

**DE GRUYTER  
MOUTON**

*Rajendra Singh (Ed.)*

**ANNUAL REVIEW  
OF SOUTH ASIAN  
LANGUAGES AND  
LINGUISTICS  
2010**

TRENDS IN LINGUISTICS

**DE  
G  
I  
M**

Annual Review of South Asian Languages and Linguistics  
2010

# Trends in Linguistics

## Studies and Monographs 238

*Editor*

Volker Gast

*Advisory Editors*

Walter Bisang

Hans Henrich Hock

(main editor for this volume)

Matthias Schlesewsky

Niina Zhang

*Founding Editor*

Werner Winter

De Gruyter Mouton

# Annual Review of South Asian Languages and Linguistics

2010

*edited by*

Rajendra Singh

De Gruyter Mouton

ISBN 978-3-11-025483-9  
e-ISBN 978-3-11-025497-6  
ISSN 1861-4302

*Bibliographic information published by the Deutsche Nationalbibliothek*

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

© 2010 Walter de Gruyter GmbH & Co. KG, Berlin/New York  
Printing: Hubert & Co. GmbH & Co. KG, Göttingen  
∞ Printed on acid-free paper  
Printed in Germany.  
[www.degruyter.com](http://www.degruyter.com)

# Contents

Editorial Preface.....	vii
------------------------	-----

## General Contributions

On the Origin of Affixal Polysemy/Homonymy in Bengali.....	3
<i>Shishir Bhattacharja</i>	

Two Types of NPIs in Hindi with One Licensing Condition.....	19
<i>Rajesh Kumar</i>	

Wh-Phenomena in Sinhala.....	49
<i>Chandrika Maliyagoda</i>	

Diverging Pathways: The Current Status of Grammaticalization of Hindi Verbs.....	75
<i>Shakthi Poornima and Robert Painter</i>	

## Special Contribution: *Hindi Pragmatics*

On Hindi Conditionals .....	107
<i>Ghanshyam Sharma</i>	

## Regional Reports

Issues in the Description of Bangla .....	137
<i>Probal Dasgupta</i>	

Europe.....	143
<i>John Peterson</i>	

## Reviews

*R. K. Agnihotri and H. K. Dewan (Eds.) – Knowledge, Language and Learning* .....167  
by *Anjani K. Sinha*

*Shreesh Chaudhary – Foreigners & Foreign Languages in India: A Sociolinguistic History*.....175  
by *Rajend Mesthrie*

## Dialogue

The Standard, (Non-) Rhoticity and Rhythm in Indian English:  
A Response to Lange .....183  
by *Pingali Sailaja*

## Appendices

Announcements: The Gyandeeep Prize/Housekeeping.....189

Notes on Contributors.....191

## Editorial Preface

*Annual Review of South Asian Languages and Linguistics (ARSALL)* is devoted to bringing out what is currently being explored in South Asian linguistics and in the study of South Asian languages in general. South Asia is home to a wide variety of languages, structurally and typologically quite diverse, and has often served as a catalyst and testing ground for theories of various kinds.

Although linguists working on South Asia have made significant contributions to our understanding of language, society, and language in society, and their numbers have grown considerably in the recent past, until recently there was no internationally recognized forum for the exchange of ideas amongst them or for the articulation of new ideas and approaches grounded in the study of South Asian languages. *The Yearbook of South Asian Languages and Linguistics*, of which this annual is a direct descendant, played that role during the last decade, but in 2007 we decided to go a bit further and incorporate a slightly modified form of such a forum into *Trends in Linguistics*. This is the fourth issue of *ARSALL* as part of the series *Trends in Linguistics: Studies and Monographs*.

Each volume of this annual has five major sections:

- i. **General Contributions**, consisting of selected open submissions that focus on important themes and provide various viewpoints.
- ii. **Special Contributions**, consisting of generally invited single or multiple contributions on important issues, ranging from the narrowly grammatical to the wide-scope socio-linguistic/socio-political. When many contributions are included, the section will in effect constitute a mini-symposium, albeit in the written form, on the issue chosen for a given year. It will serve the function of familiarizing the reader with current thinking on issues seen as salient in the study of South Asian languages.
- iii. **Reports**, consisting of reports from around the world on research on South Asian languages



iv. **Reviews and Abstracts**, consisting of reviews of important books and monographs and abstracts of doctoral theses.

v. **Dialogue**, consisting of a forum for the discussion of earlier work, preferably previously published in this annual, comments, reports on research activities, and conference announcements.

Other than excellence and non-isolationism, *ARSALL* has no theoretical agenda and no thematic priorities.

The first, general section of this, the fourth, issue of *ARSALL* contains four contributions: Bhattacharja's *On the origin of affixal polysemy/homonymy in Bengali*, Kumar's *Two types of NPI's in Hindi*, Poornima and Painter's *Diverging pathways: the current status of grammaticalization of Hindi verbs*, and Maliyagoda's *Wh-phenomena in Sinhala*.

The Special Contributions section is dedicated to Pragmatics, generally not studied with any rigour in the South Asian context. In his contribution to this section, Sharma carefully examines the pragmatics of conditionals in Hindi.

As Europe, like North America, continues to be a major center for the study of South Asian languages, we thought it was appropriate to have our Regional Reports section publish a report on European research on South Asian languages in this issue of *ARSLL*. The first issue, readers may recall, also contained a report on European research on South Asian languages. Following the tradition inaugurated in 2004, it also contains a report on what Dasgupta sees as the still unresolved problems of Bengla linguistics.

The Review section of this issue contains reviews of Agnihotri and Dewan's recent important book on knowledge, language, and learning and of Chaudhary's non-technical book on foreigners and foreign languages in India.

I am particularly happy to note that the Dialogue section of this issue contains a response to a review published in *ARSALL-2009*.

I am grateful to Prof. Hans Henrich Hock, Mr. Wolfgang Konwitschny, and Mr. Mooli Ikome for help and support in the preparation of this issue.

*Rajendra Singh*

## **General Contributions**



# On the Origin of Affixal Polysemy/Homonymy in Bengali\*

*Shishir Bhattacharja*

*In Bengali, words endowed with a variety of semantic readings and categorical affiliation end in an isophonic suffix, a phenomenon known as affixal polysemy (or homonymy) in the literature. In this article, we account for this phenomenon in the light of W(hole) W(ord) M(orphology), elaborated in Ford, Singh, and Martohardjono 1997. WWM claims that any morphological relationship between two words of a language can be described by a W(ord) F(ormation) S(trategy) licensed by a set of semantically related pairs of words showing the same formal difference and categorical affiliation. We claim that each new output of a particular WFS manifests a unique semantic relatedness with the input. Consequently, each new pair (constituted of the input and the output) may subsequently serve as a different model to form some other new words. If some of the outputs and inputs of such WFSs undergo categorical changes, new WFSs based on different categorical affiliations also come into being. These are some of the factors that motivate multiplication of WFSs that are based on the same formal difference. As other models of morphology describe certain formal differences (manifested in different sets of word-pairs) as affixes, they see the multiplication of WFSs based on those formal differences as affixal polysemy or homonymy.*

## 1. Preliminaries

As in many other human languages, different words endowed with a variety of semantic readings and categorical affiliations end in an isophonic suffix in Bengali.<sup>1</sup> For example, in (1-15), as many as 15 different nouns and adjectives endowed with different semantic readings end in the suffix {i}.<sup>2</sup> The question that naturally arises is why in human languages, isophonic affixes are found in a variety of words?<sup>3</sup> There may reasonably be a debate (see Plag 1997:236) about whether the same polysemous suffix appears in words of this kind, or if each one of them contains a different homonymous suffix.<sup>4</sup> In this article, we shall attempt to

answer this question and try to contribute to the debate in the light of W(hole) W(ord) M(orphy) (elaborated in Ford, Singh, and Martohardjono 1997).

#### NOUN-ADJECTIVE

(1) 'Who/which has X (X = base)' [- dynamic, - material]<sup>5</sup>  
*shukh* 'happiness' → *shukhi* 'happy'

(2) 'Of X' [- dynamic, - material]  
*japan* 'Japan' → *japani* 'of Japan'

(3) 'Made of X' [- dynamic, + material]  
*reshom* 'silk' → *resh(o)mi* 'made of silk/silken'

(4) 'Which has the colour of X' [- dynamic, - material]  
*golap* 'rose' → *golapi* 'rose-coloured/rose/rosy'

(5) 'administratively concerning X' [- dynamic, - material]  
*nirbacon* 'election' → *nirbaconi* 'electoral'

(6) 'Which does X' [- dynamic, - material]  
*binash* 'destruction' → *binashi* 'destructive'

(7) 'Which follows or supports X' [- dynamic, - material]  
*marksbad* 'Marxism' → *marksbadi* 'Marxist'

(8) 'Which measures or weighs X' [- dynamic, - material]  
*paNc fuT* 'five feet' → *paNc fuTi* 'five feet long'<sup>6</sup>

#### ADJECTIVE-NOUN

(9) 'X-ness' [- dynamic, - material]  
*calak* 'clever' → *calaki* 'cleverness'

## NOUN-NOUN

(10) ‘Who does X’ [+ dynamic, + material]

*khun* ‘murder’ → *khuni* ‘murderer’

(11) ‘Who uses X as an instrument’ [+ dynamic, + material]

*kOrat* ‘saw’ → *kOrati* ‘sawer’

(12) ‘Activities of X’ [– dynamic, – material]

*mashTar* ‘teacher’ → *mashTari* ‘teaching’

(13) ‘Money/present given as a token of X’ [– dynamic, + material]

*shomman* ‘respect’ → *shommani* ‘honorarium’

(14) ‘Done with X’ [– dynamic, – material]

*phaNsh* ‘a slip knot’ → *phaNshi* ‘death or practice of killing by hanging’

(15) ‘Made of X’ [– dynamic, + material]

*baNsh* ‘bamboo’ → *baNshi* ‘flute’

It does not seem to be possible to treat {i} in the same way as Plag (1997) treats the suffix {ize} in English. According to Plag, most of the words that end in {ize}, including about 300 types of neologisms he mentions, have the same underlying *lexical semantic structure*. He (1997:235) concludes from this that concatenating {ize} to different base words is a “semantically transparent polysemous process.” However, on the basis of a synchronic and diachronic analysis of a variety of Spanish words ending in the suffix {azo}, Rainer (2003:204) demonstrates that the attempt to derive all such words as contextual variants of one single abstract meaning suffers from serious defects. Similarly, the semantic and categorical diversity of the derived words in (1-15) cannot be handled with one single process, or with one single underspecified meaning like ‘related to X’ or ‘of X’.

Following Lieber (2004), although we have provided the examples (1-15) with semantic features of the affix {i}, we will not present any semantic or pragmatic analysis of this or any other affix in the present description. We shall not do this because there is no model that we know of which would allow us to do such an analysis satisfactorily. Lieber (2004) demonstrates that different existing models of semantic analysis (e.g. Jackendoff 1990, Pustejovsky 1995, Wierzbicka 1996 and Szymanec

1988) among others) are not adequate for handling affixes. However, there are examples among (1-15) which pose problems for Lieber's (2004) own model. This is because her model does not allow a particular affix to create both concrete [+ material] and abstract [- material] nouns. But, as we see in the examples above, {i} can appear in both abstract and concrete nouns ((9) and (10)), and adjectives ((6) and (3)) respectively. We find that, contrary to Lieber's (2004) predictions, both stative (13) and activity nouns (12) can end in the same suffix {i}.

## 2. Whole word morphology

We will now briefly describe the model WWM in what follows, before we move on to demonstrate how words with a variety of semantic reading and categorical affiliation happen to end in an isophonic suffix in Bengali. According to Singh (2006:578)

“All that needs to be said about word structure in any language (of any type whatsoever) can and must be said by instantiations of the schema in (S1). These instantiations are referred to as Word Formation Strategies (WFSs) because, as generalizations drawn from known particular facts, they can be activated in the production and understanding of new words. WFSs must be formulated as generally as possible, but – and this is crucial – only as generally as the facts of the matter permit.

S1.  $/X/_{\alpha} \leftrightarrow /X'/_{\beta}$  where

1.  $/X/_{\alpha}$  and  $/X'/_{\beta}$  are words and X and X' are abbreviations of the forms of classes of words belonging to categories  $\alpha$  and  $\beta$  (with which specific words belonging to the right category can be unified or on to which they can be mapped).
2. ' represents (all the) form-related differences between  $/X/$  and  $/X'/_{\beta}$  that fall outside of automatic phonology.
3.  $\alpha$  and  $\beta$  are categories that may be represented as feature bundles.
4. The  $\leftrightarrow$  represents a bidirectional implication (if X then  $/X'/_{\beta}$ , and if  $/X'/_{\beta}$ , then  $/X/_{\alpha}$ ).
5. The interpretation of  $/X/_{\alpha}$  is a semantic function of  $/X'/_{\beta}$  and vice versa.
6. ' can be null iff  $\alpha \neq \beta$ .”

As Singh (2006:578) expresses it, WWM sees morphology, “not as a combinatorics of morphs or morphemes but as a system of generalized and

abstract bidirectional correspondence among patterns instantiated by sets of whole words that exploit the same contrast.” Singh (2006:578) goes on to state that some advocates of WWM (e.g. Ford, Singh, and Martohardjono 1997) take the ‘dissociative’ view of morphology and “postulate the existence of rules of interpretation associated with WFSs”, whereas others (e.g. Neuvel 2003) subscribe to the ‘associative’ view *à la* Corbin (1987) and require the said contrast to be “both formal and semantic.”

In the present article, we will adopt the dissociative view of morphology in the sense that each WFS has to be licensed by a set of semantically related pairs of words showing the same i) formal contrast and ii) categorical affiliation. For example, (16) instantiates a WFS of English because it is licensed by a set of semantically related word-pairs which manifest the same i) formal contrast: X/Xli and iii) categorical affiliation: Noun/Adjective, but not necessarily the ‘same’ semantic contrast. Each WFS is provided with an ad hoc rule of interpretation (e.g. ‘X/-like’) in the present article. The bidirectional arrow implies that a WFS can be activated both ways by using either of the pair-mates as the input.

- (16) /X/<sub>N</sub> ↔ /Xli/<sub>Adj</sub> ‘X/-like’  
*friend* ↔ *friendly*; *man* ↔ *manly*

According to WWM, words have no internal (non-phonological) hierarchical structure. However, if a particular word is mapped onto some relevant WFS it can be analyzed into two subcomponents, a *variable* (*friend/man*) and a *constant* (/li/). Subcomponents can be represented by any phonic element: single phoneme, meaningless sound cluster, words, and discontinuous or continuous segmental as well as supra-segmental means like stress and tone (variables, however, cannot be exclusively supra-segmental). For example, if the Hebrew word /*hagdala*/ ‘enlargement’ is mapped onto (17), the variable will be represented by the discontinuous sequence of consonants: /h/-/gd/-/l/, and the constant by the discontinuous sequence of vowels: /a/-/a/-/a/. Equally, if the Chinese word /*ts<sup>h</sup>ɔ<sup>35</sup>*/ ‘a plough’ is mapped onto (18), the constant will be represented by the rising tone while the variable will be represented by the sequence of segments.

- (17) /CaCCaCa/<sub>N</sub> ↔ /CiCCiC/<sub>V</sub> ‘Action of doing /CiCCiC/’  
*/hagdala/* ‘enlargement’ ↔ */higdil/* ‘enlarge’  
*/haxtaba/* ‘dictation’ ↔ */hixtib/* ‘dictate’ (Data:Booij 2005:38)



- (18) /CV<sup>11</sup>/<sub>V</sub> ↔ /CV<sup>35</sup>/<sub>N</sub> ‘To do the action by using /CV<sup>35</sup>/’  
 /mɔ<sup>11</sup>/ ‘to grind’ ↔ /mɔ<sup>35</sup>/ ‘a grind’  
 /ts<sup>h</sup>ɔ<sup>11</sup>/ ‘to plough’ ↔ /ts<sup>h</sup>ɔ<sup>35</sup>/ ‘a plough’ (Data:Yu 2007:191)

### 3. Multiplication of WFSs based on the same formal difference

We will now move on to see how different words happen to end in an isophonetic suffix. It may be claimed that one of the reasons behind this phenomenon is that, some model WFSs are split into other different WFSs in the course of time. To begin with, one may consider (19) which had been a part of the morphological module of Bengali for quite some time but had remained quite unproductive until very recently.

- (19) /X/<sub>N</sub> ↔ /Xkheko/<sub>Adj/N</sub> ‘/X/-devourer’ [– dynamic, +/ – material]  
 /manush/ ‘human being’ ↔ /manushkheko/ ‘man-eater’  
 /kOlje/ ‘liver’ ↔ /kOljekheko/ ‘who eats (human) liver’

Between November 2006 and January 2009, (19) became comparatively more productive, possibly because of the indirect influence of socio-political incidents in Bangladesh (when the fourth care-taker government in power engaged itself in some sort of ‘corruption-cleansing’ activities). At the beginning of this period, a word with the particular semantic reading of ‘forest-devourer’ /bonkheko/ was formed, we suppose, by mapping /bon/ ‘forest’ onto (19). It was followed by the formation of another new word /nodikheko/ ‘river-devourer’. We assume that these two pairs: /bon/~ /bonkheko/ and /nodi/~ /nodikheko/ license a new WFS (20) with a rule of interpretation which is different from the rule of interpretation found in (19). Several other neologisms like /bhumikheko/ ‘land-devourer’, /shomudrotOtkheko/ ‘sea-beach devourer’ were formed with (20) in the recent times.

- (20) /X/<sub>N</sub> ↔ /Xkheko/<sub>Adj/N</sub> ‘A disgusting person who illegally takes some public property /X/ in his possession and makes every possible misuse of it’ [– dynamic, +/ – material]  
 /bon/ ‘forest’ ↔ /bonkheko/ ‘who misuses a forest by selling off trees, etc.’  
 /nodi/ ‘river’ ↔ /nodikheko/ ‘who takes a river in his possession and misuses it’

In the above, one needs to be aware of the fact that some speakers may merge (19) and (20) by adjusting their rule of interpretation to some extent. However, it is equally possible that other speakers may not do so because unlike /*manushkheko*/ ‘man-eater’, /*bonkheko*/ ‘who illegally misuses a forest’ is not an animal, but a human being, more specifically, a high government official in charge of the forest (and also because, for a tiger the act of devouring human beings cannot be considered to be illegal or disgusting). One may also argue that /*bonkheko*/ was in fact formed after /*manushkheko*/ and not necessarily by activating (19).<sup>7</sup> However, we assume that the Bengali journalist who coined /*bonkheko*/ did it by activating (19), but he could also have coined it after /*manushkheko*/ ‘man-eater’ if his lexicon lacked other words ending in [kheko], and/or if (19) was not part of his morphological module. Words can be formed through various other means, but if a particular WFS is a part of the morphological module of a speaker-hearer, it is more likely that he will activate the WFS for creating a new word. We do not claim that WFS is the only tool for creating new words. We simply claim that it is a tool among many others.

It can be claimed that each new output of a particular WFS manifests a more or less unique semantic relatedness with the input, or, to put it in Aronovian (1976) terms, each output of a WFS is a unique function of its input. For example, unlike /*bonkheko*/, a /*nodikheko*/ is not a government official, but rather a local political leader or an influential hooligan. It is quite likely that in the near future (20) will also split into two different WFSs. Each new pair has, to some extent, the potential to license a different WFS so that whenever a set of pairs with a similar semantic relatedness becomes available, the model WFS is split to create a new WFS endowed with a different rule of interpretation.<sup>8</sup>

If some of the outputs and inputs of a particular WFS undergo categorical changes, a new WFS based on different categorical affiliations can also come into being. As a result, the number of WFSs, each based on the same formal difference, but endowed with different rule of interpretation and/or categorical affiliation continues to multiply. As other models describe particular formal differences like [i] and [kheko] manifested in different sets of word-pairs as *affixes*, they see the multiplication of WFSs based on these formal differences as affixal polysemy (or homonymy). For a model that considers affixes as lexical entries (such as Lieber 1992), (1-15) or (19-20) can be seen as multiplication of the number of isophonetic affixes, each one endowed with a different shade of meaning and (probably) different argument structures.

This paper shall now proceed to focus on another case which clearly shows how splitting becomes imperative due to the diversity of semantic relatedness manifested in different word-pairs licensing a particular WFS. One can assume that (21) has been part of Bengali morphology since the time that the Bengali speaking society in conjunction with other Indic language speaking societies in South Asia, used to consider women to be the property of respective husbands while an army was the property of the commander. The reason behind such an assumption is that unless one postulates (21) it is difficult to explain how the sequence [poti] could appear in words like /shenapoti/ ‘army commander’ and /koTipoti/ ‘millionaire’.

- (21) /X/<sub>N</sub> ↔ /Xpoti/<sub>N</sub> ‘Who owns /X/’ [+ material, – dynamic]  
 /shita/ ‘Sita, heroine of the epic *Ramayana*’ ↔ /shitapoti/ ‘husband of Sita/god Rama’  
 /bhogni/ ‘sister’ ↔ /bhognipoti/ ‘husband of one’s sister/brother in law’  
 /koTi/ ‘ten million’ ↔ /koTipoti/ ‘owner of ten million rupees/millionaire’  
 /shena/ ‘soldier, army’ ↔ /shenapoti/ ‘army commander’  
 /shilpo/ ‘industry’ ↔ /shilpopoti/ ‘industrialist/owner of an industry’  
 /bicar/ ‘justice’ ↔ /bicarpoti/ ‘judge’  
 /nOgor/ ‘city’ ↔ /nOgorpoti/ ‘owner of the city/a king’

It is possible that words like /shitapoti/ ‘husband of Shita/god Rama’, /bhognipoti/ ‘brother in law’, etc. are the oldest among these words and /nOgorpoti/ ‘king’, /shenapoti/ ‘army commander’, etc. are later creations because etymologically, /poti/ is the person who gives birth to /Opotto/ ‘children/siblings’.<sup>9</sup> However, in the then Bengali speaking society /poti/ ‘husband’ was also considered the master of his wife/wives, and some speaker-hearer(s), based on this second meaning of /poti/, could have coined words like /nOgorpoti/ ‘King’, /kulpoti/ ‘a patriarch’, etc. As a consequence, (21) was established as a WFS. Be that as it may, in the course of time, the meaning/use of some of the outputs of (21), such as /bhognipoti/, /shenapoti/, /nOgorpoti/, etc. underwent changes, and consequently two other WFSs (22-23) came into being, each endowed with a different rule of interpretation. We may presume that this happened because of changes in the Bengali speaking society: a /bhognipoti/ could never own somebody’s sister anymore than a /nogorpoti/ or /bicarpoti/ could own the city, or own justice.

- (22) /X/<sub>N</sub> ↔ /Xpoti/<sub>N</sub> ‘Husband of X’ [+ material, – dynamic]  
 /bhogni/ ‘sister’ ↔ /bhognipoti/ ‘husband of one’s sister/brother in law’  
 /shoci/ ‘Shoci/wife of the god Indra’ ↔ /shocipoti/ ‘husband of  
 Shoci/god Indra’
- (23) /X/<sub>N</sub> ↔ /Xpoti/<sub>N</sub> ‘Who is in charge of /X/’ [+ material, – dynamic]  
 /shena/ ‘soldier/army’ ↔ /shenapoti/ ‘army commander’  
 /bicar/ ‘justice’ ↔ /bicarpoti/ ‘judge’  
 /nOgor/ ‘city’ ↔ /nOgorpoti/ ‘owner of the city/a mayor’

Although (21) still remains a part of Bengali morphology, it has undergone two changes: i) presently it is licensed with a subset of word-pairs that used to license it in the earlier period of time, and ii) words like /bicarpoti/ ‘judge’ or /bhognipoti/ ‘husband of one’s sister’ cannot be formed or analyzed with it anymore. We note here that a WFS can survive until there is in the lexicon, the required number of adequate pairs.

At the very beginning of the process of splitting, split WFSs may be metaphorically and metonymically linked with each other (cf. Rainer 2003). However, with the passage of time, such links are usually lost. We can note that the semantic features (*à la* Lieber 2004) that we have provided in (21-23) are the same, as it is possible to show that these WFSs manifest a common, underspecified rule of interpretation: ‘related to /X/’. Nevertheless, as we can realize, the relations of a millionaire with his money, of a husband with his wife, and of a judge with justice are semantically (and also pragmatically) too diverse to handle the outputs of (21-23) with one single WFS (24). For example, (24) will not be of much help if a speaker-hearer wants to analyze a previously unencountered word like *shOrpopoti* which means either ‘King of snakes’ or ‘who owns a lot of snakes’, both manifesting the underspecified meaning ‘related to snakes’. To analyze the word morphologically, the speaker-hearer must have (25) and/or (26) in his morphological module.

- (24) \*/X/<sub>N</sub> ↔ /Xpoti/<sub>N</sub> ‘related to /X/’ [+ material, – dynamic]  
 /koTi/ ‘ten million’ ↔ /koTipoti/ ‘owner of ten million rupees/  
 millionaire’  
 /bhogni/ ‘sister’ ↔ /bhognipoti/ ‘husband of one’s sister/brother in law’  
 /bicar/ ‘justice’ ↔ /bicarpoti/ ‘judge’
- (25) /Xo/<sub>N</sub> ↔ /Xopot/<sub>N</sub> ‘King of /Xo/s’ [+ material, – dynamic]  
 /shOrpo/ ‘snake’ ↔ /shOrpopoti/ ‘king of snakes’  
 /mOtsho/ ‘fish’ ↔ /mOtshopoti/ ‘king of fishes’

- (26) /X/<sub>N</sub> ↔ /Xpoti/<sub>N</sub> who owns a lot of /X/ [+ material, – dynamic]  
 /shOrpo/ ‘snake’ ↔ /shOrpopoti/ ‘who owns a lot of snakes’  
 /Oshsho/ ‘horse’ ↔ /Oshshopoti/ ‘who owns a lot of horses’

When a regular word (e.g. /poti/ ‘husband’) appears as the constant in some WFS, it enters the process of grammaticalization. In the course of time, it may lose its original meaning/use, and finally, its form, becoming something that other models call an *affix*. The sequence [poti] in (22-23) and (25-26) cannot be listed as a full word anymore because (27) sounds odd and (28) is simply not acceptable. In our view, the outputs of (22-23) are problematic for any model which claims to handle compounds in syntax with one of the following mechanisms: i) Movement (Roeper and Siegel 1978), ii) Incorporation (Baker 1988, Shibatani and Kageyama 1988, Kageyama 1991), iii) Argument linking principles (Lieber 1992), and iv) Coindexing (Lieber 2004).

- (27) ?/tini    amar    bhognir            poti  
           he    my    sister’s            husband  
           ‘He is my sister’s husband’
- (28) \*tini                    shilper                    poti  
           he                    industry’s                owner  
           ‘He is the owner of some industry’

If words ending in [poti] are formed in syntax, it is unclear why no other synonym of /poti/ (e.g. /shami/, /bOr/, etc.) can replace this sequence \*/shilposhami/ ‘industrialist’ or \*/koTibOr/ ‘millionaire’. Sequences like [poti] in (22-23) are not generally considered as suffixes either. It is unclear how words like /shilpopoti/ ‘industrialist’ or /bicarpoti/ ‘judge’ can be handled in morpheme-based models like that of Lieber (1992) or Kiparsky (1996) unless they can accommodate categories like *affixoids* (cf. Booij 2004) or ‘becoming affix’ and consider [poti] as such.

We claim that no such problems arise in a WWM account of the outputs of (1-15), (19-20) and (21-23) because the model attributes no status to [i], [kheko] or [poti] which appear as the constant in those words. One has the impression that [poti] in /bicarpoti/ is a separate entity, merely because this sequence is isophonetic with the full word /poti/ ‘husband’. However, the fact that the sequence [poti] is isophonetic with the regular word /poti/ or [i] is isophonetic with the constant of some other WFSs is irrelevant for morphology *à la* WWM. What is indispensable here is that, on the one

hand, there is one set of nouns in the lexicon that end in these sequences, while on the other hand, another set of nouns (or adjectives) is also there that lack these sequences. Morphology *à la* WWM will take care of the rest. Hence, we have a variety of words that end in the sequence [poti] because (21-23) and (25-26) allow us to form these words while the unavailability of words like \*/shilposhami/ or \*/koTibOr/ is due to the fact that our morphological module lacks the relevant WFSs.<sup>10</sup>

#### **4. Conclusions**

In this article, we have attempted to show how the number of WFSs based on the same formal contrast multiply in the course of time. As other models of morphology describe certain formal differences (manifested in different sets of word-pairs) as affixes, they see this phenomenon as affixal polysemy or homonymy. However, we claim that multiplication of WFSs based on the same formal contrast, but where each is endowed with a different semantic interpretation could be one of the reasons why a variety of words happen to end in a polysemous/homonymous affix in Bengali.

#### **Acknowledgements**

I am grateful to the reviewers whose thought provoking comments and suggestions have been very helpful. Usual disclaimers apply. I thank my senior colleague Dr. Dil Afroze Quader who has read several versions of this paper during the last four years and tightened up my prose. I must thank Professor Taro Kageyama, my Post-doc research supervisor in Japan and Professor Peter Hook for their helpful comments. I also thank Professor Monsur Musa, another senior colleague who sometime in the late eighties drew my attention to various {i}s in Bengali. For financial support, I gratefully acknowledge the beneficence of the Japan Society for the Promotion of Science (JSPS) and University of Dhaka, Bangladesh, while for logistic support, I am indebted to the Kwansei Gakuin University and the National Institute for Japanese Language and Linguistics, Tokyo, Japan.

**Notes**

\*An earlier version of this article was presented at *Theoretical contribution from lesser-studied languages and language varieties*, conference organized by Rice a Linguistic Society, at Rice University, USA, 21-22 February 2009.

1. Althoughs we have used traditional terms like *affix*, *suffix root* throughout this article to ease discussion, there would in fact be no need for this if we adopted the WWM framework.
2. {i} is also used as a derivational suffix for forming feminine nouns (a) and as an inflectional suffix for 1<sup>st</sup> person singular or plural (b). These two uses are not the concern of the present article.
  - (a) ‘Wife/feminine of /X/’ [- dynamic, + material]  
*mama* ‘maternal uncle’ → *mami* ‘wife of maternal uncle’
  - (b) ‘I do the action referred to in /X/’ [+ dynamic, – material]  
*pOr* ‘read!’ → *pori* ‘I read’
3. For similar cases in other languages, see Booij (1986) for Dutch, Beard (1990) for Russian, English, German and Serbo-Croatian, Lehrer (2003) for English and Rainer (2003) for Spanish.
4. Chatterji ([1926] 1970:671) (volume 2, chapter 1 entry 418) describes {i} as “a secondary affix, forming nouns and adjectives”. According to him “Three separate affixes of OIA seem to converge into this single NIA form: (i) <<-ī < -in >>; <<mālin > NIA mālī *gardener* : but the force of the nominative form seems to have been only of secondary importance in the evolution of the NIA. affix; (ii) << -īya >> : <<dēśīya > NIA. dēśī >> *native*; and (iii) <<-ika >>: <<grāmika >> > gāī >> *village name, clan name*: the feminine form of this <<-ikā >> is also <<-ī , -i >>, which is the most popular feminine affix of NIA.” Chatterjee (ibid.) also states that “Later in NIA. this affix was strengthened by the <<- ī >> of Persian.”
5. Examples in (1-15) are provided with an ad hoc semantic reading of the words along with the semantic features of the suffix {i} following Lieber (2004) who claims that most affixes can be classified by using only three features i) [+/- material], ii) [+/- dynamic] and [+/- IEPS (Inferable Eventual Position or State)]. According to her (2004:24) [+/-material] “defines the conceptual category of SUBSTANCE/THINGS/ESSENCE, the notional correspondence of the syntactic category Noun. The positive value denotes the presence of materiality, characterizing concrete nouns. Correspondingly, the negative value denotes the absence of materiality; it defines abstract nouns.” The feature [+/- dynamic] on the other hand, “signals an eventive or situational meaning, and by itself signals the conceptual category of SITUATIONS. The positive value corresponds to an EVENT or Process, the negative value to a STATE.” [+/- IEPS] denotes mainly change (s) in position or state.
6. In (9), {i} is concatenated to a phrase having the measure word /fūT/ as its head and the numeral /paNc/ as a modifier. Hence /paNc fūTi/ can be compared with

- ‘Queen of England’s’ in which the genitive suffix {s} is concatenated to a phrase and not to a word.
7. One of our reviewers reminds us that ‘no matter which morphological theory one espouses, there must be other means of creating words than those that belong to morphology proper. There may now be a ‘gate’ morphological strategy for creating names for political scandals, but the first word coined after ‘Watergate’ cannot logically be the output of morphology.’ We completely agree with the fact that the very first word formed after *Watergate* (as far as I remember it was *Irangate* formed in the mid-eighties) was not an output of morphology. However, there is a difference between *Irangate* and /bonkheko/. No WFS like /X/<sub>N,country</sub> ~/Xgate/<sub>N</sub> ‘political scandal related to /X/’ preceded the creation of the former, whereas it can be claimed, on the basis of attested pairs, that (19) was part of Bengali morphology much before /bonkheko/ came into existence.
  8. Another reviewer has found our idea of splitting of WFSs ‘troubling’ because, as (s)he expressed it: ‘Once new words are learned or coined, they become ‘alive’ to a point where it is no longer possible to generate them using the WFS that created them. If, by some luck, another word created by the same WFS were to ‘evolve’ in exactly the same way, then, yes, these new word pairs would license the creation of a new WFS. But how likely is that?’ In answer, we say that such a phenomenon could not take place randomly, which in turn could be the reason why affixal polysemy (or homonymy) does not pervade the language(s) in question.
  9. Words like /poti/, /Opotto/, /napat/ (>/nati/ in Bengali/Hindi) ‘grand children’ and English *nephew*, *nepotism*, *potent*, *power*, *possible* have derived from the same Indo-European root which probably meant ‘to give birth’ (see <http://www.etymonline.com>).
  10. One of our reviewers expresses the opinion that, for handling the problem of isophonic affixes, WWM is not fundamentally a better choice than the traditional Item-and-Arrangement analysis. According to him/her, ‘the approaches one can take are essentially the same: one can posit several isophonic affixes, rules or strategies or one can posit a single affix, rule or strategy with an underspecified meaning relation.’ If we look at (1-15) only, we cannot but agree with him/her. However, examples given from Spanish in Reiner (2003) and (24-26) convincingly show that a common underspecified rule of interpretation cannot satisfactorily account for all semantic relatednesses manifested in different pairs. Examples like (19-20) and (21-23) suggest that the splitting of some word formation process is basically triggered by the uniqueness of the semantic relatedness between two whole words whereas it is unclear whether *items* like [kheko] or [poti] or their *arrangements* have any impact on the splitting of the word formation rules which would involve those items



## References

- Aronoff, Marc  
1976 *Word Formation in Generative Grammar*. Massachusetts: The MIT Press.
- Baker, Mark  
1988 *Incorporation: a Theory of Grammatical Function Changing*. Chicago: The University of Chicago press.
- Beard, Robert  
1990 The Nature and Origins of Derivational Polysemy. *Lingua* 81:101-140.
- Booij, Geert  
1986 Form and meaning in morphology: the case of Dutch 'agent nouns'. *Linguistics* 24: 503-517.  
2004 Compounding and derivation, evidence from construction morphology. In *Morphology and its Demarcations*, Wolfgang U. Dressler, Dieter Kastovsky, Oslar E. Pfeiffer, and Franz Rainer (eds.), 109-131. Amsterdam: John Benjamins Publishing company.  
2005 *The grammar of Words*. Oxford University Press, Oxford.  
Chatterji, Sunitikumar  
1970 *The Origin and Ddevelopment of the Bengali language* (3 vols). London: George Allen and Unwin Ltd. Original edition, Calcutta University, 1926.
- Corbin, Danielle  
1987 *Morphologie Derivationelle et Structuration du Lexique*. Tubingen: Niemeyer.
- Ford, Alan, Singh Rajendra, and Gita Martohardjono  
1997 *Pace Panini, towards a Word-based theory of Morphology*. New York: Peter Lang.
- Jackendoff, Ray  
1990 *Semantic Structures*. Cambridge, MA:MIT Press.
- Kageyama, Taro  
1991 Light verb constructions and the syntax-morphology interface. In *Current English Linguistics in Japan*, Heizo Nakazima (ed.), 169-203. Berlin: Mouton de Gruyter.
- Kiparsky, Paul  
1996 Allomorphy or morphophonology. In *Trubetzkoy's Orphan*, Rajendra Singh (ed.), 13-31. Amsterdam: John Benjamins.
- Lehrer, Adrienne  
2003 Polysemy in derivational affixes. In *Polysemy, Flexible Patterns of Meaning in Mind and Language*, Brigitte Nerlich, Zazie Todd, Vimala Herman, and David D. Clarke (eds.), 217-232. Berlin: Mouton de Gruyter.

