "goal" of Ostler's system, but the definition of these features is somewhat different. Here, the feature [SRC +] means that the moving object/theme leaves this location, and [SRC –] means the moving object/theme does not leave this location. Similarly, the feature [END] for the goal has a positive value when the moving object/theme enters into the location.

According to this definition, the features [SRC] and [END] may be renamed as [ENTER] and [EXIT], respectively, but we retain the names [SRC] and [END] that are more commonly used, e.g., in Ostler (1979).

The prototypical entries of the important spatial usages are given below. We present the actual lexical entries of some spatial markers in the next section.

(60) Sample entries of pure/abstract ablative, perlative, allative or locative usages

Ablative : [PLACE X, PATH [SRC +, END –], DYN]

Allative : [PLACE X, PATH [SRC –, END +], DYN]

Perlative : [PLACE X, PATH [SRC +, END +], DYN]

Locative : [PLACE X]

The lexical entry of locative usages does not have the PATH component. The lexical entries for the ablative and allative (goal/endpoint) markers are self-explanatory. The entry for the perlative maker has both [SRC] and [END] with a positive value. It is because the moving object enters the perlative marked place and leaves it too. Hence both [SRC] (for leaving the location) and [END] (for entering the location) have positive values.

The ablative, allative and perlative markers show three out of four combinations of the two binary features [SRC] and [END]. We do not find the fourth combination [PATH [SRC –, END –]] among the abstract lexical entries. This feature set represents a marker for a place in which the moving body/theme neither enters, i.e., [END –], nor leaves, i.e.,

[SRC –], this place at the same location. This characteristic cannot be found in any marker that involves PATH. It is another description of the (static) locative in which PATH is not mentioned.

In this model, the notation of underspecification is different from the traditional notation in which the name of the underspecified feature is omitted from the lexical entry. This omission may result in less readability of the lexical entry. A reader may fail to realize that there is an underspecified feature in this entry. The valueless feature solves the problem by making it clear to the reader that an underspecified feature is present in the lexical entry. The feature [DYN] in the example (60) is an underspecified feature.

In this model, the underspecified feature means that all the possible values of that feature are possible. The feature values determine whether the marker can be used for a spatial usages with some particular value or not, i.e., the lexical entries do not provide features to some semantic structure of the whole phrase/clause. The features are there only for the identification of correct usages. Appendix B gives an alternative notation for the above-described model. Some readers may find this alternative notation more readable.

3.6. Accounting for the South Asian Data

The underspecified feature based model for spatial markers can be used to model the South Asian spatial markers. Section 3.2 described the case markers with multiple spatial usages that offer challenges to the existing spatial models. The following discussion gives the lexical entries of these markers in the proposed model. These entries show that the model provides a better solution for the modeling of South Asian spatial markers.

3.6.1. Locative markers

In many South Asian languages, the locative marker, e.g., Urdu/Hindi *par* and Pashto *pa,* are used for both the locative and the goal usages. This means that Urdu/Hindi *par* has the lexical entry [PLACE ON] for both of the following usages.

(61) a. kitaab mez=par hai
book table=LOC_on be.PRES
'The book is on the table.' <Urdu/Hindi>

b. maiN=ne mez=par kitaab rakh-ii
1SG=ERG table=LOC_on book.F.SG put-PERF.F.SG
'I put the book on the table.' <Urdu/Hindi>

The locative marker *par* used in (60b) is marking the goal. However, it does not have the feature [PATH [SRC –, END +]] in its lexical entry. The path showing the goal usage is provided by the verb as described in section 3.4.1.

There is another issue with the modeling of locative markers. Most of the surveyed languages have specific markers for marking locative-on and locative-in usages, e.g., Sindhi has locative-on *te* and locative-in *meN* markers. However, Nepali has a single form *maa* to mark both locative-on and locative-in usages.

(62) a. kitaab mez=te aahe
book.NOM table.OBL=LOC_on be.PRES
'The book is on the table.' <Sindhi>

b. paan1ii piyaali=meN aahe
Water.NOM bowl.OBL=LOC_in be.PRES
'There is water in the bowl.' <Sindhi>

(63) a. kitaab mez=maa cha
book table=LOC be.NPST
'The book is on the table.' <Nepali>

b. paani gilaas=maa cha
water glass=LOC be.NPST
'There is water in the bowl.' <Nepali>

The lexical entries for these markers are the following.

(64) a. Sindhi *te* = [PLACE ON]

b. Sindhi *meN* = [PLACE IN]

c. Nepali *maa* = [PLACE null]

Sindhi has specific markers for locative-on and locative-in usages, hence the locative feature of the lexical entries in (63a) and (63b) have the values [ON] or [IN] corresponding to the feature [PLACE]. The Nepali locative *maa* is used for both the configurations IN and ON. Hence, we use the general feature value [null] in its lexical entry. The locative markers of the other languages can be modeled in a similar way.

3.6.2. Ablative-Perlative marker

Section 3.3.2 showed that many South Asian languages use the same form for the ablative and perlative marker. See the examples of Urdu/Hindi *se,* used as an ablative (48a) and perlative (48b) marker.

(65) a. maiN g^{h}ar=se aa-yaa
3SG.M house.M.SG=ABL come-PERF.M.SG
'He came from the house.' <Urdu/Hindi>

b. haamid baaG=se guzr-aa
Hamid garden=through pass.PERF.M.SG
'Hamid passed through the garden.' <Urdu/Hindi>

The lexical entries for the pure/abstract ablative and perlative usages are the following.

(66) a. Ablative : [PLACE X, PATH [SRC +, END –], DYN]

b. Perlative : [PLACE X, PATH [SRC +, END +], DYN]

The feature [END] is the main difference between these lexical entries. If this feature is considered as underspecified, a single lexical entry for both of these usages is obtained. Hence, the lexical entry for Urdu/Hindi *se* will be the following.

(67) Urdu/Hindi *se* = [PLACE null, PATH [SRC +, END], DYN]

The polysemous ablative-perlative markers are found in Sindhi, Punjabi and some other languages too. The lexical entries for Sindhi and Punjabi ablative/perlative markers are discussed in the following section.

3.6.3. Place and Path

Section 3.4.1 showed that Sindhi, Saraiki and Punjabi has compound ablative/perlative markers that are composed of place and path components. The feature [PATH] of these ablative/perlative markers will be similar to the feature [PATH] of Urdu/Hindi *se* given in (67).

The difference between different ablative/perlative markers is the place component of these markers. See the difference of usages of Sindhi markers *khaaN*, *maaN* and *taaN*.

(68) a. ho ghar=k^{h}aaN aayo
3SG.M.NOM house=ABL come.PERF.M.SG
'He came from the house.'

b. muuN mez=taaN kitaab khan-o
1SG.M table.SG=ABL_on book.M.SG take-PERF.M.SG
'I took the book off the table.'

c. kapRaa peTii-a=maaN b1aahar kaDh
Cloth.PL box-OBL=ABL_in outside take-out
'Take the clothes out of the box.'

The lexical entries of these markers in the proposed model are the following. The difference between the semantic usages of these markers is modeled by using different values of the feature [PLACE].

(69) a. Sindhi *khaaN* = [PLACE AT, PATH [SRC +, END], DYN]

b. Sindhi *taaN* = [PLACE IN, PATH [SRC +, END], DYN]

c. Sindhi *maaN* = [PLACE ON, PATH [SRC +, END], DYN]

Punjabi too has compound ablative/perlative markers. See the following examples of Punjabi *toN* and *vicoN* and their lexical entries.

(70) a. main g^{h}ar toN aa-yaa
3P.M.Sg house.M.Sg ABL come-PERF.M.SG
'I came from the house.'

b. haamid baaG=vicoN/toN langiaa
Hamid garden=PER_in/PER pass.PERF.M.SG
'Hamid passed through the garden.' <Punjabi>

(71) a. Punjabi *toN* = [PLACE null, PATH [SRC +, END], DYN]

b. Punjabi *vicoN* = [PLACE IN, PATH [SRC +, END], DYN]

There is a difference between the Sindhi marker *khaaN* and the Punjabi marker *toN*. The marker *khaaN* cannot be used to mark the specialized usages, i.e., the ablative-on and ablative-in usages of (68b) and (68c). On the other hand, the Punjabi marker *toN* can be used to mark the usages of other markers too. To model this difference, the feature [PLACE null] is used in the lexical entry of the marker *toN* that allows it to be used in the semantic space of the marker *vicoN* too.

3.6.4. Static and Dynamic Ablatives

Nepali has two ablative markers for static and dynamic paths as discussed in sections 3.3.3 and 3.3.4. The following examples show the semantic difference between ablative-static *dekhi* and ablative-dynamic *baaTa*.

(72) us=le dilli=dek^{h}i kathmandu=samma baaTo banaa-yo
3SG=ERG Delhi=ABL Kathmandu=LOC_to road make.PST
'He built a road from Delhi to Kathmandu.' <Nepali>

(73) u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=LOC_to ran-PST
'He ran from Delhi to Kathmandu.' <Nepali>

The feature [DYN] is introduced to deal with static and dynamic paths of the above examples. The positive value of [DYN] is used for *baaTa* that marks the source from which the motion starts. The negative value of [DYN] is associated with *dekhi* that marks the static source, e.g., the starting point of the road in (72). The lexical entries of the Nepali ablative markers are the following.

(74) a. Nepali *baaTa* = [PLACE null, PATH [SRC +, END], DYN +]

b. Nepali *dekhi* = [PLACE null, PATH [SRC +, END –], DYN –]

As the marker *baaTa* has ablative/perlative usages, its lexical entry is similar to the lexical entry of Urdu/Hindi *se* presented in (67). The major difference is that *baaTa* has a feature [DYN +], which ensures that it can only be used to mark the source of (dynamic) motion. Similarly, the ablative marker *dekhi* has the feature [DYN –] that enable it to be used only as the origin of static path. (As a static perlative example has not been found, the marker *dekhi* is considered to have only the ablative usage.)

This static/dynamic contrast of ablative usages is found in Malayalam too. The ablative marker *mutal* can be used only as the ablative-static marker, while the marker *ninnə* can be used as marker for both ablative-static and ablative-dynamic usages.

(75) aa rooDə [Delhi-yil-ninnə / Delhi mutal] bombe vare uNDə
this road Delhi-LOC-ABL / Delhi ABL Bombay LOC_till be-PRES
'This road goes from Delhi to Bombay.'

(76) avan [Delhi-yil-ninn∂ / Delhi mutal*] vannu
3SG Delhi-LOC-ABL / Delhi ABL come-PST
'He came from Delhi.'

The lexical entries for Malayalam ablative markers are the following.

(77) a. Malayalam *ninn∂* = [PLACE null, PATH [SRC +, END –], DYN]

b. Malayalam *mutal* = [PLACE null, PATH [SRC +, END –], DYN –]

The lexical entry of *mutal* is identical to the lexical entry of Nepali ablative-static *dekhi*. As *ninn∂* marks both ablative-static and ablative-dynamic usages, the feature [DYN] is underspecified in its entry.

3.6.5. Locative-Perlative Marker

The Pashto marker *pa* gives another example of polysemy in spatial case markers. It is used for locative and perlative usages.

(78) a. kitaab pe mez de
book LOC-on table is
'The book is on the table.'

b. haamid pa baG tiir sho
Hamid PERL garden pass was
'Hamid passed through the garden.'

The lexical entries of the locative and perlative usages of *pa* are the following.

(79) a. Locative *pa* = [PLACE null]

b. Perlatve *pa* = [PLACE null, PATH [SRC +, END +]]

By combining these two entries, the following unified entry of the locative/perlative *pa* is obtained. This entry has an optional [PATH] feature.

(80) Pashto *pa* = [PLACE null, (PATH [SRC +, END +])]

This lexical entry of Pashto locative/perlative *pa* and the other lexical entries provided above show that the proposed model solves the issues identified in the section 3.5. It can successfully model the polysemous spatial markers used in South Asian languages. As it retains the positive features of Kracht's, Jackendoff's and Ostler's models, it can also still account for the range of phenomena these models were designed for.

Having dealt with the spatial usages of South Asian case markers at some length now, the next chapter returns to the issue of non-spatial, non-canonical marking of event participants.

Chapter 4

Non-Spatial Usages of Spatial Markers

4.1. Introduction

The main topic of discussion of the previous chapter was the polysemy related to spatial usages. It discussed the examples and semantic models for the use of a single form for more than one spatial usage. The data presented in Chapter 2 (primarily section 2.4) pointed at another aspect of polysemy related to case markers. There are many non-spatial usages that are marked by the spatial markers. The current chapter specially focuses on three examples of the use of spatial markers for non-spatial usages. The examples are: spatial markers on the second argument (or object), spatial markers on the addressee and the spatial markers used as the instrument marker.

The survey of case marking in South Asian languages shows that spatial markers are used to mark the second mandatory argument of certain verbs. Recall that the Urdu/Hindi spatial marker *par* is used as the locative-on marker. The same form is used to mark the second argument of verbs like 'trust', 'blame' and 'attack', etc.

(1) kitaab mez=par hai
book table=LOC_on be.PRES.SG
'The book is on the table.' <Urdu/Hindi>

(2) maiN=ne us=par bharosaa ki-yaa
1SG=ERG 3SG=LOC_on trust.M.SG do-PERF.M.SG
'I trusted him.' <Urdu/Hindi>

In example (2), the locative-on marker is used to mark the second mandatory argument of the verb 'trust'. Similarly, the second arguments of the verbs 'trust', 'attack', 'fear', 'love' and 'fight', etc. are marked by spatial markers in most of the surveyed languages.

As a lot of data and alternation related to the phenomenon of "non-canonical second argument" (NCSA) marking is available in South Asian languages, the following sections mainly focus on this topic. But the discussion is also relevant to the other two phenomena, i.e., addressee and instrument marking.

Spatial markers are used to mark the addressee of speech verbs. In Sindhi, the ablative marker *khaaN* is used to mark the addressees of the verb 'ask'.

(3) ho ghar=khaaN aayo
3SG.M.NOM house=ABL come.PERF.M.SG
'He came from the house/college.' <Sindhi>

(4) muuN cokre=khaaN savaal puc-yo
1SG.OBL boy.OBL=ABL question.NOM ask-PERF.M.SG
'I asked the boy a question.' <Sindhi>

Similarly, in many languages a spatial marker is also used as the instrument marker. The Punjabi marker *naal* is used as locative-beside and comitative marker. The same form is used to mark the instrument too.

(5) maiN kampiutar=de naal kitaab rak^{h}-ii
1SG computer=GEN beside book.F.SG put.PERF.F.SG
'I put the book beside/near/by the computer.' <Punjabi>

(6) o=ne caabi naal buuhaa k^{h}ol-iaa
3SG=ERG key.F.SG INST door.M.SG open-PERF.M.SG
'He opened the door with the key.' <Punjabi>

The organization of this chapter is as follows. In section 4.2, relevant literature on differential argument realization and multiple semantic usages of the form used as the instrument marker is reviewed and discussed. The theories of argument realization help in understanding the semantic reasons for NCSA marking. Section 4.3 presents the data related to non-spatial usages of spatial markers in South Asian languages. Section 4.4 tries to explain the South Asian data using the different theories and models described in section 4.2. The challenges presented by South Asian data to these theories and models are also identified and discussed there. Finally, section 4.5 discusses the semantic properties of different NCSA markers and provides a complete picture of the semantic explanation of the South Asian patterns.

4.2. Literature Review

The introduction above illustrated that there are verbs in South Asian languages that have NCSA marking. The detailed description and examples of NCSA marking are given in section 4.3. However, the presence of spatial marking on the second argument of the verb poses a question. Are South Asian languages unique in having non-canonical marking on the second arguments of some verbs? Or is this

phenomenon found crosslinguistically?

The presence of non-canonical markers for the second argument of the verb is noted for many other languages. Finnish verbs for 'love', 'admire', 'hate', 'enjoy' need obligatory partitive marking (Kratzer 2002). The verbs for 'beat', 'bite', 'expect', 'harm', 'help', 'kiss', 'look at', 'meet', 'push', 'read', 'stab', 'wait for' occur in N.W. Caucasian in the nominative-oblique construction (Catford 1975:44).

In Hungarian, *felel* 'answer', *gratulal* 'congratulate', *integet* 'greet' have dative complements (Blume 1998). The Russian verbs of authority, ruling, or disposition, e.g., *rukovodit'* 'rule/direct/manage', *upravljat'* 'govern', take instrumental complements while the verbs of the same class in Lithuanian take dative second arguments (Nichols 1975). Michaelis (1993) has analyzed non-canonical case marking in Latin and suggested that most of these verbs are cognitive.

Tiriyó and other Cariban languages have mental-state postpositions for several predicates, e.g., desiderative ('want, like'), cognoscitive ('know'), ignorative ('not know'), protective ('pity, jealous'), apprehensive ('afraid of'), superioritive ('more'), irascitive ('angry, wild'), odiative ('hate'), appreciative ('admire'), difficultative ('hard'), fidelitive ('trust, believe'), satisfactive ('enough'), etc. (Meria 2004).

English has transitive verbs whose near synonyms are intransitive. *Ask* is a near-synonym of *demand* and *request*, *look at* and *watch*, *go across* and *cross* are near-synonymous (Levin 1999).

(7) a. He demanded the book.

b. He asked for the book.

The verbs *demand* and *ask* in the above examples are synonyms. However, the verb *demand* has a canonical, while the verb *ask* has an oblique second argument. Apparently, there is no semantic reason for this object/oblique difference of these verbs.

Hence, we find that there are many languages that have non-canonical marking of the second argument. This is a problem related to argument realization. There are several theories and models related to this topic. A brief description of several of these theories is presented below. Some of the following theories explicitly discuss non-canonical marking on the second argument. The other theories do not discuss it

explicitly, but some unique properties of the verbs with non-canonical second arguments can be inferred by these theories.

The discussion begins with theories about high and low transitivity in two-participant predicates. Predicates that are low in transitivity may have non-canonical marking. Then, the event structures of the verbs with non-canonical second arguments are analyzed and the commonalities in these structures are described. Next, the localist approaches, or rather hybrid localist approaches, are presented. The discussion in section 4.4 looks at whether a localist approach can solve the puzzle of spatial markers on non-spatial arguments.

Finally, the lexical entailment (semantic features) based approaches are presented. It is important to note that all these approaches do not contain mutually exclusive semantic reasons for non-canonical marking on the second argument. The categorization made here is according to the major outlook of the theory/model.

4.2.1. High and Low Transitivity

Before discussing high and low transitivity and related semantic factors, we present the definitions of valency and transitivity used in the following discussion. Valency is the number of core or mandatory arguments of a verb. The verbs *go*, *run*, *swim* and *fall,* etc. have a valency equal to one. These verbs require only a single mandatory argument. There are others verbs, e.g., *kill*, *eat*, *see* and *find,* etc., that need two mandatory arguments. These verbs have a valency equal to two.

Sometimes, the transitive and intransitive verbs are synonymously termed as bivalent and monovalent verbs respectively. But the terms "transitivity", "high transitive verbs" and "low transitive verbs" have a slightly different but related meaning. Hopper & Thompson (1980) described transitivity as:

> "Transitivity is traditionally understood as a global property of an entire clause, such that an activity is 'carried-over' or 'transferred' from an agent to a patient. Transitivity in the traditional view thus necessarily involves at least two participants, and an action which is typically EFFECTIVE in some way."

Hence, transitivity is not only related to two participants (of bivalent verbs), but it requires an effective action as well. Both of these factors jointly determine whether a verb or clause is transitive or not.

4.2.1.1. Hopper & Thompson (1980)

According to Hopper & Thompson (1980), there is no two-way distinction between a transitive clause and a non-transitive clause. Instead, there is a scale of transitivity related to the clauses. Hopper & Thompson have identified certain factors that determine whether a clause is more or less transitive. The identified factors or parameters are listed in Table 4.1. In this table, Dixon's (1979) terms "A" (agent) and "O" (object) are used to refer to the two participants of the clause. However, Hopper & Thompson do not make any claim about the actual grammatical relation of these participants.

Table 4.1: Parameters related to high and low transitivity

	Factors/ Parameters	High Transitivity	Low Transitivity
A	Participants	2 or more participants, A and O	1 participant
B	Kinesis	Action	Non-action
C	Aspect	Telic	Atelic
D	Punctuality	Punctual	Non-punctual
E	Volitionality	Volitional	Non-volitional
F	Affirmation	Affirmative	Negative
G	Mode	Realis	Irrealis
H	Agency	A high in potency	A low in potency
I	Affectedness of O	O totally affected	O not affected
J	Individuation of O	Individuated	Non-individuated

The clauses are termed as more transitive or less transitive depending upon the number of high transitivity parameters available for the clause. See the following examples.

(8) Jerry knocked Sam down.

(9) Jerry likes beer.

The high transitivity parameters in the example (8) are:

Participants: two

Kinesis: action

Aspect: telic

Punctuality: punctual

Affectedness of O: total

Individuation of O: individuated

The high transitivity parameters in the example (9) are:

Participants: two

This shows that example (8) has a higher number of high transitivity parameters than example (9), hence the clause in (8) is more transitive than the clause in (9). Hopper & Thompson point out that in some languages sentences similar to example (9) do not have canonical marking on the first participant. For example, the Spanish sentence that is the equivalent of (9) is marked by the dative marker.

(10)	Me	gusta	la	cerveza
	1SG.DAT	pleases	the	beer
	'I like the beer.'			Hopper & Thompson (1980)

Finnish provides an example of non-canonical marking due to the aspect. See the following examples from Hopper & Thompson (1980).

(11)	Liikemies	kirjoitti	kirjeen	valiokunnalle	
	businessman	wrote	letter.ACC	committee-to	
	'The businessman wrote a letter to the committee.'				<Finnish>

(12)	Liikemies	kirjoitti	kirjettae	valiokunnalle	
	businessman	wrote	letter.PART	committee-to	
	'The businessman was writing a letter to the committee.'				<Finnish>

Both of the above sentences have the same parameters except for aspect. The example in (11) is telic. The action got completed and the object is totally affected. Hence, it is more transitive than the example in (12), which has an atelic event. As the example (12) is less transitive, the second argument is marked by the non-canonical partitive marker.

Hence, it is established that transitivity is not a binary valued property. A clause may be low in transitivity based on certain semantic properties, and a low transitivity clause may have non-canonical marking on the subject or the second argument/object.

4.2.1.2. Tsunoda's Hierarchy

Tsunoda (1981, 1985, 1999), Blume (1998) and Testelec (1998) also claim that

certain predicates are more likely to be transitive than others depending on certain semantic factors. Tsunoda identifies a cline of transitivity for predicates. The cline introduced by him is the following.

Table 4.2: Transitivity cline given by Tsunoda (1985)

Type	Name	Examples
1A	Direct affect on patient – resultative	kill, break
1B	Direct affect on patient – resultative	Hit
2	Perception	See
3	Pursuit	Search
4	Knowledge	Know
5	Feeling	Love
6	Relationship	Have
7	Ability	Capable

Tsunoda ordered these verbs on the basis of affectedness of the second argument. The object of a verb of Type 1A, e.g., *kill,* is totally affected. The affectedness decreases when we move down in the above list. The object of the verb *see* (Type 2) is less affected than the objects of Type I verbs, e.g., *kill* and *hit*. The second argument of the verbs further down in the list are lesser affected. For example, the second argument of the verb *have* (Type 6) is not affected but only involved in the event.

Tsunoda claimed that the verbs with low transitivity (or lower in the transitivity cline) have higher chances of deviating from the prototypical case frame of the language. Different languages set a different cut-off point in his transitivity cline. If a verb belongs to a type higher (a lower number) than the cut-off point, it has the canonical case frame of the language. The verbs belonging to the types lower than the cut-off point have non-canonical case frames.

4.2.1.3. Malchukov's Two-Dimensional Hierarchy

Malchukov (2005) decomposed Tsunoda's transitivity cline into finer hierarchies. He followed Givon's (1985:90) observation that the properties contributing to high

transitivity can be of the following three types.

(a) Agent related: The prototypical transitive clause has a visible, salient, volitional, controlling agent-cause that initiates the event.

(b) Patient-related: The prototypical transitive clause has a visible, salient, non-volitional, non-controlling patient-effect which registers the bulk of change associated with the event.

(c) Verb-related: The prototypical transitive clause has a compact, perfective realis verb or verbal tense-aspect-modality (TAM).

This implies that high and low transitivity depends on properties of agent, patient and TAM. The verbs in Tsunoda's hierarchy are arranged on the basis of both the decrease in agent properties and the decrease in patient properties. Malchukov proposed that we need at least two parallel hierarchies for the decrease in transitivity. One hierarchy has the verb types that are ordered with respect to the decrease of the prototypical agentive properties. The second hierarchy has the verb types ordered with respect to the decrease of the prototypical patient-hood properties.

He proposed the following two-dimensional verb type hierarchy.

(13) Transitivity hierarchy by Malchukov

Contact → Pursuit → (Motion)

Effective Action

Perception/Cognition → Emotion → (Sensation)

The verb types in the above hierarchy are the same as the verb types introduced by Tsunoda. The only exception is the verb types written in parenthesis that are introduced by Malchukov. Both of these types are lowest in transitivity as these have only one argument. The examples of motion verbs are *go* and *run,* etc. The examples of sensation verbs are *get sick* and *freeze*.

This two-dimensional hierarchy implies that languages can introduce a cut-off point for canonical case frames at two different points in the two wings of the hierarchy. A language can have different case marking strategies regarding to the decrease in agentivity and the decrease in patient-hood. The two-dimensional hierarchy allows for encoding this difference.

Furthermore, Malchukov extended the two-dimensional hierarchy into a map that has more layers of verb types. A verb type (or verb-type layer) present between the two layers is an intermediate between the two layers. For example, he introduced the verb type "interaction" (e.g., *help*, *speak*) as an intermediate between the verb type "pursuit" (e.g., *search*) and the verb type "symmetric predicates" (e.g., *marry*).

Hence, it is found that a low transitivity clause may allow non-canonical marking on the core arguments. The transitivity hierarchies help in modeling the cut-off point(s) for transitive constructions in different languages.

4.2.2. Event Structure

The above discussion illustrated that verbs low in transitivity can allow non-canonical marking. It provided some parameters and properties of the arguments of high transitivity clauses.

There is a similar explanation about non-canonical marking based on the event structure of a verb. When the event structures of a verb with canonical marked arguments and a verb with a non-canonical second argument are compared, it is found that there is a fundamental difference between the representations of the two verbs. This difference is presented in the following discussion.

4.2.2.1. Levin's Explanation

The concept of event structure and different types of verbs with respect to aspect (actually, *Aktionsart*) was introduced by Vendler (1967) and Dowty (1979). Different authors present slightly different notations and definitions for different types of event structures. Here we present the view of Levin (1999), who compared the structure of NCSA marked verbs with the structure of a prototypical verb.

Levin (1999) provides the following event structure templates for different kinds of simple events.

(14) a. [x ACT <*MANNER*>] (activity)

b. [x <*STATE*>] (state)

c. [BECOME [x <*STATE*>]] (achievement)

An example of the event structure template for a complex event is the following.

(15) [[x ACT <*MANNER*>] CAUSE [BECOME [y <*STATE*>]]] (causative)

The event structure template tells us about the aspectual properties of the verb. The structure has also a constant part that encodes the core idiosyncratic meaning of the verb. The name of the verb comes from this constant.

The verb *run* is an example of the activity. See the following example.

(16) a. Pat ran.

b. [x ACT <*RUN*>]

The core meaning of the verb, i.e., its name, is encoded as the constant in angle brackets. The runner is encoded as the structural participant x. Semalfactives also are non-durative type of activities. The verbs *cough* and *wink,* etc. have the same event structure as that of the verb *run*.

The examples of state and achievement verbs and the corresponding event structures are given in (17) and (18).

(17) a. The door is opened. (state)

b. The door opened. (achievement)

(18) a. [x <*OPEN*>] (state)

b. [BECOME [x <*OPEN*>]] (achievement)

The complex events have more than one simple event. See the example of a causative verb in (19). The first participant performs an activity that causes a change of state in the second participant.

(19) a. Jack opened the door.

b. [[x ACT] CAUSE [BECOME [y <*OPEN*>]]]

After the brief survey of different types of event structures, we return to the activity/semalfactive verbs. It is not necessary that the activity/semalfactive verbs have a single participant. The activity verb *sweep* and the semalfactive verb *hit* both have two participants, as is shown in the following event structures.

(20) a. [x ACT <*SWEEP*> y]

b. [x ACT <*HIT*> y]

Following Rappaport Hovav & Levin (1998), Levin (1999) defined two types of variables. The first type is defined as the structural requirement of the template. For

example, the activity needs a structural variable *x* for the participant that performs the activity. The constant of the verb provides this variable to fulfill the syntactical requirement of the template.

But, the constant may provide another variable if the meaning of the variable requires more than one participant. The verbs *sweep* and *hit* have two participants, hence the manner constants of these verbs provide the second variable *y* to the template. This variable *y* is termed as the pure constant variable in contrast to the structural participant *x*.

One can note the difference between the event structures of the two participant activities in (20) and two participant causatives in (19b). Both of the participants of a causative are the structural variables of the event structure. Levin termed the two participant verbs whose participants are both structural variables as core transitive verbs (CTVs). On the other hand, the two participant activities have a structural variable and a pure constant variable. The verbs with this type of event structure are termed as non-core transitive verbs (NCTVs).

Levin suggested that it is difficult to identify the semantic role of the pure constant variable because these variables lack an event structure template characterization. A language may have non-canonical marking on these participants. Languages vary as to which and how many semantic sub-classes of the NCTVs come under (language specific) oblique linking rules. Different languages have different numbers and types of the oblique rules.

Hence, Levin does not define any cut-off criteria for canonical and non-canonical object marking. She predicts that some of the verbs having pure constant variables may have non-canonical objects (second arguments) based on the linking rules specific for a particular language.

4.2.2.2. Ramchand's Rheme

Ramchand (2008) proposed that the event is decomposed into three subevents: causation, process and result. All the verbs do not have all these subevents, but the event structure of any verb can be encoded in terms of these subevents. There are a number of general primitive predicates over events corresponding to the basic subevent *e*:

(21) a. State(e) : e is a state

b. Process(e): e is a process or transition

c. Causing(e): e is an initiational process or transition

There are no thematic roles in Ramchand's structure. Instead, there are roles corresponding to the subevents that show the relation of the participant and the subevent.

(22) a. Subject (x, e) and Causing(e) entails that x is the initiator of e.

b. Subject (x, e) and Process(e) entails that x is the undergoer of the process.

c. Subject (x, e) and State(e) entails that x is the resultee of the state.

The following sentence has the causative verb *open.*

(23) Jack opened the door.

The structure corresponding to the above sentence is the following.

(24) Initiator: Jack

Undergoer: door

Resultee: door

In the above example, the door is both the undergoer and the result. There are other kinds of verbs that have the same participant as initiator and result. See the following sentence and its structure.

(25) Jack walked.

(26) Initiator: Jack

Undergoer: Jack

Ramchand also introduces another role: "rheme". See the following example.

(27) Jack walked on the trail.

Here the trail is the path traversed by Jack. The undergoer passes through the path that is a subtype of the rheme. Hence the structure of the above sentence is:

(28) Initiator: Jack

Undergoer: Jack

Rheme: trail

There are others types of rheme too. Following Saksena's (1982) definition of affected agent, Ramchand suggests that the subject of the ingestion verbs, e.g., *eat* and *drink,* etc., is both the initiator and the undergoer. The ingested object is the rheme path traversed by the undergoer. See the following sentence and its structure.

(29) Jack ate the apple.

(30) Initiator: Jack

Undergoer: Jack

Rheme: apple

Another example of rheme appears in the following sentence.

(31) Katherine fears nightmares.

(32) Initiator: Katherine

Rheme: nightmare

Hence, the rheme is a set of heterogeneous elements. The rhemes of the above three examples are very different from one another. The point relevant to our topic is that the non-canonically marked second arguments are always the rheme in the structure. It is the main difference between the non-canonical second argument of such verbs and the object of a prototypical verb.

Levin (1999) is unable to explain why some pure constant variables are realized as non-canonical objects (second arguments), while the others are realized as canonical objects. Similarly, Ramchand's model does not explain why some rhemes are realized as having non-canonical marking (or obliques), while the others are realized as canonical objects. The event-structure approach also does not help us in predicting the appropriate (non-canonical) case marker for these non-prototypical objects.

4.2.3. Localist Approach

As discussed in the previous chapter, the localist approach models all predicates in terms of spatial constructs. The different arguments of the predicate are taken to be

theme, source and goal. Hence localist approaches may give an explanation of non-spatial usages of spatial markers.

4.2.3.1. Ostler's Model

The basics of Ostler's model were presented in section 3.2.2. That section discussed the positional predicates BE_{posit} and GO_{posit}. In this section, the other types of BE and GO constructs are presented. Ostler proposed five subscripts for these constructs: "posit" (position), "ident" (identification), "poss" (possession), "cognit" (cognition) and "volit" (volitional). The examples corresponding to the first four of these types are:

(33) a. John is in the room. (BE_{posit})

b. John is tall. (BE_{ident})

c. John has money. (BE_{poss})

d. John knows the answer. (BE_{cognit})

e. John is in money. ($BE_{poss\text{-}inv}$)

The roles theme, goal, source and path were introduced in section 3.2.2 as well. The theme and goal relations of the above examples are shown in the following:

(34) Theme and goal for the sentences in (33)

	Theme	**Goal**
BE_{posit}	John	room
BE_{ident}	John	tall
BE_{poss}	money	John
BE_{cognit}	answer	John
$BE_{poss\text{-}inv}$	John	money

In addition to the spatial features [source] and [goal] corresponding to the four spatial roles, Ostler introduces other non-spatial features to elaborate the properties of different types of the spatial roles. The first important feature is "abstract" [Abs]. The construct BE_{posit} has a genuine spatial goal where the theme is physically present. The other constructs, e.g., BE_{poss}, have abstract spatial roles. In (33c), *John* is the abstract location (goal) of *money*. The goals corresponding to (33b)–(33e) thus have the

feature [+Abs].

The abstract spatial roles can be sentient too. The feature "sentient" [Se] is introduced to encode this. The goals in (33a)–(33b), i.e., *room* and *tall*, have the feature [–Se] and the goals in (33c)–(33d), i.e., *John* (in both examples), have the feature [+Se].

The example (33e) introduces another feature. If we compare (33c) and (33e), both of those have equivalent meaning. In (33c), *John* is the goal that has the theme, i.e., *money*. The situation is inverse in (33e). Here, *money* is the goal and *John* is the theme that is present in the goal. Hence, we use the feature "invers" [+Inv] in example (33e).

The features corresponding to the theme and goal of (33) are presented in (35). The spatial feature [source] and [goal] already introduced in section 3.2.2 are represented now as [So] and [Go], respectively.

(35) Features corresponding to the theme and goal for the sentences in (33)

	Theme	**Goal**
BE$_{posit}$	John	room
	[–So,–Go,+Se,–Abs,–Inv]	[–So,+Go,–Se,–Abs,–Inv]
BE$_{ident}$	John	tall
	[–So,–Go,+Se,+Abs,–Inv]	[–So,+Go,–Se,+Abs,–Inv]
BE$_{poss}$	money	John
	[–So,–Go,–Se,+Abs,–Inv]	[–So,+Go,+Se,–Abs,–Inv]
BE$_{cognit}$	answer	John
	[–So,–Go,–Se,+Abs,–Inv]	[–So,+Go,+Se,+Abs,–Inv]
BE$_{poss\text{-}inv}$	John	money
	[–So,–Go,+Se,+Abs,+Inv]	[-So,+Go,-Se,+Abs,+Inv]

For the sake of simplicity, the examples of the construct BE were used for purposes of illustration above — the construct BE has only two roles: theme and goal. The construct GO allows the other roles, i.e., source and path. The examples of the construct GO are given in (36).

(36) a. John went from Frankfurt to Munich. (GO_{posit})

b. John got the money from David. (GO_{poss})

In (36a), the theme *John* started from the source *Frankfurt* and moved to the goal *Munich*. In (36b), *money* is the theme, *John* is the goal and *David* is the *source*. The examples in (36) encode dynamic situations in contrast to the static situations of the examples in (33). Ostler introduces the feature "dynamic" [Dyn] to encode this difference.

Ostler introduced another feature "actional" [Ac] to differentiate the action by a sentient being and an event. See the following examples to understand the difference.

(37) a. The cyclone hit the harbour. ([–Ac])

b. John hit David. ([+Ac])

Another feature in Ostler's system is "external" [Ext]. The mandatory participants of a verb have the feature value [–Ext]. The arguments that are not part of the core meaning of the verb have the value [+Ext]. There is another feature "volitional" [Vol] that is related to the GO construct. The construct GO_{volit} shows the volition of the main participant who performs the task.

The above-described features are used by Ostler to model different case markers of Japanese, Turkish and Sanskrit. For example, the Turkish locative marker *-de* is represented as:

(38) *-de*: [+Go, –So, –Dy], Abs, Se, Ext

The different semantic usages of this marker with the respective features are:

(39) a.	tiyatro-**da**-dir	Location
	'He is *at the theatre*.'	[+Go, –So, –Dy]
b.	o kilik-**ta** bir adam	Attribute
	'a man *of that appearance'*	[+Go, –So, –Dy, +Abs]
c.	ben-**de** bes lira var	Possessor
	'I have got five liras *on me*.'	[+Go, –So, –Dy, +Se]
d.	tiyatro-**de** lui piyes gordum	Place
	'I saw a good play *at the theatre*.'	[+Go, –So, –Dy, +Ext]

The example (39a) has the location *theatre* marked by the locative marker. The example (39b) has an abstract (identification) location represented by [+Abs]. The example (39c) has the possessor marked by the locative marker. The possessor is the sentient [–Se] location. The example (39d) has the location that is an external participant [+Ext]. Compare it with the location mentioned in (39a) that is the mandatory (meaning non-external) participant of the predicate.

Hence, Ostler shows that different semantic usages of a case marker can be modeled by the variation of the above-mentioned features.