- Lieber, Rochelle  
 1992 *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago/London: The University of Chicago press.  
 2004 *Morphology and Lexical semantics*. Cambridge: Cambridge University Press.
- Neuvel, Sylvain  
 2003 *Metamorphology: A word-based account of polysynthesis and other multivalent morphological relations*, Ph.D. diss., Department of Linguistics, University of Chicago.
- Plag, Ingo  
 1997 The polysemy of –ize derivatives: on the role of semantics in Word formation. In *Yearbook of Morphology*, Booij Geert, and Jaap Van Marle (eds.), 219-242. Dordrecht: Kluwer.
- Pustejovsky, James  
 1995 *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Rainer, Franz  
 2003 Semantic fragmentation in word-Formation: The case of Spanish –AZO. In *Explorations in Seamless Morphology*, Rajendra Singh, and Stanley Starosta (eds.), 197-211. New Delhi: Sage Publications.
- Roeper, Thomas, and Muffy E. A. Siegel  
 1978 A Lexical transformation for Verbal compounds. *Linguistic Inquiry* 9:1999-260.
- Shibatani, Masayoshi, and Taro Kageyama  
 1988 Word formation in a modular theory of grammar: postsyntactic compounds in Japanese. *Language* 64: 451-484.
- Singh, Rajendra  
 2006 Whole word morphology. *Elsevier Encyclopedia of Linguistics*, 2nd ed., 578-579.
- Szymanek, Bogdan  
 1988 *Categories and Categorization in Morphology*. Lublin: Catholic University Press.
- Wierzbicka, Anna  
 1996 *Semantics: Primes and Universals*. Oxford: Oxford University Press.
- Yu, A. C. L.  
 2007 Understanding near mergers: the case of morphological tone in Cantonese. *Phonology* 24-1:187-214.



# Two Types of NPIs in Hindi with One Licensing Condition

Rajesh Kumar

*This paper presents a study of the syntax of two types of negative polarity items (NPIs) and their licensing in negative and non-negative contexts in Hindi. After providing an overview, it outlines the clause structure of Hindi for locating the position of negatives within that structure. It argues that the sentential negative marker heads its own maximal projection, which is immediately dominated by TP. The paper then presents the distribution of two types of NPIs in Hindi. Based on the syntactic evidence for the two types of NPIs, this paper shows that type I NPIs require a clause mate c-commanding negative licenser whereas type II NPIs do not. This paper finally shows that both type I and type II NPIs are licensed overtly at s-structure.*

## 1. Introduction

There are many studies discussing the syntax of NPIs (Linebarger 1987; Mahajan 1990; Laka 1994; Progovac 1994; Benmamoun 1996; Lahiri 1998, 2001; Vasishth 1999; Kumar 2003; Benmamoun & Kumar 2004). Progovac (1994) noted an asymmetry in the licensing of NPIs in Serbo-Croatian. Although Hindi NPIs are not uniform in their syntactic behavior (Lahiri 1998; Kumar 2003), differences in the occurrence and the positioning of the associated negative licenser in a sentence can be explained syntactically. Some NPIs, such as *ek phuuTii kauRii* ‘a red cent’, *Tas se mas honaa* ‘to budge an inch’, and *hargiz* ‘at all’, require a strictly local (clause mate) c-commanding negative licenser. By contrast, other NPIs—such as *koi bhii* ‘any’, *kisii bhii* ‘any’, and *ek bhii* ‘even one’—do not. The first grouping of NPIs obeys the locality conditions that usually constrain movement whereas the second grouping of NPIs does not. In this paper, I discuss the asymmetry outlined above in detail. I also discuss the syntactic constraints governing the two types of NPIs in Hindi.

The organization of this paper is as follows. After the introduction in section one, I outline the two types of NPIs in Hindi in section two. Section three presents a brief overview of negative markers in Hindi. Section four discusses the position of negative markers in the clause structure of Hindi. An analysis for the licensing of NPIs (Kumar 2003; Benmamoun & Kumar 2004) is presented in section five. Section six discusses the syntax of the two types of NPIs. The conclusion in section seven summarizes the paper.

## 2. Two types of NPIs

There are two distinct types of NPIs in Hindi<sup>1</sup>. I label them as type I NPIs and type II NPIs. Type I NPIs are a group of NPIs that require a clause mate c-commanding negative licenser. Type II NPIs are a group of NPIs that are permitted in long distance licensing contexts. The non-negative contexts such as questions, modality, conditionals, and adversative predicates that license type I NPIs do not license type II NPIs.

Type I NPIs are illustrated in (1) and type II NPIs are illustrated in (2).

- (1) *ek phuuTii kauRii* ‘a broken penny’  
*hargiz* ‘at all’  
*muNh kholna* ‘to say something’
- (2) *koi bhii* ‘anybody’  
*kisii bhii* ‘anybody’  
*kuch bhii* ‘any thing’  
*ek bhii* ‘even one/any’

All type I NPIs in (1) can only be licensed by sentential negation as illustrated in (3-5).

- (3) a. *maiN tum-ko ek phuuTi kauRii nahiiN*  
 I you-DAT one broken penny NEG  
*du-Ngaa*  
 give-FUT-MASC-IIISG  
 ‘I will not give you a red cent.’

- b. \* *maiN tum-ko ek phuuTi kauRii*  
 I you-DAT one broken penny  
*du-Nгаа*  
 give-FUT-MASC-IIISG  
 \* ‘I will give you a red cent.’
- (4) a. *raajiiiv hargiz nahiiN bol-egaa*  
 rajiv at all NEG speak-FUT-MASC-IIISG  
 ‘Rajiv will not speak at all.’
- b. \* *raajiiiv hargiz bol-egaa*  
 rajiv at all speak-FUT-MASC-IIISG  
 ‘Rajiv will speak at all.’
- (5) a. *raajiiiv-ne apnaa muNh nahiiN khol-aa*  
 rajiv-ERG self mouth NEG open-PERF  
 ‘Rajiv did not say a word.’
- b. \* *raajiiiv-ne apnaa muNh khol-aa*  
 rajiv-ERG self mouth open-PERF  
 ‘Rajiv said a word.’

On the other hand, type II NPIs can be licensed by sentential negation as well as other elements. Examples (6-9)<sup>2</sup> illustrate this in the context of questions.

- (6) a. *us kamre meN koi bhii sTuDeNT nahiiN thaa*  
 that room in any EMPH student NEG PST  
 ‘There was no/not even one (there wasn’t any) student in that room.’
- b. *us kamre meN koi bhii sTuDeNT thaa*  
 that room in any EMPH student PST  
 ‘Was there any student in that room?’
- (7) a. *maiN-ne kisii bhii sTuDeNT ko*  
 I-ERG any EMPH student ACC  
*nahiiN dekh-aa*  
 NEG see-PERF  
 ‘I did not see any student.’

- b. *maiN-ne kisii bhii sTuDeNT ko dekh-aa*  
 I-ERG any EMPH student ACC see-PERF  
 ‘Did I see any student?’
- (8) a. *maiN-ne kuch bhii nahiiN khaa-yaa*  
 I-ERG anything EMPH NEG eat-PERF  
 ‘I did not eat anything.’
- b. *maiN-ne kuch bhii khaa-yaa*  
 I-ERG anything EMPH eat-PERF  
 ‘Did I eat anything?’
- (9) a. *us kamre meN ek bhii sTuDeNT nahiiN thaa*  
 that room in one EMPH student NEG PST  
 ‘There wasn’t any student in that room.’
- b. *us kamre meN ek bhii sTuDeNT thaa*  
 that room in one EMPH student PST  
 ‘Was there was any student in that room?’

Before we move on to discuss the two types of NPIs and their syntax, I will present a quick overview of the three negative markers in Hindi.

### 3. Negative markers

Example (10) shows the distinct lexical negative markers in Hindi. They are *nahiiN*, *na* and *mat*. Of these three, *nahiiN* occurs most frequently (Bhatia 1973; Kumar 2003). The occurrence of the other two negatives, *na* and *mat*, is subject to restriction. *Na* occurs in subjunctives as in (10b) whereas *mat* is restricted to imperative sentences as in (10d). The negative marker *na* is preferred in the subjunctive while the use of *na* is marginal in the imperative. *Mat* is marginally allowed in the subjunctive when the sentence has a request or a prohibition of force as (10c) demonstrates.

- (10) a. *raajiiv aaj dillii nahiiN/\*/?na/\*mat jaa-egaa*  
 rajiv today Delhi NEG go-SUBJ  
 ‘Rajiv will not go to Delhi today.’

- b. *hamlog kahaaN na/nahiiN/\*mat jaa-yeN*  
 we where NEG go-SUBJ-IPL  
 ‘Where wouldn’t we go?’
- c. *mandir ke andar tasviir na/nahiiN/?mat*  
 temple of inside picture NEG  
*le-N*  
 take-SUBJ  
 ‘Please do not take picture inside the temple.’
- d. *mat/nahiiN/?na khaa-o*  
 NEG eat-EMPR  
 ‘Do not eat.’

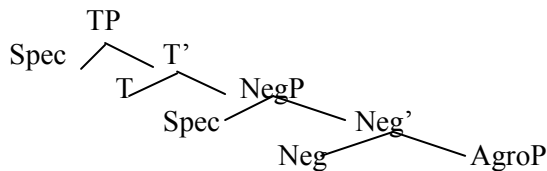
In the following section (4), I will outline the position of negation in the clause structure as discussed in Kumar (2003, 2004).

### 3.1. Sentential negative markers in the clause structure

Following Pollock (1989), Chomsky (1991) treats negation as a phrasal category where negation heads its own maximal projection and occurs in the functional domain of the clause structure below TP. In such systems, NegP occurs between TP and AgrP. In Pollock’s system NegP occurs between TP and AgrP. His argument for the split of IP into TP and AgrP is largely based on the position of negation in the clause structure that Zanuttini (1991) demonstrates to occur in Romance languages where NEG may be higher than TP.

Mahajan (1990) suggests that in Hindi negation heads its own maximal projection NegP and occurs below TP as in (11). Mahajan adopts the structure from Chomsky (1991). However, on the basis of the scrambling of DO, Mahajan (1990) suggests that the displacement of negation at LF is required for independent reasons in Hindi. For Mahajan, such an assumption is independent of the licensing of negative polarity items.

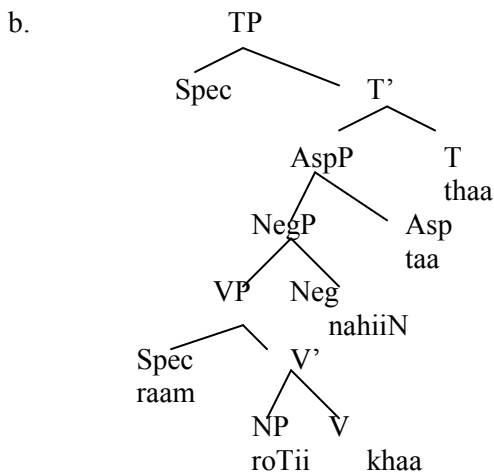
(11)





Dwivedi (1991) suggests that negation in Hindi heads NegP. She argues that the distinct syntactic behavior of Hindi negation shows that Hindi negation is unlike other modifiers. While modifiers are syntactically adjuncts that adjoin to the phrases they modify, negation is a head that takes a complement. Dwivedi suggests that sentential negation in Hindi selects VP as its complement and occurs in the clause structure below AspP. This is demonstrated in (12b).

- (12) a. *raam roTii nahiiN khaa-taa thaa*  
 raam bread NEG eat-HAB PST  
 ‘Ram did not use to eat bread.’



Dwivedi's suggestion is based on the adjacency requirement between the negation and the verb. Both Mahajan and Dwivedi agree that NegP occurs below TP in the functional domain.

I suggest that sentential negatives are not inside VP in Hindi. Although on the surface it appears that there is an adjacency requirement between the sentential negation and the verb, negation is outside VP in the functional domain. The evidence comes from the order of words in Hindi. Consider the following examples.

- (13) a. *raajiiv-ne anpnii kitaab [aisaa mujhe lag-taa*  
 rajiv-ERG self book so I-DAT seem-HAB  
*hai ki ] rekhaa ko nahiiN dii*  
 PRES COMP rekhaa-DAT NEG give-PERF  
 ‘It seems to me that Rajiv did not give his book to Rekha.’

- b. ??/\* *nahiiN raajiiv-ne anpnii kitaab [aisaa mujhe*  
 NEG rajiv-ERG self book so I-DAT  
*lag-taa hai ki] rekhaa ko dii*  
 seem-HAB PRES COMP rekhaa-DAT give-PERF  
 ‘It seems to me that Rajiv did not give his book to Rekha.’

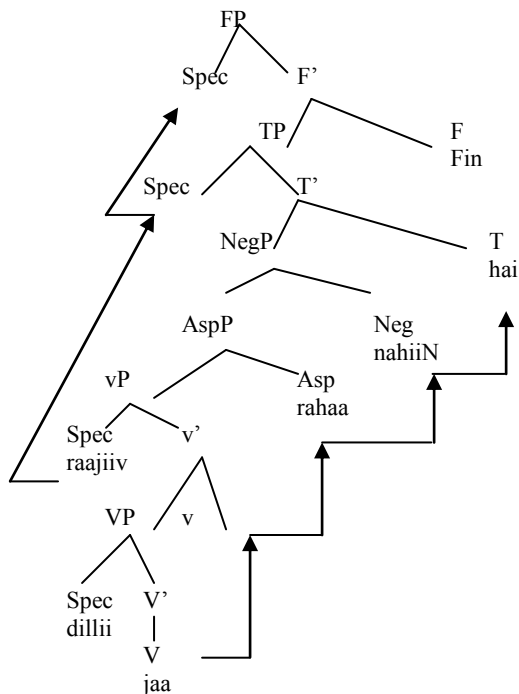
The example in (13b) clearly shows that the NEG is certainly not in the spec position. If we assume NEG to be in the spec position, we have to stipulate a movement that puts every thing from the left of NEG in (13a) to the right of NEG as in (13b). Such a movement makes the phrase ungrammatical. In concurrence with most of the proposals, it appears convincing that NEG heads its own phrase and occurs in the functional domain in the clause structure.

Other facts related to word order in Hindi show that NegP occurs below TP and above items such as AspP, MoodP and Modality. Following Pollock (1989), Chomsky (1989, 1995), and Mahajan (1990), I would suggest that the NegP is located below TP. Contrary to Dwivedi’s (1991) proposal, however, I suggest that NegP is higher than AspP.

- (14) a. *raajiiv dillii nahiiN jaa rahaa hai*  
 rajiv delhi NEG go PROG PRES  
 ‘Rajiv is not going to Delhi.’

- b. ? *raajiivdillii jaa nahiiN rahaa hai*  
 rajiv delhi go NEG PROG PRES  
 ‘Rajiv is not going to Delhi.’

c.



In the clause structure (14c), the V moves to *v* and then to Asp to incorporate the aspectual morpheme. Once it gets the aspectual morphology, the complex (*v* + Asp) moves to Neg and lands to the right of Neg<sup>3</sup>. From there, the verbal complex plus negation (NEG+verb) moves to T where it takes the tense morpheme. In this way, the sentence gets the desired order. Note that this derivation preserves the adjacency requirement between NEG and the verb. Dwivedi's (1991) argument that NEG is an X<sup>0</sup> category that heads its own maximal projection appears to be consistent with the empirical facts. However, NegP is higher than AspP and outside VP. If NegP occurs lower than AspP, the unacceptability of the example in (14b) remains unexplained. The occurrence of the negation marker between the verb and the continuous aspect marker is possible because the continuous aspect marker *rahaa* is not a morpheme. Rather, it is an independent word. Although (14b) is not ungrammatical it is also not the preferred choice in the intended reading of sentential negation. If we assume NegP to be below the aspect marker, (14b) is predicted to be acceptable. Thus, the unacceptability of (14b) and also the

aspectual morphology on the verb show that NegP is located outside VP and above AspP in the clause structure.

Following Bhatt (1999) and Rizzi (1997), I propose the structure for the Hindi sentence outlined in (14c). From Rizzi (1997), I borrow the concept of FP. The FP (Finiteness Phrase) in my system is equivalent to IP. In this system, the subject of the sentence originates in the spec, vP position. It then moves to the Spec of TP for the nominative case. Once it checks the nominative case, it moves from spec TP to spec FP to satisfy the EPP. The object stays within VP where it originates in the spec of VP. The verb (V) assigns an accusative case to its object under the Spec-head relationship. In the case of non-nominative subjects such as ergative and dative subjects, the subject moves directly to the spec of FP for EPP requirements. This move does not appear to violate any condition as the Spec of TP is not available. The Spec of TP is already filled with the NP that agrees with the verb. In these instances, where subjects are already marked with the non-nominative cases, the object, which agrees with the verb, moves to the spec TP position and gets the nominative case. The tense is located in T at the head of TP and the F of FP indicates the finiteness or non-finiteness of the sentence.

Thus far, I have discussed the two types of NPIs in Hindi, the three negative markers, and their position in the clause structure. In the following section, I will discuss the distribution of NPIs. I will also present the argument that NPIs in Hindi are licensed overtly in the course of derivation as discussed in Kumar (2003) and Benmamoun & Kumar (2004).

#### 4. The Syntactic Distribution of NPIs and the Licensing

In the following example NPIs (in this case *ek bhii*) occur in a variety of contexts. They can occur with subjects (15a), objects (15b), indirect objects (15c), and genitives (15d).

- (15) a. *ek bhii laRkaa klaas meN \*(nahiiN)aa-yaa*  
 one EMPH boy class in NEG come-PERF  
 ‘None of the boys came to the class.’
- b. *raajiv ek bhii kitaab \*(nahiiN)paR-taa hai*  
 rajiv one EMPH book NEG read-HAB PRES  
 ‘Rajiv does not read any book.’
- c. *raajiv-ne ek bhii laRke-ko kitaab \*(nahiiN)*

rajiv-ERG one EMPH boy-DAT book NEG  
*dii*  
 give-PERF  
 ‘Rajiv did not give book to any boy.’

d. *raajiv ek bhii dost-ke ghar \*(nahiiN)*  
 rajiv one EMPH friend-GEN house NEG  
*ga-yaa*  
 go-PERF  
 ‘Rajiv did not go to any friend’s house.’

Interestingly, the NPI *ek bhii* does not have to be in the same clause as the negative licenser. It can even be scrambled long distance. This is illustrated in (16) where the DP containing the NPI *ek bhii* scrambles the matrix clause to a position higher than the clause containing its negative licenser.

(16) [*ek bhii laRke-ko*]<sub>i</sub> *sariitaa-ne kah-aa ki*  
 one EMPH boy-ACC sarita-ERG say-PERF COMP  
*raajiv-ne t<sub>i</sub> yahaan nahiiN dekh-aa*  
 rajiv-ERG here NEG see-PERF  
 ‘Sarita said that Rajiv did not see any boy here.’

The fact that the NPI *ek bhii* does not co-occur with its negative licenser in the same clause raises questions about how it is licensed. There are many proposals for the licensing of NPIs. Some (Ladusaw 1979; Lahiri 1998) suggest semantic accounts for the licensing of NPIs, whereas others (Linebarger 1987; Mahajan 1990; Benmamoun 1996; Kumar 2003, among others) argue for a purely syntactic account. Kumar (2003) and Benmamoun & Kumar (2004) suggest that NPIs in Hindi are licensed overtly during the course of derivation prior to scrambling (i.e. at s-structure). In this proposal for the licensing of NPIs, the clause-mate negative licenser c-commands the NPI. Two alternative possibilities for this are given.

First, consider the argument for covert movement of the negative at LF where the NPI is not overtly in the c-command domain to negative markers (Mahajan 1990). Assuming that the negative does not overtly c-command the NPI subject, the negative *nahiiN* moves at LF to license

NPI. In this scenario, the negative licensor moves at LF and adjoins to the finite IP, thereby licensing the NPI. However, as Benmamoun & Kumar (2004) argue, there are problems with any analysis that relies on the LF movement of Negation. First, the movement of the negative licensor is an instance of head movement, which is considered to be local. If the movement of negation is necessary for the licensing of NPIs, this analysis predicts that heads may move out of their clause in long distance scrambling contexts. But this movement clearly violates the head movement constraint (Travis 1984; Rizzi 1990).

Second, Mahajan's analysis makes the incorrect prediction that the NPI can be generated in a higher clause and the negative licensor generated in a lower clause. In other words, we should expect sentences such as (16) to be possible.

- (16) \* [raajiv-ne ek bhii laRke-ko kah-aa] ki  
 rajiv-ERG one EMPH boy-ACC say-PERF COMP  
 sariitaa nahiiN aa-yii thii  
 sarita NEG come-PERF PST-FEM  
 \*'Rajiv told any boy that Sarita did not come.'

Finally, it not conceptually clear what motivates the movement of the negative licensor other than the need for it to be in a position where the c-command requirement is maintained.

Another option is to bring the NPI into the c-command domain of negation to reconstruct it in its pre-scrambling position since the NPI can be scrambled. This assumes that the licensing of NPIs is an LF condition. For example the NPI in (17) can be licensed at LF through reconstruction to the pre-scrambling position.

- (17) [ek bhii kitaab]<sub>i</sub> [ramesh nahiiN soc-taa  
 one EMPH book ramesh NEG think-HAB  
 hai [ki raajiv-ne t<sub>i</sub> paRh-ii]]  
 PRES COMP rajiv-ERG read-PERF  
 'Ramesh does not think that Rajiv read any book.'

However, there is evidence that the process of reconstruction cannot apply to the licensing of NPIs. One crucial piece of evidence involves the interaction of scrambled NPIs and Binding.

It is well known that binding conditions are sensitive to reconstruction (Barss 1986). For example, in Hindi an anaphor can be scrambled long distance. If the licensing of anaphors obtains at LF then the reconstruction or the copy of the scrambled element is what saves the representation. Although reconstruction is not an obligatory process, sometimes reconstruction must be blocked in situations where it may lead to violations of binding condition (Aoun & Benmamoun 1998). Consider the following sentence.

- (18) *raajiv-ne kah-aa ki [vo pikcars jo*  
 rajiv-ERG say-PERF COMP that pictures which  
*sariitaa<sub>i</sub>-ne lii]* *usko<sub>i</sub> pasand hai t<sub>j</sub>*  
 sarita-ERG take-PERF she-DAT like PRES  
 ‘Rajiv said that she likes the pictures that Sarita took.’

Reconstruction in (18) would lead to a Principle-C violation as the pronoun *usko* c-commands its antecedent *sariitaa* as shown.

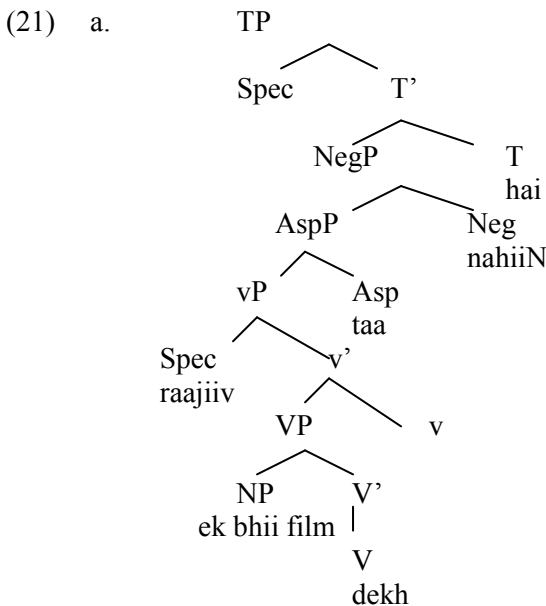
This is evident where reconstruction is blocked by cases where it must license NPIs at LF. Under such a scenario, we predict the outcome to be ungrammatical. However, this is not the case as illustrated in (19). Under the copy theory of movement, the reconstruction or access of the copy would lead to a violation of Principle-C as illustrated in (20) because the pronoun *usko* ends up c-commanding its antecedent *sariitaa*.

- (19) *raajiv-ne kah-aa ki [vo ek bhii pikcar*  
 rajiv-ERG said COMP that one EMPH picture  
*jo sariitaa<sub>i</sub>-ne lii]<sub>j</sub> usko<sub>i</sub> pasand nahiiN hai t<sub>j</sub>*  
 that Sarita-ERG took she-DAT likes NEG PRES  
 ‘Rajiv said that she does not like any picture that Sarita took.’

- (20) \* *raajiv-ne kah-aa ki usko<sub>i</sub> pasand nahiiN*  
 rajiv-ERG say-PERF COMP he-DAT like NEG  
*hai [vo ek bhii pikcar jo*  
 PRES that one even picture which  
*sariitaa-ne<sub>i</sub> lii]*  
 sarita-ERG take-PERF  
 ‘Rajiv said that she does not like any picture that Sarita took.’

Based on the evidence, it seems reasonable to conclude that under the copy theory of movement the NPI within the scrambled clause cannot be licensed through reconstruction or access to the original copy. Thus, we argue for the third option where NPIs are licensed overtly prior to scrambling (i.e., at s-structure) in the c-command domain of the negative licensor.

The licensing of NPIs in the object position is straightforward. The negative marker *nahiiN* c-commands NPIs in the object position. The same analysis carries over to datives and NPIs within complement and genitive positions.



- b. *raajiiiv ek bhii film nahiiN dekh-taa hai*  
 rajiv one EMPH film NEG watch-HAB PRES  
 ‘Rajiv does not watch any movie.’

The licensing of NPIs in the subject position raises questions. Assuming that the subject is generated in the Spec of VP, one could argue that the c-commanding negative licenses it prior to its movement to the functional layer (such as IP/TP). This raises concerns for languages such as English where A-moved elements in general (including subjects of passives



and of subject to subject raising constructions) must be licensed in the final (surface) A-position.

- (22) a. \* Any one was not arrested.  
 b. \* Any one is not likely to come.

On the other hand, since in Hindi there is plausible evidence that the negative *nahiiN* ends up on the head of IP/TP as we assumed above (see Dwivedi 1991; Kumar 2003, 2004), one could argue that the NPI in Spec IP is licensed in a Spec-head relation with the negative in I. This follows arguments that NPIs can be licensed in a Spec-head configuration (Hageman 1995; Benmamoun 1996)<sup>4,5</sup>.

In the following section, I discuss the syntax of the two types of NPIs based on the idea that they are licensed overtly and that both the types of NPIs are licensed under on licensing condition.

## 5. Syntax of the two types of NPIs

In the previous section we find that NPIs are licensed overtly. I extend the argument that both types of NPIs are licensed under the same licensing condition. Nonetheless, there are systematic syntactic differences between type I NPIs and type II NPIs. In the following sections, I discuss these differences.

### 5.1. Local vs. Long Distance Licensing of NPIs

The following data presents the systematic contrast between the two types of NPIs. Type I NPIs are not licensed long distance whereas type II NPIs can be licensed long distance. Type I NPIs in (23) such as *ek phuu-Tii kauRii* ‘a red cent’, *Tas se mas honaa* ‘budge an inch’, and *hargiz* ‘at all’ obligatorily require a local licenser.

- (23) a. \**raajiv-ne nahiiN kah-aa ki maiN tum ko*  
 rajiv-ERG NEG say-PERF COMP I you DAT  
*ek phuuTi kauRii* duu-Ngaa  
 one broken penny give-FUT-MASC-IIISG  
 ‘Rajiv did not say that I will give you a red cent.’

- b. \* *raajiv-ne nahiiN kah-aa ki vo Tas se mas*  
 rajiv-ERG NEG say-PERF COMP he deviate  
*huaa*  
 happen-PERF-MASC-IIIISG  
 ‘Rajiv did not say that he did not budge an inch.’
- c. \* *raajiv-ne nahiiN kah-aa ki ramesh hargiz*  
 rajiv-ERG NEG say-PERF COMP ramesh at all  
*bol-egaa*  
 speak-FUT-MASC  
 ‘Rajiv did not say that Ramesh will speak at all.’

On the other hand, type II NPIs such as *koi bhii* (any) and *kisii bhii*<sup>6</sup> (any) allow long distance licensors as in sentences (24) below.

- (24) a. *sariitaa-ne nahiiN kah-aa ki us kamre meN*  
 Sarita-ERG NEG say-PERF COMP that room in  
*koi bhii aadmii thaa*  
 any EMPH person PST  
 ‘Sarita did not say that there was anybody in that room.’
- b. *sariitaa-ne nahiiN kah-aa ki maiN-ne*  
 Sarita -ERG NEG say-PERF COMP I-ERG  
*kisii bhii laRke ko dekh-aa*  
 any EMPH boy ACC saw  
 ‘Sarita did not say that I saw any boy.’

In (24), the negative licensor *nahiiN* is in the matrix clause. The NPIs *koi bhii* and *kisii bhii* are in the lower clause.

There is, therefore, an asymmetry between the two types of NPIs in Hindi. To account for examples such as (24), Laka (1994) suggests that the negated verbs select for a special type of negative complementizer as the negative licensor in the matrix clause fails to license the NPI in the embedded clause. Consider the examples in (25) and (26).

- (25) a. [That anyone might do anything like that] never occurred to John.  
 b. [<sub>CP</sub> [NEG] COMP [<sub>IP</sub> ... NPI ...]]<sub>i</sub> ... V NEG ... t<sub>i</sub> ...
- (26) a. Mary does not think that Ann read any books last week.  
 b. ... V NEG ... [<sub>CP</sub> [NEG] COMP [<sub>IP</sub> ... NPI ...]]

Since the negative licensors in the matrix clauses do not license the NPIs in (25) and (26), Laka (1994) suggests that the negative COMP (selected by the negated verb in the matrix clause) licenses the NPI in the lower clause. Her analysis may work for type II NPIs in Hindi. Following her analysis, the type II NPIs in (24) can be explained as schematically represented in (27).

- (27) a. .. nahiiN (NEG) V.. [<sub>CP</sub> [NEG] COMP [<sub>IP</sub> ... *koi* *bhii* (NPI)... ]]  
 b. .. nahiiN (NEG) V.. [<sub>CP</sub> [NEG] COMP [<sub>IP</sub> ... *kisii* *bhii* (NPI) ... ]]

In (27a) and (27b) the negated verb *kahaa* selects a negative COMP. Thus, we can say that the NPIs in (24) are licensed by the negative COMP. However, Laka's assumption of a negative COMP does not work for type I NPIs as in (23). The examples in (28) illustrate this.

- (28) a. \*... nahiiN (NEG) V... [<sub>CP</sub> [NEG] COMP [<sub>IP</sub> ... *ek phuuTi*  
*kauRii* (NPI) ... ]]  
 b. \* ... nahiiN (NEG) V... [<sub>CP</sub> [NEG] COMP [<sub>IP</sub> ... *hargiz* (NPI)  
 ... ]]

All the examples in (23) have a negated verb in the matrix clause. The negative COMP in (23) does not license the NPIs in the lower clause as schematically presented in (28). Clearly, the analysis under which a negated verb selects a negative COMP does not explain the ungrammaticality of the examples in (23).

Progovac (1994) suggests a binding theoretic approach to the licensing of NPIs along the lines of the Generalized Binding of Aoun (1985, 1986). Progovac argues that NPIs in Serbo-Croatian are anaphoric. To capture the distributional properties of Serbo-Croatian NI-NPIs and I-NPIs, Progovac suggests that the NI-NPIs in (29) are subject to Principle-A and that the I-NPIs as in (30) are subject to both Principle-A and Principle-B. I-NPIs obey restrictions on Principle-B at the surface structure and obey the requirements of Principle-A at LF. As the negative licensors are not clause mates with I-NPIs at surface structure, the I-NPIs move at LF to be bound within the governing category. Serbo-Croatian NI-NPIs and I-NPIs are somewhat parallel to Hindi in expressing a similar locality restriction (local and long distance licensing). NI-NPIs occur only in the presence of an overt, clause mate negative licensor. Thus, the NI-NPIs of Serbo-Croatian are parallel to the strong NPIs of Hindi. Consider the examples below.

- (29) a. *milan ne vidi nista*  
 Milan NEG see-HAB nothing  
 ‘Milan cannot see anything.’
- b. \* *milan vidi nista*  
 Milan see-HAB nothing
- c. \* *milan ne tvrdi [da marija poznaje nikio-ga]*  
 Milan NEG see-HAB that Mary know-HAB none  
 ‘Milan does not claim that Mary knows no one.’ (From Progovac 1994)

While I-NPIs are licensed by a negative licenser, the negative licenser must be in the matrix clause. If they are in the same clause then the sentence is ungrammatical. In this case, Serbo-Croatian I-NPIs are similar to the Hindi NPIs *koi bhii* and *kisii bhii* to the extent that these Hindi NPIs can be licensed long distance as well. The following examples illustrate this.

- (30) a. \* *milan ne zna ista*  
 Milan NEG know-HAB anything  
 ‘Milan does not know anything.’
- b. *milan ne tvrdi [da marija poznaje iko-ga]*  
 Milan NEG see-HAB that Mary know-HAB none  
 ‘Milan does not claim that Mary knows no one.’ (From Progovac 1994)

Now, we can explain the long distance licensing facts of Hindi NPIs. We can say that type I NPIs in Hindi are anaphoric and that they can be licensed under Principle-A. This is so because they can only be licensed by a local negative licenser. Type II NPIs *koi bhii* and *kisii bhii*, on the other hand, move to higher clause at LF.

- (31) a. \* *raajiv-ne nahiiN kah-aa ki maiN tum ko*  
 rajiv-ERG NEG say-PERF COMP I you DAT  
*ek phuuTi kauri duu-Nгаа*  
 one broken penny give-FUT-MASC-IIIISG  
 ‘Rajiv did not say that I will give you a red cent.’