4.2.3.2. Jackendoff

Jackendoff's model was already introduced in Chapter 3 (section 3.2.1). It uses spatial constructs to model non-spatial predicates. According to Jackendoff, spatial constructs alone cannot model different types of predicates and arguments. In many situations, we need the concepts like affectedness etc. as well. He presented the following examples to explain his point.

(40) a. Sue entered the room.

b. Sue hit Fred.

In both of these examples, *Sue* is the theme that moves to goal. In example (40a), *room* is the goal, while in (40b) *Fred* is the goal. We know that there is a difference between these two goals. The *room* is just a location, but *Fred* is a patient too. The goal *Fred* is affected by the action. We need some construct to model this additional property of *Fred*.

To solve this problem, Jackendoff introduced the concept of two tiers of representation. The "thematic tier" models the predicates in terms of spatial constructs. It has the concepts like PLACE and PATH to deal with motion and location. On the other hand, the "action tier" is used to model actor and patient relations. See some examples with the corresponding thematic and action relations.

(41) a. Sue entered the room.

Theme Goal (thematic tier)

Actor (action tier)

b. Sue	hit	Fred.	
Theme		Goal	(thematic tier)
Actor		Patient	(action tier)
c. The car	hit	the tree.	
Theme		Goal	(thematic tier)
Actor		Patient	(action tier)
d. Sue	threw	the ball.	
Source		Theme	(thematic tier)
Actor		Patient	(action tier)

The difference between these examples is obvious. The goal of (41a) does not have a patient relation like the goals of (41b)–(41c). The patient of the action tier does not always correspond to the goal of the thematic tier. In (41d), the theme (moving entity) is the patient.

The representation used in (41) is not the formal representation, i.e., the conceptual structure of Jackendoff's model. The conceptual structure corresponding to example (41c) is the following.

(42) [INCH [BE ([CAR], [AT [TREE])]
AFF ([CAR], [TREE])]

The upper line is for the thematic tier and the lower line is for the action tier. INCH (inchoative) shows that *hit* is analyzed as an inchoative verb.

Jackendoff used the superscript with the constructs CS (cause) and AFF (affected) to show whether it is positive, negative or undetermined. The verbs *manage* and *succeed* in a sentence like *He managed/succeeded to do* …. have CS^{+}. The verb *fail* has CS^{-}, while the verb *try* is underspecified thus having the construct CS^{u}. Similarly for affectedness, the verb *help* has the construct AFF^{+} and the verb *let* has the construct AFF^{0}.

Jackendoff also models experiencer predicates. According to him, there are three criteria for differentiation of experiencer predicates. The first criterion is the position of experience in the sentence (Chomsky 1965, Lakoff 1970). Is the experiencer a subject (e.g., *like*) or is it an object/oblique (e.g., *please*)? The second criterion is

about the positive (e.g., *please*) or negative (e.g., *displease*) effect of the event on the experiencer. The third criterion asks whether the predicate is an event or state.

Jackendoff provided constructs for all of these criteria. The event/state difference is encoded by the constructs BE and GO. For positive, negative or undetermined effects the superscripts "+", "–" and "u" are used. When the experiencer is in an object/oblique position, the construct AFF is used.

(43) a. X pleases Y. $[_{State}\ AFF^{+}\ ([X], [Y])]$

b. X displeases Y. $[_{State}\ AFF^{-}\ ([X], [Y])]$

c. X matters to Y. $[_{State}\ AFF^{u}\ ([X], [Y])]$

But if the experiencer is in a subject position, we need the reverse of the above situation. Therefore, Jackendoff introduced the construct REACT that has patient as its first argument. The examples of the use of REACT are the following.

(44) a. Y likes X. $[_{State}\ REACT^{+}\ ([Y], [X])]$

b. Y hates/fears X. $[_{State}\ REACT^{-}\ ([Y], [X])]$

c. Y regards X as crazy. $[_{State}\ REACT^{u}\ ([Y], [X])]$

Appendix C investigates whether South Asian data follows this classification of experiencer predicates or whether it presents some challenges to it.

In sum, Jackendoff recognizes that the predicate should be modeled by using both the localist construct (thematic tier) and the agency/affectedness concept (action tier). However, the model does not discuss anything related to non-spatial usages of spatial prepositions/case markers.

4.2.3.3. Butt

Butt (2006b) presented a two-dimensional view of case. Like Jackendoff, she argued that one should consider both spatial concepts and notions of control simultaneously. As her model is about case markers, it is relevant to this discussion.

According to her, the spatial dimension is primary. This dimension shows how the arguments are placed relative to each other in a spatial relationship. But these arguments also act on each other. This interaction of the arguments is modeled by involving the concept of control/agency. The two-dimensional view of case is presented in Table 4.3.

Table 4.3: Two-dimensional view of case

More Control	**Place**	**Path**
Ergative		
Genitive		
Instrumental		
Dative		
Accusative		
Less Control		

In this table, cases are arranged on the basis of more to less control. The spatial dimension of case is also presented. Take the example of the dative case in the above table. Its spatial dimension is identical to that of the ergative (or genitive) but it has less control in the control dimension. Hence it is used to mark the experiencer subject that has less control than the canonical subject or the agent. The example of choice of ergative or dative in Urdu on the basis of control is shown in (45).

(45) a. naadyaa=ne kahaanii yaad kii
Nadya.F.SG=ERG story.F.SG memeory do.PERF.F.SG
'Nadya remembered the story.' (actively) <Urdu/Hindi>

b. naadyaa=ko kahaanii yaad aa-yii
Nadya.F.SG=DAT story.F.SG memeory come-PERF.F.SG
'Nadya remembered the story.' (memory came to Nadya) <Urdu/Hindi>

When the subject *Nadya* actively memorizes the story, as in (45a), it is marked by the ergative marker that has more control. In (45b), the memory of the story occurred to her, and hence the subject with less control is marked by the dative marker.

Butt's model has two important aspects. The first one is the interaction of space and control. The second aspect is the gradient of control of the markers that have identical spatial dimensions. Later on, we see that the choice of case markers in some alternations is used to show the gradient of affectedness or some other semantic property.

Hence, the localist approach uses both spatial and non-spatial concepts to model the semantic structure of predicates and the argument realization. The constructs introduced by the above models can be used to predict the choice of spatial marker in non-spatial usages.

4.2.4. Lexical Entailments for Core Arguments

The above discussion described the modeling of the predicate by the event structure and the localist approach. In both approaches, there are small numbers of semantic roles. The participants of the predicates have one of these roles. In the localist approach, the core roles are source, goal and theme. In the event structure approach, the position of the participant in the event structure template determines its properties.

There are other ways to classify the different types of arguments of a predicate. As described in Chapter 2, thematic roles are used to label different types of arguments. For example, the roles agent, patient, instrument and experiencer are widely used. The authors each differ about the list of thematic roles or their order in a hierarchy. Sometime it is not easy to assign an appropriate thematic role to an argument. To solve this problem, Dowty (1991) introduced the concept of "proto-roles". He reconceived the concept of thematic roles as a set of lexical entailments.

The concept of proto-roles is not directly relevant to the topic of non-canonical second arguments in South Asian languages. However, the lexical entailments or semantic properties provide a way to model different non-canonical arguments that either do not exactly match with any of the standard thematic role, or display fine grained differences among the arguments that apparently have an identical thematic role. Lexical entailments of this kind are used to model different type of arguments and case markers (e.g., Primus (1999), Ackerman & Moore (2001), Grimm (2005) and Beavers (2006), among others).

Before introducing some of these models, it is necessary to have a look at the lexical entailments as introduced by Dowty. He introduced two proto-roles: proto-agent and proto-patient. The lexical entailments of these roles are:

(46) **Proto-agent properties**

- volitional involvement in the event or state
- sentience (and/or perception)

- causing an event or change of state in another participant
- movement (relative to the position of another participant)
- exists independently of the event

(47) **Proto-patient properties**:

- undergoes change of state
- incremental theme
- causally affected by another participant
- stationary relative to movement of another participant
- does not exist independently

These are the lexical entailments associated with a prototypical agent and patient. These features are commonly found in subjects and objects corresponding to different predicates. However, it is not necessary that all the subject or object arguments must have all the proto-agent or all the proto-patient properties respectively.

To identify subjects and objects corresponding to non-prototypical arguments, Dowty proposed a rule. The argument that has the greatest numbers of proto-agent properties is lexicalized as the subject, and the argument with the greatest number of proto-patient properties is lexicalized as the direct object.

It is clear that Dowty's model cannot help us in deciding which non-canonical marking is used for which non-prototypical argument. However, the list of proto-role properties with some modification can be used to model different types of non-standard thematic roles or different variants of a particular standard thematic role. The following discussion and section 4.3.3 describes how lexical entailments can be used to model the non-canonical marking on the second argument.

Ackerman & Moore (2001) introduced another proto-patient property: bounding entity. Furthermore, they suggested that oblique marking on the argument is due to a decrease in proto-patient properties. It implies that if an object has a fewer number of proto-patient properties, it may be marked by an oblique marker. The non-canonical partitive case in the following Estonian example is due to the absence of a bounding entity that is a proto-patient property.

(48) a. Madis joob oma tee aera
Madis drink.3SG.PRES own tea.GEN/ACC PREV
'Madis will drink up the tea,' <Estonian>

b. Madis joob teed
Madis drink.3S.PRES tea.PART
'Madis is drinking tea.' <Estonian>

Similarly, Primus (1999, 2002) introduced "possession" as new proto-agent property, and "proto-recipient" as a new proto-role for modeling recipient, benefactive and addressee. These arguments possess proto-agent properties with respect to the patient, but they appear as proto-patients with respect to the agent.

Grimm (2005) models different case markers in terms of Dowty-style semantic properties. The properties used by Grimm are: instigation, motion, sentience, volition, motion and persistence. He claims that causation and control are not primitive properties, but are composed of different properties. Similarly, affectedness is not a binary property, but an entity can be totally affected, partially affected or unaffected.

The totally affected argument has the feature *existential persistence (beginning)*. As the partially affected argument exists after the event as well and there is only a qualitative change to it, it has the property *qualitative persistence (beginning)*. The unaffected argument has the property *total persistence*. Causation is modeled as [Arg X: + instigation, Arg Y: – persistence (end)].

Grimm took all possible combinations of the primitive properties and arranged the combination of properties in the form of a lattice. He termed it "agency lattice" (see Figure 4.1). The topmost node at the left end of this lattice [instigation, motion, sentience, volition, motion, total persistence] represents the maximal agent, while the node at right having the property [existential persistence (beginning)] represents the maximal patient.

For a two-argument predicate, the argument that is higher and near to the maximal agent is realized as the subject, and the argument that is lower and near to the maximal patient is realized as the object.

This agency lattice is also used to model case marking. Different case markers have different set of properties. The set of properties corresponding to a particular case marker can be shown as a region on the agency lattice. Grimm has provided examples from Urdu/Hindi, Russian and Ancient Greek to illustrated his system. His most extensive example about case marker regions comes from Ancient Greek. Figure 4.1 shows the region of Ancient Greek case markers on the agency lattice. The

instrumental dative marker has the properties [total persistence] and [total persistence, instigation]. The proper dative marker has the property [sentience, qualitative persistence (beginning)].

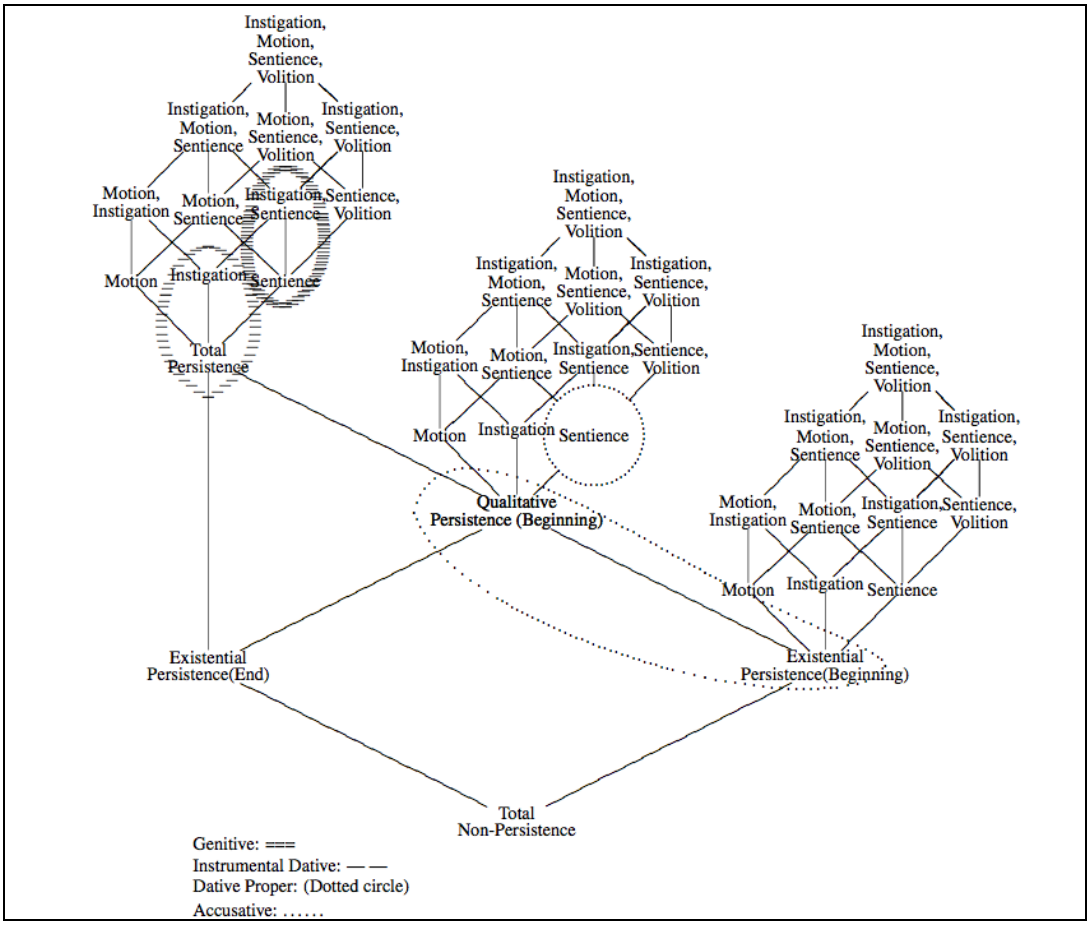

Figure 4.1: Regions of Ancient Greek case markers on agency lattice (Grimm 2005)

Do South Asian languages have regions for non-canonical second argument markers too? And do we need to introduce the spatial features (developed in Chapter 3) into the Dowty-style property set? These questions are investigated in section 4.4, as the lexical entailment approach seems to provide us tools to model different kinds of non-canonically marked arguments.

4.2.5. Linking Theories

"Linking" is a term used for generalizations that are involved in mapping predicate-argument structures to a syntactic representation. See Butt (2006a) for description and references of important linking theories. Here, a brief introduction of Lexical

Mapping Theory (LMT) is presented that is the standard linking theory of Lexical Function Grammar (LFG).

LFG has SUBJ(ect), OBJ(ect) and OBL(ique), etc., as "grammatical functions" (GF). The grammatical functions show the relation of an argument with the predicate. These are part of the syntactic inventory of every language. There are semantically restricted versions of OBJ and OBL as well. The GFs OBJθ and OBLθ can be used for the arguments with a particular thematic role θ.

The theory has two binary-value features [±r] and [±o] for "restricted" and "objective", respectively. The feature [+r] is used with an argument that is semantically restricted to a particular thematic role. The feature [+o] is used with the object-like arguments. These features are used to describe grammatical functions.

(49) Features	**Grammatical Functions**
[–o]	SUBJ, OBL
[+o]	OBJ, OBL
[–r]	SUBJ, OBJ
[+r]	OBJθ , OBLθ

The following is an example of mapping of the argument structure (a-structure) to GF. It is related to a prototypical two-argument verb, i.e., *cut*.

(50) **a-structure**	cut < agent	theme>
	[–o]	[–r]
GF	SUBJ	OBJ

Both [–o] and [–r] can be linked with SUBJ, but there is a mapping principle. It says that [-o] is mapped onto SUBJ. Only in the absence of [-o], the feature [-r] is mapped to SUBJ. Hence, [–o] is mapped to SUBJ and [–r] is mapped to the only remaining option, i.e., OBJ.

Butt (1998) proposes to incorporate aspectual affectedness into linking. Aspectually inert objects get the feature [–r] and are linked to OBJ, while aspectually active objects get the feature [+r] and are linked with OBJθ. After incorporating this proposal, the relation of thematic role and features is as follows.

(51) **Thematic Role**	**Feature(s)**
Theme	[–r] / [+r]
Secondary patient-like roles	[+o]
All others	[–o]

The LMT can be used to model alternations of case marker. The following pair of sentences has the alternation of accusative and instrument(-perlative) marker. In (52), the affectedness of the causee (*saddaf*) is emphasized, while in (52) the affectedness of the patient (*spice*) is emphasized.

(52) anjum=ne saddaf=ko masaalaa cakh-va-yaa
Anjum.F=ERG Saddaf.F=ACC spice.M.NOM taste-CAUS-PERF.M.SG
'Anjum had Saddaf taste the seasoning.' <Urdu/Hindi>

(53) anjum=ne saddaf=se masaalaa(=ko) cakh-va-yaa
Anjum.F=ERG Saddaf.F=INST spice.M.NOM/(ACC) taste-CAUS.PERF
'Anjum had the seasoning tasted by Saddaf.' <Urdu/Hindi>

The linking of the above examples (52)–(53) is shown in (54)–(55) respectively.

(54) **a-structure** CAUSE < agent theme taste < agent theme > >

	[-o]	[+r]/[–r]		[+r]/[–r]
Default	[–r]	[+o]		[+o]
GF	SUBJ	OBJ_θ/*OBJ		OBJ/*OBJ_θ
Case		ACC/NOM		NOM/ACC

(55) **a-structure** CAUSE < agent theme taste < agent theme > >

	[–o]	[–o]		[+r]/[–r]
Default	[–r]			[+o]
GF	SUBJ	OBL		OBJ/OBJ_θ
Case		INST(-ABL)		NOM/ACC

The causee (*Saddaf*) in (54) is considered to be semantically affected via the feature [+r]. Hence, it is mapped to OBJ_θ that is correlated with the accusative marker *ko*. On the other hand, the embedded patient (*spice*) in (55) is considered as semantically affected via the feature [+r]. In this case the causee is mapped to OBL that is correlated with instrument(-ablative) marker *se*.

Hence, linking theories can also be used to model the alternation of case markers due to semantic reasons.

4.2.6. Instrument Marking: Another Non-Spatial Usage

The above discussion focused on argument realization and especially on the semantic reasons for non-canonical second arguments. The material presented above will be helpful for building a model for NCSA marking in South Asian languages. However, section 4.1 mentions that there are two more non-spatial usages of the spatial makers (beside spatial marking for NCSA). The instrument and addressee in many languages are marked by the same form that has spatial usages. This section therefore focuses on what has already been written about the polysemy related to instrument markers.

In a landmark paper, Lakoff & Johnson (1980) proposed that instruments are conceptualized as companions.

(56) a. John went to the market with Michael.

b. John ate the meal with the fork.

In (56a), Michael is the companion of John. John went to the market accompanied by Michael. The companion is marked by the preposition *with*. It is interesting that the same form is used to mark the instrument in (56b). It means that the instrument is metaphorically conceived as a companion.

Lakoff & Johnson (1980) claim that almost all languages use the companion metaphor for the instrument marking. Stolz, Stroh & Urdze (2006) challenged this claim and pointed out that the companion metaphor is not found crosslinguistically. According to them, 65% of all languages (out of 323 languages surveyed) are A-type languages. They define a language as A-type if it uses different markers for accompaniment and instrument. (Stolz, Stroh & Urdze 2006:105)

A parallelism of source with agent, goal with patient and path with instrument is suggested by Anderson (1971:173). Ostler (1979) defines the instrument as a path with the features [+source,+goal]. The action passes through that path. He pointed out that the same marker is used to mark the reciprocal. The object of the reciprocal is both the actional source (agent) and the goal (patient) at the same time. According to Ostler, the feature [animate] is the difference between instrument and reciprocal. Instrument is the path with the feature [–animate], while reciprocal has the same

feature set [+source,+goal] with the feature [+animate]. Moreover, the instrumental of Butt (2006b) shown in table 4.3 corresponds to both place and path.

Jackendoff's model uses the construct BY to model the instrument. See the following example and its partial conceptual structure.

(57) Phil opened the door with the key.

(58) [BY [CS^{+} ([PHIL] , [AFF^{-} ([KEY],[DOOR])])

AFF^{-} ([PHIL], [KEY])]]

The second line shows that the actor Phil affects the key. The first line shows that the actor Phil causes the door opening (key affecting the door) by the key.

Hence, Ostler (1979), Lakoff & Johnson (1980) and Butt (2006b) introduce different metaphors for instrument markers. Jackendoff does not model the instrument in terms of spatial construct (in thematic tier). Only the action tier has constructs for the instrument.

4.3. Testing the Models with Respect to South Asian Data

After the introduction of different approaches related to spatial marking used for non-spatial usages specially for NCSA in section 4.2, this section pulls out the South Asian data presented in Chapter 2 that is relevant to this discussion. The data illustrate the spatial marking on NCSA, instruments and addressees. The following sections discuss the challenges introduced by this data to the argument realization and other approaches discussed above.

4.3.1. Non-Canonical Second Argument Marking

As already discussed in Chapter 2, in most of the surveyed South Asian languages, the canonical marking on the subject is either nominative or ergative. Similarly, the object is canonically marked either by the nominative or the accusative marker in these languages. The following examples show the canonical marking on subject and object in Urdu/Hindi.

(59)	laRkaa	ghar	taamiir	kar	rahaa	hai
	boy.NOM	house.NOM	construction	do	PROG.M.SG	be.PRES
	'A/the boy built a/the house.'			<Urdu/Hindi>		

(60) laRke=ne ghar=ko taamiir kiyaa
boy.OBL=ERG house.OBL=ACC construction do.PERF.M.SG
'A/the boy built the house.' <Urdu/Hindi>

Along with the sentences with canonical subjects and objects, we find a lot of examples of non-canonically marked subjects and objects. As explained in chapter 2, I use the term non-canonical second argument instead of non-canonical object. The non-canonical subjects are widely studied. A subject can be non-canonically marked by dative, genitive, ablative/instrumental and locative (Mohanan 1994, Butt & King 2005).

The case marking survey in Chapter 2 showed that many verbs in South Asian languages have non-canonical second arguments. Some examples of non-canonically marked objects are given in (61)–(64).

(61) jamiil=ko zaahid=par bharosaa hai
Jameel.M.SG=DAT Zahid.M.SG=LOC-on trust.M.SG be.PRES
'Jameel trusts Zahid.' <Urdu/Hindi>

(62) jamiil=ne zaahid naal gal ki-tii
Jameel.M.SG=ERG Zahid.M.SG COM talk do-PERF.F.SG
'Jameel talked with Zahid.' <Punjabi>

(63) un=le mohan=sanga bihaa gar-in
3SG=ERG Mohan.M.SG=COM marriage do-PST.M.SG
'She married Mohan.' <Nepali>

(64) jamiil saaNp=se dar-taa hai
Jameel.M.SG snake.M.SG=ABL fear-IMPF.M.SG be.PRES
'Jameel fears snakes.' <Urdu/Hindi>.

In the above sentences, the second argument of the predicate is either marked by a locative or comitative marker. Most of the verbs having non-canonical second argument arguments are psych verbs, but non-psych verbs also appear with non-canonical arguments.

(65) us=ne mulk=par hamlaa ki-yaa
3SG=ERG country.M.SG=LOC-on attack.M.SG do-PERF.M.SG
'He attacked the country.' <Urdu/Hindi>

On the basis of non-canonical subject and object marking, section 2.4.3 proposed six classes of verbs, repeated here in Table 4.4.

The proposal of verb classes on the basis of syntactic properties follows the idea introduced by Levin (1993). She claimed that the verbs that show similar syntactic

behavior constitute a coherent semantic class. Hence, the subcategorization frame and argument alternations provide information about the semantics of a verb. The classes given in Table 4.4 follow the same principle. These are obtained by considering different combinations of subject and object marking.

Table 4.4: Classes of NCSA verbs in South Asian languages

Class	Subject Marking	Object Marking	Examples
I	NOM/ERG, DAT	ABL	fear
II	NOM/ERG	ABL	Resign
III	NOM/ERG	LOC-on/DAT	bless, capture, govern, attack, sign, blame, monitor
IV	NOM/ERG, DAT	LOC-on/DAT	trust, doubt, suspect, believe
V	NOM/ERG	COM/DAT	talk, meet, marry
VI	NOM/ERG, DAT	COM	love, hate

It is important to note that the NCSA marking on the verbs in the above classes are similar but not identical crosslinguistically. For example, the object of the verb 'meet' uses oblique (used for dative) marking in Balochi, but takes the comitative marker *saaN* in Sindhi.

After this brief introduction, the following sections go into further detail about non-canonically marked objects. The data presented here is discussed in section 4.4 and the conclusion is presented in section 4.5.

4.3.1.1. Locative Marked Arguments

The locative marker is used to mark NCOs of two classes of verbs. The first class (Class IV of Table 4.4) is exemplified by the verbs 'trust', 'doubt', 'suspect' and 'believe', etc. These are psych verbs that allow both canonical and dative subjects in Urdu/Hindi and many other languages.

(66) jamiil=ko zaahid=par bharosaa hai
Jameel.M.SG=DAT Zahid.M.SG=LOC_on trust.M.SG be.PRES
'Jameel trusted Zahid.' <Urdu/Hindi>

(67) jamiil=ko zaahid=par bharosaa huaa
Jameel.M.SG=DAT Zahid.M.SG=LOC_on trust.M.SG be.PERF
'Jameel trusted Zahid.' <Urdu/Hindi>

(68) jamiil=ne zaahid=par bharosaa kiyaa
Jameel.M.SG=ERG Zahid.M.SG=LOC_on trust.M. do.PERF
'Jameel trusted Zahid.' <Urdu/Hindi>

The *noun* + *be* complex predicate in (66) means that the (experiencer) subject is in a certain state. The *noun* + *become* (*be.*PERF) complex predicate in (67) means that the (experiencer) subject achieved a state, and the *noun* + *do* complex predicate in (68) means that the (experiencer) subject also has the agentive property. The other predicates of Classes IV and VI (that allow dative marked subjects) allow similar syntactical behavior.

The predicates of Class III of Table 4.4 also have a locative marked second argument. The verbs 'bless', 'govern', 'attack', and 'blame', etc. have a locative marked second argument, but these verbs do not allow a dative subject construction.

(69) fauj=ne Seher=par hamlaa kiyaa
Army.F.SG=DAT city.M.SG=LOC_on attack.M.SG do.PERF.M.SG
'The army attacked the city.' <Urdu/Hindi>

(70) a. *fauj=ko Seher=par hamlaa huua
Army.F.SG=DAT city.M.SG=LOC-on attack be.PERF.M.SG
'The army attacked the city.' <Urdu/Hindi>

b. Seher=par hamlaa huua
city.M.SG=LOC-on attack be.PERF.M.SG
'The city got attacked.' <Urdu/Hindi>

Hence, the lexical semantics of the verb decides whether it allows a dative subject construction with the light verb *ho* 'be' or not. As the subject of 'attack' cannot be considered as an experiencer or recipient, it cannot have the dative subject.

Nepali also has these two classes of verbs with a locative marked object. But, in Nepali we find alternation of the dative/accusative with the locative marker.

(71) mai=le us=laai viswaas gar-eN
1SG=ERG 3SG=DAT/ACC trust do-PST.1SG
'I trusted him.' <Nepali>

(72) ma=laai us=maa viswaas thyo
1SG=DAT 3SG=LOC trust be.PST
'I trusted him.' <Nepali>

(73) mai=le Tren=maa viswaas gar-eN
1SG=ERG train=LOC trust do-PST.1SG
'I trusted the train.' <Nepali>

(74) ma=laai Tren=maa viswaas thyo
1SG=DAT train=LOC trust be.PST
'I trusted the train.' <Nepali>

In the above examples, the inanimate second argument is always marked by the locative marker in both canonical subject and dative subject sentences. But if an animate second argument is used in the construction that contains the light verb 'do' and a canonical subject, then the second argument can be marked by the dative/accusative marker *laai*.

Similarly, we find an alternation of dative/accusative and locative marker for the verb 'attack'. The animate object is marked by the dative/accusative marker, while the inanimate object is marked by the locative.

(75) mai=le mohan=laai aakraman gar- eN
1SG=ERG Mohan=DAT/ACC attack do-PST.1SG
'I attacked Mohan.' <Nepali>

(76) sena=le sahar=maa aakraman gar-yo
army=ERG city=LOC attack do-PST.3SG
'The army attacked the city.' <Nepali>

The semantic similarities and differences between these two locative marked classes will be discussed in section 4.6.

4.3.1.2. Ablative Marked Arguments

The verbs 'fear', 'ask', 'request', 'beg', 'resign' and 'avoid' are usually marked with an ablative marker in South Asian languages. These verbs can be divided into two classes based on whether dative subjects are allowed on the psych verbs. The verbs of Class I, e.g., 'fear', allow the dative subject in Punjabi and many other languages.

(77) maiN=nuuN saaNp=toN Dar lag-daa ae
1SG=DAT snake=ABL fear.M.SG strike-IMPF be.PRES
'I fear snakes.' <Punjabi>

(78) maiN saaNp=toN Dar-daa uuN
1SG snake.M.SG=ABL fear-IMPF.M.SG be.1SG
'I fear snakes.' <Punjabi>

The verbs of Class II, e.g., 'resign', do not allow for a dative subject. See the following example from Punjabi.

(79) maiN nokrii=toN istiifa diyaa
1SG job.F.SG=ABL resign.M.SG give.PERF.M.SG
'I resigned from the job.' <Punjabi>

As the lexical semantics of 'attack' do not allow the dative subject as shown in (70a), the lexical semantics of 'resign' do not allow the dative subject either.

Nepali uses two different ablative markers for the Class-I verb 'fear' and the Class-II verb 'resign'.

(80) u sarpa=dekhi DarauuN-cha
3SG snake=ABL fear-NPST
'He fears snakes.' <Nepali>

(81) us=le jaagir=baaTa raajinaamaa di-yo
3SG=ERG job=ABL resignation give-PST.3SG.M
'He resigned from the job.' <Nepali>

Recall that the Nepali ablatives introduced a serious challenge for existing theories within the spatial domain (cf. Chapter 3). Indeed, the same forms also introduce a challenge in the non-spatial domain as they show an alternation in this domain as well. The solution to the puzzle introduced by the data above as well as other issues related to ablatives is presented in section 4.6.

4.3.1.3. Comitative Marked Arguments

The second argument of Class VI that consists of the psych verbs 'love' and 'hate' is marked by the comitative marker. In many languages, these verbs allow a dative subject construction.

(82) raam=le siita=sanga prem gar-yo
Ram=ERG Sita=COM love do-PST.3SG
'Ram loved Sita. <Nepali>

In Malayalam, the verb *sneehik* 'love' has an accusative marker on the object, but the *noun + be* complex predicate has dative subject and comitative marked object.

(83) raaman siita-ye sneehik-unnu
Raman Sita-ACC love-PRES
'Raman loves Sita.' <Malayalam>

(84) raaman-∂ siita-ooD∂ sneeham uND∂
Raman-DAT Sita-COM love be.PRES
'Raman loves Sita.' <Malayalam>

There is another class of verbs with comitative second arguments. The verbs of Class V use the comitative to mark the second argument, but these verbs do not allow for a dative subject construction. This class includes 'fight', 'talk', 'meet', and 'marry', etc. See the Nepali example.

(85) mai=le u=sanga bihaa gar-eN
1SG=ERG 3SG=COM marriage do-PST
'I married her.' <Nepali>

It is important to note that the above examples provide an overview of the overall pattern of case marking. Different languages show minor patterns of differences in the case marking of the object. For example, Urdu/Hindi and Sindhi have the comitative marker for the second argument of 'meet', but it is marked by the dative/accusative marker in Punjabi and Nepali.

Section 4.4 discusses the data presented above and tries to identify the semantic reasons for the non-canonical marking displayed by these verbs.

4.3.2. Instrument Marking

In many of the surveyed South Asian languages, the same form is used to mark some spatial usage as well as instrument. The following table is reproduced from section 2.4.2. It shows the ploysemy related to the instrument marker.

Table 4.5: Other senses of instrument marker

Language	**Other senses of instrument marker**
Punjabi	LOC-beside, COM
Saraiki	LOC-beside, COM
Nepali	ERG
Manipuri	ERG
Pashto	LOC, PERL
Balochi	COM
Sindhi	COM
Malayalam	-
Urdu/Hindi	ABL, PERL

The table shows that the form used as instrument marker is also used as ergative, comitative or perlative markers. The companion relation marked by comitative is related to a spatial notion. The use of the same form for locative-beside and comitative in Punjabi and Saraiki provides more evidence for the connection between comitatives and a spatial relation. Hence, it can be concluded that many of the instruments are marked by the same form that is also used as a spatial marker.

The comitative-instrument and perlative-instrument patterns are discussed in detail in the following sections.

4.3.2.1. Comitative, Locative and Instrument

In Punjabi, Sindhi and Balochi, the same form is used to mark comitative and ablative usages. See the examples of comitative and ablative usage in Punjabi.

(86) maiN o=de naal baazaar gayaa
1SG 3SG=GEN COM market.M.SG go.PERF.M.SG
'I went to the market with him.' <Punjabi>

(87) o=ne caabi naal buuhaa khol-iaa
3SG=ERG key.F.SG INST door.M.SG open.PERF.M.SG
'He opened the door with the key.' <Punjabi>

The same form *naal* is also used as locative-beside marker in Punjabi and Sindhi. See the example from Punjabi.

(88) maiN kampiutar=de naal kitaab rakh-ii
1SG computer.M.SG=GEN beside book.F.SG put.PERF.F.SG
'I put the book beside/near the computer.' <Punjabi>

4.3.2.2. Perlative and Instrument

In Urdu/Hindi and Pashto, the same form is used to mark the perlative and instrument usage. See the Pashto examples.

(89) haamid pa baG tiir sho
Hamid PERL garden pass was
'Hamid passed through the garden.' <Pashto>

(90) asad paray pa chaara prekaR
Asad rope INST knife cut
'Asad cut the rope with a knife.' <Pashto>

We find examples of the use of the same form for both perlative and instrument usages in other South Asian languages as well. Torwali, an Indo-Aryan language spoken in Pakistan, uses the same form *te/de* for both of these usages.

(91) ahmad baagh te lhangu
Ahmad garden through passed
'Ahmed passed through the garden.' <Torwali>

(92) ahmad-e kera de bhowai chi ki
Ahmad-ERG knife with apple cut did
'Ahmed cut the apple with the knife.' <Torwali>

The ancient Indo-Aryan languages Sanskrit and Pali too have the same form for both instrument and perlative usages. See the Pali inflection *-ena* used in these examples.

(93) kassako sarena sigaalam vijjhati
farmer arrow.INST jackal shoots
'The farmer shoots the jackal with (an) arrow.' <Pali>

(94) Assaa maggena dhaavanti.
Horse.PL Path.PERL run
'Horses run through the path/road.' <Pali>

Frankfurter (1883:27) presents examples of perlative usage of the instrument marker for Pali. Similarly, the Sanskrit instrument inflection is used for both instrument and perlative usages (Whitney 1889:94).

Hence, South Asian languages provide many examples of the common form for both spatial and instrument usages. Some of the semantic reasons for these patterns have already been already mentioned in section 4.2.6. Further discussion on this topic is presented in sections 4.4 and 4.5.

4.3.3. Marking on the Addressee

The examples presented in the above two sections showed that some spatial markers are also used to mark NCSA and instrument. As mentioned in section 2.4.1, the addressee in some languages is marked by the same form that is used as a spatial marker.

Table 4.6 is a reproduction of the table given in section 2.4.1. It shows that addressees are marked by either of the dative, allative, comitative or ablative markers in the surveyed South Asian languages. We also find a clear pattern in Table 4.6. The semantic reasons for the use of the same form for spatial usages and addressees are again discussed in sections 4.4 and 4.5.

Table 4.6: Marking on the addressee

Language	**Say to**	**Ask (a question)**
Punjabi	DAT/ACC	ABL, DAT/ACC
Saraiki	DAT/ACC	ABL
Nepali	DAT/ACC	COM
Manipuri	LOC/DAT	LOC/DAT
Pashto	ALL	ABL
Balochi	DAT/ACC	ABL
Sindhi	DAT/ACC	ABL
Malayalam	COM	COM
Urdu/Hindi	ABL/INST/COM	ABL/INST/COM

4.4. Discussion

Section 4.2 presented different theories about argument realization and instrument marking. In section 4.3, the South Asian data that show the use of spatial markers for non-spatial usages, including non-canonical second argument (NCSA) and instrument marking were presented. One can apply these theories of argument realization to find the semantic reasons for NCSA marking. The current section points out the interesting points and challenges introduced by the South Asian data for the existing theories.

In particular, this section investigates three questions. Why do spatial markers extend their domain and start marking non-spatial usages? How can we model these non-spatial usages using semantic constructs or features? What are the conditions/semantic reasons that are responsible for NCSA marking?

The section mainly focuses on the NCSA marking. The other usage of spatial markers for other non-spatial usages, i.e., instrument and addressee is mentioned in passing. A comprehensive semantic analysis of these two usages is given in section 4.5.

4.4.1. Why Are Spatial Markers Used?

It is an interesting puzzle why spatial markers can be used to express non-spatial usages. Chapter 3 shows spatial usages of spatial markers. These spatial markers are modeled using spatial features. On the other hand, section 4.2.4 about lexical

entailments shows that different semantic roles like agent, patient, recipient and instrument, etc. are expressed by non-spatial features. However, a form used as spatial marker having spatial features can also be used to mark these non-spatial usages having non-spatial features.