- b. *sariitaa-ne nahiiN kah-aa ki us kamremeN*  
 Sarita -ERG NEG say-PERF COMP that room in  
*koii bhii aadmii thaa*  
 any EMPH person PST  
 ‘Sarita did not say that there was anybody in that room.’

In (31) negative licensors are in the matrix clause and the NPIs are in the lower clause. Following Progovac’s assumption, we have to say that the NPIs in (31b) move out of the lower clause at LF and are thus licensed. However, as we have seen earlier, the LF movement of either NPIs or NEG is problematic in Hindi. Thus, progovac’s (1994) suggestion of the movement of NPI at LF for licensing does not work for the licensing of Hindi NPIs. Furthermore, the negative licensor and the type II NPI can occur very far apart from each other as in (32). Thus, (32) shows that type II NPIs require only a c-commanding negative licensor.

- (32) *sariitaa-ne nahiiN kah-aa ki meghaa*  
 Sarita -ERG NEG say-PERF COMP Megha  
*jaantii hai ki us kamre meN*  
 knows PRES COMP that room in  
*koii bhii aadmii thaa*  
 any EMPH person PST  
 ‘Sarita did not say that Megha knows that there was anybody  
 in that room.’

## 5.2 Licensing of NPIs in non-negative contexts

Type I NPIs do not occur in the context of questions, conditionals, modals, and adversative predicates whereas type II NPIs may occur in the absence of a negative licensor. Lahiri (1998) has a detailed semantic treatment of Hindi NPIs in such contexts. Consider the following examples.

## Type- I NPIs:

## In the context of questions:

(33) a.\* *voh tum ko ek phuuTi kauRii de-gaa*  
 he you DAT one broken penny give-FUT  
*(kyaa)*  
 what  
 ‘Will he give you a red cent?’

b.\* *raajiv hargiz bol-egaa (kyaa)*  
 rajiv at all speak-FUT what  
 ‘Will Rajiv speak at all?’

## In the context of conditionals:

(34) a.\* *(agar)voh tum ko ek phuuTi kauRii de-taa*  
 if he you-DAT one broken penny give-HAB  
*hai to accha hai*  
 PRES then good PRES  
 \* ‘If he gives you a red cent, then its ok.’

b.\* *(agar) raajiv hargiz bol-egaa*  
 if rajiv at all speak-FUT  
*to maiN maan jaa-uuNgaa*  
 then I believe go-FUT-ISG  
 \* ‘If Rajiv speaks at all, then I will believe it.’

## In the context of modals:

(35) a.\* *voh tum ko ek phuuTi kauRii de sak-taa*  
 he you DAT one broken penny give can-HAB  
*hai*  
 PRES  
 ‘He can give you a red cent.’

- b. \* *raajiv hargiz bol-taa hai*  
 rajiv at all speak-HAB PRES  
 \* ‘Rajiv speaks at all.’

*In the context of adversative predicates:*

- (36) a. \* *mujhe aaScarya hai ki us-ne tum*  
 I-DAT surprise PRES COMP he-ERG you  
*ko ek phuuTi kauRii dii*  
 DAT one broken penny give-PERF  
 ‘I am surprised that he gave a red cent.’
- b. \* *mujhe aaScarya hai ki raajiv hargiz*  
 I-DAT surprise PRES COMP rajiv at all  
 bol-aa  
 speak-PERF  
 ‘I am surprised that Rajiv spoke.’

The data above in (33), (34), (35), and (36) show that type I NPIs are not permitted in the absence of a negative licenser in the context of questions, conditionals, modals, and adversative predicates.

*Type-II NPIs:*

*In the context of questions:*

- (37) a. *us kamre meN koi bhii sTuDeNT thaa (kyaa)*  
 that room in any EMPH student PST what  
 ‘Was there even one /any student in that room?’
- b. *aap-ne ek bhii sTuDeNT ko dekh-aa (kyaa)*  
 you-ERG one EMPH student ACC see-PERF what  
 ‘Did you see any student?’

*In the context of conditionals:*

- (38) a. (*agar*) *us kamre meN koi bhii sTuDeNT*  
 if that room in any EMPH student  
*aa-taa hai to maiN aap ko bataa-uuNgaa*  
 come-HAB PRES then I you- DAT tell-FUT-ISG  
 ‘I will let you know if any student comes to that room.’

- b. (agar)maiN-ne **ek bhii** sTuDeNT ko dekh-aa  
 if I-ERG one EMPH student ACC see-PERF  
 to us ko pit-uuNgaa  
 then him DAT beat-FUT-ISG  
 'If I see any student, I will beat him up.'

*In the context of modals:*

- (39) a.us kamre meN **koi bhii** sTuDeNT  
 that room in any EMPH student  
 baiTh sak-taa hai  
 sit can-HAB PRES  
 'Any student can sit in that room.'

- b. maiN **kisii bhii** sTuDeNT se mil  
 I any EMPH student with meet  
 sak-taa huuN  
 can-HAB PRES  
 'I can meet with any student.'

*In the context of adversative predicates:*

- (40) a.mujhe aaScarya hai ki kal **koi bhii**  
 I-DAT surprise PRES COMP yesterday any EMPH  
 sTuDeNT aa-yaa  
 student come-PERF  
 'I am surprised that any student came yesterday.'

- b. mujhe aaScarya hai ki kal maiN-ne  
 I-DAT surprise PRES COMP yesterday I-ERG  
**kisii bhii** sTuDeNT ko dekh-aa  
 any EMPH student DAT see-PERF  
 'I am surprised that I saw any student yesterday.'

The examples above in (37), (38), (39), and (40) show that type II NPIs such as *koi bhii*, *kisii bhii*, and *ek bhii* are permitted in the absence of negative licensors in the context of questions, conditionals, modals, and adversative predicates.

Once again, Laka (1994) assumes the selection of a negative COMP primarily to provide a syntactic account for the licensing of the NPIs in sentential subjects. Note that in such cases the NPI is not in the c-command domain of the negative licensor. Uribe-Echevarria (1994) notes that Laka's (1994) proposal accounts for the licensing of NPIs in



adversative predicates as well. On the other hand, Progovac (1994) appeals to the presence of an operator in CP for the licensing of NPIs in the context of questions, conditionals, modals, and adversative predicates. With particular reference to adversative predicates on the basis of the following examples, Progovac notes that adversative verbs are not polarity licensors. Consider the example below.

- (41) a. \* John doubts anything.  
 b. \* John denied anything.  
 c. John doubts that Mary ate anything.  
 d. John denied that he had eaten anything.  
 e. [IP ... adversative V [CP OP [IP ...NPI ...]]]

Adversative verbs such as *doubt* and *deny* do not license polarity items when polarity items appear as the complements of these verbs. Note that the NPIs are c-commanded by the adversative verbs. However, NPIs are licensed if they occur in a clausal complement, shown in (41c-d). On the basis of the contrast in (41c-d) Progovac assumes the presence of an operator in CP (41e) that licenses these NPIs.

If we look at Hindi NPIs in light of these two analyses, the fundamental assumptions of both analyses do not account for the inability of questions, modals, conditionals and adversative predicates to license type I NPIs. Therefore, only licensing conditions on the two types of NPIs make them distinct. Type I NPIs require a c-commanding negative licensor within CP, whereas type II NPIs must be in the c-commanding domain of the negative licensor.

In the following section, I discuss the locality constraints on the licensing of NPIs. The following discussion further substantiates the claim that type I NPIs require a clause mate negative licensor whereas type II NPIs do not.

### 5.3. Licensing of NPIs and Island Constraints

Type I NPIs such as *ek phuuTii kauRii* (a red cent), *Tas se mas honaa* (budge an inch), and *hargiz* (at all) do not violate island constraints. In other words, type I NPIs cannot be licensed across syntactic islands. On the other hand, type II NPIs such as *koi bhii* (any) and *ek bhii* (any) can be licensed across syntactic islands. This is to say that they violate the island conditions that usually constrain movement. I explore the validity

of this claim with respect to Adjunct Islands and later with the Complex NP Constraint (CNPC).

### 5.3.1 NPIs within Adjunct Islands

Type I NPIs such as, *ek phuuTii kauRii* (a red cent) and *Tas se mas honaa* (budge an inch) cannot be licensed across adjunct islands. In other words, they obey locality conditions as in (42).

- (42) a. \**[ki raajiv ko ghar se ek phuuTii kauRii*  
           that rajiv    DAT home from a broken penny  
           *mil-ii] aaScaryajanak nahiiN thaa*  
           get-PERF surprising    NEG PST  
           ‘That Rajiv received a red cent from home was not surprising.’
- b. \**[ki raajiv hargiz bol-aa] aaScaryajanak*  
           that rajiv    at all    speak-PERF    surprising  
           *nahiiN thaa*  
           NEG    PST  
           ‘That Rajiv budged an inch was not surprising.’

The type I NPIs are inside adjunct islands in (42). The negative licensors, however, are in the main clause. It is thus clear from the above examples that the negative licensors do not license NPIs across the adjunct island.

On the other hand, the type II NPIs *koi bhii* (any) and *ek bhii* (any) are located within adjunct clauses. Their negative licensors are outside the island in the main clause as in (43).

- (43) a. *raajiv [koi bhii baat karne ke liye] nahiiN*  
           rajiv    any EMPH talk do for    NEG  
           *aa-yaa hai*  
           come-PERF PERS  
           ‘Rajiv did not come to talk about anything.’

- b. *raajiiv [ek bhii laRke ke ghar jaane ke liye]*  
 rajiv one EMPH boy of house go for  
*nahiiN aa-yaa*  
 NEG come-PERF  
 ‘Rajiv did not come to go to any boy’s house.’

### 5.3.2 NPIs within Complex NPs

Type I NPIs obey the Complex NP constraint as in (44), where type I NPIs are not licensed in complex NPs in the absence of a clause mate (local) negative licenser.

- (44) a. \* *raajiiv ke pitaajii [jinhoN-ne raajiiv ko*  
 rajiv of father who-ERG rajiv DAT  
*ek phuuTii kauRii dii] dilli meN nahiiN*  
 one broken penny give-PERF Delhi in NEG  
*rah-te*  
 live-HAB-PL  
 ‘Rajiv’s father who gave Rajiv a red cent does not live in Delhi.’
- b. \* [*vo aadmii jis ka raajiiv hargiz kuch*  
 that person who of rajiv at all something  
*kar paa-yaa] dilli meN nahiiN rah-taa*  
 do get-PERF Delhi in NEG live-HAB  
 ‘The person who Rajiv could do anything at all does not live in Delhi.’

In (45), we find that the sentence turns out to be grammatical if there is a negative licenser present in the complex NP.

- (45) a. *raajiiv ke pitaajii [jinhoN-ne raajiiv ko*  
 rajiv of father who-ERG rajiv DAT  
*ek phuuTii kauRii nahiiN dii] dilli meN*  
 one broken penny NEG give-PERF Delhi in  
*rah-te haiN*  
 live-HAB-PL PRES  
 ‘Rajiv’s father who did not give Rajiv a red cent lives in Delhi.’

Contrary to the syntactic behavior of the type I NPIs, type II NPIs *koi bhii* (any) and *ek bhii* (any) behave similarly with respect to the Complex NP Constraint as they did with the adjunct island. They are licensed across complex NPs, as well. In other words, they violate island constraints. This is demonstrated in (46).

- (46) a. *raajiiv-ne [uskii maaN-ne jo **koi bhii** baat*  
 Rajiv-ERG his mother-ERGthat any EMPH thing  
*kah-ii thii] vo **nahiiN** man-ii*  
 tell-EMPH PST that NEG obey-PERF  
 ‘Rajiv did not obey anything that his mother told him.’
- b. *raajiiv [un aadmioN se jis-ne **ek bhii***  
 rajiv those people with which-ERG one EMPH  
*bhaaShn di-yaa] **nahiiN** mil-aa*  
 speech gave-PERF NEG meet-PERF  
 ‘Rajiv did not meet anybody who gave any speech.’

From the examples above, it is clear that the type I NPIs do not violate island constraints. At the same it is also clear that the type II NPIs do not obey them. We have established that the Hindi NPIs are licensed overtly by a c-commanding negative licenser. This explains why the NPIs in adjunct islands and in complex NPs are not licensed.

## 6. Conclusion

In this paper, I demonstrate that there are two distinct types of NPIs in Hindi. This is established on the basis of local vs. long distance licensing, the occurrence of NPIs in non-negative contexts, and the licensing of NPIs within the syntactic island. I show, however, that there is one licensing condition that governs them as NPIs are licensed overtly at the s-structure prior to scrambling (Kumar 2003; Benmamoun & Kumar 2004). Following these arguments, I show that the assumption of negative comp (Laka 1994) as well as the binding theoretic approach for the licensing of NPIs (Progovac 1994) do not account for the unification of the two types of NPIs in Hindi.

**Notes**

1. This dichotomy echoes the typology of Serbo-Croatian NPIs discussed in Progovac (1994).
2. See Lahiri (1988) for a detailed analysis of the semantic contexts where Hindi NPIs are allowed.
3. I do not have an explanation for why the verb occurs to the right of the NEG in sentential negation when other heads appears to the left.
4. This would imply that in English the negative is not in I/T at the stage where the NPI subject must be licensed. Another important issue is whether one can dispense with one of the structural conditions, Spec-head or c-command (Agree). I put this issue aside as it is not relevant to main topic of the paper. However, it is possible that both may be needed (see Benmamoun 1996). This has its own implications in light of recent discussions of agreement such as Chomsky (2000).
5. Another option would be to generate sentential negation in a projection higher than the projection that hosts subjects. This would obviously have implications for head movement and the ordering of functional categories above the VP (see Ouhalla 1990, Zanuttini 1997, and Cinque 1999 for discussions of possible sentential negative head positions).
6. *kisii* is the oblique form of *koi*. In other words, *kisii* always occurs followed by a postposition.

**References**

- Aoun, Joseph  
1985 *A Grammar of Anaphora*, Massachusetts: MIT Press
- Aoun, Joseph  
1986 *Generalized Binding*, Foris: Doedrecht.
- Aoun, Joseph, and Elabbas Benmamoun  
1998 'Minimality Reconstruction and PF Movement' in *Linguistic Inquiry* 29 p 569-597
- Barss, Andrew  
1986 *Chains and Anaphoric Dependence*, MIT Doctoral Dissertation
- Benmamoun, Elabbas  
1996 'Licensing of Negative Polarity Items in Moroccan Arabic' in *Natural Language and Linguistic Theory*, 15: 263-287

- Benmamoun, Elabbas and Rajesh Kumar  
2004 'NPI Licensing: Covert or Overt?' paper presented at the 78<sup>th</sup> Annual meeting of the Linguistic Society of America at Boston, U.S.A.
- Bhatia, Tej Krishan  
1973 'On the Scope of Negation in Hindi', in *Studies in the Linguistic Sciences*, Vo. 3:2. Urbana: University of Illinois
- Bhatt, Rakesh  
1999 *Verb Movement and the Syntax of Kashmiri*, Dordrecht: Kluwer Academic Publishers.
- Chomsky, Noam  
1991 'Some Notes on the Economy of Derivation and Representation', in Freidin, R. (ed.), *Principles and Parameters in Comparative Grammar*, pp. 417-454. Cambridge: MIT Press.
- Chomsky, Noam  
1995 *The Minimalist Program*, Massachusetts: MIT Press
- Chomsky, Noam  
2000 'Minimalist Inquires: The Framework', in *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, Roger Martin, David Michaels and Juan Uriagereka (eds.), pp. 89-155. Cambridge, Mass.: MIT Press
- Cinque, Guglielmo  
1999 *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, Oxford.
- Dwivedi, Veena  
1991 'Negation as a Functional Projection in Hindi' in the Proceedings of the Western Conference on Linguistics 1991, vol 4, pp 88-101
- Haegeman, Lilian  
1995 *The Syntax of Negation*, Cambridge: Cambridge University Press
- Kumar, Rajesh  
2003 *The Syntax of Negation and the Licensing of Negative Polarity Items in Hindi*, Ph.D. Dissertation, University of Illinois at Urbana-Champaign
- Kumar, Rajesh  
2004 'Negation in the Clause Structure of Hindi', a paper present at CLS 40 (to appear in CLS 40 Proceedings) at the University of Chicago

- Ladusaw, William  
 1979 *Polarity Sensitivity as Inherent Scope Relations*, Doctoral Dissertation, The University of Texas at Austin
- Ladusaw, William  
 1992 'Expressing Negation' in C. Baker and D. Dowty (eds.) *Proceedings of SALT 2*, pp. 237-59. The Ohio State University, Columbus
- Lahiri, Utpal  
 1998 'Focus and Negative Polarity in Hindi' in *Natural Language Semantics* 6:57-123
- Lahiri, Utpal  
 2001 'Even-incorporated in NPIs in Hindi definites and correlatives' in Hoeksema, J. (et al) Ed. *Perspectives in Negation and Polarity Items*. Amsterdam/Philadelphia: John Benjamins Publishing Company
- Laka, Itziar  
 1994 *On the Syntax of Negation*, New York and London: Garland Publishing, Inc.
- Linebarger, Maria Christine  
 1987 'Negative Polarity and Grammatical Representation' in *Linguistics and Philosophy* Vol. 10, pp. 186-95
- Mahajan, Anoop Kumar  
 1990 'LF Conditions on Negative Polarity Item Licensing', *Lingua* 80: 333-348
- Ouhalla, Jamal  
 1990 'Sentential Negation, Relativized Minimality and Aspectual Status of Auxiliaries', *Linguistic Review*, 7: 183-231
- Pollock, J.-Y  
 1989 'Verb Movement, UG and the Structure of IP', *Linguistic Inquiry*, 20: 365-424
- Progovac, Ljiljana  
 1994. *Negative and Positive Polarity*, Cambridge: Cambridge University Press
- Rizzi, Luigi  
 1990 *Relativized Minimality*, Cambridge: MIT Press.
- Rizzi, Luigi  
 1997 'The Finite Structure of the Left Periphery' in Haegeman, Lillian (ed.) 1997 *Elements of Grammar*, pp 281-337. Netherlands: Kluwer Academic Publishers
- Travis, Lisa  
 1984 *Parameters and effects of word order variation*. PhD Thesis, MIT

Uribe-Echevarria, Miriam

1994 *Interface Licensing Conditions on Negative Polarity Items: A Theory of Polarity and Tense Interactions*. Ph.D. Thesis, University of Connecticut

Vasishth, Shrawan

1999 'Surface Structure Constraints on Negative Polarity and Word Order in Hindi and English' Paper presented at the 11<sup>th</sup> European School in Logic

Zanuttini, Raffaella

1997 *Negation and Clause Structure: A Comparative Study of Romance languages*, New York: Oxford University Press





# Wh-Phenomena in Sinhala\*

*Chandrika Maliyagoda*

*This paper claims that contrary to what is commonly accepted Sinhala has a rightward overt syntactic movement. The study argues that in Sinhala the uninterpretable feature is placed on the question particle  $d\bar{a}$   $Q$  and that thus, the moving element in Sinhala is not a wh-phrase but  $d\bar{a}$ . The question particle  $d\bar{a}$  in this case then has an overt syntactic movement to the right end of the clause fulfilling the function of a wh-phrase. On this basis, it is proposed that in Sinhala, wh-stem + trace of the question particle  $d\bar{a}$  is weak and the wh-stem +  $d\bar{a}$  is strong. It is claimed that the moving element is language dependent, and is not necessarily a wh-phrase in all languages. The evidence is that there is an additional landing site below the subject in addition to the traditional landing site for wh-phrases. The additional landing site, it is argued, is not visible when it is not needed, in order to conform with the economy principle and for the reasons of convergence at C-I interface. The evidence shows that Sinhala has multiple specifiers. It is assumed that the movement of one element to Spec-CP is a Universal phenomenon.*

## 1. Introduction

The Sinhala language, which is of Indo-Aryan origin, is most widely spoken in Sri-Lanka. Though Sinhala is said to belong to the class of Nominative-Accusative languages in Typology, it exhibits default agreement in complex predicates as well as agreement with the noun that occurs in a conjunctive verb, as in Hindi. However, present-day Sinhala, interestingly, shows a mix of shared characteristics with Tamil.

It is generally accepted that Sinhala wh-phrases are in-situ and they do not show any syntactic movement. This study raises questions in regard to many of these conventional views on Sinhala. The conventional views on Sinhala conclude that Sinhala has only wh-in-situ cases, and Sinhala wh-phrases/ $d\bar{a}$  have no overt syntactic movement. In the dissertation submitted to the Department of Linguistics at the University of Delhi in 2009, I have

dealt with some of these issues in Sinhala. Since then I have modified my views on some of these matters. The present study sets my current views down.

In this paper, I specifically argue that Sinhala has two types of overt syntactic movement. One exhibits the overt syntactic movement of the question particle *də* and the other demonstrates the overt movement of the question particle *də* + *wh-stem*. This means, first, that the Sinhala question particle *də* moves overtly to Spec-CP in all *wh*-in-situ cases. Secondly, that Sinhala, like English permits only one element to move overtly to Spec-CP, a phenomenon which I assume to be Universal. I propose that the element that moves overtly to Spec-CP is language dependent and it can be either a *wh*-phrase as in English or a question particle as in Sinhala. (In this case, I do not exclude the possibility of the moved element being any other phrase/particle than a *wh*-phrase/a question particle, in languages in the world). Based on data, it is possible to show that the Sinhala *wh*-elements *wh-stem*+Q in simple interrogatives exhibit overt movement to the right end of the clause. The study concludes that the Sinhala question marker/particle *də* is a clitic and that an uninterpretable feature [uF] is placed on it. On this basis I suggest that in Sinhala *wh-stem* + *də* is strong [uF] and *wh-stem* + *t* [iF] is weak. I point out that Sinhala respects the locality principle strictly. Further, this study claims that the clause bound-nature of Sinhala results from the strict locality principle. The evidence shows that Sinhala *wh*-elements have an additional landing site below the subject in addition to the traditional landing site and that the former is visible only when needed for the purpose of conforming to the economy principle and for the reasons of convergence at C-I interface. Evidence shows the existence of multiple specifiers in Sinhala.

The first part of this paper discusses rightward syntactic movement and short movement in Sinhala, and then in the section 4 of the paper, the discussion moves to the long movement in Sinhala. Then I shall look-at the two types of syntactic movements and the moving element and its characteristics. Final sections of the paper examine the possibility of having an additional landing site for *wh*-phrases in Sinhala and the availability of multiple specifiers in Sinhala.

**2. Rightward syntactic movement and short movement in Sinhala**

In this present section, I shall discuss rightward syntactic movement and short movement in Sinhala, which involves a wh-stem + question particle of Sinhala.

As we can see, the wh-phrases *ka-tə-də* (1) in simple interrogative sentences in Sinhala exhibit rightward movement to the end of the clause (1 - 5), in contrast to the pattern that occurs in English where all wh-phrases are fronted to a clause initial position with leftward movement. I propose that Sinhala wh-phrases have an overt syntactic movement to the right end of the clause in which they occur. In Sinhala, rightward movement of the wh-elements occurs only when there is a single wh-phrase in the simple interrogatives, as in the following examples (1-5)<sup>1</sup>.

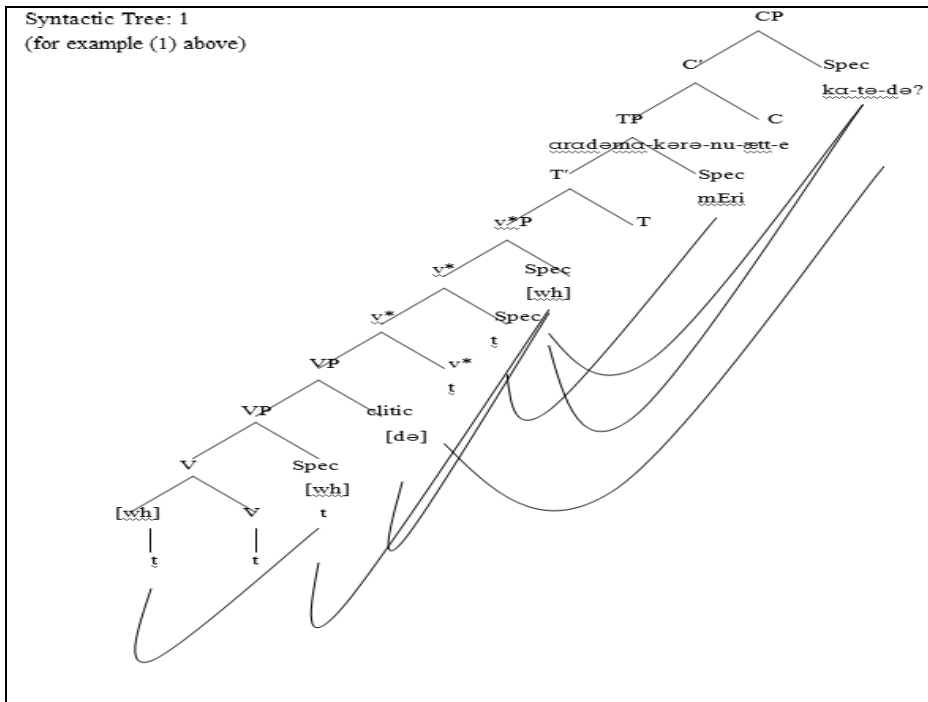
1. [TP mEri t<sub>i</sub> aradəna-kərə-nu-ætt e [CP ka-tə-də<sub>i</sub>?]]  
 Mary t invitation-do-PRT-has-Fut-Foc who-Acc-Q  
 ‘Whomi will Mary t<sub>j</sub> invite t<sub>i</sub>?’
2. t<sub>j</sub> t<sub>i</sub> mEri-tə aradəna-kərə-nu-ætte kawu-dək?  
 t t Mary-Acc invitation-do-PRT-has-Fut who-Q  
 ‘Who t t will invite Mary?’
3. mEri t<sub>i</sub> kanu-ætt-e monəwa-də?  
 mary t eat-PRT-has-Fut-Foc what-Q  
 ‘Whati will Mary t<sub>j</sub> eat t<sub>i</sub>?’
4. mEri t<sub>i</sub> aradəna-kərə-nu-ætt-e [kumənəi-t<sub>i</sub>-rahəs-parlksəkəyā-tə-də]<sub>i</sub>  
 Mary t invitation-do-PRT-has-Fut-Foc whci secret-examiner-Acc-Q  
 ‘Which<sub>i</sub> detective will<sub>j</sub> Mary t<sub>j</sub> invite t<sub>i</sub>?’ |
5. mEri u:ran-tə t<sub>i</sub> kəmə-de-nu-ætt-e kesE-də?  
 Mary pig-PI-Acc t food-give-PRT-has-Fut-Foc wh-Q  
 ‘How will Mary t<sub>j</sub> feed pigs t?’

I have indicated the syntactic movements of the interrogative sentence (1) in the tree diagram below (Syntactic tree 1). In Sinhala, it is assumed that, as illustrated there, V moves to v\* due to the affixal nature of the verb. I propose, in Sinhala, that the movement to v\*/v is obligatory due to this affixal nature of VP. It follows from Boškovič (2007) that the EPP is placed on the moving element, and thus the element moves looking for a

Spec to match and delete its uK feature. Based on this assumption, I argue that *mEri* has the uK feature which drives the movement from the base position to Spec-TP (Tree 1). In this case, the movement is driven by the moving element and therefore the question whether T is defective or not does not arise here. (Boškovič, 2007, helps me to avoid the inconclusive controversy whether T is defective or not because the movement is driven by the moving element and T does not raise the nominative subject *mEri* in this case. Tista Bagchi has pointed out in a discussion that some of the studies done on Bangla Tense and Aspect shows that the T is not defective). As the movement and the EPP are placed on the moving element (Boškovič, 2007), the question whether T without C is incomplete or not does not arise here.

6. [VP t t [v\* p t t t [TP mEri [CP  $\alpha$ radəna-kəra-nu-ət-te ka-tə-də?]]]]]

The *wh*-phrase *ka-tə-də* in the syntactic tree (The example 1, repeated as 6 above) which has question-particle *də* (uF) attached to it makes a rightward movement to Spec-CP to get its uninterpretable feature (uF) deleted (This is considered a formal inadequacy of the moving element, Boškovič, 2007).



The inadequacies induce movement of the elements which is essential to save the structure, and Chomsky (2007) concludes that inadequacy makes a language perfect (Chomsky in his early work assumed that inadequacy is a imperfection feature of Language). The movement of the *wh+stem-Q* occurs in the same clause and does not cross any barriers so that the occurrence of short movement in Sinhala is similar to that in English. The question particle *də*, I assume, is internally merged IM (Move) at the root of the tree to any phrase that it moves up with. Internally merged question particle *də* leaves copies for reconstruction and moves up to Spec-CP to mark Scope. The question particle *də* in Sinhala, I argue, cannot be externally merged which generates argument structure. It may be noted here that the Merger can take place only in overt syntax. The clitic *də*, in that case, has to be merged internally to the *wh-stem kə-tə-* in order to make the movement of the *wh-stem* to Spec-CP possible. The copies *t* left are not pronounced, but are needed for semantic interpretations at C-I. In the examples (1-5), we can see that the *wh-stem* moves up to Spec-CP with *də Q* for scope. The movement of the residue is permitted for the purpose of convergence at PF.

Based on these assumptions, I claim that the *wh-stem kə-tə-* in the tree diagram merges with the question particle *də* and moves to the outer Spec of *v\**, and then from there to Spec-CP. It is argued that the question particle *də* in Sinhala is base generated in a position internal to VP and merges with the element that it gets attached with. Merge cannot take place at an embedded phase, and thus the *də* has to be merged at the root of the tree, a position which is internal to VP.

The question particle *də* (cf. syntactic tree 1 above), which is base generated in a position internal to VP, I propose, can use the outer Spec of *v\*P/vP* as a landing site in its movement to Spec-CP to respect MCLP. Hence it follows that the question particle *də* has to be internally merged to *wh-stem kə-tə-* (syntactic tree 1) before its movement to outer Spec of *v\*P*, and I argue that the movement of the question particle *də* to Spec-CP in Sinhala is obligatory.

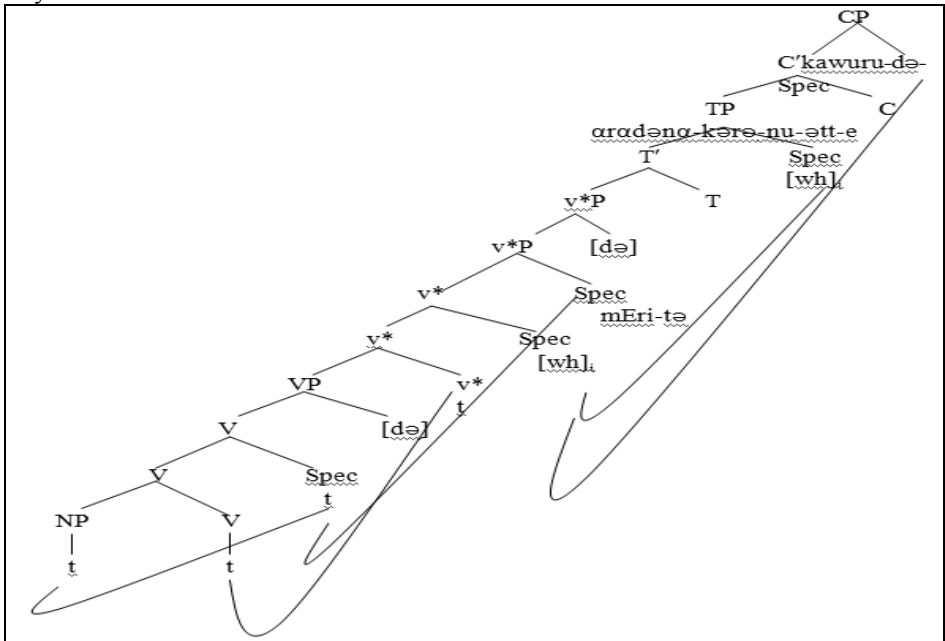
In the syntactic tree (1), the Accusative Case is checked at the outer Spec-*v\**. The rightward movement of the *wh-phrase kə-tə-də* for scope (syntactic tree 1) occurs after assignment of the accusative Case by the verb *aradəna-kərə-nu-ətt-e* (discussion with Nirmalngshu Mukherji). The lower copy *t* of the *wh-phrase* binds the moved *wh-phrase kə-tə-də*, hence it is a variable as in sentence [*mEri aradəna-kərə-nu-ətt-e ohu-tə; [... kə-tə-də]*] [whom<sub>i</sub> is such that [Mary will invite him<sub>i</sub>]]. The *wh-phrase* that is moved to the right end of the clause in the example (6) cannot be assigned

a Case as the copy *t* left by the wh-phrase *ka-tə-də* is already Case marked by the transitive verb *aradəna-kərə-nu-ətt-e*. The transitive verb *aradəna-kərə-nu-ətt-e* moves up to C after assigning the accusative case to the wh-phrase *ka-tə-də*. The movement of the verb *aradəna-kərə-nu-ətt-e* to C is obligatory as the participial *-nu* on the verb needs to be checked and deleted (6).

Contrary to the view held by Chomsky until 2006, he proposed in 2007 that the A and A-bar movements are uniform. Following Chomsky, I have shown in the (syntactic tree 2, the example 2 repeated as 7 below) that there is no relationship between *kawuru-də<sub>j</sub>* ‘who’ and *kawuru-də<sub>k</sub>*, but there is a direct relationship between ‘who<sub>j</sub>’ and ‘who<sub>k</sub>’ and between ‘who<sub>i</sub>’ and ‘who<sub>j</sub>’. The movement to Spec-TP (who<sub>i</sub> who<sub>j</sub>) in syntactic tree (2) is invisible, but is visible at the C-I interface for semantic interpretations. When the object NP *mEri-tə* is present as in the syntactic tree (2), the landing site for the clitic *də*, I assume, would be above the landing site of the object NP *mEri-tə*. The movement of *də* to this position is obligatory to satisfy MCLP.

7. mEri-tə aradəna-kərə-nu-ətt-e kawu-də?

Syntactic Tree: 2



The example (8) below shows that Sinhala is not permitted to leave the residue behind in its movement to Spec-CP as in English. In the ungrammatical example (9) the residue *basə-ye* is left behind. In the example (8), the clitic *də* ‘Q’ moves from the base position and gets attached to the residue NP *rahas-parlkšəkəyā-tə* leaving a copy *t* of *də* at the wh-stem *kumənə-t* for the semantic interpretations at C-I. Also, the whole phrase, wh-stem *kumənə-t<sub>j</sub>* and the residue *rahas-parlkšəkəyā-tə-də* ‘which detective’, moves rightward as a unit and, in this case, nothing can be left behind. The copy *t* left with wh-stem *kumənə-t* is not pronounced at the SM interface but it is needed for the semantic interpretations at the C-I interface.

8. mEri t aradənā-kərə-nu-ætt-e [kumənə-t-rahas-parlkšəkəyā-tə-də?]  
 mary invitation-Pl-do-PRT-has-Fut-Foc wh-t-secret-Pl-examiner-Defi-  
 -Ac-Q  
 ‘Which detective will Mary invite?’
9. \*mEri basə-yə t pitatwə-giyə [kumənə-də?]]  
 Mary bus-Obl. t out-went wh-Q  
 ‘In which bus did Mary leave?’