The reason for the extension of spatial features to non-spatial usages is partially provided by the localist approach. This approach models the predicates in terms of a theme present at a location or moving from one location to another. Hence, the localist model for a predicate considers its argument as theme, source or goal, etc. It is possible and predictable that these abstract source or abstract goal arguments are marked with the markers for the real source and real goal.

The ablative marked arguments give good examples in this regard. The ablative marked argument or adjunct of the verbs 'take' and 'go' is marked by the ablative marker.

(95) jamiil=ne bilaal=se kitaab lii
Jameel=ERG Bilal=ABL book.F.SG take.PERF.F.SG
'Jameel took the book from Bilal.' <Urdu/Hindi>

(96) jamiil karaaci=se aa-yaa
Jameel Karachi=ABL come-PERF.M.SG
'Jameel came from Karachi.' <Urdu/Hindi>

The ablative marked entity in the above example is the source from which the theme starts its movement. The source, *Bilal* in (95) and *Karachi* in (96), is marked by the source marker, i.e., the ablative marker. Similarly, the object of the verb 'resign' is also marked by the ablative marker.

(97) bilaal=ne nokrii=se istiifa di-yaa
Bilal=ERG job.F.SG=ABL resign.M.SG give-PERF.M.SG
'Bilal resigned from the job.' <Urdu/Hindi>

The above sentence can be modeled as following. In the event of 'resign', *Bilal* (the theme) moves away from the *job* (the source). Hence, the source argument is marked by the ablative marker. Similarly, the use of some other spatial markers in non-spatial domains can also be explained in terms of the localist approach.

However, the above explanation does not answer all of the questions related to the usages of spatial markers in non-spatial usages. Section 4.4.3 discusses why some non-spatial arguments are marked by spatial markers, while others are not.

The explanation provided by the localist approach is incomplete because of another reason. There is spatial marking on many arguments that cannot be explained by the localist approach. The instrument is marked by the locative-beside/comitative marker in Punjabi as shown in section 4.3.2. Ostler and Jackendoff did not mention any explanation about the metaphorical use of the marker for locative-beside to mark the accompaniment (comitative marked).

Similarly, section 4.3.3 showed that the addressee of the verb 'ask' in some languages is marked by the ablative marker. The addressee can be considered as the goal (of the theme 'question'), but the marking suggests that it is the source. The reason is that another metaphor, discussed in section 4.5, is responsible for the ablative marking.

Hence, it is concluded that metaphors are responsible for the usages of spatial markers for non-spatial usages. Some metaphors are provided by the localist model, but there are other metaphors as well. Section 4.5 discusses the possible semantics and related metaphorical usages for all the spatial markers for non-spatial usages listed in section 4.3 (the South Asian data).

4.4.2. How Can Non-Spatial Usages Be Modeled?

I suggested above that different metaphors are responsible for the use of spatial markers for non-spatial usages. The next question concerns the semantic modeling of these non-spatial usages. Suppose we are using semantic features to model different kinds of arguments. In this case, we need to ask ourselves which type of semantic features should be used to represent the spatially marked argument.

Take the example of the instrument marker. We can list the (non-spatial) lexical entailments for the instrument argument. If Dowty's set of proto-role properties is considered, a prototypical instrument has the following properties: "causing an event or state of change in another participant", "exists independently of the event" and "causally affected by another argument". In Grimm's system, the properties will be [instigation] and [total persistence].

None of the above listed agency properties are related to spatial features discussed in Chapter 3. However, the Punjabi locative-beside *naal*, Pashto locative/perlative *pa* and Urdu/Hindi ablative/perlative *se* are used as the instrument marker. Should the list

of agency features for the instrument include the spatial features corresponding to these spatial markers?

The inclusion of spatial features into the set of agency properties/features is not a good option. There are many spatial markers corresponding to the instrument usage. If we have to add the spatial features in the feature set of the instrument usage, then we have more than one potential spatial feature set. The Punjabi locative-beside *naal* has the feature [PLACE BESIDE]. This form is used to mark the instrument usage too. The companion metaphor (Lakoff & Johnson 1980) gives the semantic reason for this pattern.

Should we include the feature [locative-beside] or [companion] in the agency feature list of the prototypical instrument? Hence, should Grimm's system now include [companion] along with [instigation] and [total persistence]? This is not a good option because the Nepali ergative/instrument marker *le*, Pashto locative/perlative/instrument marker *pa* and instrument markers of some other languages do not allow it. These markers are used to mark the instrument usage, but are not related to the feature [companion]. Furthermore, Pashto *pa* would demand an introduction of the feature [path-through] in the list of agency features corresponding to the instrument usages. So, we cannot decide on a unique spatial feature set for the instrument usage and would therefore need to extend the proto-role entailments in an unsystematic and therefore unacceptable manner.

This issue is not a problem specific to the instrument usage. The addressee of the verb 'ask' is marked by either of the dative, comitative or ablative markers in South Asian languages. These case markers are roughly correlated to the features [goal/location], [companion] and [source], respectively. But, do we need to assign all these three features to every addressee argument of the verb 'ask'? The answer clearly is "no" because it will result in confusion and proliferation of features in the agency domain.

Hence, we do not introduce the spatial features in non-spatial domain. Jackendoff (1990) and Butt (2006b) followed a similar approach when they introduced two different layers for spatial and non-spatial constructs. There is an interaction and correlation of features between these layers, but these are not mixed with each other to the point of being one system.

4.4.3. Why Are NCSA Markers Used in Place of Canonical Markers?

The above discussion discussed the reasons of spatial marking for non-spatial usages, and posed the question of how we can represent the interaction of spatial and non-spatial domains. The current section focuses on the reasons for non-canonical marking in South Asian languages. These languages have canonical marking for the second argument (i.e., the object) of some verbs, and non-canonical marking for the second argument of other verbs. Is there a criterion to decide whether a non-canonical marker will be used with certain verbs? The following discussion investigates whether the argument realization theories, mentioned in section 4.2, can answer this question.

4.4.3.1. Event Structure Approach

The event structure approach provides the reason why some second arguments are marked by non-canonical markers. As discussed in section 4.2.2, Levin says that the second argument of many predicates is not a structured variable, but it is introduced by the predicate constant. These verbs, called non-core transitive verbs (NCTV), can be marked by non-canonical markers. The verbs of the South Asian NCSA classes are NCTVs. Similarly, all the NCSAs are rhemes in Ramchand's model. Hence, we have a partial explanation of the non-canonical marking. But these approaches do not explain why some constant variables/rhemes are marked by non-canonical markers, while others are not.

4.4.3.2. Transitivity Hierarchy

The transitivity gradient and hierarchy tries to answer the question whether we have a criterion to predict that a certain verb has non-canonical marking on its arguments. The discussion in section 4.2.1 presented the concept of scale and gradience of transitivity. If we look at the examples of the verbs with NCSA marking in Table 4.4, we find that all these verbs have a lesser number of high transitivity properties, i.e., these verbs are less transitive. Hence, the non-canonical marking on these verbs fulfills the predictions of Hopper & Thompson.

Tsunoda introduced the concept of cut-off point for the verb types that allow non-canonical marking. In this hierarchy, the verb type "pursuit", e.g., *search*, follows the verb type "perception", e.g., *see*. It implies that if pursuit verbs do not allow non-canonical marking in language, then the verb type preceding it, i.e., perception verbs will do the same. The South Asian data provides a negative example for this claim.

The verb type "pursuit" in Urdu/Hindi and some other South Asian languages allows canonical marking only. See the following examples.

(98) a. maiN=ne kitaab DhuunD-ii
1SG=ERG book search-PERF.F.SG
'I searched for a/the book.'

b. maiN kitaab DhuunD rahaa huuN
1SG book search PROG.M.SG be.PRES.1.SG
'I am searching for a/the book.'

c. maiN=ne apne dost=ko DhuunD-aa
1SG=ERG self friend=ACC search-PERF.M.SG
'I searched for my friend.'

d. maiN apne dost=ko DhuunD rahaa huuN
1SG self friend=ACC search PROG.M.SG be.PRES
'I am searching for my friend.'

All of the above four sentences have canonical nominative/ergative marking on the subject and canonical nominative/accusative marking on the object. Hence, the verbs of the pursuit class do not allow for the non-canonical marking, but the verb types following and preceding the pursuit class in Tsunoda's hierarchy allow for the non-canonical marking.

However, the verb *dikh* 'see' of the verb type "perception" has non-canonical marking on the subject. Similarly, the complex predicates *sunaaii de* 'hear' and *xuSbuu/aavaaz aa* 'hear'/'smell' also have dative subjects.

(99) bilaal=ko aik pahaaR dikh-aa
Bilal=DAT one mountain appear-PERF.M.SG
'Bilal saw a mountain.'

It is in contradiction of Tsunoda's claim that all the verb types preceding a canonical marked verb type do not allow for the non-canonical case marking. It is important to note that Tsunoda claimed for non-canonical marking on both the subject and the object/second argument.

Malchukov tried to solve this problem by introducing a two-dimensional hierarchy. Each wing of the hierarchy deals either with a non-canonical subject or a non-canonical second argument. Hence, if we are studying NCSA marking, we need to concentrate only on the part of the hierarchy that is related to the non-canonical patients (i.e., the second argument).

(100) Transitivity Hierarchy by Malchukov

Contact → Pursuit → (Motion)

Effective Action

Perception/Cognition → Emotion → (Sensation)

The verb type "perception" is less transitive because the subject has a fewer number of prototypical agentive properties. It is the reason why Malchukov put it on the decrease in agentivity part of the hierarchy.

On the other hand, the verb type "pursuit" is less transitive because the second argument is less affected. It is placed in the decrease in patient-hood part of the hierarchy. As both parts of the hierarchy have different cut-off points, the Urdu/Hindi data presented above do not pose a problem for Malchukov's transitivity hierarchy.

Hence, Malchukov's hierarchy predicts the non-canonical marking in South Asian languages better than Tsunoda's hierarchy. But it has problems as well. In the decrease in patient-hood part of the hierarchy, the verb type "pursuit" is preceded by the verb type "contact", which allows for non-canonical marking as well. See the following examples.

(101) maiN=ne mez=par mukka maar-aa
1SG=ERG table=LOC_on punch.NOM hit-PERF.M.SG
'I punched at the table.'

(102) maiN=ne bilaal=ko mukka maar-aa
1SG=ERG Bilal=ACC/DAT punch.NOM hit-PERF.M.SG
'I punched Bilal.'

The contact verb 'punch' in the example (101) has the non-canonical locative marking on the second argument. Hence, we find a negative example for Malchukov's hierarchy.

Hence, the behavior of the verbs 'hit'/'punch' do not fit in Tsunoda's and Malchukov's hierarchy. Moreover, the transitivity approach only predicts the cut-off point between canonical and non-canonical marking. For the choice of which specific marker is used, we need to consult ideas from other approaches.

4.4.3.3. Localist Approach

The localist approach is the most useful in predicting the appropriate non-canonical case marker for an argument. In section 4.4.3.2, we discussed the verbs 'hit' and

'punch'. The second arguments of these verbs are the goal locations. As these arguments must be modeled as goal in the localist approach, these are marked by the locative marker.

Similarly, section 4.4.1 discussed the semantic reason for the NCSA marking of the verb 'resign' (cf, example (97)). As the theme (*resigning person*) moves away from the location (*job*), the ablative marker can be used to mark the non-canonical (source) second argument, i.e., *job*.

However, a fundamental question is not explicitly answered by the localist approach. The source of the verbs 'take', 'come' and 'resign' are marked by the ablative marker (cf. section 4.3.1), but we do not find this marker on the source of the verb 'give'.

(103)	jamiil=ne	bilaal=ko	kitaab	dii
	Jameel=ERG	Bilal=DAT	book.F.SG	give.PERF.F.SG
	'Jameel gave the book to Bilal'.			<Urdu/Hindi>

In this sentence, *book* is the theme, *Jameel* is the source and *Bilal* is the goal. However, the source is marked by a canonical marker (in place of the ablative marker), because it is a prototypical subject (or prototypical agent). The prototypical subjects and objects are marked by the canonical markers. When an argument deviates from the prototypical image and the clause become less transitive then that argument may take a spatial marker.

The reasons for the deviation from the prototypical images are found using the other approaches. If a non-prototypical argument needs non-canonical marking, then the reason for choosing a particular spatial marker may come from the localist approach.

4.4.3.4. Lexical Entailments

The feature-based models try to provide the reason for the choice of appropriate case marker. These models do not introduce a transitivity hierarchy, but they assign features to all the arguments including the object. If an object has certain features, it will be marked by a certain spatial or non-spatial marker.

Dowty pioneered the concept of decomposing the thematic role into simpler semantic lexical entailments, but his proposal is mainly to identify agent and patient

that in turn predict the subject and object of the clause. There is no discussion of non-canonical case marking in his proposal.

However, other proposals using features based on Dowty-like semantic properties try to answer our question. Section 4.2 showed how different regions of Grimm's agency lattice can account for the distribution of Greek case markers. Similarly, the spatial markers for NCSA may have particular regions on the agency lattice, which determine their use and distribution.

There are two potential problems to this approach. The first one is that any proposed feature set starts to grow slowly. Primus (1999) and Ackerman and Moore (2001) added new properties to Dowty's property set. Grimm (2005) proposed a new property set (by decomposing Dowty's set). Later on, however, he introduced the feature [potency] to model instrument subjects of English (Grimm 2007). The discussion in section 4.5 will introduce some more features to model the South Asian NCSA markers.

Thus, modeling more phenomena means blowing up the size of the feature set. Do we really want to do it? It seems that there may be a big universal set of features and every language uses a small subset of this universal set. So, if someone is modeling case markers of a particular language, they will only be concerned with a small subset of the semantic features from the universal set and the size of complete feature set will not be a problem.

The second issue is fundamental. Does the case marking on an argument solely depend on its semantic features, or can some other features of the clause influence on it? See the following Nepali examples.

(104) mai=le us=laai viswaas gar-eN
1SG=ERG 3SG=DAT/ACC trust do-PST.1SG
'I trusted him.' <Nepali>

(105) ma=laai us=maa viswaas thyo
1SG=DAT 3SG=LOC trust be.PST
'I trusted him.' <Nepali>

When the predicate has the agentive-experiencer subject and the light verb 'do', the object is considered as the affected argument and is marked by the dative/accusative marker. On the other hand, in case of the light verb 'be' and the experiencer subject, the object is considered to be a location and is marked by the locative.

It is important to note that Hopper & Thompson (1980) associated the low transitivity with the semantic properties of the clause, and not with the individual argument. Hence, aspect can be responsible for the non-canonical marking. An example of the role of telicity/boundedness is present in the following alternation of accusative/partitive marker.

(106) a. Madis joob oma tee aera
Madis drink.3SG.PRES own tea.GEN/ACC PREV
'Madis will drink up the tea,' <Estonian>

b. Madis joob teed
Madis drink.3S.PRES tea.PART
'Madis is drinking tea.' <Estonian>

Hence, aspectual factors are responsible for case marker choice for the object/ second argument. Hopper & Thompson (1980) and Kiparsky (1998) gave the explantion of similar alternations in terms of telicity/boundedness. However, Ackerman & Moore (2001) introduced a proto-patient lexical entailment "bounding entity" to model this alternation. Hence, we can claim that the object of (106a) has the feature [+bounding entity] and the object of (106b) has the feature [–bounding entity].

However, it is not easy to invent a proto-patient property for the Nepali alternation shown above. Malchukov (2006) discussed a similar issue related to accusative/instrument alternation in Russian.

(107) a. On krutil rulj
3SG rotate wheel.ACC
'He rotated the wheel consciously.' <Russian>

b. On krutil rulj-om.
3SG rotate wheel-INST
'He rotated the wheel unconsciously.' <Russian>

In this example, the volition feature of the subject is responsible for the case marking on the object. This cannot be modeled by proposing a new feature for the object. On the contrary, we need to consider semantic features of the whole clause to determine the case marking of a particular argument.

Under my view, the lexical entailment-based approach is the best among the approaches discussed above. It can incorporate solutions from different approaches about the reasons of particular marker or alternation by introducing an entailment for the proposed reason. However, the approach needs to be extended so that features

from the clause are allowed to influence the case marking of an argument, i.e., there should be no claim that all the features responsible for a particular case marker on an argument are only related to the semantic properties of that argument.

4.4.3.5. Linking Theories

The linking theories provide an explanation why we have oblique marking on some arguments. We can use these theories to explain some NCSA and alternations presented in section 4.3. The thematic roles corresponding to the argument can predict the appropriate case marker for a given usage. However, the linking theories usually have a small set of thematic roles. The previous section has shown that sometimes we need to decompose monolithic thematic roles into a set of lexical entailments. Something similar may be required for some thematic roles related to the NCSA. Zaenen (1993) introduced this approach, when she used Dowty's proto-role properties to identify [–r] and [–o] features for the argument.

In summary, none of the approaches can provide a full solution for predicting a non-canonical correct case marker on an argument. We therefore need to borrow the successful parts from all of the different approaches. A semantic feature-based model can incorporate different solutions in it. A comprehensive explanation of the South Asian data presented in section 4.3 is provided in section 4.5.

4.5. Towards Modeling South Asian Data

The important theoretical issues related to spatial marking on non-spatial usages have already been discussed. This section provides a brief summary of the big picture painted above and then focuses on the finer details. It provides an explanation for all of the phenomena listed in section 4.3. The next two sections then provide a review of our questions about the semantic reasons for spatial markers being extended to non-spatial usages, and the semantic reasons for marking non-canonical second arguments with particular case markers. After that, the spatial markers used to mark non-spatial usages are considered one by one and the semantic reasons and features corresponding to those usages are explained.

4.5.1. Spatial Markers for Non-Spatial Usages

The data in section 4.3 shows that many non-spatial usages are marked by the same form that is primarily used to mark a spatial usage. There are spatial metaphors that

are responsible for the use of spatial markers in non-spatial agency domains. The abstract concepts of agency domain can be conceptualized by using spatial metaphors. For example, the instrument is considered as a companion that is present beside the agent, or it can be considered as a path through which action traverses from the agent to a patient. As we conceptualize the instrument in terms of spatial roles, the markers of those roles are also used to mark the instrument.

However, we decided not to unnecessarily introduce spatial features into the agency domain. The two domains remain separate. Instrument is metaphorically imagined as a spatial concept, but it does not have any spatial feature. It is modeled using agency features in the same way as the other arguments (marked by non-spatial markers) are modeled.

It is important to note that the metaphorical extension of the usage of spatial markers is different from the real use of space in the agency domain. We do not in fact use spatial features in the agency domain. In contrast, in the spatial domain, modeling the recipient of the verb 'give' and the source of the verb 'take' must involve the features [goal] and [source]. We cannot model these usages without the use of spatial features.

4.5.2. Reasons for NCSA Marking

The verbs with non-canonical second arguments (NCSA) are semantically different from the verbs with canonical objects. These bivalent verbs are lower in transitivity and their subjects and objects have a lesser number of prototypical agent and prototypical patient properties. A non-prototypical argument can be marked by a non-canonical marker.

The data from South Asian languages show that spatial markers are used as non-canonical markers. The selection of an appropriate spatial marker for a particular NCSA depends on the use of an appropriate metaphor. Sometimes localist models help to provide the appropriate metaphor (and thus the appropriate spatial case marker).

There is no universal rule to predict the boundary line between verbs with NCSA and other verbs. Transitive hierarchies break down for the verb 'hit'. The lexical entailment approach has its own shortcomings discussed in 4.4.3.4, however, it is the

best approach to model NCSA marking and the related alternations because we can introduce a new or modified lexical entailment for a new phenomenon.

4.5.3. Ablative marked argument

The ablative marker is used to mark the arguments that can be thought of as a source. It is used to mark the object of Class-II verbs, e.g., 'resign'. The job/position argument for the verb 'resign' is the source from which the subject moves away. Hence, it is marked by the ablative marker.

The stimulus of the verb 'fear' (Class I of NCSA verbs) is the origin of the 'fear' and hence it is marked by the ablative marker. However, there is a difference between these two ablative marked arguments. Nothing actually or metaphorically travels from stimulus to experiencer in the verb 'fear'. This difference is encoded by the use of two different ablative markers in Nepali corresponding to the objects of these two verbs.

(108) u sarpa=dekhi DarauuN-cha
3SG snake=ABL fear-NPST
'He fears snakes.' <Nepali>

(109) us=le jaagir=baaTa raajinaamaa di-yo
3SG=ERG job=ABL resignation give-PST.3SG.M
'He resigned from the job.' <Nepali>

Compare the non-spatial usages of *dekhi* and *baaTa* with the spatial usages.

(110) us=le dilli=dekhi kathmandu=samma baaTo banaa-yo
3SG=ERG Delhi=ABL Kathmandu=LOC_to road make.PST
'He built a road from Delhi to Kathmandu.' <Nepali>

(111) u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=LOC_to ran-PST
'He ran from Delhi to Kathmandu.' <Nepali>

The object/stimulus of 'fear' is marked by the static ablative *dekhi* that is used to mark the static origin (e.g., starting point of a road). The dynamic ablative marker *baaTa* is used with the object of 'resign' because it shows the abstract, yet dynamic motion away from a place.

The ablative marker is also used to mark the addressee of the verb 'ask' in many South Asian languages, as shown in Table 4.6. It is important to note that the addressee of the verb 'say' does not have the ablative marking in the same languages. The addressee can be understood as the recipient of speech and that is why many

languages use the dative marker to mark the recipient of the verb ‘say’. However, when a question is asked, the addressee is the recipient of the speech as well as the potential source of the reply. This is why the addressee of ‘ask’ is marked by the ablative marker (for potential source) in many languages. Thus languages can prefer any of these two roles for the selection of case markers.

We learn an important lesson from the above discussion. An argument, e.g., addressee of ‘ask’, can be understood either as a source or as a goal, depending on the metaphor that is chosen. Similarly, a predicate can have more than one source argument. In the case of ‘ask’, the speaker is the source of the question and the addressee is the potential source of the reply.

4.5.4. Dative Marker

The dative marker not only marks the experiencer subjects of Class I, III and V of NCSA verbs, but it can also alternate with the locative marker. Here, I am interested in this alternation because it helps to understand the semantics of locative marked arguments. See an example of dative/locative alternation corresponding to the verb ‘punch’/’hit’.

(112)	main=ne	mez=par	mukka	maar-aa
	1SG=ERG	table=LOC_on	punch.NOM	hit-PERF.M.SG
	‘I punched at the table.’			<Urdu/Hindi>

(113)	main=ne	bilaal=ko	mukka	maar-aa
	1SG=ERG	Bilal=DAT	punch.NOM	hit-PERF.M.SG
	‘I punched Bilal.’			<Urdu/Hindi>

If the target of ‘punch’ or ‘hit’ is animate, it is marked by the dative marker, while the inanimate target is marked with the locative marker. It makes clear that the dative marker is used as a special goal marker that is related to some special semantic property.

We claim that the reason for dative vs. locative marking on the goal of ‘hit’ is due to affectedness/change or perception of affectedness. When an inanimate object is hit, there is no change in it. Hence, it is treated as a canonical goal and marked by the locative marker. On the other hand, when an animate, i.e., a sentient being is hit, there

is no visible change,[1] but it is psychologically affected. Hence the dative marker marks the affected goal.

Both Urdu/Hindi and Nepali allow the alternation shown in (112)–(113). However, we find a different behavior of both languages in some other constructions. Nepali has a dative-accusative/locative alternation of the second argument of 'trust' or 'attack' as shown in (114)–(115). Urdu/Hindi uses the locative marker in the equivalent of both of these sentences.

(114) mai=le mohan=laai aakraman gar- eN
1SG=ERG Mohan=DAT/ACC attack do-PST.1SG
'I attacked Mohan.' <Nepali>

(115) sena=le sahar=maa aakraman gar-yo
army=ERG city=LOC attack do-PST.3SG
'The army attacked the city.' <Nepali>

My analysis is that Nepali considers the animate goal as the affected goal and uses the dative marker for it. The inanimate goal is considered as the unaffected goal and is marked by the locative marker.

As mentioned earlier, Urdu/Hindi uses the locative marker for both animate and inanimate goals of the verb 'attack'. It is different from the dative/locative alternation for the 'hit'/'punch' shown in (112)-(113). There is some difference between the objects of 'hit' and 'attack'.

The difference between these two can be understood in terms of the achieved vs. intended goal. If you hit someone (say with a stick), it means that the stick successfully touched him/her (and reached the target). On the other hand, attacking someone with the stick does not always mean that the stick touched the target. The attack can be stopped. So, the goal of 'hit' is always achieved, but the goal of 'attack' is not. This difference can help us understand the Nepali vs. the Urdu/Hindi data. The Nepali dative marker is used with all kinds of affected goals. It does not distinguish between achieved vs. intended goal. Urdu/Hindi uses the dative marker with the achieved affected goal, while the intended affected goal is marked by the locative marker. The semantic difference between the usages of Urdu/Hindi and Nepali dative markers is shown in Table 4.7.

[1] The animate being could get wounded, and an inanimate thing could get cracked or broken. However, the verb 'hit' does not require that the object get wounded or cracked.

Table 4.7: Case marking patterns for the verbs 'hit' and 'attack'

	Hit (achieved goal)	**Attack** (intended goal)
Urdu/Hindi	Anim : DAT Inanim : LOC	Anim : LOC Inanim : LOC
Nepali	Anim : DAT Inanim : LOC	Anim : DAT Inanim : LOC

Hence, we can conclude that the Nepali dative marker marks the (achieved/intended) affected goal, while Urdu/Hindi dative marker marks the achieved affected goal. The dative experiencer subjects (of Class-I, -III and -V verbs) are the achieved affected goal, and hence marked by dative marker by both Urdu/Hindi and Nepali. The dative marked arguments of Class-III verbs are the intended affected goal, hence these are marked by dative marker only in Nepali.

4.5.5. Locative marked argument

The locative marker is used to mark the NCOs of verbs of Class III, e.g., 'trust', and Class IV, e.g., 'attack'. The semantic reasons for the locative marker on the object of the verb of 'attack' were discussed above in section 4.5.4. In Urdu/Hindi, the dative marks more specialized goals than the goals marked by the locative. If a goal is achieved and affected, it is marked by the dative marker. All the other goals are marked by the less semantically specific locative marker.

In Nepali, all unaffected goals are marked by locative, and all the affected goals are marked by the dative. Here, the term "affected" includes both potentially and actually affected arguments.

Now, we need an explanation of locative marking on the verb 'trust'. The object of 'trust' is not similar to the experiencer. You can trust someone without his/her knowledge or experience. Hence, the object of 'trust' is also an intended goal like the object of 'attack'. It is why both objects are marked by the locative marker. For the same reason, the locative marker has a similar (to ((114)–(115)) alternation with the

dative marker in Nepali. Besides, there is another interesting locative/dative alternation for the verb 'trust' in Nepali.

(116) mai=le us=laai viswaas gar-eN
1SG=ERG 3SG=DAT trust do-PST.1SG
'I trusted him.' <Nepali>

(117) ma=laai us=maa viswaas thyo
1SG=DAT 3SG=LOC trust be.PST
'I trusted him.' <Nepali>

The animate object of the 'do' light verb construction is marked by the locative dative marker, while the animate object of the 'be' light verb construction is marked by the locative marker. According to our previous analysis, the object of (117) is an affected goal, and it must be marked by the locative marker. However, there is a difference in agentivity in both examples. The example (116) is an event, and the ergative marked subject of the *noun* + *do* predicate is considered as agentive. This agentive subject can be responsible for the affectedness of the object. However, example (117) is a state that has an experiencer subject. Hence, the object of this construction can be considered to be an unaffected goal and is marked by the locative marker.

Given the data and the discussion, I conclude that the locative marker is related to the unaffected or intended goal. It is in competition and alternation with the dative marker.

4.5.6. Comitative Marked Argument

The comitative marker, like the dative marker, is not a core spatial marker, but it is used to model spatial relations too. The Punjabi examples in section 4.3.2.1 show that the same form is used for both the locative-beside and the comitative marker.

The comitative marker is used with the arguments that are involved in the event. The marker for accompaniment provides a good metaphor for the other involved argument. The Class-VI verbs are reciprocals, e.g., 'meet' and 'fight', etc. The second argument of these verbs is involved in the action and hence marked by the comitative marker.

Similarly, the addressee argument is usually marked by the dative marker, as it is the recipient of the speech. But there is a difference between the recipient of the verb 'give' and the recipient/addressee of speech verbs. As the addressee is less affected

than the canonical recipient, some languages use the comitative marker to mark it. It shows that the comitative marked addressee is involved in the communication act.

The objects of Class-V verbs 'love' and 'hate' are also marked by the comitative marker. The loved/hated one argument is involved in the action. Hence it is marked by the comitative marker. However, there is an open question left. The verbs 'love'/'hate' seem similar to the verb 'trust', but many South Asian languages select different markers for the second argument of these verbs. The fine grained semantic difference between these two needs to be investigated.

The instrument marker in many languages is marked by the same form that is used as the comitative marker. This can be argued to be due to the companion metaphor, according to which the instrument is involved in the event along with the agent.

Hence, the analysis of all the listed semantic usages of the comitative marker shows that the comitative marker is used to mark those arguments that are *involved* in the event.

4.6. Conclusion

This chapter presented non-spatial usages of spatial markers. The South Asian data presented some new observations and challenges for the existing models. The analysis of this data finds that no single model can explain all the issues related to these non-spatial usages especially non-canonical second argument marking. Concepts from different approaches (high and low transitivity, localist, agency/features) must all form part of a complete explanation.

I provided a partial explanation for NCSA marking crosslinguistically by looking at six classes of South Asian verbs that have NCSA marking. I investigated the case marking alternations and possibilities of these verb classes in some detail and provided an explanation of the distribution and usage of the NCSA marking. The analysis is based mainly on semantic features and implicates the use of different metaphors for allowing for an extension from spatial to non-spatial meanings. In particular, I showed that different semantic usages of a particular marker are systematically related to each other.

Chapter 5

Conclusion

This dissertation discussed the relation between spatial expressions and case in South Asian languages. It analyzed the diachronic development of case markers and conducted a synchronic study of non-spatial usages of spatial forms to investigate this relation.

The data for the dissertation came from the survey of case marker usages in ten South Asian languages. The surveyed languages belong to four language families: Indo-Aryan (Haryani, Nepali, Punjabi, Saraiki, Sindhi and Urdu/Hindi), Iranian (Balochi and Pashto), Dravidian (Malayalam) and Tibeto-Burman (Manipuri). Beside these languages, data from Old Urdu was extensively explored. South Asia is considered as a "linguistic area" or *Sprachbund* (Emeneau 1956) that depicts convergence of linguistic features due to language contact. The data patterns emerging from this survey confirms the observation that areal features are found common among the languages of different families.

During the collection of the data, grammar books were consulted, but it was not the only source of information. Native speakers and books/newspapers were more important source for the data. Most of the informants could be contacted through internet, which enabled me to obtain any missing piece of data or clear up any confusion. This approach gave first hand access to the raw data that allowed for a better analysis.

The survey listed all the usages of a case marker. This approach provided interesting observations. For example, *da/ta* is used as a locative marker in Manipuri. However, it is also used to mark the recipient, addressee, purpose, causee and object of the verb 'hit'. This and other data called for an explanation why the domain of a locative is extended to these non-spatial usages. Similarly, the Pashto form *pa* is used to mark locative as well as instrument usage. The Haryani form *nae* and the Balochi form *-ara* (found in a dialect of Balochi) are used as dative and ergative markers. This data confirms Butt's (2006b) proposal that ergative and dative markers may originate from the same source.

Another important issue is the use of multiple markers in the experiencer subject construction of Pashto. These experiencer subjects can be marked by allative *ta*, dative *la* or genitive *da*. A significant amount of this data (specially related to spatial markers) was analyzed in detail. However, some interesting patterns that were only identified as part of the survey fell outside of the scope of this dissertation and await further inquiry in future work.

The survey investigated the diachronic development of case markers as well. It found that many core case markers like ergative and accusative originated from spatial terms. The Sindhi accusative marker *khe*, the Saraiki accusative marker *kuuN* and the Urdu/Hindi accusative marker *ko* all appear to originate from the Sanskrit locative *kakSe* meaning 'in the armpit' or 'at the side' (Beames 1872). This data has set a background for the investigation about the synchronic relation of spatial expressions and case.

The study focused on the polysemy and fine grained differences among different spatial markers in the surveyed languages. These problems were already addressed by Ostler (1971), Jackendoff (1990) and Kracht (2002). However, the models proposed by them cannot explain all the issues raised by the South Asian data. This dissertation pointed out merits and shortcoming of each of these models with reference to the South Asian data. It proposed the following alternative underspecified feature-based model for spatial markers.

(1) Underspecified feature based model for spatial markers

- Every spatial marker has three primary features: PLACE, PATH and DYN(amic).
- Each of the above features may have a set of features as the value.
- The feature PLACE may have following values:

  null, ON, AT, IN, BESIDE, …

- The feature PATH may have the following features as values:

  S(ou)RC(e) : the theme leaves the place

  END : the theme enters the place

- The feature DYN(amic) indicates whether an action/activity is performed or it is a static situation.
- The features SRC, END and DYN have a positive (+) or negative (–) value or they can be underspecified. The underspecified feature is represented by the feature's name only.
- The underspecified PATH is shown by a parenthesis around the structure.

Some sample entries of pure/abstract spatial markers according to this model are given below.

(2) Sample entries of pure/abstract ablative, perlative, allative or locative usages

a. Ablative : [PLACE X, PATH [SRC +, END –], DYN]

b. Allative : [PLACE X, PATH [SRC –, END +], DYN]

c. Perlative : [PLACE X, PATH [SRC +, END +], DYN]

d. Locative : [PLACE X]

In particular, the concept of PLACE and PATH allows modeling of Sindhi, Punjabi and Sariaki ablative markers whose usages differ in the configuration of the theme with respect to the source (see the examples given in section 3.3.1 for the usages of these ablative markers). The features [SRC] and [END] of the feature [PATH] allows the modeling of the polysemous markers for different types of paths (i.e., usages marked by ablative, perlative and allative markers). Hence, the ablative-perlative marker found in many of the surveyed languages is modeled easily (see section 3.3.2 for examples).

The most important discovery related to the spatial usages is the presence of dynamic and static ablative markers in Nepali and Malayalam. As the example in (3) shows, the Nepali static ablative marker *dekhi* is used to mark the static origin (e.g., starting point of the road) and the dynamic ablative marker *baaTa* is used to mark the source of the dynamic motion (e.g., starting point of the runner).

(3) us=le dilli=dek^{h}i kathmandu=samma baaTo banaa-yo
3SG=ERG Delhi=ABL Kathmandu=LOC_to road make.PST
'He built a road from Delhi to Kathmandu.' <Nepali>

(4) u dilli=baaTa kathmanDu=samma kud-yo
3SG Delhi=ABL Kathmandu=LOC_to ran-PST
'He ran from Delhi to Kathmandu.' <Nepali>

The difference between these two markers is modeled by the feature [DYN]. The entries of these markers in the proposed model are the following.

(5) a. Nepali *dekhi* = [PLACE null, PATH [SRC +, END –], DYN –]
b. Nepali *baaTa* = [PLACE null, PATH [SRC +, END], DYN +]

The main difference between these entries is the [DYN] feature that has a negative value for the static marker and positive value for the dynamic marker. Otherwise, the entry of Nepali *dekhi* is similar to the pure/abstract ablative in (2a). However, the entry of Nepali *baaTa* is even more different from (2a). The form *baaTa* is used as both the ablative and the perlative marker that has [END –] and [END +] features, respectively. Hence, the underspecified feature [END] is used that is compatible with both of these usages. Thus, we find a single lexical entry covering two different usages.

Another interesting topic of investigation in this dissertation was the usage of the same form for spatial and non-spatial usages. There are many examples in which the form that is used to mark a spatial usage is also used to mark non-spatial usages. The Nepali ablative markers *dekhi* and *baaTa* provide an example of the non-spatial usage of a spatial form.

(6) u sarpa=dek^{h}i DarauuN-cha
3SG snake=ABL fear-NPST
'He fears snakes.' <Nepali>

(7) us=le jaagir=baaTa raajinaamaa di-yo
3SG=ERG job=ABL resignation give-PST.3SG.M
'He resigned from the job.' <Nepali>

Comparing the non-spatial usages (6)–(7) with the spatial usages (3)–(4) provided some important results. It showed that the use of spatial markers in a non-spatial domain is not idiosyncratic, but systematic. The spatial domain provides metaphors to a non-spatial argument domain. The *snake* in (6) is the source of *fear*. Similarly, the *job* in (7) is the abstract location from which the subject (i.e., the resigned person) moved away. Thus

both of these source arguments are marked by the source marker, i.e., the ablative marker.

However, the more interesting point is the choice of the appropriate ablative marker in the non-spatial domain. The stimulus of the verb 'fear' is not an instigator from which some abstract entity moves to the subject. Hence it is marked by the static ablative marker *dekhi*. However, in the case of the verb 'resign', the subject moves away from the abstract location, hence it is marked by the dynamic ablative marker *baaTa*.

The non-spatial usage of *dekhi* and *baaTa* is not a unique example. The data obtained from the survey shows that spatial forms are used to mark the instrument, the addressee as well as the non-canonical second argument (NCSA) of some verbs like 'fear' and 'trust', etc. Different semantic usages of the form used as the instrument marker are shown in Table 5.1.

Table 5.1: Multiple usages of instrument markers

Language	**Other usages of instrument marker**
Punjabi	LOC-beside, COM
Saraiki	LOC-beside, COM
Nepali	ERG
Manipuri	ERG
Pashto	LOC, PERL
Balochi	COM
Sindhi	COM
Malayalam	-
Urdu/Hindi	ABL, PERL

The table shows that the form marking the instrument marker is also used as the comitative marker. This pattern is already predicted by Lakoff & Johnson (1980) who claimed that almost all languages use the companion metaphor for the instrument marking. They explain that instrument is considered as a companion in the action and hence is marked by the same form that marks accompaniment usage (i.e., the comitative marker). The data in Table 5.1, however, provided two important results. The form used

as the comitative marker is also used as the locative-beside marker in Punjabi and Saraiki. Hence, the spatial concept "locative-beside" is the basis of the companion metaphor for the instrument, at least in some languages.