The examples discussed so far in this paper involve a one wh-phrase. In the next section, I shall discuss the characteristics of the wh-phrases/wh-stems when two wh-phrases occur in simple interrogatives and the status of the question particle *də* in such cases. There is required in establishing some critical evidence for my arguments on Sinhala.

### 3. Interrogatives with two wh-phrases

In Sinhala, when two wh-elements are involved in simple interrogatives, the language behaves differently from what I have discussed so far, and in this section, I shall discuss some of the differences, which are important to explain the behaviour of Sinhala language with wh- in a full manner.

As shown in the simple interrogative sentences (10 & 11), the wh-stems *kawuru-t*, ‘who’ *kā-tə-t*, ‘whom’ *kumak-t* ‘what’ that are left with a copy *t* of the question particle *də* do not move up to Spec-CP, but stay in-situ in contrast to the Sinhala simple interrogatives with one wh-phrase that show movement to the right end of the clause in which they occur (12). At the same time, in all these cases, we can see that the question particle *də* moves to the right end of the clause leaving a copy *t* with the relevant wh-



stems + *t*. The evidence shown so far concludes that the question particle *də* in Sinhala has the uninterpretable feature. The uninterpretable feature uF of *də* forces the overt syntactic movement. The question particle *də* is allowed to carry an extra baggage with it because of the overt syntactic movement (10 & 11). The covert movements are not permitted to carry any extra baggage with them. The question particle *də* that carries extra baggage in its movement as in the example (10) favours my argument that claims Sinhala has an overt syntactic movement.

10. kawuru<sub>j</sub>-t<sub>i</sub>    kα-t<sub>α</sub><sub>k</sub>-t<sub>i</sub>    aradənα-kərə-nu-æ<sub>t</sub>i-[d<sub>ə</sub>]<sub>i</sub>?  
             who-t            who-Acc-t    invitation-do-PRT-has-Fut-Q  
             ‘Who will invite whom?’
11. kawuru<sub>j</sub>-t<sub>i</sub>    kumak<sub>k</sub>-t<sub>i</sub>    ka-nu-æ<sub>t</sub>i-də?  
             who-t            what-t        eat-PRT-has-Fut-Q  
             ‘Who will eat what?’
12. mEri    t<sub>j</sub>    t<sub>i</sub>    aradənα-kərə-nu-ætt-e                            kα-t<sub>α</sub><sub>j</sub>-d<sub>ə</sub>?  
             mary invitation-Pl-do-Fut-PRT-has-Fut-Foc    who-Acc-Q  
             ‘Whom will Mary invite?’

In Sinhala, when more than one wh-element occurs in a sentence, the movement of the wh-stem + *t* is barred: but I argue that the question particle *də* is permitted to move up to Spec-CP. The syntactic movement of the question particle *də* in these instances, I hypothesise, is overt. The question particle *də* in Sinhala, I claim, has the uninterpretable feature uF necessary for its movement. The overt syntactic movement of *də* to Spec-CP will match and delete the uF of the question particle. The wh-stem + *t* which is in-situ, it is proposed, does not have the uninterpretable feature uF which is required for the movement. The copy *t* left by *də*, I propose, does not have the uF property which drives the movement. Hence the wh-stem + copy *t* of the question particle *də* cannot drive the movement. The movement, I assume, following Boškovič (2007), is driven by the moving element. In the case of Sinhala, looking at the data discussed in the section 2 & 3, it is evident that in all the examples the question particle *də* moves to right end of the clause irrespective of the fact whether simple interrogatives involve a single wh- phrase or more. On this ground, I argue that the clitic *də* in Sinhala has the uF property which is needed to drive the movement to Spec-CP. The copy *t* left by the question particle *də* is required for

the semantic interpretations at the C-I interface. The evidence suggests that the clitic *də* in Sinhala has the uF property which is needed to drive the movement. In the examples (10 & 11) above, the question particle *də* moves up to Spec of the CP overtly with the verb that it is attached with. The question particle *də*, I propose, spreads its uninterpretable features uF to any word that it gets attached to, licensing the movement. The verb *aradāna-kərə-nu-æti-* (10) is required to be heard at the PF.

Following the conclusion drawn above, it is evident that in the example (12) which contains a single wh-phrase, the Sinhala question particle *də*, which has the uninterpretable feature (uF) attached to it, licenses the movement of wh-stem *kə-tə-* to Spec-CP by spreading its features on to the wh-stem *kə-tə-*. Consequently, the wh-stem *kə-tə-* gets activated as a goal G. The wh-stem *kə-tə-*, it appears, does not have the uF that permits the movement, and therefore, it cannot move up to Spec-CP without the question particle *də* which does carry the uninterpretable feature uF. The question particle *də* needs to get strong Q checked at Spec-CP and the wh-stem *kə-tə-* and the verbs in the examples (10 & 11) in this case carry along with it and the question particle *də* is permitted to carry extra baggage with it only if the syntactic movement is overt.

The Sinhala wh-stems, as we can see in the section (2), examples (1-5), have the question particle *də* attached to them: *kawuru-də*, ‘who’, *kumak-də*, ‘what’, etc, and in the examples (10 & 11) it is attached to the verb. The question marker *də* in Sinhala can get attached to any phrase that it occurs with as pointed out earlier in this paper. It is on these grounds that I argue that the Sinhala question marker *də* is a clitic and I suggest that the Sinhala question particle *də* has the uninterpretable feature uF placed on it. It appears that the question particle *də* can spread its uninterpretable feature uF to any word/phrase that it gets attached to. Without the clitic *də*, ‘Q’, the wh-stems *kawuru-*, *kə-tə-*, etc. cannot move up to Spec-CP by itself. I argue, therefore, that it is the question particle *də* that licenses the movement of the wh-stem to Spec-CP in Sinhala. In the following section, I shall bring in evidence from colloquial Sinhala to support the conclusion drawn on wh-stems and the question particle *də*.

On the surface structure, it appears that the following colloquial examples in Sinhala raise questions for the unmarked written Sinhala. In the colloquial examples (13 & 14), as we can see, the wh-phrases wh-stems+Q: *kawu-də*, *monəwa-də* and *kə-tə-də*, do not move to the right end of the clause but stay in a position below the subject (I shall be dealing with colloquial in-situ wh-phrases later in this paper). The colloquial examples in (13 & 14), I argue, do not contradict the assumptions made above. The

Sinhala marked sentences (13 & 14) are examples of deletion. In those examples (13 & 14), the verb is being deleted in fast speech. The examples (15 & 17) exhibit that they are two clauses each containing a *wh*-phrase<sup>2</sup>. The example (16) is an instance of deleting the head noun or the subject in the speech.

13. kawu-də t monəwa-də kə:we?  
 Who-Q what-Q eat-Pst  
 ‘Who ate what?’
14. kawu-də t kə-tə-də entə-kivv-e  
 wh-Q t wh-Acc-Q come-Pst-told  
 ‘Who asked to come whom?’
15. kawu-də [kə:we], monəwa-də kə:we  
 wh-Q [ate], wh-Q ate  
 ‘Who ate what?’
16. Ø monəwa-də kə:we?  
 [Sub] wh-Acc-Q eat-Pst  
 ‘What did you eat?’
17. kəwe kawu-də, kəwe monəwa-də?  
 ate wh-Q ate wh-Q  
 ‘Who ate what?’

The discussion so far was limited to short movement in Sinhala. In the next section, I shall examine whether long movement exists in Sinhala, and if it does, is it similar to languages like English.

#### 4. Long movement in Sinhala

The data discussed in the sections (2 & 3) above (I exclude the Colloquial data here) illustrate the short movement of the *wh*-phrases to the right end of the clause. The pattern that occurs in Sinhala *wh*-phrases is similar to that involved in the case of short *wh*-movement in English, with parametric differences. The English-like languages have long *wh*-movement, in addition to short *wh*-movement. However, it is interesting to note that, on the basis of the data looked at so far, such long *wh*-movements are absent in Sinhala in similar situations. However, Contrary to the claim just made that

there is no long wh-movement in Sinhala, it seems to be evident from the following data, that Sinhala does indeed exhibit long movement, below, this paper will examine the type of long movement that occurs in Sinhala.

The behaviour of the following set of examples is clearly different from the simple interrogatives which have one/more wh-phrase and which have been discussed above. The single wh-phrase *ka-tə-t*, *kawu-də* wh stems + traces in the unmarked examples (18 & 19) stays in-situ in contrast to the short wh-movement in the Sinhala simple interrogatives. The clitic *də* which has the uninterpretable feature uF moves to the right end of the main clause.

18. mEri *ka-tə-t*      *aradəna-kərə-nu-ætəi*      *kiya t obə*  
 Mary wh-Acc-t invitation-PI-do-PRT-has-Fut that t you  
*višvasə-kəra-nne-hi-də?*  
 belief-PI-do-Pre-2Sg-Q  
 ‘Whom do you believe that Mary will invite?’
19. mEri *kawə-də-t* *pəmine-nu-ætəi* *kiya obə* *višvasə-kəran-ne-hi-də?*  
 mary wh-time-t come-Fut-PRT-has that you believe-do-2Sg-Q  
 ‘When do you believe that Mary will come?’

The question particle *də* (18 & 19) which shows movement uses the intermediate landing site in its movement to Spec-CP. I propose that the wh-phrase *ka-tə-t* with the copy *t* of the question particle *də* moves covertly to a Spec-CP below the subject for the C-I interpretations. The question particle *də*, which has the uninterpretable feature uF, moves up overtly to Spec-CP in the highest position. In this case, wh-stem+ trace *ka-tə-t* with an interpretable feature –uF is forced to stay in-situ. The evidence leads to the conclusion that the wh-stem+trace does not have uninterpretable feature uF which is necessary for movement. On this basis, it is argued that the *wh-stem+də* is strong in Sinhala and that *wh-stem+trace of the question particle də* is weak.

Thus far, I have discussed short and long movement in simple interrogatives and direct questions and the involvement of wh-phrases and the question particle in this regard. From now onwards, this paper will examine the behaviour of embedded/indirect questions in Sinhala.

### 5. Embedded or indirect questions in Sinhala

The embedded/indirect questions in Sinhala exhibit a unique behaviour that differentiates them from simple interrogatives with one *wh*-phrase and direct questions. However, they share some similarities with simple interrogatives that have two or more *wh*-elements in a sentence.

In direct questions above (18 & 19), the question particle *də* moves to the end of the main clause. It may be noted that both types of questions contain two clauses: higher and lower. In embedded question (20), the question particle *də* does not make a movement to end of the main clause, but moves to end of the subordinate clause leaving a copy *t* with *wh*-stem. The behaviour of the question particle *də* in this instance (20) conforms to the pattern that occurs in simple interrogatives with two or more *wh*-phrases. The question particle *də* in (20) moves overtly to the right end of the clause that it occurs in, but does not cross the phrase *kiya*. The weak *wh*-phrase *ka-tə-t* with the interpretable feature *iF* stays in-situ. Evidence from the embedded/indirect questions also leads to the conclusion that; in Sinhala, the moving element is the question particle *də*. The movement of the question particle *də*, in this case (20), is similar to the movement of the *wh*-phrase *whom* in English (21).

20. [mEri ka-təi-t<sub>i</sub> aradəna-kəra<sub>vi</sub>-də kiya] oɦu kalpəna-kəra-i  
 Mary who-Acc invitation-do-Fut-Q that he thoughts-do-Pre-3Sg-Be  
 ‘He wonders whom will Mary invite’

21. He wonders whom<sub>i</sub> will Mary invite t<sub>i</sub>?

In contrast to the embedded questions (20 & 23), the Sinhala example (22) which incorporates a direct question exhibits similarities with the English example (26) which also incorporates a direct question. In (22), the question particle *də* in Sinhala moves to the highest CP exhibiting long movement as the English (26) does. The phrase *kiya*, which corresponds to *that* in English, does not act as a barrier. The embedded questions in Sinhala (20 & 23) share the same characteristics with the English (25).

**Sinhala**

22. mEri t<sub>i</sub> k $\alpha$ -t $\acute{a}$ i-t<sub>j</sub> kawad $\alpha$ -t<sub>j</sub> arad $\acute{a}$ n $\alpha$ -k $\acute{e}$ ra-nu- $\acute{a}$ t $\acute{a}$ e $\acute{a}$ i [CP t kiya  
Mary t who-Acc-t wh-t invitation-do-has-Fut that  
ob $\acute{a}$  vi $\acute{s}$ vAs $\acute{a}$ -k $\acute{e}$ rann-e-hi-d $\acute{a}$ ?]  
you belief-PI-do-Pre-2Sg-Q  
'Whom do you believe that Mary will invite when?'
23. mEri k $\alpha$ -t $\acute{a}$ -t arad $\acute{a}$ n $\alpha$ -k $\acute{e}$ r $\acute{e}$ -nu- $\acute{a}$ t $\acute{a}$ i-d $\acute{a}$ [CPkiya mam $\acute{e}$  kalp $\acute{e}$ n $\alpha$ -k $\acute{e}$ r $\acute{e}$ -mi]  
mary wh-Ac-invitation-do-PRT-has-Q that I wonder-do-1Sng-Pre  
'I wonder whom Mary will invite?'
24. mEri k $\alpha$ -t $\acute{a}$ -t arad $\acute{a}$ n $\alpha$ -k $\acute{e}$ r $\acute{e}$ -nu- $\acute{a}$ t $\acute{a}$ e $\acute{a}$ i [CPkiya ob $\acute{a}$  kalp $\acute{e}$ n $\alpha$ k $\acute{e}$ ran-ne-hi-d $\acute{a}$ ?]  
mary wh-Ac-invitation-do-PRT-has that you wonder-do-3Sg-Pre-Q  
'Whom do you think that Mary will invite?'

**English**

25. I wonder whom do you believe that Mary will invite?  
26. Whom do you believe that Mary will invite t when?

The embedded question (25) has a long wh-movement in English but Sinhala wh-stems+t in (23 & 27) stay-in-situ, in contrast. The Sinhala wh-stem+t k $\alpha$ -t $\acute{a}$ -t in (27), in this case, behaves in the same manner as an interrogative with two or more wh-phrases. The question particle *d $\acute{a}$*  which moves to Spec-CP overtly has short movement. In Sinhala, long movement is absent in embedded questions but it must be noted that in all such cases, the question particle *d $\acute{a}$*  exhibits overt movement to the right end of the clause. Below, I have considered other possible theoretical hypotheses about the Sinhala data discussed, and attempt to show that they could be falsified.

27. mEri *kα-tə-t*<sub>j</sub> aradənα-kəra-vi-də<sub>j</sub> kiya owun siṭu-de mama  
 Mary who-Acc invitation-do-Fut -Q that they think-Pst-thing I  
 anumənə-kəle-m-i  
 guess-do-Pst-1Sg-Be

‘I wonder whomi they believe t that Mary will invite t’

For cases such as (27), another possible line of argument is that, since the Spec-CP is already occupied by *-də*, the *wh-stem* + trace *kα-tə-t* in the embedded sentence is unable to move to the right end of the clause, so that it is, therefore, forced to stay in-situ. But it must be noted that the *wh-stem* + trace *kα-tə-t* does not have the uF to drive the movement to Spec-CP; the argument, therefore, fails to stand its ground. It is also possible to argue that the word *kiya/kiyəla* together with *də* blocks the movement of *kα-tə-t* to Spec-CP. As we can see, there are some problems with this line of argument as well, because the claim that the phrase *kiya* blocks the movement of *də* is shown not to be valid by such examples as (22), in which the question particle *də* has moved to the lower clause crossing the phrase *kiya*. The movement, it is assumed, is driven by the moving element that has the uninterpretable feature uF. On the other hand, it is obvious that in the example (22), the phrase *kiya/kiyəla* does not block the movement of *də*. Thus, it is evident that this line of argument does not stand its grounds.

Based on the Sinhala data considered, I claim that the overt movement of the *wh-stem*+*Q* to the right end of the clause is allowed only in the simple interrogatives with a single *wh-phrase*. However, the data reveal that the question particle *də* in Sinhala moves overtly to the right end of the clause in all the cases discussed so far in this paper. The examples above show evidence of long movement in Sinhala. In those examples, the question particle *də* moves to the right end of the upper clause demonstrating similarities to the movement of the *wh-phrase* in English-like languages. On this basis, I conclude that the question particle *də* has overt syntactic movement to Spec-CP. The question particle *də* in Sinhala fulfils the requirements of a *wh-phrase*. In the case of Sinhala, the weak *wh-phrase* with the interpretable feature iF is left in-situ, with a copy *t* of the question particle *də*, while the question particle *də* which has the uF moves to Spec-CP. The *wh-stem* which has the iF is not permitted further movement. The data discussed so far in this paper show that the element that moves to Spec-CP is language dependent and that it is not always a *wh-phrase* in all languages including Sinhala. Thus, the evidence leads to the conclusion that it is the question particle *də* which moves to Spec-CP (28). Based on such evi-

dence, I argue that the question particle *də* in Sinhala fulfils the same function as that of a *wh*-phrase in English-like languages. The movement of the question particle *də* to Spec-CP is overt and it is similar to that of the *wh*-phrase in English. Watanabe (2001) and Richards (1999) propose that the languages allow only one *wh*-phrase to be moved overtly. I suggest that in Sinhala, overt movement is allowed only for the question particle *də*. Agreeing with Watanabe (2001) and Richards (1999), I argue that Sinhala also permits only one element to move overtly to Spec-CP, as English too does.

28. mEri kə-tə-t kawa-də arədənə-kərə-nu-ætəi- kiyə ohu  
 Mary wh-Acc-t wh-time invitation-Pl-do-PRT-has-Fut-Q that he  
 kalpənə-kəran-ne-hi-də?  
 think-Pl-do-Pre-3s-PRT-Q  
 ‘Whom does he wonder that Mary will invite t when?’

Watanabe (2001) discusses the possibility of having English *wh*-phrases licensed by some kind of unselective binding, instead of undergoing LF movement, which I assume is another possibility to explore for the Sinhala *wh*-in-situ cases.

The Sinhala in-situ *wh*-phrases, I claim, bear the interpretable feature *iF* like Korean *wh*-phrases. Thus, the *wh*-phrases that are left with a copy *t* of *də* cannot drive the movement to Spec-CP at the end of the clause. However, *wh*-elements are not permitted to stay in-situ as C-I interface requires them to move to a *wh*-landing site covertly for semantic interpretations. The only possible candidate available for the covertly moved *wh-stems+t* would be the landing site immediately below the subject. I suggest that the *wh*-phrases *wh-stem + t* with an interpretable feature *iF* in Sinhala example (28) move covertly to a landing site below the subject for C-I interpretations, when they are in-situ.

In section (4), I shall examine Sinhala multiple *wh*-phrases and their nature when they occur in multiple clauses in comparison with Hindi and Iraqi-Arabic.

## 6. The behaviour of *wh*-phrases in multiple clauses

The nature of Sinhala multiple *wh*-phrases with multiple clauses below differ from the data discussed above. The multiple interrogative clauses



each contains a *wh*-phrase (29). The behaviour of Sinhala multiple *wh*-phrases in this case is similar to that of the simple interrogatives discussed above in this paper. The strong *wh*-phrases *wh stem*+ *də* exhibit rightward movement to the end of the clause in which they occur (29). However, when two *wh*-phrases occur, as shown in the example (30) then the behaviour of the *wh*-phrases changes. In the example (30), *wh-stems*+*t* stay in-situ in contrast to (29) which has one *wh*-phrase in each clause. The question particle *də* moves overtly to right end of the clause in which it occurs. Based on evidence from Sinhala data, I suggest that Sinhala multiple interrogative clauses require either a *wh*-phrase or the question particle to occur in each clause for scope. This leads to the conclusion that Sinhala multiple interrogatives respect the locality principle. Dayal (1996) and Watanabe (2001) point out that the *wh*-in-situ in Hindi (32 & 33) and Iraqi-Arabic (34 & 35), respects the locality principle (The examples cited here are from Watanabe, 2001). Dayal (1996) points out that the clause-boundness of Hindi *wh*-phrases results from the extraposed finite complement clauses that act as islands for movement. In Sinhala, we must note that the moving element is question particle *də* and not the *wh*-phrase as in (31). Sinhala multiple clauses, it appears do not permit long movement as they respect strict locality. Comparing to Hindi and Iraqi-Arabic, Sinhala in the same environment requires two clauses to occur each containing a *wh*-phrase/*də*. This requirement, I suggests, results from scopal properties that forces multiple scope markers in multiple clauses as multiple clauses do no permit long movement for *wh*-phrases/*də*.

### *Sinhala*

29. ohu milə-dI gatt̪-e                      kumak-də kiya obə  
 he price-give-Pst-take-Pst-Foc what-Q that you.  
 Qsu-we kα-gen-də?  
 ask-Pst-wh-from-Q  
 ‘Whom did you ask whether he bought what?’
30. kawuru-tj kumak-tj kohen-tj      milə-dI-gatt̪ə də                      kiya  
 who-t what-t where-from-t price-give-Pst-take-Pst-Q that  
 matəkə                      kα-tə-də  
 remembrance wh-Acc-Q  
 ‘Who remembers bought what from where?’

31. mEri k $\alpha$ -t $\bar{a}$ -t  $\alpha$ r $\bar{a}$ d $\bar{a}$ n $\alpha$ -k $\bar{e}$ r $\bar{e}$ -nu- $\bar{a}$ t $\bar{a}$ i kiya ob $\bar{e}$  sitan-ne-d $\bar{e}$   
 mary wh-Ac-t invitation-do-PRT-has that you think-Fut-Sg.-Q  
 ‘Whom do you that that Mary will invite?’

**Hindi**

32. raam-ne puuchaa [ki mohan-ne kis-ko dekhaa?]  
 Ram-Erg asked that Mohan-Erg who saw  
 ‘Ram asked who Mohan saw?’
33. \*raam-ne socaa [ki kOn aayaa hEI?]  
 Ram-Erg thought that who come has  
 ‘Who did Ram think had come?’

**Iraqi-Arabic**

34. \*Mona tsawwarit [Ali ishtara shenoI?]  
 Mona thought Ali bought what  
 ‘What did Mona think Ali bought?’
35. Mona se?lat Ali IRo?a ishtarat shenoI  
 Mona asked Ali Ro?a bought what  
 ‘Mona asked Ali what Ro?a bought?’

The unique behaviour exemplified by Sinhala wh-phrases right along in this discussion, calls for further examination of wh-in-situ cases in Sinhala. Below, I shall look-at the data from both colloquial and standard written Sinhala.

**7. An additional landing site for wh-phrases in Sinhala**

Hence forth in this paper, I shall bring in critical evidence from Sinhala to establish the fact there is an additional landing site below the subject for wh-phrases in Sinhala. For Sinhala, so far, I have argued that there is overt syntactic movement of wh-phrases/*d $\bar{a}$*  as well as wh-in-situ cases. Sinhala colloquial examples and relative clauses demand further investigation with regard to the status of wh-in-situ. The question arises: do wh-phrases in Sinhala stay in-situ?

The colloquial example (36) demonstrates that Sinhala permits wh-phrases *wh-stem+Q* to stay below the subject without moving to the right end of the clause and this contradicts the unmarked written Sinhala in which the movement to the right end of the clause is obligatory for the wh-

phrase, as in (38). The question phrase in the example (36), I argue, does not stay in-situ but moves overtly to a landing site below the subject. The native speaker of Sinhala uses both types of sentences in their speech freely. If the language is rule governed then such flexibility is granted to the speaker due to the rules in operation and the same explains why the native speaker does not accept the topicalized question (40) and its answer (41) as natural. The acceptability of the sentence (36) is due to the existence of this additional landing site below the subject which permits *wh*-phrases to be moved below the subject. The unacceptability of (40 & 41) in turn shows that lack of evidence of any possibility of postulating a landing site to left of the clause. The unacceptability of the example (42), I propose, is due to the fact that Sinhala does not permit movement of two *wh*-phrases *wh-stem+də* to right of the clause. However, it must be noted, for the native speakers of Sinhala, the examples (36 & 38) are equally acceptable. In the example (39), the head noun is deleted (which would be allowed only in a discourse). This should not be taken as an instance in which the *wh*-phrase is topicalized. The example (37) shows deletion of verbs in fast speech. The unmarked written example (42) is not acceptable to the native speaker- that is because Sinhala does not permit movement of two *wh*-phrases to the right end of the clause.

36. oya monəwa-də kə:we?  
 [Sub] wh-Acc-Q eat-Pst  
 ‘What did you eat?’
37. kawu-də t monəwa-də kə: we?  
 Wh-Q t wh-Q eat-Pst  
 ‘Who ate what?’
38. Ø monəwa-də kəwe?  
 [Sub] wh-Q eat-Pst wh-Q  
 ‘What did you eat?’
39. oya kə:we monawa-də?  
 [Sub] eat-Pst wh-Q  
 ‘What did you eat?’

40. Q.?? monəwa-də mələ kə:we?  
 wh-Q mala eat  
 ‘What did Mala eat?’
41. A. ?? baṭ mələ kəwe  
 rice mala eat  
 ‘Mala ate rice’
42. \*obə kəwe monəwa-də kiyətə-də?  
 You eat wh-Q wh-time-Q  
 ‘What did you eat when?’

The in-situ wh-phrases in the colloquial examples above support my proposal that argues for an additional landing site for wh-phrases in Sinhala. In this case, I suggest that the wh-phrases are allowed to stay in a position below the subject because of this additional landing site below the subject. As shown above, the evidence from Sinhala suggests that only available position for such wh-phrases is below the subject and this argument would be supported further in the discussion below on relative clauses and multiple specifiers.

In addition to the Sinhala colloquial examples discussed, relative clauses in Sinhala show further evidence for an additional landing site below the subject (43 & 44). In the colloquial sentence (47), the wh-phrase is allowed to stay in a position below the subject. The Sinhala relative clause in the unmarked written sentence (43) corresponds with the colloquial example (47) in this respect. The wh-phrase *kawu-də* which has the uninterpretable feature uF moves overtly to a landing site below the subject *minis-α* ‘the man’. The noun *minis-α* and the wh-phrase *kawu-də* bear the nominative case. The wh-phrase *kawu-də* and the head noun *minis-α* are closely connected in their semantics. The discourse properties of the wh-phrase D-linked it with the head noun *minis-α* and therefore the wh-phrase *kawu-də* has a definite description. For the wh-phrase *kawu-də* in the relative clause (43), the landing site below the subject *minis-α* is the only available position. The example (46) also shows evidence that there is a landing site below the subject *mEri*. In the example *mEri* and *john* both occupy subject positions.

The relative clause examples (44 & 45) show the existence of multiple specifiers in Sinhala. In (44), as we can see, there is [pro] below the subject *John* and (45) exemplifies that the [pro] in (44) is pronoun *ohu* that occu-

pies a position below the subject *John*. Sinhala accepts both (44 & 45) as grammatical. Both [pro] and pronoun show evidence of another subject position below TP. The evidence from relative clauses leads to the conclusion that Sinhala has multiple specifiers. The existence of multiple specifiers in Sinhala in turn supports the claim that there is an additional landing site below the subject for wh-phrases. Based on evidence, I propose, an additional position for a subject can be either Case or non-Case position, for example, position for pro or a pronoun, depending on the subject to be moved.

43. mEri aradənə-kərə-nu-æti minis-α kawu-də kiya mamə dani-m-i  
 Mary invitation-PI-do-PRT-has-Fut. man-Defi. wh-Nom-Q that I know--  
 1Sg-Pre  
 ‘I know who the man is that Mary will invite’
44. [TP jɔ:n [?? pro [VP aradənə-kərənə-wə yæ-i kiya-pəwəs-u minisə  
 John pro invitation-do-Fut-PRT-that that say-Pst-PRT man-Defi.  
 mo-hu-yə]]]]  
 this-he-Be  
 ‘This is the man whom John said that [he will invite]’
45. [TP jɔ:n [?? ohu [VP aradənə-kərənə-wə-yæi-kiya-pəwəs-u  
 John he invitation-do-Fut-that that say-Pst-PRT  
 minis-α mo-hu-yə]]]]  
 man-Defi. this-he-Be  
 ‘This is the man whom [John said that he will invite]’
46. mEri [CP jɔ:n giya yæi kiya] pəwa-su-we kə-tə-də?  
 mary john went that that say-Pst-PRT wh-Ac-Q  
 ‘Whom did Mary say that John left?’
47. oyA monəwa-də kə:we?  
 [Sub] wh-Acc-Q eat-Pst  
 ‘What did you eat?’

Based on the data I have given in the above section, I claim that (a) in Sinhala, there is an additional landing site below the subject for wh-phrase,

and (b) Sinhala has multiple specifiers. Simpson & Bhattacharya (2003) argue that in Bangla (48 & 49), the *wh*-landing site is always not a clause initial (S) position but there is also a position lower than the subject (48 & 49). They also argue that the Bangla has an overt *wh*-movement in all *wh*-in-situ questions (Simpson & Bhattacharya, 2003). Watanabe (2001) points out that the Japanese (50) *wh*-phrases involve movement in overt syntax. However, in Sinhala, the movement of question particle *də* to the clause end position, I claim, is obligatory in all the cases. Sinhala question particle *d'* is not permitted to stay below the subject in its movement. The additional landing site below the subject is available only for the *wh-stem+də*. On that basis, I suggest that the additional landing site is not available for the overt syntactic movement of *də*. Therefore I propose that the additional landing site below the subject is not visible if it is not needed like the *v\*P/vP* in order to respect the economy principle.

### **Bangla**

48. jɔ:n ([CP meri cole gæche]) bollo ([CP meri cole gæche])  
 John Mary left gone said Mary left gone  
 'John said that Mary left'
49. ora [CP ke aʃbe] ʃuneche Sub [CP .... wh ....]  
 they who come will heard  
 a. Who have they heard will come?  
 b. They have heard will come.

### **Japanese**

50. Boku-wa [CP [IP John-ga nani-o katta] ka] shiritai  
 I-Top John-Nom what-Acc bought Q want-to-know  
 'I want to know what John bought'

## **8. Conclusion**

For Sinhala *wh*-elements, I have argued for an overt syntactic movement. The Sinhala *wh*-phrases demonstrate movement in overt syntax to the right end of the clause which is a rare phenomenon among languages that have been researched so far. The *wh*-phrases in Sinhala exhibit rightward movement only when there is one *wh*-phrase involved in simple interroga-

tives. Even when there is one wh-phrase in a clause, if the question particle *də* is not attached to the wh-stem, the movement of the wh-stem + trace *t* is not possible. When there is more than one wh-phrase in a clause or a complement clause in a complex sentence, I have shown that there is no movement of wh-phrases in Sinhala. Hence, I conclude that the wh-stem with the trace *t* of the question particle *də* is weak, as it does not have the uninterpretable feature uF necessary for the movement to take place. However, based on the evidence it is shown that in all these cases, the question particle *də*, which has the uninterpretable feature uF, has an overt movement to the right end of the clause. Thus, I argue that it is the question particle *də* that bears the uninterpretable feature in Sinhala and not the wh-stems, further, the evidence indicates that the element that undergoes overt syntactic movement to Spec-CP does not have to be a wh-phrase in all languages and that, in fact, the moving element is language dependent. I have also argued that the Sinhala *wh-stem + t* is weak and the *wh-stem + də* is strong, and that the findings lead to the conclusion that the Sinhala wh-stem does not have the uninterpretable feature which requires for the movement. Sinhala wh-phrases and the question particle *də*, it is claimed, respect the locality principle in certain environments.

I have proposed an additional landing site for Sinhala wh-phrases below the subject, and make the assumption that the landing site below the subject is visible (like v\*P/vP) whenever it is needed, in order to conform to the economy principle and for the purpose of convergence at C-I interface.

I also have shown that there are multiple specifiers in Sinhala and the additional subject position can either be a Case or non-Case position depending on the element to be moved.

The question particle *də* in Sinhala, it is shown, is a clitic and I argue that the uninterpretable feature uF is placed on the question particle *də* 'Q', which is necessary for movement, and that therefore *də* is strong in Sinhala. The Sinhala clitic in this case is different in its nature from the clitic in other similar cases, because it is allowed to bear the uninterpretable feature uF.

The evidence leads to the conclusion that irrespective of the fact that Sinhala and Bangla, English, and Japanese belong to three different language families historically; all four languages have, in respect of the syntactic phenomena dealt with in this paper, one element either a wh-phrase or a question particle moving overtly to Spec-CP. The findings permit the assumption that this behaviour reflects something universal among the languages that have been researched to date.

## Notes

\*I am deeply grateful to Professor K.N.O Dharmadasa for helping me with Sinhala data and to Professor Tista Bagchi, Professor Thiru Kandiah, Professor Nirmalangshu Mukherji, for their suggestions and guidance and to Professor Noam Chomsky for his advice and support. In finishing this paper, I am particularly grateful to Professor Thiru Kandiah for his encouraging notes and discussions.

1. The English examples (1-5) on which the Sinhala ones are based are taken from Haegeman (1991).
2. The questions and answers of the native speakers stated below (1-10) strongly support my argument for verb deletion in colloquial sentences. The answers (4, 5, 6, 7, & 10) given by the native speakers (who belong to different stratus of the society), exhibit the verb deletion in the questions. Some of the native speakers preferred not to delete the verb in their questions (4). In (1, 9 & 10), the subject is being dropped by the speaker (Sinhala is a pro-drop language) when talking to the 2<sup>nd</sup> person but not when referring to a 3<sup>rd</sup> person in their speech. The informants, it may be noted, do not have any Linguistic background (unlike Professor Dharmadasa who is a well known scholar and a Sociolinguist). The example (11), which was given by me, was not acceptable to the native speakers as they all found it unnatural- unless in a discourse- The answers of the native speakers also illustrate the fact that pair list answers are possible in Sinhala. The wh-phrases that are in-situ in colloquial examples, on the other hand, support the proposal for a wh-landing site below the subject.

1. ∅        monəwa-də gatt-e?  
       [Sub] wh-Acc-Q    buy-Pst  
       ‘What did you buy?’
2. kawuru-t    ekkə-də Av-e?  
       wh-t        with-Q    come-Pst  
       ‘With whom did you come?’
3. kawu-də        monəwa-də gatt-e  
       wh-Nom-Q wh-Acc-Q    buy-Pst  
       ‘Who bought what?’
4. kawu-də        gatt-e        monəwa-də gatt-e?  
       wh-Nom-Q buy-Pst    wh-Acc-Q    buy-Pst  
       ‘Who bought what?’
5. kawu-də        giye        e: gollə        monəwa-də gatte?  
       Wh-Nom-Q go-Pst    those-people wh-Acc-Q    buy-Pst  
       ‘Who bought what?’



6. mamə pota-k gatta saraṭ saḡəra-wa-k gatta  
 I book-Indef. bought Saraṭ magazine-Indef. bought  
 ‘I bought a book and Saraṭ bought a magazine’
7. A pota-k gatta B pQ:nQ-k gatta C p[ns]la-k gatta  
 A book-Indef. bought B pen-Indef. bought C pencil-Indef. bought  
 ‘A bought a book, B bought a pen and C bought a pencil’
8. saraṭ monəwa-də gatt-e?  
 Saraṭ wh-Acc-Q buy-Pst  
 ‘What did Saraṭ buy?’
9. Ø monəwa-də gatt-e?  
 [Sub] wh-Acc-Q buy-Pst  
 ‘What did you buy?’
10. Ø ædum gatta waləlu gatta palətu ru gatta  
 [Sub] clothe-Pl bought bangles bought fruit-Pl bought  
 ‘I bought clothes, bangles and fruits’
11. Ø monəwa-də gatt-e  
 [Sub<sub>3re</sub>] wh-Acc-Q buy-Pst  
 ‘What did she buy?’

## References

- Bošković, Željko  
 1999 On Multiple Feature Checking: Multiple *Wh*-Fronting and Multiple Head Movement. In Samuel David Epstein and Norbert Hornstein (eds.). *Working Minimalism*. 159-187. MIT Press, Cambridge MA.
- Bošković, Željko  
 2002 On Multiple *wh*-Fronting. *Linguistic Inquiry* 33, 351-383.
- Bošković, Željko  
 2007 On the Locality and Motivation of Move Agree: and Even More Minimal Theory. *Linguistic Inquiry* 38, 589-644.
- Chomsky, Noam  
 1994 Bare Phrase Structure. In G Webelhuth (ed.). 1995. *Government and Binding Theory and the Minimalist Programme*. 383-439. Blackwell, Oxford.
- Chomsky, Noam  
 1995 *The Minimalist Programme*. MIT Press, Cambridge. MA.
- Chomsky, Noam  
 2000 Minimalist Inquiries: The Framework. In Roger Martin, David Michaela, and Juan Uriagereka. *Step by Step: Essays on Minimalist Syntax in Honour of Howard Lasnik*. 89-155. MIT Press, USA.

- Chomsky, Noam  
1999 Derivation by Phase. Ms. MIT.
- Chomsky, Noam  
2001 Beyond Explanatory Adequacy. Ms. MIT.
- Chomsky, Noam  
2006 On Phases. In Robert Freidin, Carlos Otero and Maria-Luisa Zubizarreta (eds.) *In fundamental Issues in Linguistic Theory*. 1-25. MIT Press, Cambridge, MA.
- Chomsky, Noam  
2007 Approaching UG from Below. In Uli Sauerland and Hans Martin Gartner (eds.) *Interfaces + Recursion = Language?: Chomsky's Minimalism and the view from the syntax – semantics*. 1-29. Mouton de Gruyter, Berlin, New York.
- Davison, Alice  
2001 April, 2009. Internet. [www.benjamins.com/cgi-bin/t\\_articles.cgi?bookid=LAP](http://www.benjamins.com/cgi-bin/t_articles.cgi?bookid=LAP). *Word Order, Parameters and the Extended COMP Projection*.
- Dayal, Veneeta  
1996 *Locality in Wh-Quantification: Quantification and Relative Clauses in Hindi*. Kluwer Academic Publishes, Boston.
- Haegeman, L  
1991 *Introduction to Government and Binding Theory*. Basil Blackwell Ltd. Oxford.
- Hagstrom, P  
1999 April, 2009. Internet. [www.bu.edu/linguistics/UG/hagstrom/papers/Q-movement-LSA-1999.pdf](http://www.bu.edu/linguistics/UG/hagstrom/papers/Q-movement-LSA-1999.pdf). *Q-movement*.
- Hinzen, Wolfram  
2006 *Mind Design and Minimal Syntax*. Oxford University Press, New York.
- Kayne, Richards  
1994 *The Antisymmetry of Syntax*. MIT Press, Cambridge MA.
- Maliyagoda, C.K  
2006 *Relative Clauses in Tulu*. Unpublished M.A. Dissertation. University of Delhi, India.
- Maliyagoda, C.K  
2009 Wh-Phenomena and Relative Clauses in Sinhala. Unpublished M.Phil Dissertation. University of delhi, India.
- Nuhman, M.A  
2003 *A Contrastive Grammar of Tamil and Sinhala Noun Phrases*. University of Peradeniya, SriLanka.

Richards, Norvin

- 1999 Feature Cyclicity and Ordering of Multiple Specifiers. In Samuel David Epstein & Norbert Hornstein (eds.). *Working Minimalism*. 127-158. MIT Press, Cambridge MA.

Simpson, Andrew, Tanmoy Bhattacharya

- 2003 Obligatory Overt *Wh*-Movement in a *Wh*-in Situ Language. *Linguistic Inquiry* 34, 127-142.

Watanabe, Akira

- 2001 Wh-in-situ Languages. In Mark Baltin & Chris Collins (eds.). *The Handbook of Contemporary Syntactic Theory*, 203 – 225. Blackwell, Oxford, Malden, MA.

# Diverging Pathways: The Current Status of Grammaticalization of Hindi Verbs

*Shakthi Poornima and Robert Painter*

*This paper argues that collocations of main verb and light verb in Hindi can be input for either grammaticalization or lexicalization. We examine the type and token frequencies of twelve verbs which can be identified as 'light' in a random sample taken from the EMILLE corpus and show that these verbs exhibit various degrees of on-going grammaticalization. Certain combinations of main verb and light verb have also come to be lexicalized, forming idiomatic phrases. The data presented here is analyzed via an integrated model of grammaticalization and lexicalization presented in Brinton and Traugott (2005).*

## 1. Introduction

This paper explores the premise, first described by Hook (1974), that constructions formed with Hindi light verbs can be input for processes of both grammaticalization and lexicalization to varying degrees. Using extensive synchronic data extracted from the twelve million word Hindi *EMILLE* corpus, we observe that Hindi light verbs indeed exhibit features of on-going grammaticalization, without necessarily taking on the formal and functional properties of the distinct class of auxiliaries. This is an important finding since research in grammaticalization theory often describes light verbs as occupying an intermediate position on an unidirectional cline of development from full verb to auxiliary to affix (see e.g. Hopper and Traugott 2003:111). Recent work on the diachronic development of complex predicates in general, and of light verbs specifically, have been increasingly critical of the idea that light verbs belong on the typical verb-to-affix grammaticalization clines which are well-attested for developments elsewhere, e.g. in Romance (see Bower 2008). At the same time, we show that some combinations of main verb and light verb have come to be lexicalized, fossilizing into idiomatic phrases. This, too, was described earlier by Hook (1974); however, the use of the corpus data allows us to show that the instances of idioms involving light verbs are perhaps more widespread in the modern language than suggested by earlier studies. While not denying

that the use of corpora is a means towards extracting language data for testing theoretical questions empirically, the use of the *EMILLE* corpus to cull instances of complex predicate constructions in a systematic fashion took on a special significance in the present study. This is because nearly all previous theoretical work on Hindi light verbs has been based on relatively limited sources of data. For instance, the often-cited discussion of light verbs in Hook (1991), a study which forms the basis of Hopper and Traugott's treatment of the grammaticalization of light verbs (2003: 112-114), consists of only two sets of a hundred sentences containing complex predicates from Marathi and Hindi-Urdu, obtained from a set of short stories written by two authors, S.V. Chirmule and K.B. Vaid (see Hook 1991: 81). This dataset is arguably too small both in overall size and number of representative sources for formulating strong conclusions about the distribution of light verbs in Hindi and the patterns of combination between particular light verbs and main verb in the formation of complex predicates. By using the *EMILLE* corpus to extract a hundred sentences for each of twelve verbs, commonly listed in other works as 'light' (see Section 2), the present work hopes to reevaluate the claims, and potentially revalidate the important contribution, of studies such as Hook 1974 and 1991, but with a much greater amount of empirical data. In undertaking this research, our main goals were to obtain a clearer sense of how prevalent light verb constructions are in modern Hindi in terms of both type and token frequency; what their pattern of distribution is, in the sense of positional occurrence, i.e. standard versus reverse predicate constructions; what patterns of restrictions (if any) exist in the collocation of specific main verbs with light verbs; and, finally, to learn to what extent light verbs can be described as being involved in processes of grammaticalization and lexicalization.

The framework for our analysis of the light verb data is the model presented in Brinton and Traugott (2005), which integrates grammaticalization and lexicalization into a unified continuum of related processes. Following Bowerman 2008 in seeing light verb constructions as a stage of development resulting from various processes of grammatical change, our analysis describes light verbs in modern Hindi as partially grammaticalized in some cases, lexicalized in others, and yet we avoid taking the strong position presented in Hopper and Traugott (2003) that light verbs occupy an intermediate stage of development between full verb and auxiliary.

This study is outlined as follows: we revisit the debate in the literature on light verbs and grammaticalization in Sections 1.1 and 1.2. We present extensive evidence of our analysis on the grammaticalization and lexicali-

zation of Hindi light verbs using data from the Hindi EMILLE corpus in Section 2. Discussion and conclusions about the implications of viewing instances of Hindi complex predicates as alternatively both lexicalizing and grammaticalizing are presented in Section 3.

### 1.1. Defining Hindi Light Verbs

Like other South Asian languages, Hindi has a rich set of complex predicates that are formed when two verbs comprise the predicate of a single clause (Hook 1974; Butt 1995; for further cross-linguistic examples of complex predicates, cf Bowerman 2008). Syntactically, the ‘main’ verb encodes the main predication while the ‘light’ verb inflects for number and gender, and agrees with the unmarked argument in the clause. Semantically, the specific contribution of a light verb is hard to pin down, but it is fairly well known that, in Hindi, light verbs contribute an aspectual meaning such as perfectivity (Hook 1991, 1993; Butt 1994). In slightly different terminology, light verbs have been argued to affect the *Aktionsart* of the main predication by creating an accomplishment reading (Butt 2002:14). Other subtle semantic notions, such as a nuance of suddenness or benefaction, which are not readily analyzable as aspectual in nature, are also observable in complex predicate constructions. By and large, light verbs in Hindi accord with syntactic and semantic properties of light verbs found in other languages, such as Farsi and Uzbek (cf Bowerman 2008:163).

A typical predicate construction in Hindi with a single finite main verb is illustrated in (1).<sup>1</sup> By way of comparison, (2) shows complex predicate constructions with the non-finite main verb and finite light verbs *le* ‘take’ (2a) and *de* ‘give’ (2b), respectively:

- (1) shyam=ne      makaan      banaa-yaa  
 shyam.M=ERG house.M.SG:MV make-M.SG:LV  
 ‘Shyam made the house.’
- (2) a. shyam=ne      makaan      banaa<sub>MV</sub> li-yaa<sub>LV</sub>  
 shyam.M=ERG house.M.SG make take-M.SG  
 ‘Shyam made the house (for himself).’
- b. shyam=ne      makaan      banaa<sub>MV</sub>      di-yaa<sub>LV</sub>  
 shyam.M=ERG house.M.SG make give-M.SG  
 ‘Shyam made the house (for someone else).’

Note that, traditionally, the glossing of light verbs in the examples throughout this paper indicates the lexical citation meaning of the corresponding main verb. The actual lexical semantics of a given light verb in each construction is variable. For instance, the light verb *le* ‘take’ in (2a) adds a notion of benefaction for oneself, i.e. Shyam made the house for his own use. Likewise, the light verb *de* ‘give’ in (2b), imparts a sense of benefaction for others, i.e. Shyam made the house for someone else’s benefit. Additionally, the examples in (2) also show that in a complex predicate construction, it is the light verb that inflects for number and gender. Since Hindi is a head-final language, the light verb usually follows the main verb (but see Section 2.3 on the *reverse* complex predicate construction, in which the light verb precedes the main verb). This suggests a degree of fusion between the main verb and the light verb, which is further supported by the fact that the number of particles that can appear between the two verbs in a complex predicate construction is quite restricted. For instance, the emphatic marker in (3a) is one of a limited number of elements which can appear between the main verb and the light verb in order to modify the main verb; other elements, such as a negative particle, cannot come between the two verbs, as shown in (3b). A rigid word order is also seen between the complex predicate and auxiliaries, with auxiliaries in Hindi coming phrase finally, as in (3c).

- (3) a. shyam=ne      makaan      banaa<sub>MV</sub>      hii      di-yaa<sub>LV</sub>  
 shyam.M=ERG house.M.SG make                  emph give-M.SG  
 ‘Shyam **really** made the house (for someone else).’
- b. \*shyam=ne      makaan      banaa<sub>MV</sub>      nahii      di-yaa<sub>LV</sub>  
 shyam.M=ERG house.M.SG make                  neg give-M.SG  
 \*Intended: ‘Shyam **did not** make the house (for someone else).’
- c. shyam=ne      makaan      banaa<sub>MV</sub>      di-yaa<sub>LV</sub>      th-aa<sub>AUX</sub>  
 shyam.M=ERG house.M.SG make      give-M.SG be-PAST.3.SG  
 ‘Shyam **has** made the house (for someone else).’

It should be noted that the light verbs and auxiliaries are distinct syntactic categories in Hindi. Hindi auxiliaries distinguish themselves from light verbs by not making any other semantic contributions (volitionality, benefaction, etc.) to the argument structure. Auxiliaries instead concatenate with either the verbal root or verbal inflected forms and they serve to signal distinctions of tense, aspect, mood, and voice. Auxiliaries and light verbs also

show distinct syntactic behaviors with regard to case marking, word order, reduplication, and topicalization (see Butt 2002:11-12).

## 1.2. Diachronic Development and the Question of Clines

The general consensus over the diachronic development of the complex predicate construction in Hindi is that the modern construction has its origin in the Sanskrit gerund/absolute, *-tva/tya*, or *-ya* (cf. Tikkanen 1987, Butt 2002:16-25, Butt 2005:13-18). We show in (4) a hypothesized source construction from Old Indo-Aryan where the verb ‘go’ is not the main predication of the sentence, but rather this verb modifies the event semantics of the participle ‘fly’ (Tikkanen 1987, cited in Butt 2002:23).

- (4)   tato   maksiko-ddi-ya           gataa  
       then   fly/insect-fly-GD       go  
       ‘Then the fly flew away.’

In the above example, although the gerund appear to be used in connection with the perfective auxiliary *gam-* ‘go’, Tikkanen explains this as follows: “[I]n Vedic and Classical Sanskrit, the gerund was frequently construed with stative or habitual light verbs, but only in and after Pali do we find the gerund to be construed with perfective light verbs that often add some additional shade of meaning indicating the point of view or role of the speaker, e.g., *gam-* ‘go all the way/away with’” (1987:262). This meaning of the verb *jaa* ‘go’ persists in modern day Hindi, as in (5).

- (5)   tab   paanchii           uD<sub>MV</sub> ga-yaa<sub>LV</sub>  
       then   bird.F.SG       fly   go-M.SG  
       ‘Then the bird flew away.’ (first author’s dialect)

The modern complex predicate construction is considered to be nonexistent in Sanskrit and an innovation in Indo-Aryan languages by Hook (1991:61) and Tikkanen (1987:262). However, Butt (2002) cites various counterexamples to demonstrate that the complex predicate construction has existed in the language at least since Old Indo-Aryan (c. 1200-200 BC). As she points out in a later work, the construction becomes easier to identify in texts starting from Middle Indo-Aryan (1200-200 BC) and Old Hindi (200 BC-1100 AD) (Butt 2005). There is clearly still a need for a definitive historical study on the emergence of these forms in the textual record. Cloud-



ing the issue, as Bowerman has recently argued, there is compelling evidence that light verb constructions in many languages may emerge through borrowing: either the predicate construction with light verbs itself is borrowed across languages or a new light verb construction develops when speakers of one language borrow verbs from another language as preverbs (Bowerman 2008:173). Along similar lines, Hook (1991) suggests that light verbs constructions may have been borrowed into Hindi from a Dravidian language as an areal feature.

Despite uncertainty as to the exact pathway of development of the Hindi main verb-light verb construction, earlier research in grammaticalization analyzed the construction as a rare example of the gradual emergence of aspectual meaning or “aspectogenesis” (cf Hook 1991, 1993). The light verbs, according to Hook, exhibit a degree of semantic bleaching, having lost contentful lexical meaning but acquiring ‘functional’ grammatical meaning, e.g. perfective aspect (1991:59). As the light verbs come to be obligatory markers of perfective aspect, the corresponding constructions with only the main verb are reanalyzed to mark imperfectivity; as Hook states: “thus, we are dealing not just with the grammaticalization of a set of overt linguistic items but with the complementary grammaticalization of zero” (Hook 1991:78). This observation is corroborated by Butt: “Native speakers will insist that the action seems *incomplete* or unsituated when only a single verb is used” (1994: 84, original italics). This contrast between complex predicate (perfective) and single predicate (imperfective) is illustrated in the examples in (6) and (7) (modified from Hopper and Traugott 2003:112):

(6)      mǎi=ne    das    baje      aap=ko    fon      kar<sub>MV</sub>    li-yaa<sub>LV</sub>  
           I=ERG    10    o'clock you=DAT phone make    take-M.SG  
           ‘I telephoned you at 10 o’clock.’ (and we actually spoke)

(7)      mǎi=ne    das    baje      aap=ko      fon      ki-yaa<sub>MV</sub>  
           I=ERG    10    o'clock you=DAT      phone make-M.SG  
           ‘I telephoned you at 10 o’clock.’ (the phone rang with no answer)

In (6), where the verb phrase is a complex predicate formed by the main verb *kar* ‘make in this context’ and the light verb *le* ‘take’, the telephoning-event was successfully completed; in comparison, the preferred interpretation in (7) with the main verb *kar* ‘make’ in the verb phrase is that the call was unsuccessful or incomplete.

This apparent emerging ability of light verbs to mark perfectivity distinctions led first Hook (1974), followed by Hopper and Traugott, to conclude that a subset of all light verbs are developing functional properties intermediate between those of full verbs and auxiliaries. This evidence suggested to these researchers that light verbs are an optional stage on a kind of verb-to-affix grammaticalization cline from ‘less grammatical’ to ‘more grammatical’, as in (8) (cf Hopper and Traugott 2003: 111-116).

(8) full verb > (light verb) > auxiliary > verbal clitic > verbal affix

Citing just the data from Hook (1991), Hopper and Traugott observe that some light verbs appear to be emerging as more grammaticalized than others, a fact which is taken to mean that these verbs are further advanced on the cline toward more auxiliary-like status. They point out that the five most frequent light verbs in the Hindi-Urdu text sample accounted for 92% of the total number of light verbs, without any reference to the form of these verbs (Hopper and Traugott 2003: 116). They conclude that “in Hindi a handful of light verbs is gaining the ascendancy in the competition for auxiliary status” (ibid.). This can be conceptualized as the diagram in (9).

(9) full verb > (light verb) > auxiliary  
   Most light verbs                                 5 light verbs

Other lines of research on the Hindi-Urdu verbal system refute this idea that light verbs, as a class or as a subset, are diachronically moving toward more auxiliary-like status (see Butt and Geuder 2001; Butt 2003; Bower 2008). The observed diachronic stability of a light verb construction in languages which possess one is supported by Bower’s cross-linguistic study (2008:175). For Hindi, the claim is that since light verbs have been highly stable historically, involve no phonological loss and demonstrate no observable shift toward more auxiliary-like functions, they should not be seen as an intermediate stage on the grammaticalization cline between full verb and auxiliary (Butt 2003:16). Light verbs in Butt’s analysis are viewed as “intimately connected to their main verb counterpart in the lexicon” (ibid.), with just one lexical entry which to her suggests that the light verb does not enter the grammaticalization cline. The relationship of the light verb to the type of grammaticalization cline given in Hopper and Traugott (2003: 111) is presented in (10) (from Butt 2003: 16):

- (10) main verb > auxiliary > clitic > affix(es)  
       |  
       light verb

Studies such as Butt and Bowerman are consistent with more recent, and often critical, work on grammaticalization theory which avoid making quasi-teleological statements about directionality and clines of the type in (8) (cf. Norde et al 2004). In our opinion, the strict reading of the question of whether or not the light verb is on the verb-to-affix grammaticalization cline is, in essence, comparatively uninteresting. More fascinating is that the development of light verbs, both in Hindi and cross-linguistically, seem to follow its own pathway of diachronic development, emerging from a number of source constructions, particularly hypotactic or embedded structures or developing through the borrowing of preverbs (see Bowerman 2008:166ff). A much needed line of research is to trace in detail what stages of grammaticalization, if any, can be traced for the complex predicates cross-linguistically, and more specifically, in Hindi. Particularly of interest would be to examine whether the diachronic development of complex predicates is unidirectional, a feature of grammaticalization which is often criticized, as well as to better isolate and follow what stages of development (what “cline”) are commonly observed. Whereas the typical verb-to-affix cline is largely defined by loss on several parameters – e.g. semantic bleaching, phonological erosion, morphological fusion to a stem – it is not clear that light verbs exhibit such properties, with the possible exception of a greater degree of fusion to an external host (see Section 1.2).

If we start with the simple question of “what is the best evidence for grammaticalization phenomena in respect to light verbs in Hindi?”, an admittedly simplistic but useful answer is to see this data in terms of divergence. Originally conceived as a kind of heuristic to detect instances of incipient or on-going grammaticalization, divergence is a natural outcome of grammaticalization processes, whereby “when a lexical form undergoes grammaticalization, for example to an auxiliary...the original form may remain as an autonomous lexical element and undergo the same changes as any other lexical items” (Hopper 1991:24). Divergence can result in multiple forms which have a common etymology, but diverge functionally. From this starting point, the recognition of homonymic but functionally-diverse lexemes, we realize that any study of Hindi light verbs in the context of grammaticalization (or lexicalization) needs a framework which allows for degrees of ‘less grammatical’ to ‘more grammatical’ to be determined in a rigorous and consistent fashion. One such approach within

the grammaticalization literature is the integrated model of Brinton and Traugott (2005). Since this model provides a framework for our analysis, we present a brief overview below.

## 1.2. An Integrated Approach to Grammaticalization and Lexicalization

It is important to consider what we mean by grammaticalization; the definition of grammaticalization we adhere to is presented below:

Grammaticalization is the change whereby in certain linguistic contexts speakers use part of a construction with a grammatical function. Over time the resulting grammatical item becomes more grammatical by acquiring more grammatical functions and expanding its host-classes (Brinton and Traugott 2005:99).

Studies on grammaticalization, word-formation and lexicalization have recognized much overlap between the two processes (Lehmann 2002, 2005; Brinton and Traugott 2005). Lehmann argues, “[w]e may say that grammaticalization pushes a sign into the grammar, while lexicalization pushes it into the lexicon” and goes on to say that “a conceptualization of the relationship of the two processes therefore presupposed an account of the relationship between lexicon and grammar” (2005:13). Brinton and Traugott (2005) present an integrated model of grammaticalization and lexicalization, which recognizes that the relationship between lexicon and grammar is more of a continuum than a strict dichotomy and state of their model:

It is not heavily theory-dependent but is dynamic, allows for constructions, gradience, and degrees of productivity...the nonproductive to productive continuum correlates in general with a continuum from lexical to grammatical items, from open to closed classes, from major to minor classes, from items that are maximally free (...) to maximally obligatory, and from contentful (conceptual, relatively concrete, usually referential) to functional (relatively abstract, often indexical) (2005:91-92).

The continuum on which grammaticalizing or lexicalizing forms are distributed is given in Table 1 (from Brinton and Traugott 2005:92):

*Table 1. Schema of correlations of categories along continua*

<b>Level</b>	<b>Continuum</b>		
Lexicon	Lexical		Grammatical
Category	Open/Major		Closed / Minor
Syntax	Free		Obligatory
Semantics	Contentful		Functional
Morphology	Nonproductive	Semi-productive	Productive

On a certain level of abstraction, Brinton and Traugott view both grammaticalization and lexicalization as processes by which a form takes on categorial properties consistent with either end of the continuum above; and they see the two processes as having a significant area of overlap synchronically in the semi-productive area (2005: 92). The concept of productivity in modern linguistics is generally vague, meaning different things to different linguists (for an excellent overview, see Bauer 2001), and Brinton and Traugott understand productivity as “the ability of word-forming elements to be used to form new linguistic expressions” and they view productivity as a gradient phenomenon (2005:92).

To summarize briefly, Brinton and Traugott see independent lexical items as entirely non-productive in that they are made up of morphemes that never combine, or combine very infrequently, such as the fossilized morpheme *-t* in *theft*, to create new words. In comparison, items on the grammatical end of the continuum are highly productive in that they combine by ‘default’ in the morphosyntax, e.g. obligatory inflections (cf Brinton and Traugott 2005: 92). The ‘semi-productive’ area of the continuum is where items combine “less or more regularly to recategorize the lexical items with which they combine (e.g. grammatical derivational forms like *-ness*, *-ly*) and those that combine with relatively few constraints to change the semantics of the lexical items with which they combine (e.g. semantic derivational *un-* in *unhappy* or *-ful* in *hopeful*)” (ibid.). This definition of productivity assumes that as items become increasingly lexicalized, their constituent morphemes become fused and are thus less free to be analyzed as individual morphemes. Likewise, as items become increasingly grammaticalized, the assumption is that the constituent morphemes will be used in a greater range of morphosyntactic constructions than before. The two processes are correlative developments along a single gradient continuum, represented in Table 2, taken from Brinton and Traugott (2005:94).

*Table 2. Synchronic clines of lexicality and grammaticality*



The three levels of grammaticality (e.g. G1, G2, G3) presented by Brinton and Traugott above are identified with respect to degrees of fusion with external elements, whereby,

G1 = periphrases, e.g. *be going to, as far as, in fact* (in their early stages)

G2 = semi-bound forms like function words and clitics, e.g. *must, of, 'll*

G3 = affixes that changes the grammatical class of the stem, e.g. adverbial – *wise* (fairly productive); most especially inflectional morphology (very productive; sometimes default), including zero inflection (2005: 93).

Brinton and Traugott stress that the three levels are not intended to be discrete, and there is obviously fine-grained syntactic and semantic gradience which is observed between these levels. However, the cline G1-G3 is determined by the degree of fusion with the external host: periphrases are relatively free positionally and affixes are always bound. Moving from the semi-productive area to the left in Table 2, they propose a cline of lexicality and postulate that “[s]ince lexical items are representations of major categories, and relatively free, they form a cline defined primarily with respect to degree of fusion in internal structure. Lexical items range from fully transparent to less transparent to idiosyncratic” (2005:94). Their cline of lexicality is as follows:

L1 = partially fixed phrases, e.g., *lose sight of, agree with*

L2 = complex semi-idiosyncratic forms, e.g., *unhappy, desktop*

L3 = simplexes and maximally unanalyzable idiosyncratic forms, e.g., *desk, over-the-hill*

Importantly, the Brinton-Traugott model outlined above also makes no claim that an item will necessarily move further along the continuum in either direction. In this sense, the continuum presented in Table 2 is inherently less teleological, or at least less prone to be interpreted as teleological in nature than the grammaticalization cline in (8) which predicts that

full verbs, if they do grammaticalize further at all, will become auxiliaries, and then eventually affixes (cf Hopper and Traugott 2003). Instead, the model allow us to observe that if an item does undergo grammaticalization, then the process can be viewed as including a greater degree of fusion with its external host, an increase in productivity, and a shift toward more functional semantics. Likewise, an item that begins to lexicalize should exhibit a lesser degree of internal structure (i.e. less distinct morpheme boundaries) and a semantic development of more idiosyncratic, contentful meaning, which may not be compositional. However, the model as we view it places no predictions as to which morphosyntactic category this grammaticalizing, or lexicalizing, form belongs to. That is to say, a light verb that grammaticalizes further does not necessarily join the class of auxiliary verbs; we restrict ourselves to claiming that the light verb is simply more grammaticalized as a light verb in the sense of greater fusion to the host, more productive in occurrence, with an increased shift to functional semantics.

The relationship between grammaticalization and lexicalization presented in the above model explicitly recognizes that the input for any process of grammaticalization and lexicalization is going to be a form or forms which can be viewed as situated at the semi-productive middle of a continuum between lexical and grammatical categories. This model is therefore ideal for discussing the Hindi light verbs whose “predicational power ... lies somewhere between that of a *lexical and a functional element*” (Butt 2003:11; emphasis added). Given the range of complex predicate constructions in Hindi, if the model proposed by Brinton and Traugott reflects linguistic reality, we should see certain light verbs which are grammaticalized on the one hand, and forming constructions which are (somewhat) lexicalized on the other hand. The remainder of this paper is devoted to testing this hypothesis, using data from the EMILLE corpus.

## 2. The EMILLE Corpus

In order to collect a representative sample of instances of Hindi aspectual light verbs, the present study used synchronic data from the EMILLE corpus, which is a major corpus of Hindi online. This corpus contains over 12 million words, comprising of articles from India’s national news websites, 20th c. literature, and miscellaneous documents written in Hindi.<sup>2</sup> We extracted a hundred sentences for each of the following twelve verbs that are commonly identified as ‘light’ (Abbi 1991; Butt 1994):

(11)	<i>aa</i> ‘come’	<i>jaa</i> ‘go’
	<i>baiTh</i> ‘sit’	<i>Daal</i> ‘put’
	<i>de</i> ‘give’	<i>le</i> ‘take’
	<i>paD</i> ‘fall’	<i>maar</i> ‘hit’
	<i>uTh</i> ‘rise’	<i>mar</i> ‘die’
	<i>nikal</i> ‘leave’	<i>nikaal</i> ‘remove’

Some note on the methodology is in order. As the focus of this paper is to investigate main verb - light verb combinations, any instances of light verb plus light verb complex predicates were excluded (cf. Mohanan 1994). Also excluded were data involving the verb *kar* ‘do/make’, which can occur both as an adverbial modifier or a light verb, depending on the context. The resulting sample is six times larger than the sample in Hook (1991), and represents light verb usage in a much wider cross-section of spoken and written discourse contexts. Our sample of twelve hundred sentences was coded for whether the particular verb in each sentence appeared as a light or a main verb. If the verb appeared in the corpus as light, a note was made of the main verb with which the light verb co-occurred. Next, we calculated the tokens (single instances) and types (set of tokens) frequencies of main verbs for each verb when it appeared as a light verb. The type frequency gives an indication of the range of verbs that the light verb selects for, while the token frequency shows which main verbs are common with a given light verb. Our results contain a total of seventy-two unique main verbs, of which several appear with multiple light verbs. It should be noted that this study does not claim to be fully representative of every possible main verb-light verb combination in the language and our analysis and discussion is restricted to the data available in our sample. The extracted data from the EMILLE corpus support the hypothesis that some light verbs are grammaticalizing, i.e. taking on specific grammatical functions in the language, and that others light verbs are lexicalizing, i.e. becoming idiomatic phrases or new lexical items. The type and token verbs for ten light verbs are presented in Table 3; the remaining two light verbs – *jaa* ‘go’ and *paD* ‘fall’ – are singled out in Table 4.



*Table 3. Frequency of light verbs in the sample data*

<b>LIGHT VERB</b>	<b>Tokens</b>	<b>Types</b>
<i>de</i> 'give'	27	19
<i>le</i> 'take'	16	15
<i>Daal</i> 'put'	35	11
<i>baiTh</i> 'sit'	12	9
<i>aa</i> 'come'	10	9
<i>uTh</i> 'rise'	6	5
<i>nikal</i> 'leave'	13	5
<i>nikaal</i> 'remove'	5	3
<i>maar</i> 'hit'	0	0
<i>mar</i> 'die'	0	0

*Table 4. Frequency of verbs *jaa* 'go' and *paD* 'fall' in the sample*

<b>VERB</b>	<i>jaa</i> 'go'		<i>paD</i> 'fall'	
	<b>Light verb</b>	<b>Passive</b>	<b>Light verb</b>	<b>Modal marker</b>
<b>Tokens</b>	28	69	2	44
<b>Types</b>	10	19	2	24

The light verbs in Tables 3 and 4 are all known to form complex predicate constructions in Hindi; patterns of grammaticalization for some of these light verbs are discussed in the order of their token frequency in Section 2.2 and 2.3, and evidence of lexicalization involving some of these light verbs is discussed in Section 2.4. The two verbs presented in Table 4, *jaa* 'go' and *paD* 'fall', seem to exhibit comparatively advanced degrees of grammaticalization, as in certain instances *jaa* 'go' can appear as a light verb and in others as a passive auxiliary; while *paD* 'fall' can likewise appear either as a light verb or a modal auxiliary. The case of these two light verbs is discussed below in Section 2.1.

### 2.1. Grammaticalization and the Light Verbs *jaa* 'go' and *paD* 'fall'

Some general evidence for grammaticalization of all main verb-light verb combinations has already been presented in Section 1, e.g. marking perfectivity, maintaining a rigid word order, etc. The verbs *jaa* 'go' and *paD* 'fall' exhibit all these types of indicators, suggesting that these verbs have undergone grammaticalization to a degree more than most verbs that have a form-identical light verb. The different usage of these two verbs as full

lexical items, light verbs, and auxiliaries is a classical example of divergence (Hopper 1991:25). We will show that in the course of their development, these two verbs seem to have diverged along (at least) two pathways of grammaticalization, one as a light verb which enters into complex predicates and therein develops the grammatical function of marking perfectivity, and another where *jaa* ‘go’ or *paD* ‘fall’ develops into an auxiliary. In the latter case, the verbs acquired the grammatical function of passivization or modality, respectively.

The verb *jaa* ‘go’ can function as a main verb, a light verb, or as a passive-marking auxiliary. The light verb form of this intransitive verb functions as a perfective marker and denotes completion of the event (cf. Butt 1994:148). An example in Hook (1993) taken from an excerpt of Ramayana that was composed “late in the first millennium of the present era” (c. 900-1000 AD) suggests that the verb *jaa* ‘go’ has maintained its light verb status for over 1000 years (from Hook 1993: 98):

- (12) chuDu ekku gaasu mahu ho-u<sub>MV</sub> ga-u<sub>LV</sub>  
 still one bite me-DAT became went  
 ‘Yet (a whole elephant) is (just) one bite for me.’

In our contemporary data, it is clear that the light verb in this construction is used as a perfective marker. Contrast the complex predicate in (13) from the EMILLE corpus with the main verb construction in (14):

- (13) bhavishyaa anischiit ho<sub>MV</sub> ga-yaa<sub>LV</sub> hai<sub>AUX</sub>  
 future.M uncertain be go-M.SG be-PRES.3.SG  
 ‘The future has become uncertain.’ (ID: 720)

- (14) bhavishyaa anischiit hu-aa<sub>MV</sub> hai<sub>AUX</sub>  
 future.M uncertain be-M.SG be-PRES.3.SG  
 ‘The future is uncertain.’

The event in (13) appears incomplete without the light verb *jaa* ‘go’ as shown in (14); *jaa* adds perfectivity to the verb *ho* ‘be’. However, the main verb-light verb combination has not developed an idiosyncratic meaning; *jaa* ‘go’ in this construction is a light verb and the two verbs do not form a lexical item.

As a passive auxiliary, *jaa* ‘go’ can appear with any transitive verb in the language which can be passivized. There is a difference in the morpho-syntactic form of the passive marker as opposed to the light verb. As men-

tioned in Section 1.1, the main verb in most complex predicate constructions is non-finite and the light verb is marked for inflection. This is shown in example (15) with the light verb *jaa* ‘go’.

- (15) Shyam=kaa haat kaT<sub>MV</sub> ga-yaa<sub>LV</sub>  
 Shyam=GEN hand.M. SG cut go-M.SG  
 ‘Shyam’s hand was cut.’

Observe that the semantics of the light verb *jaa* ‘go’ in (15) has been completely bleached, i.e. there is no motion component in the sentence. In contrast to (15), when *jaa* is used as a passive marker, the inflection must be carried by the main verb also, as shown in (16).