Beside comitative-instrument usages, the South Asian data displays two other such patterns, i.e., ergative-instrument and perlative-instrument. This is in contradiction of Lakoff & Johnson's claim that almost all languages use the companion metaphor for the instrument. The exploration of case marker usages of some related languages provided further examples of perlative-instrument markers. Two Old Indo Aryan languages, Sanskrit and Pali, and a New Indo Aryan Language, Torwali, also have the same form for perlative-instrument usages. Hence, the study showed that that the "instrument as path" metaphor (corresponding to the common form for perlative and instrument markers) is also an important metaphor for the instrument marker.

The addressee argument of communication verbs is also marked by the spatial forms. The marking of addressee arguments of the verbs 'say' and 'ask' in different languages is shown in Table 5.2.

Table 5.2: Marking on the addressee argument

Language	**Say to**	**Ask (a question)**
Punjabi	DAT/ACC	ABL, DAT/ACC
Saraiki	DAT/ACC	ABL
Nepali	DAT/ACC	COM
Manipuri	LOC/DAT	LOC/DAT
Pashto	ALL	ABL
Balochi	DAT/ACC	ABL
Sindhi	DAT/ACC	ABL
Malayalam	COM	COM
Urdu/Hindi	ABL/INST/COM	ABL/INST/COM

The addressee of the verb 'say to' is marked by the dative marker that usually marks the recipient. As the addressee is the recipient of speech, the use of the dative marker is understandable. Similarly, there are languages that mark the addressee with the

comitative marker. The reason is that the addressee can be considered as a companion that is involved in the event. The investigation on the semantic reasons of NCSA marking, presented later on, shows that the comitative marker is correlated with the feature [involved].

Another interesting phenomenon is the use of the ablative marker to mark the addressee argument of the verb 'ask (a question)'. We have already seen that an addressee can be conceptualized as a recipient. How can a recipient be marked by the ablative (source) marker? The South Asian data introduced an interesting challenge to our concept about the addressee. The solution of the puzzle is that the addressee of 'ask' is the potential source of the answer, and hence it may be marked by the source marker as well.

The multiple case marking patterns for instrument and addressee usages teach an important lesson. There can be more than one metaphor for the same argument. The language may choose one or another metaphor from the inventory of possible metaphors.

Next to the development of a new model for spatial relations, another very important finding of the dissertation is the identification of new classes of verbs on the basis of NCSA marking in South Asian languages. The analysis of the case marking pattern in the constructions with NCSA provides the following classes of verbs.

Table 5.3: Classes of NCSA verbs in South Asian languages

Class	Subject Marking	2nd Arg. Marking	Examples	Sem. Feature
I	NOM/ERG, DAT	ABL	fear	source
II	NOM/ERG	ABL	resign	source
III	NOM/ERG	LOC-on/ DAT	attack, bless	default goal / specialized goal
IV	NOM/ERG, DAT	LOC-on/ DAT	trust, doubt	default goal / specialized goal
V	NOM/ERG	COM/DAT	meet, marry	involved
VI	NOM/ERG, DAT	COM	love, hate	involved

These classes confirm Levin's (1993) claim that the verbs that show similar syntactic behavior constitute a coherent semantic class. The verbs in these South Asian classes are semantically similar.

The application of different theories of argument realization on the above mentioned South Asian data demonstrated that no single theory can provide a comprehensive explanation of all dimensions of the problem. The study provided the following general conclusions about NCSA marking. The second argument of less transitive clauses may allow NCSA marking. These non-canonical second arguments are different from the prototypical object.

The semantic features corresponding to NCSA markers are presented in Table 5.3. The discussion about the Nepali ablative markers in (6)–(7) has already illustrated the feature [source] responsible for the ablative marker on the NCSA of Class-I and Class-II verbs. The feature [involved] corresponding to the comitative marker was also introduced in the course of the discussion about comitative marked addressee arguments shown in Table 5.2. We found alternations of the dative and locative markers for the verbs 'attack' (Class IV), 'trust' (Class III) and 'hit'. The analysis of these alternations established that the Nepali dative marker marks the achieved/intended affected goal, while the Urdu/Hindi dative marker marks the achieved affected goal. As the dative marker is used for these specialized usages, the locative marker is used, as a default option, for

unaffected or intended goal (see sections 4.5.4 and 4.5.5 for the argumentation related to this result) .

The dative/locative alternation provided another important finding. In most of the examples, lexical entailments of the argument determines the choice of case marker on it, but we find an example in which the case marking on the argument is determined by the nature of the subject and the light verb used in the construction.

(8) mai=le us=laai viswaas gar-eN
1SG=ERG 3SG=DAT/ACC trust do-PST.1SG
'I trusted him.' <Nepali>

(9) ma=laai us=maa viswaas thyo
1SG=DAT 3SG=LOC trust be.PST
'I trusted him.' <Nepali>

The stative light verb 'be' and the experiencer subject give the impression of an unaffected goal, and hence it is marked by the default locative marker.

The verb Classes I, III and V are related to the experiencer subject constructions for the psych verbs. These classes demonstrate that the psych verbs can also be classified on the basis of the marking on the non-experiencer argument.

Thus the dissertation discussed different issues about the relation of spatial expressions and case. The diachronic and synchronic evidence demonstrated the systematic extension of spatial markers into non-spatial domains. This work is the first serious contribution in terms of lexical semantics in the domain of South Asian linguistics. It provided answers to many puzzling facts, however, there is a wide scope for further inquiry.

Appendix A

Survey sentences

Most of data (example sentences) analyzed in this dissertation is obtained as the result of a survey in which informants of each language are given the following sentences. The sentences are in Urdu/Hindi, as all of the informants understand Urdu/Hindi. However, if required, I consulted the informants for some other sentences (in addition to the following sentences).

(1) jamiil bilaal=se mil-aa
Jameel.M.SG Bilal.M.SG=COM meet-PERF.M.SG
'Jameel met Bilal.'

(2) lahar caTaan=se Takraa gaii
wave.F.SG rock.F.SG=COM collide PERF.F.SG
'The wave hit the rock.'

(3) gaae rassi=se bandh-ii hai
cow.F.SG rope.F.SG=INST tie-PERF.F.SG be.PRES
'The cow is tied with the rope.'

(4) gaae khonTii=se bandh-ii hai
cow.F.SG peg.F.SG=COM tie-PERF.F.SG be.PRES
'The cow is tied to a peg.'

(5) bilaal jamiil=se miltaa_jul-taa hai
Bilal.M.SG Jameel.M.SG=COM resemble-IMPF.M.SG be.PRES
'Bilal resembles Jameel.'

(6) bilaal=ne iiNtoN=se makaan ban-vaa-yaa
Bilal=ERG brick.M.PL=ABL house.M.SG make-ICAUS-PERF.SG
'Bilal caused the house built by the bricks.'

(7) bilaal=ne iiNtoN=se makaan ban-aa-yaa
Bilal=ERG brick.M.PL=ABL house.M.SG make-CAUS-PERF.SG
'Bilal caused the house built by the bricks.'

(8) bilaal=ne jamiil=se makaan ban-vaa-yaa
Bilal=ERG Jameel=ABL house.M.SG make-ICAUS-PERF.SG
'Bilal caused Jameel to build the house.'

(9) vo uchal_uchal kar gaa rah-aa thaa
3SG jump_jump do sing PROG-M.SG be.PAST.M.SG
'He sang while jumping.'

(10) jamiil=se cal-aa nahiN jaa-taa
Jameel=ABL walk-PERF.M.SG not go-IMPF.M.SG
'Jameel is not able to walk.'

(11) ye film [bilaal=ki dekh-ii hoi] hai
this film.F.SG Bilal.M.SG=GEN see-PERF.F.SG be.PERF be.PRES
'Bilal has watched this film.' Lit: 'This film is Bilal's watched.'

(12) bilaal=ne jamiil=par bharosaa kiyaa
Bilal.M.SG=ERG Jameel.M.SG=LOC_on trust.M.SG do.PERF.M.SG
'Bilal trusted Jameel.'

(13) jamiil=ko bilaal=par bharosaa hai
Jameel.M.SG=DAT Bilal.M.SG=LOC_on trust.M.SG be.PRES
'Jameel trusts Bilal.'

(14) bilaal jamiil=se mil-ne=ke lie aa-ya
Bilal.M.SG Jameel.M.SG=COM meet-INF=GEN BEN come-PERF.M.SG
'Bilal came to meet Jameel.'

(15) peR kaT-aa
tree.M.SG (get) cut-PERF.M.SG
'The tree got cut.'

(16) us=ne bilaal=par ilzaam lag-aa-yaa
3SG=ERG Bilal=LOC_on blame.M.SG touch-CAUS-PERF.M.SG
'He blamed Bilal.'

(17) kalenDar diivaar=par lag-aa hai
calendar.M.SG wall.F.SG=LOC_on touch-PERF.M.SG be.PRES
'The calendar is hanging on the wall.'

(18) bilaal haNs-aa
Bilal.M.SG laugh-PERF.M.SG
'Bilal laughed.'

(19) bilaal pahaaR=par caRh-aa
Bilal.M.SG mountain.M.SG=LOC_on climb-PERF.M.SG
'Bilal climbed on the mountain.'

(20) bilaal kitaab paRh rah-aa hai
Bilal.M.SG book.F.SG read PROG-M.SG be.PRES
'Bilal is reading a/the book.'

(21) jamiil=ne gilaas toR-aa
Jameel.M.SG=ERG glass.M.SG break-PERF.M.SG
'Jameel broke the glass.'

(22) jamiil=ne aik seb khaa-yaa
Jameel.M.SG=ERG one apple.M.SG eat-PERF.M.SG
'Jameel ate an apple.'

(23) jamiil aur bilaal baazaar gaye
Jameel.M.SG and Bilal.M.SG market.M.SG go.PERF.M.SG
'Jameel and Bilal went to the market.'

(24) jamiil=ko saaNp=se Dar lag-taa hai
Jameel=DAT snake=ABL fear.M.SG touch-IMPF.M.SG be.PRES
'Jameel fears snakes.'

(25) bilaal=ko laahoar jaa-naa hai
Bilal.M.SG=DAT Lahore.M.SG go-INF be.PRES
'Bilal has/wants to go to Lahore.'

(26) bilaal=ko bhuuk lag-ii
Bilal.M.SG=DAT hunger.F.SG touch-PERF.F.SG
'Bilal got hungry.'

(27) jamiil=ko khaaNsii hai
Jameel.M.SG=DAT cough.F.SG be.PRES
'Jameel has a cough.'

(28) jamiil=ko xazaanaa mil-aa
Jameel.M.SG=DAT treasure.M.SG meet-PERF.M.SG
'Jameel got the treasure.'

(29) bilaal=ne jamiil=ke liye aik kitaab kahriid-ii
Bilal=ERG Jameel=GEN BEN one book.F.SG buy-PERF.F.SG
'Bilal bought a book for Jameel.'

(30) jamiil=se/davaare kitaab paRh-ii gayii
Jameel.M.SG=ABL/through book.F.SG read-PERF.F.SG go.PERF.F.SG
'The book was read by Jameel.'

(31) us=ne jamiil=se shaadii kii
3SG=ERG Jameel=COM marriage.F.SG do.PERF.F.SG
'She married Jameel.'

(32) vo bilaal=se piyaar kar-tii hai
3SG Bilal.M.SG=COM love.M.SG do-IMPF.F.SG be.PRES
'She loves Bilal.'

(33) gaaRi peR=se Takraa-ii
vehicle.F.SG tree.M.SG=ABL collide-PERF.F.SG
'The vehicle collided with the tree.'

(34) bilaal jamiil=se aqalmand hai
Bilal.M.SG Jameel.M.SG=ABL intelligent be.PRES
'Bilal is more intelligent than Jameel.'

(35) bilaal sab=se aage thaa
Bilal.MSG all=ABL front be.PST.SG
'Bilal was ahead of all.'

(36) bilaal tiren=se aa-ya
Bilal.M.SG train=INST come-PERF.M.SG
'Bilal came by train.'

(37) vo hedaraabaad=se aa-yaa
3SG Hyderabad=ABL come-PERF.M.SG
'He came from Hyderabad.'

(38) jamiil ghar=ke andar gayaa
Jameel.M.SG house=GEN inside go.PERF.M.SG
'Jameel went inside the house.'

(39) jamiil=se gilaas TuuT gayaa
Jameel.M.SG=ABL glass.M.SG break go.PERF.M.SG
'The glass got broken by Jameel./ Jameel broke the glass.'

(40) sigret pii-ne=se bilaal bimaar ho gayaa
cigarette drink-INF=INST Bilal ill be go.PERF.M.SG
'Bilal got ill by smoking cigarette.'

(41) jamiil=ne caabi=se darvaazaa khol-aa
Jameel.M.SG=ERG key.F.SG=INST door.M.SG open-PERF.M.SG
'Jameel opened the door with the key.'

(42) jamiil=ne camce=se khaanaa khaa-yaa
Jameel.M.SG=ERG spoon.M.SG=INST food.M.SG eat-PERF.M.SG
'Jameel ate the meal with the spoon.'

(43) jamiil=ne sone=se haar ban-aa-yaa
Jameel=ERG gold.M=INST necklace.M.SG make-CAUS-PERF.M.SG
'Jameel made the necklace.'

(44) vo mere saath baazaar gayaa
3SG 1.GEN COM market.M.SG go.PERF.M.SG
'He went to the market with me.'

(45) bilaal jamiil=se laR-aa
Bilal.M.SG Jameel.M.SG=COM fight-PERF.M.SG
'Bilal fought with Jameel.'

(46) bilaal=ne jamiil=se kah-aa
Bilal.M.SG=ERG Jameel.M.SG=COM say-PERF.M.SG
'Bilal said to Jameel.'

(47) bilaal=ne jamiil=se baat kii
Bilal.M.SG=ERG Jameel.M.SG=COM talk.F.SG do.PERF.F.SG
'Bilal talked with Jameel.'

(48) jamiil=ne bilaal=se puuch-aa
Bilal.M.SG=ERG Jameel.M.SG=COM ask-PERF.M.SG
'Bilal asked Jameel.'

(49) bilaal jamiil=se naaraaz hai
Bilal.M.SG Jameel.M.SG=COM angry be.PRES
'Bilal is angry with Jameel.'

(50) bilaal=ko jamiil=se Sikaayat hai
Bilal.M.SG=DAT Jameel.M.SG=COM complaint.F.SG be.PRES
'Bilal has complaints against Jameel.'

(51) bilaal=ko jamiil=par ghussa aa-yaa
Bilal=DAT Jameel=LOC_on anger.M.SG come-PERF.M.SG
'Bilal got angry with Jameel.'

(52) jameel=ko bilaal=se piyaar hai
Jameel=DAT Bilal=COM love.M.SG be.PRES
'Jameel loves Bilal.'

(53) us=ne mez=par kitaab rakh-ii
3SG=ERG table=LOC_on book.F.SG put-PERF.F.SG
'He puts the book on the table.'

(54) bilaal ghoRE=par caRh-aa
Bilal.M.SG horse.M.SG=LOC_par climb-PERF.M.SG
'Bilal climbed on the horse.'

(55) bilaal tezii=se bhaag-aa
Bilal.M.SG fastness=INST run-PERF.M.SG
'Bilal ran fast.'

(56) jamiil=ne xudaa=se maaNg-aa
Jameel.M.SG=ERG God.M.SG=ABL ask-PERF.M.SG
'Jameel asked from God.'

(57) jamiil=ne xushii=se kaam kiaa
Jameel=ERG happiness.F.SG=INST work.M.SG do.PERF.M.SG
'Jameel worked happily.'

(58) bilaal baaG=se guzr-aa
Bilal.M.SG garden=PERL pass-PERF.M.SG
'Bilal passed through the garden.'

Appendix B

An Alternate Notation for the New Spatial Model

Section 3.5 presented an underspecified feature-based model for the spatial markers of South Asian languages. This appendix introduces an alternative notation for the proposed spatial model. The model presented here is equivalent to the model presented in section 3.5. The only difference is that this alternate notation may provide a better readability to some readers. The following is the description of the proposed model in the alternate notation.

- Every spatial marker has three primary features: PLACE, PATH and DYN(amic).
- Each of the above features may have a set of features as the value.
- The feature PLACE may have following values:

  null, ON, AT, IN, BESIDE, ….
- The feature PATH may have following features as values:

  S(ou)RC(e) : The object/theme enters the location

  END : The object/theme leaves the location
- The feature DYN(amic) indicates that an action/activity is performed.
- The features SRC, END, DYN and PATH can be declared as optional by putting parenthesis around the feature.

The template for a feature-based semantic model entry of the spatial marker is:

[

PLACE null / IN / ON / AT / BESIDE / ….

(PATH [(SRC) , (VIA), (END)]),

(DYN)]

The above description is the same as the description given in section 3.5. The only difference is the notation of binary valued features.

In the notation given in section 3.5, certain features may have positive or negative values. The different values of the features [SRC] and [END] show whether the lexical entry can be used for a semantic usage or not. For example, [SRC –] means that the theme does not leave this location. It is made clear in the discussion in the section 3.5 that these features are used only to identify different kinds of the spatial marker. The features, e.g., [SRC +] or [SRC –], are not passed on to some (grand) feature structure of the whole phrase or the clause.

This alternate notation is introduced to avoid this confusion. It does not have the negative value for the features. The absence of a feature [F] is equivalent to the negative value of the feature. Hence [F +] is equivalent to [F] in this new notation. [F –], as mentioned earlier, is equivalent to [ ], and the underspecified (or optional) feature is equivalent to [(F)]. In this notation, the underspecification (or optionality) of both the binary valued features, e.g., [SRC], and the complex feature, i.e., [PATH], are shown using the same notation, i.e., the parenthesis.

The prototypical entries of the important spatial usages using this new notation are given below.

(1) Sample entries of pure/abstract ablative, perlative, allative or locative usages

a. Ablative : [PLACE X, PATH [SRC], (DYN)]

b. Allative : [PLACE X, PATH [END], (DYN)]

c. Perlative : [PLACE X, PATH [SRC, END], (DYN)]

d. Locative : [PLACE X]

Compare it with the sample entries of the same prototypical usages given in section 3.2.

(2) Sample entries of pure/abstract ablative, perlative, allative or locative usages

a. Ablative : [PLACE X, PATH [SRC +, END –], DYN]

b. Allative : [PLACE X, PATH [SRC –, END +], DYN]

c. Perlative : [PLACE X, PATH [SRC +, END +], DYN]

d. Locative : [PLACE X]

It is clear that the new notation may provide a shorter lexical entry if the entry has a negative value of one or more features. The lexical entries (1a) and (1b) are more clearly arranged and more easily understandable than the equivalent entries (2a) and (2b) respectively.

The lexical entries of some spatial markers in this alternate notation are the following. The equivalent lexical entries and the discussion/explanation are given in section 3.6.