- (16) Shyam=kaa haat kaT-aa<sub>MV</sub> ga-yaa<sub>AUX</sub>  
 Shyam=GEN hand.M cut-.M.SG go-M.SG  
 ‘Shyam’s hand was cut (by a covert agent).’

The inflection in (16) reveals that the verb *jaa* ‘go’ is a passive auxiliary. Beyond the morphological difference, native speakers understand there to be suggestion of a covert agent semantically in (16), which is not present in (15). Predictably, *jaa* ‘go’ appears most frequently in our dataset as a passive auxiliary (n=69), less often as a light verb (n=28), and seldom as a main verb (n=3). Interestingly, the light verb ‘*jaa*’ appears in our sample with the main verb *ho* ‘become’ in 13 instances, i.e. nearly 50% of the time. Within Brinton and Traugott’s model, it could be argued that *jaa* as a passive auxiliary could be assigned a very advanced stage of grammaticalization, say G3; whereas *jaa* as a light verb is only partially grammaticalized in the environment of complex predicates, approximating a G2 stage. This model refers to degrees of grammaticalization based on particular observed properties but does not claim that the light verb usage of *jaa* at the G2 stage will pass on to a G3 stage as an auxiliary. Rather, a light verb at a G2 stage of grammaticalization can become further grammaticalized to G3 *as a light verb*, i.e. further fusion to the host, increased loss of transparency, etc. One can argue that there are different paths of development: one for full verbs to auxiliaries to affixes (G1 > G2 > G3); one for full verbs to light verbs to further grammaticalized light verbs (G1 > G2 > G3).<sup>3</sup>

One could also say that the full verb *jaa* ‘go’ is at the L3 position, as it is a lexical item, but to be clear, it should be stressed that *jaa* has always been a full lexical item in Hindi, i.e. it did not undergo stages of lexicalization (e.g. L1 > L2 > L3), as its position at L3 might imply.

A similar pattern of grammaticalization can be viewed for the verb *paD* ‘fall’. We see only two instances of true light verb usages in our data, noted in (17) and (18) below, marked by the fact that the main verb is infinitive and the light verb is inflected.

- (17) divaar            chal<sub>MV</sub>            paD-ii<sub>LV</sub>  
 wall.F.SG        walk                fall-M.SG  
 ‘The wall began to walk.’ (ID: 280)

- (18) galba (...)        gir<sub>MV</sub>            paD-aa<sub>LV</sub>  
 galba                fall                fall-M.SG  
 ‘Galba fell down.’ (ID:59)

In both examples, the light verb *paD* ‘fall’ adds a sense of completion and does not refer to any event of falling. That is, in (17), the light verb indicates only that the act of walking has been initiated. Moreover, in (18), a different lexeme *gir* ‘fall’ is used to express the action of falling and the light verb *paD* ‘fall’ contributes a sense of suddenness. The verb *gir* denotes the act of falling down on the same spot, whereas *paD* as a main verb has no such restriction.

Elsewhere in our sample, *paD* ‘fall’ is used primarily as a modal auxiliary. Similar to the case of the passive auxiliary *jaa* ‘go’, where both the main verb and the passive are inflected, the modal form of *paD* ‘fall’ requires the preceding infinitive to be inflected. When the verb *paD* is a modal marker, what precedes it surfaces in the so-called ‘*to*-infinitive’ (stem followed by the suffix ‘-*naa/nii*’ for masculine or feminine gender respectively). Hook suggests that in such cases *paD* ‘fall’ always expresses “completion (...) of the action expressed by the verbs it governs” (1974: 303).<sup>4</sup> In such constructions, the verb *paD* always adds the deontic meaning of the action having to be done without any choice; as in (19) and (20).

- (19) galba=ko        vaapas            lout-naa<sub>MV</sub>            paD-aa<sub>DEO</sub>  
 galba=DAT        back                return-INF.M.SG        fall-M.SG  
 ‘Galba had to return back.’ (ID:879)

- (20) unhe apne vivahit jeevan=ko goodbye kah-naa<sub>MV</sub>        paD-aa<sub>DEO</sub>  
 3.PL self married life=DAT goodbye say-INF.M.SG fall-M.SG  
 ‘They had to say goodbye to their married life.’ (ID:417)

In the examples (19) and (20), the lexical semantics of ‘to fall’ are completely bleached, and the deontic modal meaning, i.e. the grammatical meaning, has been correspondingly strengthened. Moreover, *paD* ‘fall’ can appear with a greatly expanded host-class as a deontic modal. In our dataset, the verb *paD* ‘fall’ has a higher token frequency as a modal (n = 30) than as a light verb (n = 2). As in the case of *jaa* ‘go’, there has been divergence indicative of multiple pathways of grammaticalization. As a light verb, *paD* ‘fall’ may be at a G1/G2 stage – perhaps on analogy with other main verb-light verb combinations – however, as a deontic modal auxiliary, *paD* ‘fall’ is more likely analyzed as being at G3. Together the verbs *jaa* ‘go’ and *paD* ‘fall’ thus show indications that synchronically they function not only as light verbs, but also as auxiliaries. In comparison to their usage as a light verb, when used as a passive auxiliary (*jaa*) or as a deontic modal auxiliary (*paD*), both exhibit a greatly expanded host-class; are completely bleached semantically; and have the highest type and token frequency, indicating productivity which is more consistent with a grammatical formative. Brinton and Traugott suggest, “items that grammaticalize are used in more contexts and for a larger set of lexical items” (2005: 109). This statement reflects the passive and the modal marking usages of *jaa* ‘go’ and *paD* ‘fall’. Charting the degree of grammaticalization of these verbs along the continuum discussed in Section 1.2, we arrive at Figure 1.

As illustrated in Figure 1, the verbs *jaa* ‘go’ and *paD* ‘fall’ thus have three different entries and each of these falls on a different position on the continuum. The main verb entries are part of the lexicon (L3) whereas the light verb entries fall between the semi-productive and productive (G1-G2) region of the gradient continuum. According to our data, the light verb *jaa* ‘go’ is more productive (n=28) than the light verb *paD* ‘fall’ (n=2) and therefore has been placed further on the G2 position. The auxiliary forms of these verbs are the most productive and are fully grammaticalized, i.e. they are part of the grammar of Hindi.

In the next section, we turn to light verbs that are also moving towards the G2 position, but show different types of development than the light verbs *jaa* ‘go’ and *paD* ‘fall’.

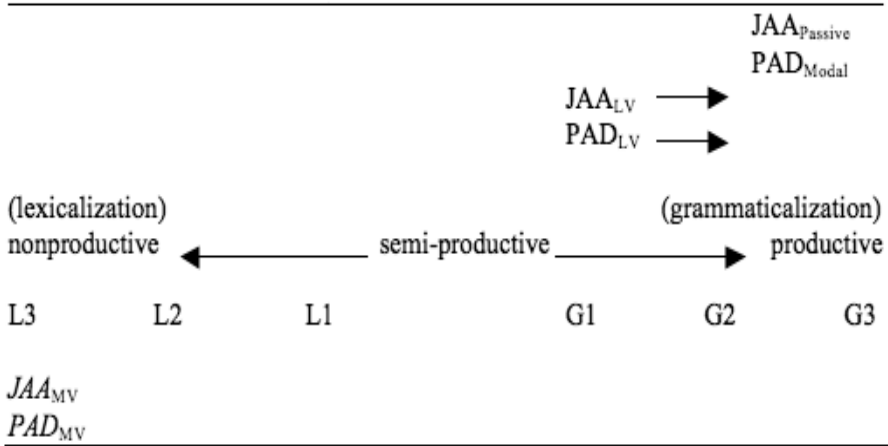


Figure 1. Grammaticalized Light Verbs on the Continuum

## 2.2. G2 position: Light verbs *de* ‘give’ and *le* ‘take’

The verbs *de* ‘give’ and *le* ‘take’ have the next highest type frequency as light verbs (n=19 and n=15, respectively). The two verbs seem to be forming a paradigmatic opposition: *de* ‘give’ is coming to be the default means of marking an action which is done for the benefaction for others; and *le* ‘take’ is coming to mark benefaction for oneself. Our data shows that these two light verbs select for various types of main verbs, i.e. they are highly productive in terms of their host class. However, *de* ‘give’ and *le* ‘take’ retain more of their lexical semantics than either *jaa* ‘go’ or *paD* ‘fall’; their semantics of benefaction restrict them from selecting for all types of verbs. For instance, as a light verb *de* ‘give’ does not appear with verbs where the action is only done for oneself. This is illustrated below where the light verb *de* ‘give’ cannot select for the main verb *cheen* ‘snatch’ in (21), whereas the light verb *le* ‘take’ can form a complex predicate (22).

- (21) \*munda=ne congress=se muddaa bhii cheen<sub>MV</sub> di-yaa<sub>LV</sub>  
 \*munda=ERG Congress=INST issue also snatch give-M.SG  
*Intended*: ‘Munda stole even an issue from Congress (for someone else).’
- (22) munda=ne congress=se muddaa bhi cheen<sub>MV</sub> li-yaa<sub>LV</sub>  
 munda=ERG Congress=INST issue also snatch take-M.SG  
 ‘Munda stole even an issue from Congress (for himself).’ (ID: 288)

The act of stealing cannot be understood to benefit someone else in (21) where *de* ‘give’ is the light verb, because the lexical semantics of ‘giving’ conflicts with the sense of ‘taking for oneself’ denoted by *cheen* ‘snatch’. With the exception of this example in the EMILLE corpus, both light verbs *de* ‘give’ and *le* ‘take’ can occur with every main verb which can occur with the other. However, *de* ‘give’ is coming to be the default means of marking an action done for the benefaction for others; and *le* ‘take’ is coming to mark benefaction for oneself. Some process of grammaticalization is underway because the functional semantics of benefaction has been enhanced at the expense of more concrete semantics of ‘giving’ and ‘taking’ encoded by each verb. Although this division of the semantics does place some observable restrictions on the kind of main verbs that both light verbs can select for – as in *cheen* ‘snatch’ with *le* ‘take’ in (21) above – in general, *de* ‘give’ and *le* ‘take’ are more productive than any of the other Hindi light verbs. As seen from Tables 3 and 4, these verbs have a higher type and token frequency as light verbs than *jaa* ‘go’ and *paD* ‘fall’, however, they are less semantically bleached since the ‘other/self’ dichotomy of benefaction meanings is conditioned by the main verb semantics of *de* ‘give’ and *le* ‘take’. In the Brinton-Traugott model, these light verbs are therefore characteristic of the G2 position as semi-bound forms (fixed elements within a complex predicate construction) with partially functional meaning (such as benefaction).

### 2.3. G1 Position: Other Light Verbs

We now describe light verbs with lower token frequencies. Of these the light verb *aa* ‘come’ does not present a clear pattern except that it mostly selects for verbs of motion in our sample, e.g. *ghus* ‘enter’, *nikal* ‘leave’. An example is shown in (22).

- (22) *vimaan barathiyaa seema*=mein *ghus*<sub>MV</sub> *aa-yaa*<sub>LV</sub> *th-aa*<sub>AUX</sub>  
 airplane indian border=LOC enter come-M.SG be-PAST.M.SG  
 ‘(An) airplane entered the Indian border.’ (ID: 404)

Aside from *aa* ‘come’, the remaining light verbs restrict themselves to fewer semantic types. Indeed, the fuller lexical meaning of each light verb seems to constrain the types of main verb it can select for. The most straightforward example of this is the light verb *Daal* ‘put’, which, despite a high token frequency, has a low type frequency. This construction also appears frequently with the verb *maar* ‘hit/kill’ to form a lexical item,

which will be discussed in Section 2.4. Typical verbs that combine with the light verb *Daal* ‘put’ are verbs of creation or destruction e.g., *ban* ‘make’, *badal* ‘change’, *toD* ‘break’. Representative examples are shown below in (23) and (24), where the light verb also contributes a nuance of unexpectedness. Moreover, *Daal* ‘put’ also occurs with verbs of communication such as *kah* ‘say’ and *puuch* ‘ask’.

- (23) sevakar=ne film udhyoog=ki kamaar hii toD<sub>MV</sub>  
 sevakar=ERG film industry=GEN back even break  
 Daal -ii<sub>LV</sub> hai<sub>AUX</sub>  
 put-F.SG be-PRES.3.SG  
 ‘Sevakar has broken the film industry’s back itself (surprisingly).’  
 (ID: 1161)

- (24) mao-tse-tung=ne samaaj bhii badal<sub>MV</sub> Daal -aa<sub>LV</sub>  
 mao-tse-tung=ERG society also change put-M.SG  
 ‘Mao Tse Tung also changed the society (surprisingly).’ (ID: 1022)

Some light verbs are very constrained in the semantic types of main verbs they select for. A clear pattern can be seen with the light verb *uTh* ‘rise’ which only appears with verbs requiring sound emission e.g., *guunj* ‘roar’, *bol* ‘say’, etc. This is shown in (25).

- (25) george=ke istifee=kii maang<sub>MV</sub> uTh-ii<sub>LV</sub> th-ii<sub>AUX</sub>  
 george=GEN resignation=GEN demand rise-F.SG be-PAST.F.SG  
 ‘George’s resignation was demanded (*lit.* a demand rose)’ (ID: 817)

Another constrained light verb is *nikal* ‘leave’ which in our data selects only for main verbs that involve motion e.g., *chal* ‘walk’ as in (26).

- (26) gandhi=kaa dhandhaa chal<sub>MV</sub> niklaa<sub>LV</sub>  
 gandhi=GEN business walk leave-M.SG  
 ‘Gandhi’s business took off (*lit.* walked).’ (ID: 272)

Furthermore, in cases where the main verb does not entail motion, the entire construction with *nikal* ‘leave’ as a light verb forms an idiosyncratic meaning which does entail motion. Such constructions are discussed in Section 2.4. The light verb *nikaal* ‘remove’ (n=5), which only has constructions with an idiosyncratic meaning, will also be discussed there. Most light verbs thus seem to have stronger restrictions on the kind of verbs that they



will appear with and this decision seems to be conditioned by their full verb semantics. We were unable to make any claims about two light verbs in Table 3, *maar* ‘hit’ and *mar* ‘die’, as they have no instances as a complex predicate construction in their representative hundred sentences. Nonetheless, these verbs have been often cited as occurring in complex predicate constructions in the literature (e.g. Hook 1974, Hook 1993).

So far, we have discussed structures where the light verb usually follows the main verb. However, Hindi complex predicates also involve another structure where the light verb precedes the main verb (i.e., a *reverse* complex predicate construction). The reverse construction is fairly restricted with respect to the range of main verbs and light verbs that can appear in them, as well as in terms of the auxiliaries that can take this construction as an argument (cf. Poornima and Koenig 2008). Within our dataset, we find examples of reverse construction with the light verb *maar* ‘hit’ as in (27).

- (27) us=ne    apnaa sir    divaar=par    de<sub>LV</sub>    maar-aa<sub>MV</sub>  
 3.SG=ERG self    head    wall=LOC    give    hit-M.SG  
 ‘He hit his head on the wall.’ (ID:1023)

Recall that in a reverse complex predicate construction, the order of the main verb and the light verb is switched i.e., the light verb, now non-finite, precedes the finite main verb. As the translation in (27) indicates, *de* ‘give’ cannot be the main verb and is instead a light verb. Such constructions are fairly infrequent; out of every instance in the EMILLE corpus with *maar* ‘hit’ as a finite masculine verb (n=1041), only thirteen were complex predicate constructions; six of which were the reverse predicate construction. The reverse complex predicate construction is fairly restricted with respect to the range of main verbs and light verbs that can appear in them as well as with respect to the auxiliaries that can enter take this construction as an argument (cf. Poornima and Koenig 2008). The presence of a reverse complex predicate construction further suggests a high degree of fusion between the light verb and its main verb host, but how this construction emerged diachronically is still unclear. Therefore, in terms of an analysis within the Brinton-Traugott model, we surmise that these two light verbs, in addition to the rest of the light verbs discussed in this section, are still at the early stages of grammaticalization and fall at the G1 stage of the continuum. Positing this stage of grammaticalization synchronically is warranted due to the partial bleaching of full verb semantics when a verb is used as a light verb.

## 2.4. L1 Position: Light Verbs and Lexicalization

Because grammaticalization and lexicalization are seen as complementary processes in the integrated model (Brinton and Traugott 2005), one might expect instances of both lexicalization and grammaticalization to arise from similar input constructions. In fact, there is good evidence in the extracted EMILLE corpus data that in some cases, certain main verb-light verb combinations are undergoing lexicalization. For example, as indicated in Section 2.3, the light verb *Daal* ‘put’ is relatively frequent in our dataset (n=35), but it does not co-occur with as many different main verbs as some light verbs such as *de* ‘give’. In fact, *Daal* ‘put’ occurs 22 times with the main verb *maar* ‘hit’; the combination of the two verbs seem to be lexicalizing. As a main verb, *maar* can mean either ‘hit’ (28) or ‘kill’ (29).

(28) tab laam=ne us=ke baayee pããv mein baan maar-aa<sub>MV</sub>  
 then lam=ERG his=GEN left leg in arrow hit-M.SG  
 ‘Then Lam hit an arrow on his left leg.’ (ID:226)

(29) ek bacchaa maar-aa<sub>MV</sub> ga-yaa<sub>AUX</sub>  
 one child hit-M.SG go-M.SG  
 ‘One child was killed.’ (ID: 283)

In contrast to the above examples, when the main verb *maar* ‘hit/kill’ appears with the light verb *Daal* (put), the act of killing is understood as being volitional, and the resulting complex predicate has the somewhat obscure and idiosyncratic meaning, ‘to kill volitionally/murder’ as in (30).

(30) ugravaadiyoon=ko graaminoo=ne maar<sub>MV</sub> Daal -aa<sub>LV</sub> th-aa<sub>AUX</sub>  
 terrorists=DAT villagers=ERG hit put-M.SG be-PAST.M.SG  
 ‘The villagers murdered the terrorists.’ (ID: 103)

That this meaning is not obviously derived from the semantics of the light verb *Daal* ‘put’ is clear from (31) which illustrates its main verb usage.

(31) is udyoog=ko sab=se jyaadaa musiibat mein  
 this business=DAT all=INST more trouble in  
 Daal -aa<sub>MV</sub> hai<sub>AUX</sub>  
 put-M.SG be-PRES  
 ‘(he has) put this business in the most trouble.’ (ID: 127)

It seems that the collocation of the main verb *maar* ‘hit/kill’ and the light verb *Daal* ‘put’ is becoming lexicalized inasmuch as the string has become fixed (in the sense that the main verb and light verb are both obligatory in the complex predicate when used with this ‘murder’ meaning) and the meaning itself is somewhat idiomatic and not obviously derivable from the semantic components of either ‘put’ or ‘hit’, although a forceful volitional hitting-event may lead to murder. Brinton and Traugott’s conceptualization of lexicalization concerns both “adoption into the inventory” and “modification of the inventory” (2005:96), and in this case, the learner of Hindi learns that this combination has the idiomatic meaning of ‘murder’. In terms of the cline of lexicality, this construction is a fixed phrase, e.g. L1, such as *lose sight of* in English. The light verb *Daal* ‘put’ still inflects for gender and number, while the main verb *maar* ‘hit/kill’ is fixed, appearing only in the infinitive form. This is comparable to the inflection in English *has lost* {sight of}; *losing* {sight of}) discussed in Section 1.2.

Similar cases of main verb-light verb combinations which can be seen as emerging fixed (L1) expressions with light verbs which otherwise have a quite low token frequency. Examples from our dataset are as follows:

- (32) *bhaag* ‘run’ + *nikal* ‘leave’ = ‘escape’  
 vahaa=se      bhaag<sub>MV</sub>      nikl-aa<sub>LV</sub>  
 there=INST      run      leave-M.SG  
 ‘(He) escaped from there.’ (*lit.* ran-leave) (ID:1501)
- (33) *khoj/DhoonD* ‘search’ + *nikaal* ‘remove’ = ‘find’  
 ek      chasmaa      dhoond<sub>MV</sub>      nikaal-aa<sub>LV</sub>  
 one      spectacles      search      remove-M.SG  
 ‘(He) found spectacles.’ (*lit.* search-remove) (ID:166)
- (34) *nichoD* ‘squeeze’ + *nikaal* ‘remove’ = ‘extract’  
 manushyaa=ne jivaan=kaa arth      nichoD<sub>MV</sub>      nikaal-aa<sub>LV</sub>  
 hai<sub>AUX</sub>  
 man=ERG      life=GEN      meaning      squeeze      remove-M.SG  
 be-PAST.M.SG  
 ‘Man has extracted out the meaning out of life.’ (*lit.* squeeze-leave). (ID: 74)

Interestingly, the light verbs *nikal* ‘leave’ in (32) and *nikaal* ‘remove’ in (33)-(34) combine with very few other main verbs in the extracted data. A search through the entire EMILLE corpus of the above combinations, with

the same inflection, revealed only four additional instances of the construction in (32), twelve of (33) and just the one of example (34). But an increased token frequency is not expected for lexicalizing items anyway, since they do not involve either an “expansion of hosts” or a “generalization across context” (Brinton and Traugott 2005: 109). Moreover, note that the semantics of a light verb such as *nikaal* ‘remove’ has been enriched; contentful, specific meaning of ‘to extract’ has been added to the original lexical semantics of ‘to remove’: this is good evidence for lexicalization.

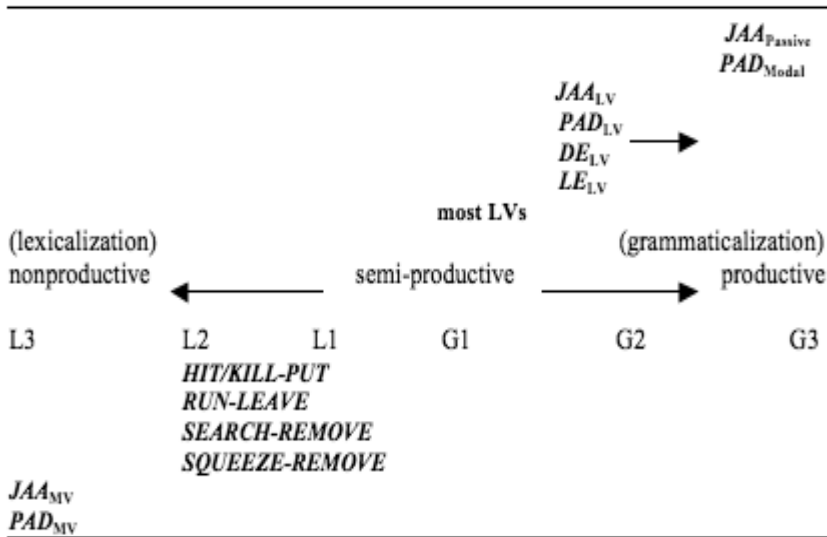


Figure 2. Hindi Light Verbs on the Lexical-Grammatical Continuum

### 3. Discussion

The integrated approach of Brinton and Traugott allows us to view certain main verb-light verb combinations as synchronically more grammaticalized or lexicalized, without making any sort of claim that light verbs, as a class, are advancing toward the status of auxiliaries. An abstract mapping of the light verbs to the lexical-grammatical continuum is illustrated in Figure 2. As the figure suggests, instead of analyzing light verbs as an intermediate stage with indeterminate properties, research in grammaticalization ought to view the relative stages of different Hindi verbs functionally as either less grammatical (marking perfectivity) or more grammatical (e.g. marking benefaction), and investigate how this distinct pathway of development

works. The case of *jaa* 'go' and *paD* 'fall' shows that certain verbs in the Hindi verbal system exhibit diverging pathways of grammaticalization: for instance, under one set of developments *jaa* has become a light verb, and under another a passive auxiliary. Importantly, the passive auxiliary did not develop from the light verb and there is no need to make reference to a standard grammaticalization cline.

By examining twelve light verbs in a random sample extracted from the EMILLE corpus, a clearer picture emerges of this area of the Hindi verbal system. While some of the results reported here corroborate findings of early work such as Hook (1973, 1991), such as the tendency of *aa* 'come' to collocate with verbs of motion or *uTh* 'rise' to occur with verbs of sound emission, other patterns emerge in this dataset which warrant further explanation. For example, more work needs to be focused on the apparent paradigmatic opposition which seems to condition the grammatical development in the light verbs *le* 'give' and *de* 'take'. In this case, it is intriguing that the original lexical semantics of the light verbs seems to play a role as these constructions take on a more functional meaning of benefaction to others (*le*) and benefaction to oneself (*de*). It is also interesting to note that light verbs whose full verb semantics are more general, e.g. *le* 'give', *de* 'take', have a higher type and token frequency than the light verbs whose full verb semantics are more specific, e.g. *nikal* 'leave', *nikaal* 'remove'. One interpretation is that the more general semantically general or bleached a light verb's lexical semantics are, the more frequent its usage, i.e. the more prone it is to grammaticalize in the first place. Another interpretation, however, is that light verbs with more specific semantics enter into complex predicates on analogy with a core set of grammaticalized light verbs. These analogically-formed constructions with light verbs that have retained more specific lexical semantics may then take on idiosyncratic meanings in context; that is, they may be prone to lexicalize. This might explain the high frequency of cooccurrence of certain light verbs with particular main verbs such as *Daal* 'put' and *maar* 'hit' (in the idiomatic sense of 'murder') and *nikaal* 'remove' and *khøj/DhoonD* 'search' (in the idiomatic sense of 'find'). The data collected in the current work suggests that once a light verb grammaticalizes in Hindi, some collocations of main verb and light verb continued to grammaticalize, whereas others began to lexicalize. Further research is needed to explore on-going developments in Hindi complex predicates both from a synchronic and diachronic perspective.

## Acknowledgements

We would like to thank Jeff Good and David Fertig for looking over earlier drafts of this paper. All errors remain ours.

## Notes

1. Abbreviations are as follows: MV = main verb, LV = light verb, F = feminine, M = masculine; ERG = ergative, NOM = nominative, GEN = genitive, DAT = dative, INST = instrumental, LOC = locative; INF = infinitive; PRES = present; PRON = pronoun. The marker ‘-’ indicates a morpheme boundary, ‘=’ separates a clitic from a lexical item.
2. Available at: [www.ling.lancs.ac.uk/corplang/emille/](http://www.ling.lancs.ac.uk/corplang/emille/)
3. Various researchers have argued that the infinitive is actually not a verb but either an infinitival complement (Kachru 1980:40) or a verbal noun (Mohan 1990:99, Butt 1994:56-66), which is irrelevant to our present discussion.
4. We would like to thank an anonymous reviewer for suggesting we make the discussion in this section clearer.

## References

- Abbi, Anvita and Devi Gopalakrishnan  
 1991 “Semantics of Explicator Compound Verbs in South Asian Languages”. *Language Sciences* 13(2). 161–180.
- Bussmann, Hadumod  
 1996 *Routledge Dictionary of Language and Linguistics*. Gregory Trauth and Kerstin Kazzazzi (eds.) London and New York: Routledge.
- Bauer, Laurie.  
 2001 *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bowern, Claire  
 2008 “The Diachrony of Complex Predicates”. *Diachronica* 25(2). 161–185.
- Brinton, Laurel J. and Elizabeth C. Traugott  
 2005 *Lexicalization and Language Change*. Research Surveys in Linguistics. Cambridge: Cambridge University Press.
- Butt, Miriam and Tracy King  
 2005 The Status of Case. In *Clause Structure in South Asian Languages*. Veneeta Dayal and Anoop Mahajan (eds.) Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer. 153-198.

- Butt, Miriam  
 1994 *The Structure of Complex Predicates in Urdu*. Ph.D. thesis, Stanford University.
- Butt, Miriam  
 2005 “Light Verb Jungle”. *Harvard Working Papers in Linguistics*, Volume 19. 1–49.
- Butt, Miriam and Wilhelm Geuder  
 2001 On the (semi)lexical status of light verbs. In Norbert Corver and Henk van Riemsdijk, ed. *Semi-lexical categories : the function of content words and the content of function words*, 323–369. Amsterdam: Mouton.
- Butt, Miriam and Aditi Lahiri  
 2002 “Historical Stability vs. Historical Change”. Available at: <http://ling.uni-konstanz.de/pages/home/butt/stability.pdf>.
- Hopper, Paul  
 1991 “On Some Principles of Grammaticalization”. In E.C. Traugott and B. Heine, ed. *Approaches to Grammaticalization: Volume I*, 17-36. Amsterdam: John Benjamins.
- Hopper, Paul and Elizabeth C. Traugott  
 2003 *Grammaticalization – 2<sup>nd</sup> Edition*. Cambridge: Cambridge University Press.
- Hook, Peter E.  
 1974 *Compound Verbs in Hindi*. Unpublished Ph. D. thesis, University of Michigan, Ann Arbor.
- Hook, Peter  
 1991 “The Emergence of Perfective Aspect in Indo-Aryan Languages”. In E.C. Traugott and B Heine, eds. *Approaches to Grammaticalization: Volume II*, 59-89. Amsterdam: John Benjamins.
- Hook, Peter  
 1993 “Aspectogenesis and the compound verb in Indo-Aryan.” In M. K. Verma ed. *Complex Predicates in South Asian Languages*. New Delhi: Manohar Publishers and Distributors.
- Hook, Peter and Prashant Pardeshi  
 2006 “Are Vector Verbs Eternal?” Eleventh Biennial Symposium: Inter-theoretical Approaches to Complex Verb Constructions. Rice University, Texas.
- Kachru, Yamuna  
 1980 *Aspects of Hindi Grammar*. New Delhi: Manohar Publishers and Distributors.
- Lehmann, Christian  
 2002 “New Reflections on Grammaticalization and Lexicalization” In I. Wischer and G. Diewald, eds. *New Reflections on Grammaticalization*, pp 1-18. Amsterdam: John Benjamins.

Lehmann, Christian

2005 "Theory and Method in Grammaticalization". *Zeitschrift für Germanistische Linguistik*, 32 (2) 152-187.

Norde, Muriel, Olga Fischer, and Harry Perridon

2004 *Up and Down the Cline: the Nature of Grammaticalization*. Amsterdam: John Benjamins.

Poornima, Shakthi and Jean-Pierre Koenig

2008 "Reverse Complex Predicates in Hindi". In *Proceedings of the 24<sup>th</sup> Northwest Linguistic Conference*, Univ. of Washington, Seattle, WA.

Singh, Mona

1998 "On the Semantics of the Perfective Aspect." *Natural Language Semantics*, pp 171-199. Springer: Netherlands.

Tikannen, Bertil

1987 *The Sanskrit Gerund: A Synchronic, Diachronic, and Typological Analysis*. Studia Orientalia, vol.62. Helsinki.





**Special Contribution: *Hindi Pragmatics***



# On Hindi Conditionals

*Ghanshyam Sharma*

*In logic it is commonly believed that a conditional sentence is an aggregate of two propositions, namely the antecedent  $p$  and the consequent  $q$ , combined together exactly in the same manner as in the case of conjunction and disjunction. From our analysis of Hindi data, it emerges that the antecedent, or protasis, is not a real proposition, but rather a propositional function and thus carries a speaker's split modal meaning. In addition, this paper claims that the notion of 'conditional clause inversion' should be looked at from the pragmatic point of view. Hindi data suggest that it is not the protasis which moves rightward as a result of speaker's 'after-thought', but rather the proposition employed in apodosis which is dislocated leftward for different pragmatic reasons, leaving the apodosis marker in its canonical place. If this hypothesis proves to be true in other languages as well, then it can be safely concluded that the clause order universal ( $p \rightarrow q$ ) remains unaltered in all languages. Such a hypothesis may also controvert the thesis according to which protases are adverbial clauses and may serve to advance the idea that conditionals are one of many other topics where pragmatic intrusion into syntax can be easily attested.*

## 1. Introduction

The present study primarily aims at classifying different varieties of Hindi conditionals, looking at underlying principles which are responsible for the selection of a particular tense-aspect-mood element (TAM hereafter) rather than others for employment in one of the two propositions that make up a conditional statement. The paper takes a pragmatic approach to conditionals to characterize an arranged distribution of TAM elements among different classes of Hindi conditionals. As a consequence, it does not aim at presenting any truth-functional account of Hindi conditionals, nor does it intend to dredge up any older versions of philosophical theories of conditionals in general, in order to initiate a theoretical debate. However, any discussion on the pragmatic principles underlying a TAM element selection

in two propositions in a conditional statement will inevitably be drawn into, or in passing refer to, some earlier philosophical discussion as well. Where necessary, we will not desist from making passing remarks about syntactic analyses in the field, although this is not the goal of the paper. The paper asserts that various linguistic theories of conditionals which strive to analyze and establish the distribution of TAM elements in the protasis<sup>1</sup> (i.e. the ‘*if*-clause’ or the ‘antecedent’) and the apodosis (i.e. the ‘*then*-clause’ or the ‘consequent’) are inadequate in that, contrary to widely held belief in most of the philosophical discussion on conditionals, the two propositions which make up a conditional are not of the same semantical value. As will become clear from the discussion in the following sections, the proposition employed in protasis is essentially a propositional function (and thus with a speaker’s split modal meaning!) rather than a real proposition, and is thus prone to take different values, some of which are true and some false. As a consequence, while it is useful to look at both the protasis and the apodosis from the point of view of ‘time of reference’ —as it helps to establish the time to which a conditional statement as a whole pertains or refers— it is nevertheless not a reliable method considering them exclusively in terms of tenses marked within them. The term *if* employed to introduce a protasis into a conditional construction divides the whole semantical world into two parts: one in which the state of affairs described in protasis is to be true, and the second in which it is to be false. The speaker does not attach a modal meaning to the protasis in the same way as he does to the apodosis. In making a conditional statement, he seems to adhere to an inherent logical disjunction in the protasis. For example, the circumstances in which an indicative conditional ‘ $P \rightarrow Q$ ’<sup>2</sup> (e.g. ‘If today is Sunday, the priest will be in church’)<sup>3</sup> could be asserted are the following: ‘Either  $\neg P$  or  $P$  and  $Q$ ’ (i.e., Either today is not Sunday or, it is Sunday and the priest is in church).<sup>4</sup> Therefore, it is inappropriate to consider a hidden subordinate indicative proposition encountered in a protasis such as ‘If John lives in Venice...’ equal to an independent indicative proposition such as ‘John lives in Venice’. In the first case, the speaker does not commit himself to the veracity of the state of affairs described therein, nor is he in a position to take any illocutionary stand thereon. That is so because the speaker neither knows nor believes whether or not John lives in Venice. He simply considers the truth-value of the proposition ‘John lives in Venice’ as a possible case for the apodosis to be true, and does not express his opinion about the case in which  $P$  is to be false. In the second case, on the other hand, for all the speaker knows, the proposition ‘John lives in Venice’ is to be necessarily true (or, to put it in a logical notation,  $K \Box P$ ). That

is to say, according to the speaker, John necessarily belongs to the class of those people who live in Venice. Failure to realize this fundamental difference between two propositions —i.e. ‘If John lives in Venice, ...’ and ‘John lives in Venice’— has led to some major misunderstandings in many spheres of both philosophy and linguistics. The paper argues that for a deeper understanding of conditional constructions it is crucial to maintain the fundamental distinction between the semantical values of protasis and apodosis and the modal meanings therein.

**2. Relationship between material implication and conditional**

Natural language conditionals have for millennia been viewed in relation to ‘material implication’<sup>5</sup> —the orthodox term in the field of logic and philosophy— and considered to have exactly the same logical structure of the latter. However, we believe that for different pragmatic reasons natural languages do not allow for certain inferences to take place in everyday communication even if they seem otherwise quite congruous in the case of ‘material implication’.<sup>6</sup> One of the most striking features of a conditional utterance (or statement) is that, although at first glance it appears to be consisting of two propositions, the first P and the second Q, it is not a semantical aggregate of the two —many suggestions to this effect from ancient philosophy notwithstanding. As has been widely reported, according to ancient Greek philosophy ‘implication’ is a relationship between two propositions, wholly analogous to ‘conjunction’ and ‘disjunction’, as can be demonstrated in table 1.

*Table 1.* The following truth-table shows the way in which a comparison was envisaged in the ancient Greek philosophy between conjunction, disjunction and conditional (or what has come to be known as ‘material implication’ in logic after Russell), all having two propositions, namely P and Q, and two values: T (true) and F (false).

			Conjunction	Disjunction	Implication
	P	Q	P and Q	P or Q	P implies Q
(1)	T	T	T	T	T
(2)	T	F	F	T	F
(3)	F	T	F	T	T
(4)	F	F	F	F	T

Thus, according to logicians, material implication is an aggregate of two propositions in a similar way as conjunction and disjunction. As can be

easily illustrated through the ‘truth table’ presented in table 1, if P is true (say, ‘Paris is in France’) and Q is true as well (say, ‘Rome is in Italy’), then the conjunction ‘P and Q’ and the disjunction ‘P or Q’ will be true, as will be the material implication ‘P implies Q’. The second row has P true (say, ‘Paris is in France’), but Q false (say, ‘London is in Italy’). Thus, the disjunction is true but both the conjunction and the material implication false. In the third row, P is false (say, ‘Paris is in England’) and Q true (say, ‘Rome is in Italy’). Apparently, the conjunction is false but disjunction and the material implication true. In the fourth row in the diagram, both P (say, Paris is in England) and Q (say, London is in France) are false. Consequently, both the conjunction and the disjunction are obviously false, but the implication turns out to be strikingly true. On this account, the material implication, to which conditionals have been linked in logic, is false only when P is true but Q false. In all other circumstances it turns out to be surprisingly true. Needless to say, natural languages tend not to accept the validity of a conditional construction which has a false antecedent but a true consequent. So, why it is that in logic a conditional can be true even when it has a false antecedent and a true consequent? It has been argued that “the material implication interpretation of the conditional, assigning ‘true’ to a conditional with a false antecedent, was justified by denying that universally quantified sentences have existential presuppositions, i.e. admitting an empty set to be the interpretation of the antecedent A and adhering to bivalence.” (Traugott et al. 1986: 15) In simple terms, although the proposition reported in a protasis can have both ‘true’ and ‘false’ values, the relation between the protasis and the apodosis should hold even if the proposition introduced in the protasis turns out to have a ‘false’ value, for it is only the ‘true’ value of the proposition in the protasis which is the requirement for the apodosis to be true. Thus, a protasis having a proposition with a false value does not cause (or require) the apodosis to have the same false value. In other words, a conditional statement does not provide any logical grounds to exclude the truthfulness of the apodosis in case where the proposition of the protasis should turn out to be false. However, as mentioned above, natural languages tend to not interpret conditionals in this way. As has been widely reported, in natural language communication, speakers generally associate the false value of the protasis with a false value of the apodosis. For example a sentence such as (1) ‘If John comes tomorrow, Mary would be happy’ is generally taken to implicate that: ‘If John does not come tomorrow, Mary would not be happy.’ Obviously, this implicature can be cancelled by adding to the sentence ‘But she would be happy anyway.’ A conditional utterance, therefore, requires a pragmatic

interpretation and cannot be studied solely as an aggregate of the truth-conditions of two propositions since they have different semantical values as well as roles in a conditional utterance.

### 3. Classification of Hindi conditionals according to time reference

In this section we present a tentative classification of Hindi conditionals according to the point in time a conditional statement as a whole refers to. Or, to put it in simple terms, the point in time the utterer of a conditional statement has in mind. Thus, the point in time of a conditional may or may not correspond to the time expressed in either the protasis or the apodosis, although, as indicated above, the tenses in the apodosis have to be taken into consideration.<sup>7</sup> The tenses marked in the protasis merely indicate the degrees of hypotheticality which serve to lay the ground for the apodosis to be true. Since the time reference of a conditional does not necessarily correspond to the overt tenses in the clauses of a conditional statement, it is necessary to maintain a distinction between the time reference of a conditional and the tenses encountered in the protasis and apodosis. In what follows, we will try to discuss and classify some types of conditionals only according to the point in time they refer to. Thus, we will not be concerned with an overall survey of Hindi conditionals.

#### 3.1. Conditionals with present time reference

##### 3.1.1. Habitual aspect in both the protasis and the apodosis

A Hindi conditional statement with present time reference may exhibit a habitual aspect in both protasis and apodosis, as in (1) and (2). In such cases, the conditional statement gets an ‘every-time-events’ rather than a ‘particular-time-single-event’ interpretation and thus the two propositions which constitute such a conditional statement can also be rendered by a non-conditional statement which has a universal time quantifier, namely ‘Whenever...’, without bringing any significant changes to the meaning.<sup>8</sup>

- (1) *agar*<sup>9</sup> *bāriś* *hotī* *hai* *to* *āgan* *mē*  
 if rain-F<sup>10</sup> be-IMPV.F AUX-PRES.3SG then courtyard in  
*kīcar* *ho jātā* *hai*  
 mud-M become-IMPV.M.SG AUX-PRES.3SG  
 ‘If it rains, then there is mud in the courtyard.’



- (2) *agar vo mujhe bulātī hai to māĩ*  
 if she I-ACC invite-IMPFV.F AUX-PRES.3SG then I  
*uske ghar calā jātā hũ*  
 her house go-IMPFV.M AUX-PRES.1SG  
 ‘If she invites me, I go to her house.’

As said above, conditionals in (1) and (2) have an every-time-events reading, namely ‘Whenever it rains there is mud in the courtyard’ and ‘Whenever she invites me, I go to her place’, respectively. In Hindi these propositions can be linked together by a time adverb and rendered through a relative-correlative Hindi construction ‘jab-jab... tab...’ (whenever...), etc.<sup>11</sup> Thus, a conditional statement having the habitual aspect in both the protasis and the apodosis carry zero hypotheticality. Similarly, when a protasis in habitual aspect is followed by an apodosis having aspects other than the habitual, it introduces a meaning similar to ‘Given that...’, ‘If it is the case/fact that...’, etc.

### 3.1.2. Protasis containing aspects other than habitual

Any aspectual change in the verb with a present time reference necessarily brings about changes in the degree of hypotheticality in the entire conditional statement. Accordingly, unlike the habitual aspect, the progressive and perfective aspects, respectively in the protasis of (3) and (4), make these statements single-event-conditionals rather than all-time-events conditionals.

- (3) *agar bāris ho rahī hai to bāhar*  
 if rain-F be-PROG.F AUX-PRES then outside  
*kīcaṛ hogā*  
 mud-M be-PRESM.M.3SG  
 ‘If it is raining, then there must be mud outside.’
- (4) *agar rāt-ko bāris hui hai to*  
 if night-during rain-F be-PFV.F AUX-PRES.SG then  
*bāhar kīcaṛ hogā*  
 outside mud-M be-PRESM.M.3SG  
 ‘If it has rained during the night, then there must be mud outside.’

As illustrated above, conditionals in (1) and (2) have an all-time reference whereas in (3) and (4) they have a single-event present time reference. The future form of the verb *honā*, ‘to be’, in the apodosis in (3) and (4) ex-

presses a presumptive modality on the part of the speaker rather than a future tense.<sup>12</sup> Thus, whether a conditional statement having a present time reference gets an all-time or a single-event interpretation depends on the type of verbal aspect of the protasis. In addition to the presumptive modality, some other TAM elements, including the future tense, can be attested in the apodosis. However, all such conditionals carry a zero hypotheticality and have present time reference.

### 3.2. Conditionals with future time reference

#### 3.2.1. Conditionals expressing possibility

Hindi conditionals with future time reference are those statements in which the action reported in the apodosis has to take place at a time later than the time of utterance. The proposition reported in the protasis can exhibit either future tense, as in (5), subjunctive mood, as in (6) and (7), or perfective aspect, as in (8), forms of the verb.

- (5) *agar vo bulāegā to maĩ uske ghar jāũgā*  
 if he invite-FUT.M.3SG then I her house go-FUT.M.1SG  
 ‘If he invites me, I will go to his house.’
- (6) *agar vo bulāe to maĩ uske ghar jāũ*  
 if he invite-SUBJ.3SG then I her house go-SUB.1SG  
 ‘If he invites/ Should he invite me, I will/would go to his house.’
- (7) *agar vo bulāe to maĩ uske ghar jāũgā*  
 if he invite-SUBJ.3SG then I his house go-FUT.M.1SG  
 ‘If he invites/ Should he invite me, I will go to his house.’
- (8) *agar usne mujhe bulāyā to maĩ uske ghar jāũgā*  
 if he-ERG I-ACC invite-PFV.3SG then I his house  
 go-FUT.M.1SG  
 ‘If he invited me, I would go to his house.’

As stated in previous sections, the protasis prepares the ground for the apodosis to be true. The protasis may or may not contain a clear indication of time in it. However, the tense morphology once employed in the protasis loses its general meaning and gets a conditional interpretation. Thus the tenses encountered in protasis do not carry a speaker’s modal meaning. They should not be confused with the tenses attested in independent propositions.

### 3.2.2. Conditionals expressing impossibility – counterfactuals with future time reference

As we shall see in the next section, Hindi conditionals having imperfective morphology in both the protasis and apodosis generally refer to an event which was scheduled to take place at a point in time prior to the utterance time. Thus, the imperfective participles in both the protasis and apodosis are generally thought to be synonymous with counterfactuals which refer to past unrealized events. However, the same counterfactual conditional with imperfective morphology can also be employed to refer to those ‘impossible’ actions which would have taken place at a point in time later than utterance time had the condition envisaged in the protasis been met. But, since according to the speaker the conditions described in the protasis are not to be met for various reasons, he can make a counterfactual statement which refers to a future event, as in (9a) and (9b). In this sense, such a usage of the counterfactual conditional with a future time reference is totally analogous to the counterfactual with a past time reference—the only difference being that the counterfactual with past time reference is impossible for temporal reasons, the counterfactuals with future reference are viewed as impossible by the speaker on grounds other than that of time. In both cases, however, the impossibility of the propositions in protasis and apodosis is presupposed and is based on the on the piece of knowledge the speaker has about the event reported in the protasis.

- (9) a. *agar vo mujhe agle hafte-kī pārṭī-mē*  
 if he I-ACC next week-of party-in  
*bulātā to maī us-mē zarūr*  
 invite-IMPFV.M.SG then I that-in certainly  
*śāmil hotā*  
 participate-IMPFV.M.SG  
 ‘If he had invited me to the next week’s party, I would certainly have participated in it.’
- b. *agar usne mujhe agle hafte-kī pārṭī-mē*  
 if he-ERG I-ACC next week-of party-in  
*bulāyā hotā to maī us-mē*  
 invite-PFV.M.SG AUX-IMPFV.M.SG then I that-in  
*zarūr śāmil hua hotā*  
 certainly participate-PFV.M.SG AUX-IMPFV.M.SG  
 ‘If she had invited me to the next week’s party, I would certainly have participated in it.’

## 3.3. Conditionals with past time reference

## 3.3.1. Habitual aspect in both the protasis and the apodosis

As in the conditionals with present time reference, the habitual aspect can be employed in conditional statements with past reference to obtain zero hypotheticality. Such conditionals thus get an ‘all-time-events’ rather than a ‘single-time-event’ interpretation. For instance, the examples illustrated above in (1) and (2) in the context of present time reference can be made to have a past time reference by changing the auxiliary, as in (10) and (11):

- (10) *agar bāriś hotī thī to āgan-mē*  
 if rain-F be-IMPV.F AUX-PST.F then courtyard-in  
*kīcar ho jātā thā*  
 mud-M become-IMPV.M.SG AUX-PST.M.3SG  
 ‘If it rained (in those days), then there was mud in the courtyard.’
- (11) *agar vo mujhe bulātī thī to māī*  
 if she I-ACC invite-IMPV.F AUX-PST.F.SG then I  
*uske ghar calā jātā thā*  
 her house go-IMPV.M AUX-PST.SG  
 ‘If she invited me (in those days), I would go to her house.’

Apparently, as (1) and (2) get an all-time interpretation in the present, so do (10) and (11) in the past, and similarly can be rendered by a ‘Whenever...’ construction. Thus, these conditionals carry a zero hypotheticality in the past. As illustrated above, Hindi conditionals with a habitual aspect in both the protasis and the apodosis consist of an imperfective participle and an auxiliary. However, in the narration of past events, it is quite common to encounter habitual past sentences without an auxiliary.<sup>13</sup> For example, in (12) both the protasis and apodosis are introduced without an auxiliary and thus do not have a tense marker. Such examples, however, indicate a habitual aspect and get from the context the past time reference rather than a counterfactual interpretation.

- (12) *bacpan-mē agar koi ham-ko paise detā*  
 childhood-in if someone we-DAT money give-IMPV.M.SG  
*to ham roz bājār jāte*  
 then we everyday market go-IMPV.M.PL  
 ‘In our childhood, If someone gave/ were to give us money, we would (= used to) go to the market everyday.’

3.3.2. *Protasis containing aspects other than habitual*

Hindi conditionals with past reference can have the protasis in aspects other than the habitual. For example, (13) and (14) have progressive and perfective aspects in the protasis respectively, and the apodosis carries the speaker's presumptive modal meaning:

- (13) *agar us vaqt bāriś ho rahī thī to bāhar*  
 if that time rain-F be-PROG.F AUX-PST.F.SG then outside  
*kīcaṛ ho gayā hogā*  
 mud-M become-PFV.M AUX-PRSUM.M.SG  
 'If it was raining that time, then it must have been muddy outside.'
- (14) *agar us rāt bāriś hui thī to bāhar*  
 if that night rain-F be-PFV.F AUX.PST.F.SG then outside  
*kīcaṛ ho gayā hogā*  
 mud-M become-PFV.M.SG AUX-PRESM.M.3SG  
 'If it had rained that night, then it must have been muddy outside.'

3.3.3. *Imperfective participle in both the protasis and the apodosis: counterfactuals*

As indicated above, Hindi does not possess any separate verb forms that could be compared with conditional verb forms<sup>14</sup> such as those attested in many Romance languages. However, it has at its disposal other morphological devices to mark the counterfactuality. The perfective-imperfective aspectual divide throughout the TAM system is one of them. In fact, Hindi makes use of this distinction in different syntactic contexts, including conditional clauses. Thus, Hindi conditionals with past time reference are those counterfactuals which exhibit imperfective verb forms both in the protasis and the apodosis simply to mark the non-completion of actions or events reported in the two propositions that constitute a conditional.

- (15) *agar us-ne mujhe bulāyā hotā to*  
 if she-ERG I-ACC invite-PFV.M.SG AUX-IMPV.M.SG then  
*mañ kal uske ghar gayā hotā*  
 I yesterday her house go-PFV.M.SG AUX-IMPV.M.SG  
 'If she had invited me, I would have gone to her house yesterday.'

- (16) *agar vo mujhe bulātī to maĩ kal*  
 if she I-ACC invite-IMPV.F.SG then I yesterday  
*uske ghar calā jātā*  
 her house go-IMPV.M.SG  
 ‘If she had invited me, I would have gone to her house yesterday.’

#### 4. Types of Hindi conditionals according to modal meaning they carry

In this section, we make an attempt to classify Hindi conditionals according to the modal meaning they carry. By modal meaning we mean that semantic element which is to be attached by the utterer either overtly or covertly to every proposition in order for it to carry the meaning that it does in a natural communication setting.<sup>15</sup> From this standpoint, many categories based on various illocutionary forces can be envisaged: epistemic conditionals and deontic conditionals, commissive conditionals, etc.

##### 4.1. Epistemic conditionals

Epistemic conditionals are those types of statements which have apodosis with a speaker’s epistemic modal meaning. Present habitual, future, subjunctive, perfective participle, imperfective participle and a perfective participle followed by the imperfective participle can appear in the protasis of an epistemic conditional. The following examples show the increasing hypotheticality of Hindi conditionals. Thus, (17) convey zero hypotheticality whereas (22) carries the highest degree of hypotheticality:

- (17) *agar chātr mehnat karte haĩ*  
 if students extortion-F do-IMPV.M.3PL AUX-PRS.3PL  
*to safal hote haĩ*  
 then successful be-IMPV.M.3PL AUX-IMPV.M.3PL  
 ‘If students work hard, they (generally) succeed.’
- (18) *agar vah mehnat karegā to safal*  
 if he extortion-F do-FUT.M.3SG then successful  
*hogā*  
 be.FUT.M.3SG  
 ‘If he works hard, he will succeed.’

- (19) *agar vah mehnat kare to safal ho*  
 if he extortion-F do-SUBJ.3SG then successful be-SUBJ.3SG  
 ‘If he works/ Should he work hard, he will succeed.’
- (20) *agar usne mehnat kī to safal hogā*  
 if he-ERG extortion-F do-PFV.F.3SG then successful  
*hogā*  
 be-FUT.M.2PL  
 ‘If he worked hard, he would succeed.’
- (21) *agar vah mehnat kartā to safal hotā*  
 if he extortion-F do-IMPV.M.3SG then successful  
*hotā*  
 be-IMPV.M.SG  
 ‘If he had worked hard he would have succeeded.’
- (22) *agar usne mehnat kī hotī to safal huā hotā*  
 if he-ERG extortion-F do-PFV.F.SG AUX-IMPV.F.SG  
 then successful be-PFV.M.SG AUX-IMPV.M.SG  
 ‘If he had worked hard he would have succeeded.’

#### 4.2. Deontic conditionals

This class consist of those conditional statements in which the apodosis carries a speaker’s deontic modal meaning. Hindi deontic conditionals can display all degrees of hypotheticality except the counterfactuality in protases and all kind of deontic devices in apodoses, as illustrated in the following examples:

- (23) *agar usne bulāyā hai to uske ghar jāo*  
 if she-ERG invite-PFV.SG AUX-PRS.SG then her  
 house go-IMP.2PL  
 ‘If she has invited you, go to her house.’
- (24) *agar vo bulāe to uske ghar jāo*  
 if she invite-SUBJ.3SG then her house go-IMP.2PL  
 ‘If she invites, then go to her house.’

In the case of the highest degree of hypotheticality, deontic modality cannot be attached through an imperative to the apodosis since the protasis

carries a counterfactual meaning. However, in such cases deontic wishes instead of an imperative can be employed, as in (25).

- (25) *agar vo bulātī to tum-ko uske ghar jānā*  
 if she invite-IMPV.F then you-DAT her house go-INF  
*cāhie thā*  
 must AUX-PST.M.SG  
 ‘Had she invited, you must have gone to her house.’

#### 4.3. Commissive conditionals

In this class of conditional, the speaker’s commitment to the hearer is expressed in the apodosis. Almost all the varieties of conditional can express a speaker’s commitment. (26), (27) and (28) are all examples of realis conditionals.

- (26) *agar vah bhārat āegā to māi use*  
 if you India-F come-FUT.M.2PL then I he-DAT  
*tājmahal dikhāūgā*  
 Tajmahal show-FUT.M.1SG  
 ‘If he comes to India I will show him the Tajmahal.’
- (27) *agar vah bhārat āe to māi use*  
 if you India.F come-SUBJ.3SG then I he-DAT  
*tājmahal dikhāūgā*  
 Tajmahal show-FUT.M.1SG  
 ‘If he comes/ Should he come to India, I will show you the Tajmahal.’
- (28) *agar vah bhārat āyā to māi use*  
 if you India-M come-FUT.M.PL then I he-DAT  
*tājmahal dikhāūgā*  
 Tajmahal show-FUT.M.1SG  
 ‘If he came to India I would show him the Tajmahal.’

From the point of view of an inherent epistemic modal meaning, protases may convey speaker’s belief only (i.e. Speaker believes but does not know that P). Hence the protasis is introduced into a conditional statement through ‘if’ marker. Protases cannot introduce speaker’s knowledge (i.e. Speaker knows that P) into a conditional, although, in the case of counterfactuals, speaker’s ‘knowledge that P’ plays an important role. In fact, in



counterfactuals the protasis is grounded on the contrary knowledge or belief of the speaker. In addition to the knowledge of belief element, protasis may also convey volitional meanings of the speaker through a subjunctive. However, protases do not carry presumptive modal meanings (i.e. Speaker knows that necessarily P). Furthermore, protases do not carry deontic modal meanings.

### 5. Overt marking of conditionality in Hindi

For the most part, both the protasis and apodosis are marked overtly in Hindi. However, the overt markers of protasis and apodosis (namely *agar/yadi...*, *to...*)<sup>16</sup> show a pattern which is to a great extent different from the one attested in English. In fact, unlike English and many other European languages, Hindi requires the apodosis to be obligatorily marked by the marker *to*, as being part of a conditional when it follows a protasis—irrespective of whether the protasis is marked or not—as can be seen from a comparison of (29a) and (29b) with (29c) and (29d). The absence of the apodosis marker *to* makes (29c) semantically odd,<sup>17</sup> and leaves (29d) a mere sequence of two propositions rather than a de facto conditional statement.<sup>18</sup> The apodosis in Hindi may be unmarked only in particular spoken forms where it has to precede the protasis, as can be seen in (30b).<sup>19</sup> Even in such cases, though, the marker of the apodosis may show up in its canonical place, after the protasis, as is evident in (30c). Such an appearance of the apodosis marker at the end of a conditional statement indisputably proves the obligatory nature of apodosis marking in Hindi. The protasis, on the other hand, is not marked obligatorily.<sup>20</sup> In effect, while it is necessary in Hindi for the apodosis to be marked, it is not obligatory for the protasis to be marked overtly, at least in a canonical conditional statement—i.e., when protasis precedes the apodosis—as is apparent in (29b). That said, the protasis is to be marked obligatorily should it be preceded by the apodosis, as can be seen from the ungrammaticality of (30a), (30d) and (30e).

- |      |    |             |                   |           |                  |                 |
|------|----|-------------|-------------------|-----------|------------------|-----------------|
| (29) | a. | <i>agar</i> | <i>Rām āyā,</i>   | <i>to</i> | <i>maĩ us-se</i> | <i>pūchhūgā</i> |
|      |    | if          | Ram come-PFV.M.SG | then      | I he-ABL         | ask-FUT.M.1SG   |
|      | b. | —           | <i>Rām āyā,</i>   | <i>to</i> | <i>maĩ us-se</i> | <i>pūchhūgā</i> |
|      |    | —           | Ram come-PFV.M.SG | then      | I he-ABL         | ask-FUT.M.1SG   |

- c. ?*agar Rām āyā*, — *maĩ us-se pūchũgā*  
 if Ram come-PFV.M.SG — I he-ABL ask-FUT.M.1SG
- d. \*— *Rām āyā*, — *maĩ us-se pūchũgā*  
 — Ram come-PFV.M.SG, — I he-ABL ask-FUT.M.1SG  
 ‘If Ram comes/came, I will/would ask him.’
- (30) a. \**to maĩ Rām-se pūchũgā agar vo āyā*  
 then I Ram-ABL ask-FUT.M.1SG if he come-PFV.M.SG
- b. ?— *maĩ Rām-se pūchũgā agar vo āyā*  
 — I Ram-ABL ask-FUT.M.1SG if he come-PFV.M.SG
- c. *maĩ Rām-se pūchũgā agar vo āyā to*  
 I Ram-ABL ask-FUT.M.1SG if he come-PFV.M.SG then
- d. *to maĩ Rām-se pūchũgā — vo āyā*  
 then I Ram-ABL ask-FUT.M.1SG — he come-PFV.M.SG
- e. \*— *maĩ Rām-se pūchũgā — vo āyā*  
 — I Ram-ABL ask-FUT.M.1SG — he come-PFV.M.SG  
 ‘I will/would ask Ram if he comes/ came.’

To recapitulate the above discussion, then, the characteristic features of conditionality markers in Hindi can be presented as in table 2.

Table 2. The distribution of overt markers of conditionality in Hindi between protasis and apodosis. The dash (—) indicates absence of the marker.

	Protasis	Apodosis	
(1)	[ <i>agar</i> -P →	<i>to</i> -Q]	a well-formed Hindi conditional
(2)	[— -P →	<i>to</i> -Q]	a well-formed Hindi conditional
(3)	?[ <i>agar</i> -P →	—Q]	a not-well-formed Hindi conditional
(4)	*[— -P →	—Q]	an ill-formed Hindi conditional
	Apodosis	Protasis	
(5)	*[ <i>to</i> -Q ←	<i>agar</i> -P]	an ill-formed Hindi conditional
(6)	*[— -Q ←	<i>agar</i> -P]	a not-well-formed Hindi conditional
(7)	[— -Q →	<i>agar</i> -P... <i>to</i> ]	a well-formed Hindi conditional
(8)	*[ <i>to</i> -Q ←	— -P]	an ill-formed Hindi conditional
(9)	*[— -Q ←	— -P]	an ill-formed Hindi conditional

As can be seen from table 2, in Hindi the apodosis is obligatorily marked whereas the protatis marker can be dropped. An unmarked apodosis renders a conditional statement unacceptable. Even in the cases where an inversion between the protasis and the apodosis is to take place, the apodosis marker is seen in its canonical place.

## 6. The order of protasis and apodosis in Hindi conditional statements

In his pursuit of universals derived from generalizations about human languages, Greenberg (1963: 84-85) made an important claim about clause ordering in conditionals saying that: “Universal of Word Order 14: In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages”. According to his observation, this linguistic universal is due to iconicity which has the sequential order principal at its heart. He attempted to establish that the sequential order of events is mirrored in a conditional statement. In other words, the clause ordering in a *If P then Q* construction exactly maps the parallels between the order of elements in language and the order of elements in experience, including the order of reasoning.<sup>21</sup> According to various truth-functional accounts of conditionals, this ordering can be explained away in the sense that the ‘implied facts’ reported in the apodosis—whether from a real or similar-to-real world—are to be true only in a possible world in which the ‘implying facts’ described in the protasis are to be true. Thus, it is not only logical but also obligatory to introduce the ‘implying elements’ reported in the protasis before the ‘implied elements’ described in the apodosis. This linguistic universal about ordering of clauses in a conditional statement, that is *If P then Q*, is fully supported by the data from Hindi, as it requires the finite verb of the main clause to stand in sentence-final position, thus requiring the apodosis to be placed after the protasis. There are some other promising proposals about the canonical protasis-apodosis ordering which all go in the direction of sustaining the above-mentioned linguistic universal.<sup>22</sup> In a nutshell, temporal reference theory suggests that the protasis-apodosis order reflects the temporal reference of two clauses,<sup>23</sup> the cause and effect theory claims that the observed linear order reflects the cause and effect relation between the two clauses.<sup>24</sup> According to another proposal, the protasis has to come first because it prepares common ground for the communicability of the apodosis.<sup>25</sup> As reported by Haiman, protases are topics and thus tend to occur sentence-initially.<sup>26</sup> Though not in total agreement with the above-mentioned proposals and various analyses, the present paper considers the suggestion put forward by various truth-functional theories to be reasonably convincing: the antecedent (i.e. protasis) has to come first in view of the fact that in a conditional statement it introduces the state of affairs of a possible world which serves as the basis for constructing a world in which the state of affairs described in the consequent will take place. Recall our previous discussion about the nature of the two propositions in a conditional statement. We noticed that, although

both the protasis and the apodosis are bivalent propositions (i.e. they can be either true or false), in a conditional statement, the truthfulness of an apodosis is viewed in relation to the truthfulness of the protasis, not the other way around. A conditional statement is false only when the protasis is true and the apodosis false. A true protasis requires the apodosis to be true as well, in order to make the conditional statement true. Furthermore, if the protasis is to be false, then the apodosis may be either true or false, without affecting the truthfulness of a conditional statement as a whole. The conditional statement will nonetheless remain true. In other words, to see whether a conditional statement is true or false depends on the truthfulness of the apodosis in relation to the truthfulness of the protasis. Hence, the universality of *if*-protasis  $\rightarrow$  *then*-apodosis order.

Despite this quasi unanimous, albeit variegated, consensus among different schools of thought about this linguistic universal, none of the suggestions put forward by scholars about ‘conditional clause inversion’ (i.e. from protasis-apodosis to apodosis-protasis) seems to be satisfactorily convincing. For example, a protasis encountered after the apodosis in a conditional statement has been generally considered to be a result of ‘conditional clause inversion’ which supposedly takes place simply due to an afterthought of the speaker. Without undertaking the task of presenting any detailed syntactic analysis of apodosis-protasis ordering in Hindi, we nonetheless intend to advance a completely different hypothesis in this direction. Our data suggest that the phenomenon of supposed ‘conditional clause inversion’ is conceived improperly, at least in the case of Hindi. We argue that the ‘conditional clause inversion’ should be viewed not as a rightward movement of the entire conditional clause (i.e. protasis) —as has been assumed by different syntactic analyses— but rather as a leftward dislocation of factual elements described in the apodosis which takes place owing to preposing (or fronting) and some other pragmatic principals. If we are prepared to look at the so-called phenomenon of inversion of clauses in conditional statements from a different point of view, we will notice that the universal order *If P then Q* does not undergo any major alterations. Consequently, we believe that the canonical protasis-apodosis ordering of clauses in conditionals, as shown in (31a), is indeed a linguistic universal, and it is not the protasis which moves rightward, as is seen at a surface level in almost all natural languages. Accordingly, the phenomenon of ‘conditional clause inversion’ in conditional statements should look as in (31b), not as in (31c), contrary to widely held belief.

- (31) a. [*if*-protasis proposition → *then*-apodosis proposition]  
 b. [apodosis proposition [*if*-protasis proposition → *then*- —]]  
 c. [*then*-apodosis proposition → *if*-protasis proposition]

We argue that for any Hindi conditional construction to be acceptable, it is necessary to keep rigidly to the ‘if-then’ order, even in those cases where the apodosis proposition has to move leftward for any pragmatic reasons. In fact, in Hindi even in cases where the apodosis proposition is dislocated to the left, the apodosis marker remains in its canonical place, as illustrated in (32a). Dislocating the entire apodosis with its marker renders the conditional construction totally ungrammatical in Hindi, as demonstrated in (32c) and (32d). The absence of an apodosis marker, as in (32b), cannot be considered a counterexample to the rigid rule of apodosis marking, as spoken Hindi does allow for the marker to be dropped, but only if the protasis is marked.

- (32) a. *maĩ Rām-se pūchūṅgā agar vo āyā to*  
 I Ram-ABLask-FUT.M.1SG if he come-PFV.M.SG then  
 b. ?— *maĩ Rām-se pūchūṅgā agar vo āyā*  
 — I Ram-ABL ask-FUT.M.1SG if he come-PFV.M.SG  
 c. *\*to maĩ Rām-se pūchūṅgā agar vo āyā*  
 then I Ram-ABL ask-FUT.M.1SG if he come-PFV.M.SG  
 d. *\*to maĩ Rām-se pūchūṅgā — vo āyā*  
 then I Ram-ABL ask-FUT.M.1SG — he come-PFV.M.SG  
 e. *\*— maĩ Rām-se pūchūṅgā — vo āyā*  
 — I Ram-ABL ask-FUT.M.1SG — he come-PFV.M.SG  
 ‘I will/would ask Ram, if he comes/ came.’

In order to check our claim, we have tried to see the data situation of protasis and apodosis across some world languages, relying exclusively on some typological studies.<sup>27</sup> Data from over 50 language suggest that it is not the protasis that moves rightward, but rather the propositional elements of the apodosis are preposed. In fact, none of the languages described in three typological studies provides ascertainable proofs of the rightward movement of the protasis in a conditional statement.

## 7. The relationship between the *if*-clause and the *then*-clause in Hindi

In previous sections we have argued that Hindi data confirm the hypothesis formulated by Greenberg in his Universal of Word Order 14, according to which the *if*-clause is to precede the *then*-clause. We have, furthermore, advocated that it is inappropriate to assume that there is a rightward dislocation of the *if*-clause in a Hindi conditional statement. As a consequence, we have advanced a hypothesis that in Hindi it is the proposition introduced by a *then*-marker which is preposed by the speaker in order to achieve different pragmatic goals, leaving the *then*-marker in its canonical position, as demonstrated in (33).

(33) [apodosis proposition [*agar*-protasis proposition → *to*- —]]

At this point it becomes necessary to see what kind of relation holds between the protasis and the apodosis. In particular, it would be interesting to see how the *if*-clause is attached to the *then*-clause. In a succinct survey of distinct syntactic theories of conditionals, Bhatt and Pancheva (2004) classify and critically examine diverse proposals put forward in this direction. As reported by the authors, in traditional grammars the two clauses are coordinated syntactically through *if* conjunction. Thus, according to traditional grammarians, the *if*-clause is equivalent to an adverbial clause. This view is further supported by many syntacticians on the grounds that VP ellipsis phenomena in conditionals goes against the hypothesis that conditionals are merely coordinated constructions. Hence conditional clauses (i.e. protases) are nothing but adverbial clauses.<sup>28</sup> Without undertaking any detailed syntactic analysis of the *if*-clause in Hindi, we consider this idea somewhat superficial in that any adverbial complementizer can be deleted without bringing much syntactic change into a sentence whereas an *if*-clause cannot, at least not semantically. Even in the cases of reduced conditionals in Hindi where only the *then*-clause can be attested in a natural discourse, the presence of an *if*-clause is presupposed. According to another proposal, conditionals are similar to correlative constructions. In the case of Hindi, we find the idea proposed by Dayal (1996) quite convincing. Schlenker (2001) has put forth a new idea for the semantic treatment of conditionals, arguing that protases are definite plural descriptions and thus subject to Condition C of the Binding Theory. However, despite numerous attempts made by diverse schools of thought, the syntax of conditional sentences remains a challenging topic. The fact that none of the theories put forward by scholars can indisputably claim to provide solutions for the

complexity of the syntax of conditionals is yet another proof of the peculiarity of conditional reasoning. We believe that the *if*-clause is the foundation of a conditional statement as a whole and cannot be considered a mere complementaizer. The two propositions employed in a conditional statement are intrinsically linked and are interdependent: none can exist without the either overt or covert (for example, in the case of reduced conditionals) presence of the other. In a conditional statement (for example, 'If John lives in Venice, he lives in Italy'), the truth of 'P implies Q' remains true due to some other argument (that is, living in Venice necessarily implies living in Italy), irrespective of whether P is true or false and whether Q is true or false. In other words, the above conditional remain true even if John does not live in Venice.