(3) Lexical entries for some spatial markers of South Asian languages

a. Sindhi *te* = [PLACE ON]

b. Sindhi *meN* = [PLACE IN]

c. Nepali *maa* = [PLACE null]

d. Urdu/Hindi *se* = [PLACE null, PATH [SRC, (END)], (DYN)]

e. Sindhi *khaaN* = [PLACE AT, PATH [SRC , (END)], (DYN)]

f. Sindhi *taaN* = [PLACE IN, PATH [SRC , (END)], (DYN)]

g. Sindhi *maaN* = [PLACE ON, PATH [SRC, (END)], (DYN)]

h. Punjabi *toN* = [PLACE null, PATH [SRC, (END)], (DYN)]

i. Punjabi *vicoN* = [PLACE IN, PATH [SRC , (END)], (DYN)]

j. Nepali *baaTa* = [PLACE null, PATH [(SRC), END], DYN]

k. Nepali *dekhi* = [PLACE null, PATH [(SRC)]]

l. Malayalam *ninnɔ̂* = [PLACE null, PATH [SRC], (DYN)]

m. Malayalam *mutal* = [PLACE null, PATH [SRC], (DYN)]

n. Pashto *pa* = [PLACE null, (PATH [SRC, END])]

The lexical entries given above may be more readable. As most of the entries model syncretism, the length of those entries and their equivalent entries in section 3.6 are almost equal. However, if we compare the lexical entries (3k)–(3m) with the equivalent entries in section 3.6.4, we find that the entries in this new alternate notation are more compressive. Hence, one may prefer to use this notation.

Appendix C

Classes of Psych Verbs

Chapter 4 provided six classes of verbs based on non-canonical marking on the subject and the object.

Table 1: Classes of NCO verbs in South Asian languages

Class	Subject Marking	Object Marking	Examples
I	NOM/ERG, DAT	ABL	fear
II	NOM/ERG	ABL	resign
III	NOM/ERG	LOC-on/ DAT	attack, bless
IV	NOM/ERG, DAT	LOC-on/ DAT	trust, doubt
V	NOM/ERG	COM/DAT	meet, marry
VI	NOM/ERG, DAT	COM	love, hate

The analysis of this table shows that here are two classes each of ablative, locative and comitative marked objects. The reason for introducing two classes for each NCO marker is the difference of the subject marking. There are verbs that allow (non-canonical) dative subjects along with the canonical subjects. These verbs corresponding to each NCO are placed in separate classes in Table 1.

Hence, we find three classes of verbs that have the subjects marked by the dative markers and the objects marked by either of the ablative, locative or comitative marker. The inspection of the verbs in these dative subject classes shows that these classes have psych or experiencer verbs. Section 2.5.3 briefly introduced the verbs with experiencer subjects. As the experiencer is an abstract goal of the experience, it is marked by the dative marker. Hence, we find three distinct classes of the psych verbs. The important point is that this classification of psych/experiencer verbs was not predicted by any earlier theory.

There is a lot of literature available on different types of experiencer. Section 4.2.3 presented Jackendoff's (1990) summary. He proposed the following three criteria for differentiation in experiencer predicates.

- Position of experience in the sentence (subject vs. object/oblique)
- Positive vs. negative (e.g., 'please' vs. 'displease')
- State vs. event

These factors are not responsible for the three South Asian classes. Haspelmath (2001) took a closer look at the criteria for the position of the experiencer (in a sentence). He differentiated between "agent-like experiencer subject" (having canonical subject marking) and "dative experiencer subject" (having dative or related marker). However, he has a single category "patient-like experiencer" for the experiencer realized as object.

The South Asian data introduces another dimension in the classification of experiencer constructions. These constructions can be classified on the basis of marking on both the experiencer and the stimulus/target of sentience.

If all the predicates related to experiencer are considered, we find the following classes of psych predicates in Urdu/Hindi.

A: Dative Subject, no second argument

(1) bilaal=ko buxaar hai
Bilal=DAT fever.M.SG be.PRES
'Bilal has fever.'

B: Canonical Subject, Canonical Object

(2) bilaal ye baat jaan-taa hai
Bilal.NOM this matter know-IMPF.M.SG be.PRES
'Bilal knows this matter.'

C: Dative/Canonical Subject, Nominative Object

(3) bilaal=ko kahaanii yaad hai
Bilal=DAT story.F.SG memory be.PRES
'Bilal remembers the story.'

D: Dative/Canonical Subject, Ablative Object

(4) bilaal=ko saaNp=se Dar lag-aa
Bilal=DAT snake=ABL fear touch-PERF.M.SG
'Bilal feared snakes.'

E: Dative/Canonical Subject, Locative Object:

(5) bilaal=ko jamiil=par bharosaa hai
Bilal=DAT jameel=LOC trust be.PRES
'Bilal trusts Jameel.'

F: Dative/Canonical Subject, Comitative Object:

(6) bilaal=ko saaNp=se nafrat hai
Bilal=DAT snake=COM hate be.PRES
'Bilal hates snakes.'

Class A has only one argument. Hence it is different from other classes. The semantic reasons for Classes D, E and F were presented in section 4.6.

The canonical subject marking in Class B cannot be explained easily because the near-synonym of the verb *jaan* 'know' is *maloom ho* 'know + be' that allows a dative subject. Hence, it is possible that the verb *jaan* 'know' does not allow a dative subject (like Class B) for some idiosyncratic reason. However, the semantic difference between a nominative object (Class B and C) and non-canonical objects (class D, E and F) needs further investigation.

Hence, the South Asian data provides another dimension for categorizing psych or experiencer verbs. Following Levin (1993), it is predicted that syntactic differences between the objects (stimulus/target of sentience) of different classes are due to semantic factors. Most of these semantic factors have already been identified in Chapter 4, but a more detailed examination of experiencer verbs in light of the discussions and proposals in this dissertation remains to be done as future work.

Bibliography

Primary Resources:

Ahmed, Ashfaq. 2005. *Taahli thalle.* Lahore: Sang-e-meel Publications.

Baloch, Khair M. Nadvi. 1981. *balochi muallim*. Karachi: Nadvi Publications.

Dard, Khawaja Mir. 1996. *Divaan-e-Dard*, ed. Rasheed Hasan Khan. Lahore: Khaiyam Publishers. [Reprint]

Dehalvi, Mir Amman. 1804. *Bagh-o-Bahaar*. Calcutta: Fort William College.

Insha, Syed Insha-ullah Khan. 1808. *daryaa-e-lataafat*. Karachi: Anjuman-e-Taraqqi-e-Urdu. [Translation published 1988]

Jalibi, Jameel. 1971. *Divan-e-Hasan Shauqi*. poet: Hasan Shauqi, compiled by Jameel Jalibi. Karachi: Anjuman-e-Taraqqi-e-Urdu.

Naqvi, Syed Qudrat. 1988. main ne lahore jaanaa hai (aik lisaanii tehqiiq). In *lisaanii maqaalaat*, 199–229. Isalamabad: National Language Authority.

Narang, Gopi Chand. 2007. karbal kathaa kaa lisaaniyaati tajziyaa. In *urdu zaban aur Lisaniyaat*, x–xx. Lahore: Sang-e-Meel Publications.

Sindhi, Yousuf. 1999. *Nelson Mandela ji kahaani*. Hyderabad: New Fields Publications.

Yad, Mansha. 1997. *TaaNvaaN TaaNvaaN taaraa*. Islamabad: Dost Publications.

Secondary Resources:

Ackerman, Farrell & John Moore. 2001. *Proto-properties and Grammatical Encoding: A Correspondence Theory of Argument Selection*. Stanford: CSLI Publications.

Ahmed, Tafseer. 2006. Spatial, temporal and structural usages of Urdu *ko*. In Miriam Butt & Tracy Holloway King (eds.), *The Proceedings of the LFG06 Conference*. [http://csli-publications.stanford.edu/LFG/11/lfg06ahmed.pdf]

Allana, Ghulam Ali. 1999. *Sindhi Linguaphone*. Hyderabad: Sindhi Language Authority.

Anderson, John M. 1971. *The Grammar of Case: Towards a Localistic Theory*. Cambridge: Cambridge University Press.

Asher, Ronald & T.C. Kumari. 1997. *Malayalam*. London: Routledge.

Bashir, Elena. 1999. The Urdu and Hindi ergative postposition *ne*: Its changing role in the grammar. In Rajendra Singh (ed.), *The Yearbook of South Asian Languages and Linguistics*, 11–36. New Delhi: Sage Publications.

Beames, John. 1872. *A Comparative Grammar of the Modern Aryan Languages of India*. Delhi: Munshiram Manoharlal. [Republished 1966]

Beavers, John. 2006. *Argument/Oblique Alternations and the Structure of Lexical Meaning*. Stanford, CA: Stanford University dissertation.

Bhatt, Rajesh. 2003. Topics in the Syntax of the Modern Indo-Aryan Languages – Causativization. Handout. [http://web.mit.edu/rbhatt/www/24.956/caus.pdf]

Blake, Barry. 2001. *Case*. Cambridge: Cambridge University Press.

Blume, Kerstin. 1998. A contrastive analysis of interaction verbs with dative complements. *Linguistics* 36. 253–280.

Bulchand, Dulamal. 1901. *A Manual of Sindhi*. Hyderabad: The Kaiseria Press. [Reprinted 2003]

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. Stanford: CSLI Publications.

Butt, Miriam. 1998. Constraining argument merger through aspect. In Erhard Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex Predicates in Nonderivational Syntax*, 73–113. New York: Academic Press.

Butt, Miriam. 2006a. *Theories of Case*. Cambridge: Cambridge University Press.

Butt, Miriam. 2006b. The dative-ergative connection. In Patricia Cabredo-Hofherr & Oliver Bonami (eds.), *Empirical Issues in Formal Syntax and Semantics 10*, 69–92. The Hague: Thesus.

Butt, Miriam & Tracy Holloway King. 1991. Semantic case in Urdu. *Chicago Linguistic Society (CLS)* 27. 31–45.

Butt, Miriam & Tracy Holloway King. 2005. The status of case. In Veneeta Dayal & Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, 153–198. Dordrecht: Kluwer Academic Publishers.

Butt, Miriam & Tikaram Poudel. 2007. *Distribution of the ergative in Nepali*. Unpubl. ms., Universität Konstanz / Universität Leipzig. [http://ling.uni-konstanz.de/pages/home/tafseer/leipzig07-hnd.pdf]

Catford, John C. 1975. Ergativity in Caucasian languages. *NELS* 6. 37–48.

Chelliah, Shobhana. 1990. Experiencer subjects in Manipuri. In Manindra K. Verma & Tara Mohanan (eds.), *Experiencer Subjects in South Asian Languages*, 195–211. Stanford: CSLI Publications.

Creissels, Denis. 2008. Spatial cases. In Andrej Malchukov & Andrew Spencer (eds.), *The Handbook of Case*, 609–625. Oxford: Oxford University Press.

Davison, Alice. 1998. Structural case, lexical case and complex predicates. *South Asian Language Review* 8.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1979. Ergativity. *Language* 55. 59–138.

Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Emeneau, Murray. 1956. India as a linguistic area. *Language* 32. 3–16.

Fillmore, Charles. 1968. The case for case. In Emmon Bach & R.T. Harms (eds.), *Universals in Linguistic Theory*, 1–90. New York: Holt, Rinehart and Winston.

Frankfurter, Oscar. 1883. *Handbook of Pali: Being an Elementary Grammar, a Chrestomathy, and a Glossary*. London, Edinburgh: Williams and Norgate.

Gehrke, Berit. 2007. On directional readings of locative prepositions. In Sylvia Blaho, Luis Vicente & Erik Schoorlemmer (eds.), *Proceedings of ConSOLE XIV*, 99–120.

Givon, Talmy. 1985. Ergativity morphology and transitivity gradients in Newari. In Frans Plank (ed.), *Relational Typology*, 89–107. Berlin: Mouton de Gruyter.

Grimes, Joseph E. & Barbara F. Grimes (eds.). 2000. *Ethnologue.* Volume 1: *Languages of the World;* Volume 2: *Maps and Indexes*. 14th edn. Dallas: SIL International.

Gordon, Raymond G. & Barbara F. Grimes (eds.). 2005. *Ethnologue: Languages of the World*. 15th edn. Dallas: SIL International.

Grimm, Scott. 2005. *The Lattice of Case and Agency*. Amsterdam: Universiteit van Amsterdam MSc thesis.

Grimm, Scott. 2007. The bounds of subjecthood: Evidence from instruments. *Proceedings of 33rd Meeting of Berkeley Linguistic Society (BLS)*. Berkley.

Gruber, Jeffrey. 1965. *Studies in Lexical Relations*. Cambridge, MA: MIT dissertation.

Gruber, Jeffrey. 1976. *Lexical Structure in Syntax and Semantics*. Amsterdam: North Holland.

Haspelmath, Martin. 2001. Non-Canonical marking in European languages. In Alexandra Aikenvald et al. (eds.), *Non-canonical Marking of Subjects and Objects*, 53–84. Amsterdam: John Benjamins Publications.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hewson, John & Vit Bubenik. 2006. *From Case to Adposition: The Development of. Configurational Syntax in Indo-European Languages*. Amsterdam: John Benjamins Publications.

Hoernle, Augustus. 1880 A comparative Grammar of the Gaudian Language, London: Trubner and Company.

Hopper, Paul & Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1987. The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18. 369–411.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.

Kachru, Yamuna. 2006. *Hindi*. Amsterdam: John Benjamins Publications.

Kellogg, Samuel H. 1893. *Grammar of the Hindi Language*. Delhi: Munshiram Manoharlal Publishers Pvt. Ltd. [2nd edn., reprinted 1990]

Kiparsky, Paul. 1998. Partitive case and aspect. In Miriam Butt & Wilhelm Geuder (eds.), *The Projection of Arguments*, 265–307. Stanford: CSLI Publications.

Korn, Agnes. 2009. The ergative system in Balochi from a typological perspective. *Iranian Journal of Applied Languages Studies* 1. 43–78.

Kracht, Marcus. 2002. On the Semantics of Locatives. *Linguistics and Philosophy* 25. 157–232.

Kratzer, Angelika. 2002. *Telicity and the meaning of objective case*. Unpubl. ms., University of Massachusetts.

Lakoff, George & Mark Johnson. 1980. *Metaphors We Live By*. Chicago: Chicago University Press.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: The University of Chicago Press.

Levin, Beth. 1999. Objecthood: An event structure perspective. *Chicago Linguistic Society (CLS)* 35 (1). 223–247.

Mahajan, Anoop. 1989. Agreement and agreement phrases. *MIT Working Papers in Linguistics* 10: 217– 252.

Malchukov, Andrej. 2005. Case pattern splits, verb types and construction competition. In Mengistu Amberber & Helen de Hoop (eds.), *Competition and Variation in Natural languages: The Case for Case*, 73–117. Amsterdam: Elsevier.

Malchukov, Andrej. 2006. Transitivity parameters and transitivity alternations: Constraining co-variation. In Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.), *Studies on Case, Valency and Transitivity*, 329–359. Amsterdam: John Benjamins Publications.

Masica, Colin P. 1991. *Indo Aryan Language*. Cambridge: Cambridge University Press.

Meira, Sèrgio. 2004. Mental state postpositions in Tiriyó and other Cariban languages. *Linguistic Typology* 8. 213–241.

Michaelis, Laura A. 1993. On deviant case-marking in Latin. In Robert D. van Valin (ed.), *Advances in Role and Reference Grammar*, 311–371. Amsterdam: John Benjamins Publications.

Mohanan, Tara. 1994. *Argument Structure in Hindi*. Stanford: CSLI Publications.

Moravcsik, Edith. 2003. Inflectional morphology in the Hungarian noun phrase. In Frans Plank (ed.), *Noun Phrase Structure in the Languages of Europe*, 113–252. Berlin: Mouton de Gruyter.

Nichols, Johanna. 1975. Verbal semantics and sentence construction. *Berkeley Linguistic Society (BLS)* 1. 343–353.

Ostler, Nicholas. 1979. *Case-linking: A theory of case and verb diathesis, applied to Classical Sanskrit*. Cambridge, MA: MIT dissertation.

Poudel, Tikaram. 2008a. Ergativity and Stage/Individual Level Predications in Manipuri. *Workshop on Differential Case Marking*. Ms. Stuttgart. [http://ling.uni-konstanz.de/pages/home/tafseer/manipuri_08.pdf]

Poudel, Tikaram. 2008b. Nepali ergativity: A historical perspective. Paper presented at the *Workshop on Case and Alignment in Indo-European*, December 10–11, 2008, University of Bergen.
[http://ling.uni-konstanz.de/pages/home/tafseer/bergen_erg.pdf]

Primus, Beatrice. 1999. *Cases and Thematic Roles – Ergative, Accusative and Active*. Tübingen: Niemeyer.

Primus, Beatrice. 2002. Proto-roles and case selection in Optimality Theory. *Arbeiten des SFB 282 "Theorie des Lexikons"*, Nr. 122. [http://www.uni-koeln.de/phil-fak/idsl/dozenten/primus/publikationen/Primus_Proto-Roles_2000.pdf]

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase. Syntax*. Cambridge: Cambridge University Press.

Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, 97–134. Stanford: CSLI Publications.

Saksena, Anuradaha. 1982. *Topics in the Analysis of Causatives with an Account of Hindi Paradigms*. Berkeley: University of California Press.

Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2006. *On Comitatives and Related Categories: A Typological Study with Special Focus on the Languages of Europe*. Berlin: Mouton de Gruyter.

Singh, Jag Deva. 1970. *A Descriptive Grammar of Bangru*. Kurukshetra: Kurukshetra University.

Subbarao, Karumuri Venkata, Upen Rabha Hakacham & Thokchom Sarju Devi. 2007. Case-marked PRO: Evidence from Rabha, Manipuri, Hindi-Urdu and Telugu. In Roland Bielmeier & Felix Haller (eds.), *Linguistics of the Himalayas and Beyond*, 291–322. Berlin: Mouton de Gruyter.

Tegey, Habibullah & Barbara Robson. 1996. *A Reference Grammar of Pashto*. Washington, DC: Center for Applied Linguistics.

Testelec, Yakov G. 1998. On two parameters of transitivity. In Leonid Kulikov & Heinz Vater (eds.), *Typolgy of Verbal Categories*, 29–45. Tübingen: Niemeyer.

Thomas, Emma. 2003. Manner-specificity as a factor in the acceptance of in and on in directional contexts. *Essex Graduate Student Papers in Language and Linguistic* V. 117–146.

Trask, Robert. 1997. *The History of Basque*. London: Routledge.

Tsunoda, Tasaku. 1981. Split case-marking in verb-types and tense/aspect/mood. *Linguistics* 19. 389–438.

Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21. 385–396.

Trumpp, Ernest. 1872. Grammar of the Sindhi Language. London: Trubner and Co. and Leipzig: FA. Brockhaus.

Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 56.143–160.

Verma, Manindra K. & Karuvannur P. Mohanan (eds.) 1990. *Experiencer Subjects in South Asian Languages*. Stanford: CSLI Publications.

Whitney, William. 1889. Sanskrit Grammar: Including Both the Classical Language

and the Older Dialects, of Veda and Brahmana. Cambridge, MA: Harvard University Press.