### 8. Pragmatic intrusion on the structure of a conditional statement

As mentioned in previous sections, the canonical structure of a conditional, namely  $P \rightarrow Q$ , has to sustain various types of pragmatic intrusions. This happens partly because of the inherent logical structure of a conditional statement but also because of its discourse boundedness. The underlying logical structure of a conditional,  $\neg P \vee P$  and  $Q$  (i.e. 'Either not-P or P and Q') induces a conditional statement to pragmatically implicate  $\diamond\neg P$ , and thus an epistemic scenario of a conditional statement becomes as follows:

$$(34) \quad K_s \diamond P \vee \diamond\neg P \text{ and } \diamond Q \vee \diamond\neg Q$$

i.e., for all the speaker knows it either is or is not Sunday and, thus, possibly it is the case that the priest is in church or possibly it is the case that the priest is not in church.<sup>29</sup> It has been widely argued that natural languages impose restriction on this epistemic scenario by excluding  $\neg P$  and  $Q$  (i.e. It is not Sunday and the priest is in church). In fact, due to a pragmatic intrusion on conditional statements (34) takes the form of (35):

$$(35) \quad B_s P \text{ (i.e. } P \equiv \neg\neg P) \rightarrow Q$$

In other words, for all the speaker believes, it is Sunday (i.e. it is not the case that not-Sunday) and the priest is in church. This epistemic scenario may undergo further restrictions as soon as P is employed in a conditional statement. Thus, what syntactic form a conditional statement is going to take in real speech depends exclusively on the context in which it is to be

made. For example, if both the speaker and the hearer mutually share a piece of knowledge, namely ‘It is Sunday’, then a conditional statement in this circumstance would most likely take the following form: ‘Since/Given that it is Sunday the priest will be in church’. If, on the other hand, the speaker is not sure whether it is Sunday or not, then what he asks the hearer to find out whether it is in effect a Sunday. It is also possible that this piece of information is provided by the hearer and the speaker neither knows nor believes it and thus merely affirms the proposition in the apodosis. In a nutshell, then, the overall structure of a conditional statement is determined by the contextually-given knowledge and beliefs of both the hearer and the speaker. Similarly, a protasis is always discourse bound and can take different degrees of hypotheticality according to the context. Conditionals do not make sense without their discourse context.<sup>30</sup> In the context of pragmatic intrusion, we can also take into consideration the topic of reduced conditionals. We argue that the form of reduced conditionals depends exclusively on the shared knowledge between the speaker and the hearer. The speaker selects the reduced conditional if the elliptical elements are already part of the shared knowledge between the speaker and the hearer.

## 9. Conclusion

At the outset of the present paper, we began by analyzing and discussing different varieties of Hindi conditionals according to the distribution of TAM elements in the protasis and the apodosis. All along our discussion, we have maintained a clear distinction between protasis and apodosis. We have tried to show that the TAM elements attested in the protasis are hypothetical and not real. They serve merely to prepare the ground on which the apodosis has to be true. We have made a brief attempt to identify those underlying pragmatic principles which are responsible for the leftward movement of the proposition in the apodosis. From our analysis of Hindi data, it emerges that the *if*-clause found after the apodosis at the surface level of a conditional does not violate the linguistic universal according to which a conditional construction should follow protasis-apodosis order. We have argued that in Hindi it is not the protasis which moves rightward but rather the propositional elements of the apodosis that are preposed by the speaker in order to accomplish some pragmatic goals, leaving the apodosis marker in its place. Thus, we have tried to advance the rigid ‘prota-



sis-apodosis' order hypothesis which, if supported by the data from other languages, may initiate a new way of looking at pragmatic intrusion into the syntax of the if-clause which has so far received a different treatment in various syntactic analyses. Furthermore, it has been argued that, contrary to the widely held belief in the domain of logic, conditionals are not a semantical aggregate of two propositions and that the protasis and the apodosis are not of the same nature.

## Notes

1. In order to draw the line between a conditional statement and other types of statements consisting of two propositions, we shall be using throughout the paper the term pair *protasis-apodosis* introduced by traditional grammarians to name the two clauses, being quite aware of the much wider acceptability of the pair *antecedent-consequent*, notably in philosophical tradition. We believe that conditionals are completely different from, and should not be confused with, other types of statements which are made of two propositions. Some other names given to the *if*-clause and the *then*-clause are also misleading. For example, the team of researcher under the editorship of Xrakovskij (2005) have used the terms 'dependent clause' (DC) for the *if*-clause and 'main clause' (MC) for the *then*-clause. We would argue that if clause is not a dependent clause at all. On the contrary, the *then*-clause seems to be dependent on the *if*-clause which is the basis of a conditional statement.
2. In order not to confuse the reader with different types of notations, we are using the material conditional notation ' $P \rightarrow Q$ ' without making any difference between this notation and others such as 'If A, B', 'If A, C' 'If A then C' or 'If..., then...', etc.
3. Example borrowed from Comrie (1986: 78).
4. For a detailed discussion on the topic, see for example Jackson (1991: 111).
5. The term 'material implication' is known also as 'material conditional' or 'truth-functional conditional' and represented through different notations or symbols in philosophical literature.
6. If we follow the way in which material implication was conceived by Greek Philosophers, we find a proposition such as "If moon is made of chocolate (F), then it is made of cheese (F) and that it is made of volcanic rock (T)" congruous since material implication allows a false antecedent to be the basis of either false or true apodosis. Obviously, that natural language communication does not allow such inferences to take place. We believe that even the introduction of 'relevant logic' to get rid of such anomalous inferences does not solve the problem.

7. In an ambitious attempt to bring out typological characterizations of conditional constructions across world languages, a team of linguists at St. Petersburg University has come up with various generalizations about conditional constructions in 24 languages. (Xrakovskij 2005) Though admirable, any such endeavour is bound to mix up things because it is impossible to be consistent throughout in the use of terminology due in part to the divergent linguistic traditions of each language, and in part also to different backgrounds of the authors. The distribution of TAM elements between the protasis and apodosis across different languages, for example, has been classified and explained using different terminologies. This has led to some inaccuracy in the analysis of a number of languages.
8. Thus, in reference grammars of various languages this type of conditional is generally dubbed ‘zero-conditional’, since the conditionals of this category contain a zero degree of hypotheticality.
9. For some technical reasons, a mixed-type of transliteration system (which is popular among scholars working on Indian languages) rather than IPA diacritic is adopted in this paper.
10. Abbreviations: 1 = first person, 2 = second person, 3 = third person, AUX = auxiliary, ABL = ablative, IMPFV = imperfective, FUT = future, PFV = perfective, PST = past, CONT = continuous, PRES = present, PRESM = presumptive, SUBJ = subjunctive, CFV = counterfactive (contrafactive), SG = singular, PL = plural, M = masculine, F = feminine, OBL = oblique, ACC = accusative, HON = honorific, PASS = passive, IMP = imperative.
11. For example, (1) and (2) can be rendered through a relative-correlative Hindi construction in the following examples:
- jab-jab      bāriś      hoī                              hai                              to              āṅan-mē*  
whenever    rain-F    be-IMPFV.F    AUX-PRES.3SG    then    courtyard-in
- kīcaṛ      ho      jātā                              hai*  
mud-M    become-IMPFV.M.SG    AUX-PRES.3SG  
‘Whenever it rains, there is mud in the courtyard.’
- jab              vo              mujhe              bulāī                              hai                              to              māī*  
when    she    I-ACC    invite-IMPFV.F    AUX-PRES.3SG    then    I
- uske      ghar              calā      jātā                              hū*  
her    house    go-IMPFV.M    AUX-PRES.1SG  
‘When she invites me, I go to her house.’
12. We have discussed this elsewhere: (Sharma 2008).
13. See McGregor (1995: 187) and Montaut (2004: 240–242) for further details.
14. Some authors maintain that the Hindi imperfective participle is counterfactual. See, for example, Oranskaya (2005: 235). However, we believe that it would not be appropriate to link the Hindi imperfective participle with any particular category as it is used throughout the Hindi tense-aspect system (for example in habitual present, habitual past) as well as to construct adjectival and adverbial complementizers. We argue that Hindi conditionals make use of the perfec-

- tive-imperfective divide to indicate various degrees of hypotheticality. Thus, the imperfective participle in Hindi is employed in both the protasis and the apodosis of counterfactuals to indicate that none of the actions reported in the protasis and the apodosis were accomplished.
15. We have discussed this in detail in Sharma 2000 and Sharma 2008.
  16. In addition to ‘agar’ and ‘yadi’, some other forms are also attested to mark the protasis: *jo* (a relative pronoun), *kahĩ* (if ever, in case, etc.), *kāś* (mostly in counterfactuals), *kadācit* (if ever; in Sanskritized Hindi). For further details see Oranskaya (2005: 222).
  17. It must be admitted though that, as a new trend—due in part to an inevitable influence of English on Hindi—some speakers tend not to mark the apodosis always, particularly in spoken Hindi. However, the standard written register of Hindi requires of the apodosis to be marked obligatorily, irrespective of whether the protasis is marked or unmarked.
  18. We do not consider such a sequence of propositions a conditional statement. However, we will not go here into the question of why and when a sentence sequence such as *You don’t want to go to the market? I will* or *Open the window and I will kill you* or *Rām āyā. maĩ us-se pūchūgā* (i.e. Ram came. I will ask him) in a particular context can have a conditional tone in normal speech.
  19. Even in these cases, the acceptability of unmarked apodosis is highly controversial among speakers of Hindi. We will return to the question of ‘conditional clause inversion’ in the next section where we advance the hypothesis that in reality there is no such thing as ‘inversion of protasis-apodosis order’ in Hindi conditional statements, but rather a displacement of the apodosis clause furthest to the left, leaving the marker of apodosis at its canonical place and thus maintaining the protasis-apodosis order intact.
  20. Thus, the claim made by Comrie (1986: 96) stating that “Although it is possible to have conditionals where neither protasis nor apodosis is explicitly marked as being part of a conditional, it is usual for the protasis to be overtly marked; marking of the apodosis is less common, and marking of the apodosis alone is particularly rare.” does not seem to be totally correct. In Hindi the apodosis is marked obligatorily, regardless of its position in the conditional (i.e., either before or after the protasis), and irrespective of whether the protasis is marked overtly or not.
  21. For a succinct discussion of the universalistic approach to conditionals, see Overview by Charles A. Ferguson et al. in *On Conditionals*, Traugott et al. (eds.), 1986, 8-10.
  22. For a summary of other proposal see Comrie (1986: 83-86).
  23. Like Comrie (1986: 85), we do not consider this hypothesis to be valid. However, we believe that the counter-example discussed by him: “If it will amuse you, I’ll tell you a joke” is inappropriate. The will modal employed in protasis in this example provides no time reference but the speaker’s modal meaning.

24. Like Comrie we do not support this hypothesis. For a summary of other proposal, see Comrie (1986: 83-86).
25. Lehman, Christian (1974) quoted by Comrie (1986: 86).
26. “A conditional clause is (perhaps only hypothetically) a part of the knowledge shared by the speaker and his listener. As such, it constitutes the framework which has been selected for the following discourse.” (Haiman 1978: 583) “The topic represents the entity whose existence is agreed upon by the speaker and his audience. As such, it constitute the framework which has been selected for the following discourse.” (Haiman 1978: 585).
27. Our observation is based exclusively on the following three works: (1) *Typology of Conditional Constructions*, Victor S. Xrakovskij (ed.), (2005) for Bulgarian (Rousselina Nicolova), Armenian (Natalia A. Kozintseva), Dari (Boris Ya. Ostrovsky), Greek (Tatayana I. Oranskaya), Early Latin (Margarita K. Sabaneyeva), French (Elena E. Kordi), German (Svetlana M. Kibardina), English (Tatiana G. Akimova, Natalia A. Kozintseva), Finnish (Hannu Tommola), Estonian (Irina P. Külmoja), Hungarian (László Jaszai, Ethelka Tóth), Hausa (Myrrah A. Smirnova, Nikolaj A. Dobronravin), Klamath (Viktor A. Stegnyy), Indonesian (Alexander K. Ogloblin), Cambodian (Natalia M. Spatari), Vietnamese (Igor S. Bystrov, Nonna V. Stankevič), Chinese (Tamara N. Nikitina), Even (Andrej L. Malchukov), Evenki (Igor V. Nedjalkov, Nina Ya. Bulatova), Eskimo (Nikolaj B. Vaxtin), Aleut (Evgeniy V. Golovko), Yukaghir languages (Elena S. Maslova) and Japanese (Vladimir M. Alpatov, Vera I. Podlesskaya); (2) *The semantics of Clause Linking*, R. M. W. Dixon, and Alexandra Y. Aikhenvald (eds.), (2009) for Akkadian, Galo, Khan, Manambu, Iquito, Aguaruna, Ojibwe, Fijian, Toqabaqita, Martuthunira, Korean, Goemai, Konso and Mali; (3) *The Indo-Aryan Languages*, Colin P. Masica, (1991) for various Indo-Aryan languages. None of the languages discussed in above mentioned works seem to have ‘apodosis-protasis’ ordering.
28. For a critical assessment of diverse syntactic theories (i.e. conditional-correlative link theory, conditional-interrogative link theory, etc.) see Bhatt and Pancheva (2004).
29. We would like to clarify Levinson’s point (2000: 109) where he illustrates the scenario of clausal implicature employing epistemic modifier K (knowledge) rather than B (belief). Although one can know that it is either P or  $\neg$ P, as soon as this piece of knowledge becomes part of a protasis, it takes the form of B (belief). The protasis, we claim, always carries speaker’s belief rather than knowledge. A piece of knowledge in the protasis would be tantamount to “Given that...”, “Since/as...” rather than to “If...”. For further discussion on the topic of implicature, see also Fauconnier (1985: 109).
30. See Akatsuka (1986: 349) for further details.

## References

- Akatsuko, Noriko  
 1986 Conditionals are discourse-bound. In *On Conditionals*, Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly, and Charles A. Ferguson (eds.), 333–351. Cambridge: Cambridge University Press.
- Bhatt, Rajesh and Roumyana Pancheva  
 2006 Conditionals. In *The Blackwell Companion to Syntax*, Vol. 1, Martin Everaert, and Henk van Riemsdijk (eds.), 638–687. Oxford: Blackwell.
- Comrie, Bernard  
 1986 Conditionals: A typology. In *On Conditionals*, Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly, and Charles A. Ferguson (eds.), 77–99. Cambridge: Cambridge University Press.
- Dayal, V.  
 1996 *Locality in Wh-quantification: Questions and Relative Clauses in Hindi*. (Studies in Linguistics and Philosophy 62) Dordrecht: Foris Publications.
- Declerck, Renaat and Susan Reed  
 2001 *Conditionals, A Comprehensive Empirical Analysis*. Berlin: Mouton de Gruyter.
- Dixon, R. M. W. and Alexandra Y. Aikhenvald (eds.)  
 2009 *The Semantics of Clause Linking. A Cross-Linguistic Typology*. (Explorations in Linguistic Typology) Oxford: Oxford University Press.
- Edgington, Dorothy  
 1991 Do conditionals have truth-conditions? In *Conditionals*, Frank Jackson (ed.), 176–201. Oxford: Oxford University Press. [First published in *Critica* XVIII, 52, 3–30, 1986].
- Fauconnier, Gilles  
 1985 *Mental Spaces. Aspects of meaning construction in natural language*. Cambridge, Mass.: The MIT Press.
- Greenberg, Joseph H.  
 1963 Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, Joseph H. Greenberg (ed.), . Cambridge, Mass.: MIT Press.  
 1986 The realis-irrealis continuum in the classical Greek conditional. In *On Conditionals*, Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly, and Charles A. Ferguson (eds.), 247–264. Cambridge: Cambridge University Press.

- Grice, Paul  
 1989 *Studies in the Way of Words*, Cambridge Mass.: Harvard University Press.
- Haiman, John  
 1978 Conditionals are topics. *Language* 54: 565–589.  
 1986 Constraints on the form and meaning of the protasis. In *On Conditionals*, Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly, and Charles A. Ferguson (eds.), 215–227. Cambridge: Cambridge University Press.
- Iatridou, Sabine  
 1994 On the contribution of conditional *then*. *Natural Language Semantics* 2: 171–199.
- Jackson, Frank  
 1991 On assertion and indicative conditionals. In *Conditionals*, Frank Jackson (ed.), 111–135. Oxford: Oxford University Press. [First published in *Philosophical Review*, 88, 565–89, 1979].
- Johnson-Laird, P. N.  
 1986 Conditionals and mental models. In *On Conditionals*, Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly, and Charles A. Ferguson (eds.), 55–75. Cambridge: Cambridge University Press.
- König, Ekkehard  
 1988 Concessive connectives and concessive sentences: cross-linguistic regularities and pragmatic principles. In *Explaining Language Universals*, John A. Hawkins (ed.), 145–166. Oxford: Basil Blackwell.
- Levinson, Stephen C.  
 2000 *Presumptive Meanings – The Theory of Generalized Conversational Implicature*, Cambridge, MA : MIT Press.
- Masica, Colin P.  
 1991 *The Indo-Aryan Languages*. (Cambridge Language Surveys) Cambridge: Cambridge University Press.
- McGregor, Ronald Stuart  
 1995 *Outline of Hindi Grammar*. Oxford: Oxford University Press.
- Monatut, Annie  
 2004 *A Grammar of Hindi*. (LINCOS Studies in Indo-European Linguistics) Munich: Lincom Europa.
- Oranskaya, Tatyana I.  
 2005 Conditional constructions in Hindi. In *Typology of Conditional Constructions*, Victor S. Xrakovskij (ed.), 218–245. (LINCOS Studies in Theoretical Linguistics 25.) Munich: Lincom Europa.

Portner, Paul

- 2009 *Modality*. (Oxford Surveys in Semantics and Pragmatics) Oxford: Oxford University Press.

Sharma, Ghanshyam

- 2000 On the modal meanings of the subjunctive in Hindi. In *The Yearbook of South Asian Languages and Linguistics*, Rajendra Singh (ed.), 177–198. New Delhi: Sage Publications.
- 2008 A pragmatic account of the Hindi presumptive. In *Annual Review of South Asian Languages and Linguistics*, Rajendra Singh (ed.), 83–113. Berlin: Mouton de Gruyter.

Schlenker, Paul

- 2001 *A referential analysis of conditionals*. Unpublished manuscript, MIT.

Traugott, Elizabeth Closs, Alice ter Meulen, Judy Snitzer Reilly and Charles A. Ferguson (eds.)

- 1986 *On conditionals*. Cambridge: Cambridge University Press.

Van der Auwera, Johan

- 1986 Conditionals and speech acts. In *On Conditionals*, Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly, and Charles A. Ferguson (eds.), 197–214. Cambridge: Cambridge University Press.

Xrakovskij, Victor S. (ed.)

- 2005 *Typology of Conditional Constructions*. (LINCOS Studies in Theoretical Linguistics 25.) Muenchen: Lincom Europa.

# Regional Reports





# Issues in the Description of Bangla

*Probal Dasgupta*

## Abstract

*The list of unsolved problems in Bangla grammar provided here covers phonology and the grammar of nominal and verbal forms. A nasalized vowel cannot be followed by [g] – such a sequence surfaces as [Vŋ] instead. An intervocalic double yy is possible, but not a double ww. Animate nouns have no plural locative form. Genitives can appear in an inanimate plural format \_\_gulo but not in an animate plural format \_\_raa. Future perfect forms always have an epistemic, never a root interpretation. Conditional perfect forms do not exist. Passives do not support the obligational construction.*

## 1. Introduction

This list of unanswered questions in Bangla grammar does not represent a hidden agenda. I do not foresee either that any particular approach to linguistics is likely to come up with viable answers to all of them or that such answers, however obtained, will offer coherent support to some particular view of grammatical or other matters. These are problems awaiting solution, and may prove to be of interest.

## 2. Phonology

Foremost on my phonology list is a problem that has to do with prenasal nasalized vowels. Such vowels in Bangla, even when the following consonant is a voiced occlusive (plosive or affricate), are not realized as nasalized vowel plus homorganic prenasal plus occlusive in the manner familiar from Hindi-Urdu, where *karũũgaa* is not a possible sequence, and where only *karuungaa* '(I).will.do' is available. Bangla has near-minimal pairs such as *baandaa* 'slave' vs *cããdaa* 'subscription', *paanjaa* 'arm

wrestling’ vs *šããjaa* ‘starter (for yoghurt)’; such pairs do not exist in Hindi-Urdu.

One important task for linguists dealing with Bangla is to devise a way to describe this feature that makes sense of the specific facts in the velar subdomain. We would expect pairs like *caangaa* ‘fit, well’ vs *\*rããgaa*. But in fact nasalized vowel plus [g] is excluded by the phonotactics of Bangla: the place of such a sequence is always occupied by vowel plus [ŋ], as in *raangaa* ‘glowing, pink’.

In word final position, the pattern reduces to *cããd* ‘moon’, *bhããj* ‘fold’, *gaang* ‘stream’ (cf. *\*gããg*); word final sequences of the form nasal plus occlusive do not occur in Bangla. In word initial position the issue does not arise.

Second on my phonology list is a hole in the pattern of vocalic sequences. Postvocally, Bangla distinguishes between the high semivowels *y*, *w* and the mid semivowels *Y*, *W*, as in *jaay* ‘(I).go’ vs *jaay* ‘(s/he).goes’, *jaaw* ‘soup’ vs *jaaw* ‘(you).go’. Intervocally, the height distinction is neutralized, and the realizations come out mid: *paaYaa* ‘leg (of furniture)’, *haaWaa* ‘wind’.

However, a tense – and extended – high front intersegment is possible, as in *gaayye* ‘singer’, *khaayyo* ‘(you).feed (a future imperative form)’. There is a hole in this pattern, in that no tense and extended high back intersegment occurs: no *ww* exists – not even in nonsense words in the paralinguistic penumbra of Bangla. It is not clear whether the high front cases should be viewed in terms of an extra syllable – *gaaiye*, *khaaiyo*; what we need is an account that would make such a choice meaningful at the level of making distinct predictions elsewhere in the system, and would help us to choose between a *yy* account and an *iy* account.

The orthography in Bangla is relatively opaque and does not help address the semivowel issue. At the level of velar nasals, Bangla orthography distinguishes the velar nasal as in *raangaa* from nasal plus plosive sequences as in *caangaa* (glosses given above).

### 3. Nominal Forms

In this problem listing exercise, it helps if we leave open the issue of whether it is up to morphology or syntax to address the issues at stake. We begin with nominal forms. The phenomena in Bangla that correspond to the tidier case and number systems in other languages are poorly understood.

One unsolved problem has to do with the locative case. Inanimate nouns can be said to exhibit a locative case: *aalu* ‘potato’, *aalugulo* ‘the potatoes’, *begun* ‘eggplant’, *begungulo* ‘the eggplants’ are nominative forms that have corresponding locatives *aalute* ‘in a/the potato’, *aalugulote* ‘in the potatoes’, *begune* ‘in an/the eggplant’, *begungulote* ‘in the eggplants’. An inanimate noun with a classification formatted numeral clitic such as *aalu-duTo* ‘the two potatoes’ also occurs in the locative: *aalu-duTote* ‘in the two potatoes’.

Animate nouns occur in the locative only in fixed locutions. The animate locatives *maanuše* ‘in man’ or *iššare* ‘in God’ or *debotaay* ‘in a/the deity’, while they do not occur freely, do occur in the environment *Aami \_\_\_ biššaaš kori* ‘I believe \_\_\_’. But animate plurals such as *maanušeraa* ‘men, people’ or *debotaaraa* ‘gods’ have no locative form at all. When animate nouns take a classification formatted numeral clitic, as in *chele-duTo* ‘the two boys’, the rest of the system enables us to imagine that its locative would look like *\*chele-duTote* ‘in the two boys’, analogical to *aalu-duTote* ‘in the two potatoes’. This conceivable animate locative form is sharply excluded. For further details, see Dasgupta, Ford & Singh (2000: 157-8), where it was first pointed out that animate plural locatives do not exist in Bangla. To this day, it is not clear just how this gap in the paradigm is to be stated.

A second unsolved problem in the morphology of Bangla has to do with the genitive. Consider (1)-(4) below, where the genitives *aamaar* ‘my’ and *aamaader* ‘our’ are able to appear in the classification format – note the forms *aamaarTaa* ‘mine’ and *aamaarTaar* ‘mine.Gen’ in (1) and (3) respectively.

1. eTaa to tomaar TikiT, aamaarTaa koy?  
this Prt your ticket, mine where  
‘This is your (Sg) ticket, where is mine?’
2. tomaar khejurgulo poRe gElo, aamaargulo naaW!  
your dates fall Aux, mine take.Imp!  
‘Your (Sg) dates have fallen down, take mine!’
3. eTaa to tomaader TEksi, aamaarTaar ki holo?  
this Prt your taxi, mine.Gen what Cop  
‘This is your (Pl) taxi; what about mine?’

4. tomaader khejurgulo paaooa gEche, aamaadergulo ki holo?  
 your dates found have.been, ours.Gen what Cop  
 ‘Your (Pl) dates have been found; what about ours?’

The availability of (1)-(4) makes it natural to ask why (5)-(8) are ill-formed. The fact is that the forms *aamaarraa* ‘mine.Pl’ and *aamaaderraa* ‘ours.Pl’, as well as their genitives *aamaarder* ‘mine.Pl.Gen’ and *aamaaderder* ‘ours.Pl.Gen’, are unavailable. We provide sentences (5)-(8) to make it clear that even in a plausible sentential context these forms sound so bizarre that a native speaker has trouble imagining what one could possibly be trying to say:

5. \*tomaar šoᅅgiraa phire ešeche, aamaarraa bepaattaa  
 your companions back have.come, mine.Pl missing  
 intended reading ‘Your (Sg) companions have come back but mine are missing’
6. \*tomaader šoᅅgiraa phire ešeche, aamaaderraa bepaattaa  
 your companions back have.come, ours.Pl missing  
 intended reading ‘Your (Pl) companions have come back but ours are missing’
7. \*tomaar šoᅅgiraa phire ešeche, aamaarder ki holo?  
 your companions back have.come, mine.Pl.Gen what Cop  
 intended reading ‘Your (Sg) companions have come back; what about ours?’
8. \*tomaader šoᅅgiraa phire ešeche, aamaaderder ki holo?  
 your companions back have.come, ours.Pl.Gen what Cop  
 intended reading ‘Your (Pl) companions have come back; what about ours?’

A linguist who wishes to see the phenomenon in context needs to note that adjectives pattern differently. Consider (9)-(12), where it is shown that adjectives can appear in the formats *\_\_Taa*, *\_\_gulo*, *\_\_raa*, *\_\_der* quite freely:

9. šobujTaa kinle kEno?  
 green.Taa you.bought why  
 ‘Why did you buy the green one?’

10. šobujgulo kinle kEno?  
green.Gulo you.bought why  
'Why did you buy the green ones?'
11. lambaaraa kothaay gElo?  
tall.Pl where have.gone  
'Where have all the tall people gone?'
12. bēTeder bujhi šaat khun maap?  
short.Pl.Gen Prt seven murders forgiven  
'Do the short people get away with murder then?'

The fact that genitives can appear in the classification formats, singular *\_\_Taa* and plural *\_\_gulo*, but not in the animate plural format *\_\_raa* and the animate plural genitive format *\_\_der*, must therefore receive an account that squares with these properties of adjectives.

#### 4. Verbal Forms

In the case of verbs, linguists neglected the Tense-Aspect-Mood system for decades. Consequently, it is unsurprising that we do not understand why Bangla future perfect forms such as *aapni aamaar ciThi peye thaakben* 'you my letter received will.have, = You will have received my letter' can only have the epistemic reading 'You have probably received my letter' and never the root reading available in, say, English, along the lines of 'Your receiving my letter is an event occurring prior to some future time point otherwise specified in the discourse'. The facts in the sister language Hindi-Urdu are similar – and have not, to the best of my knowledge, been addressed in any insightful account.

But the facts of Hindi-Urdu, where it is possible to say *agar us ne mujhe bulaayaa hotaa* 'if s/he Erg me called had, = if s/he had called me', do not prepare us for the discovery that Bangla is more restrictive. Bangla does not allow its users to say *uni jodi aamaake Deke thaakten* 'if s/he me called had, = if s/he had called me', with the copular auxiliary in the conditional form *thaakten*. This is an independently puzzling fact about Bangla that will continue to exercise linguists.

In another sector of the grammar of verbs, English allows a single verbal complex to carry both passive and obligational markers, as in *Three letters have to be written*. But Bangla prohibits this; *\*tinkhaana ciThi lek-*

*haa hote habe* ‘three letters written be have.to, = three letters have to be written’ and its variants are robustly ill-formed. The matter is discussed, and a tentative solution proposed, in Dasgupta (2010), but that text is in Bangla. That the speakers of Bangla are among the six most numerous speech communities of the world is apparently a fact that does not signify – so I have been given to understand by eminent colleagues in the community of linguists. Given the conventions that the systems of publication and accreditation have imposed on willing and unwilling participants, it thus becomes my duty to state that this problem has not been addressed in any ...*academic* publication and thus counts as unsolved.

## References

- Dasgupta, Probal, Alan Ford and Rajendra Singh  
2000 *After Etymology: Towards a Substantivist Linguistics*. Muenchen: Lincom Europa.
- Dasgupta, Probal  
2010 Bhaabbaaccer kaache-piThe. [In Bangla: ‘In the vicinity of the bhaava voice’.] *Janapadaprayaas* 7:2-3.25-28.

# Europe

*John Peterson*

## 1. Introduction

Work is currently being conducted at European universities and research institutes on virtually all aspects of South Asian languages and linguistics. Although the emphasis of these studies is constantly changing with respect to the topics receiving the most attention, at present the field as a whole is flourishing in Europe.

While the continent is still rightfully famous for its historical studies, especially on etymological aspects of the Indo-Aryan languages, other aspects of all South Asian languages have been claiming an ever-increasing share of attention over the past several decades, a development which is in no way unique to Europe but which reflects more general trends in the field. Instead, there is a clear tendency at these institutions in three other directions, namely descriptive studies, computational linguistics, and the South Asian grammatical traditions, often in conjunction with computational methods. In fact, these three areas together account for the large majority of titles mentioned in the following pages, although works on etymology and linguistic theory are also found, as well as a considerable number of sociolinguistic studies and works designed for language learners.

With respect to the languages discussed, the picture which emerges is quite a "mixed bag", with all of the major families receiving attention at European universities, from the Indo-Aryan languages of Pakistan and Rajasthan to the Munda languages of eastern central India, and from the Indo-Aryan and Tibeto-Burman languages of Nepal and Tibet to the Dravidian languages in the south. In addition, a number of works on Sri Lanka Malay will be mentioned below as well as work on "Pidgin Madam", a Sinhalese-Arabic pidgin language spoken in Lebanon.

The present study does not claim to be exhaustive with respect to these studies in Europe for the years 2007-2009. Rather, it is only intended to give the reader a general overview of the broad spectrum of work in South Asian linguistics currently being conducted at European universities and research institutes.



## 2. Historical (socio-)linguistics / history of linguistics / the Indian grammatical traditions

We begin our brief discussion with an overview of some of the work conducted in the past three years on historical aspects of the languages of South Asia, including not only the more traditional etymological works on these languages but also works dealing with the ancient grammatical traditions of South Asia (except those dealing with computational aspects, see section 4) and also historical sociolinguistics.

Jean-Luc Chevillard has published a number of studies related to the history of the Indian linguistic traditions. For example, Chevillard (2007a), written in French, deals with the use of examples in grammatical treatises from the subcontinent while Chevillard et al. (2007), also in French, deals with examples in a number of different grammatical traditions, including the Tamil tradition. Chevillard (2008a) is a companion volume to the French translation of the *Ceneavaraiyam*, containing a glossary of all technical words and phrases in Cenavaiyar's commentary of the *Collatikaram*. Other studies by this same author include Chevillard (2008b), which discusses the term *tikai-c-col* or 'regional words' as used by Tamil grammarians and their commentators and its place in Tamil literature, Chevillard (2009a), which discusses a motif in the mythology of Akattiyān, better known by his Sanskrit name *Agastya*, associated among other things with the Tamil language and grammar, and finally, Chevillard (2009b), which explores the way in which 32 technical terms were translated from Sanskrit into Tamil and some of the issues involved in translating terminology from one language into another.

Aussant (2009) is a monograph-length study (166 pages), written in French, dealing with referential aspects of proper nouns in the three traditional Indian disciplines of *vyākaraṇa*, *nyāya* and *mīmāṃsā*. It consists of seven chapters, beginning with the emergence of the idea of convention and the notion of cause of application, followed by five chapters, each dedicated to a particular thesis: that of the generic property as cause of application of the proper noun, the proper form as cause of application of the proper noun, identity as cause of application of the proper noun, direct reference, and the individual as cause of application of the proper noun. Freschi (2008), on the other hand, investigates *mīmāṃsā* hermeneutics, focusing on a passage from the *Tantrarāhasya* by Rāmānujācārya and concentrating on the principles underlying the structuralization of Vedic texts as depicted in this text.

A number of works by Pontillo as well as one work co-authored by her and Maria Candotti also deal with various issues within the Indian grammatical tradition. Candotti & Pontillo (2007) discuss the imagery of the tree and its branches in Patañjali's *Mahābhāṣya*, arguing that it is used as a metaphor to denote a specific part-whole relationship with respect to linguistic segmentation. Similarly, Pontillo (2008a) looks at another image of Patañjali's taken from the natural world, also to denote a kind of part-whole relationship, namely the rice plant and its different parts. In her Italian-language article, Pontillo (2009a) deals with the notion of *kāmacāra*- in Pāṇinian commentaries. Finally, a number of other articles by the same author combine linguistic and philological methods, such as Pontillo (2009b), which deals with the relationship between two ancient Indian traditions, that of *vyākaraṇa*- and *alaṃkāra-śāstra*, and her two-part study, Pontillo (2007) and Pontillo (2008b), an exhaustive lexical analysis of the attributes of the sea as a theme in the *Rāmāyaṇa* and the *Raghuvamśa*.

Khokhlova (2007), in Russian, deals with various stages in the development of western New Indo-Aryan languages from ergative/absolute back to nominative/accusative alignment, while Vasilyeva (2008), also in Russian, deals with the history of the Urdu language, its origin, the origin of its name, and its further development up to present times.

Three articles by Herman Tieken have also appeared since 2007 which deserve mention here. Tieken (2008a), a 46-page review article of Pollock (2006), takes a detailed look at "vernacularization", "Sanskritization" and the relationship between the two in pre-modern India. Tieken (2008b) then takes a detailed look at vowel harmony in written Tamil and Malayalam, both classical and modern. Finally, Tieken (2009) deals with the particle *kira/kiri* in Apabhraṃśa, suggesting that the uses of this particle in this language are testimony to the "artificial nature" of Apabhraṃśa.

### 3. Grammatical descriptions

The second major area to be discussed here is that of grammatical description, which continues to be a major field of research at European universities. We begin with a number of studies by Annie Montaut dealing with various aspects above all of Hindi/Urdu. Montaut (2007a), a French-language article, takes a closer look at the morpheme *-kar* in Hindi/Urdu with respect to its status as coordination or subordination, while Montaut (2007b), also in French, deals with the notions of reflexive, transitive and intransitive in Hindi/Urdu. Two further studies by this same author deal

with issues related to reduplication in these languages: Montaut (2008) deals with the various meanings of reduplication as a linguistic operation in Hindi/Urdu, while Montaut (2009c), in French, investigates total reduplication, "echo word" expressions (also known as melodic overwriting), as well as pairs of synonyms and antonyms in these languages. Finally, two further articles are dedicated to ergativity in Indo-Aryan: Montaut (2009a) discusses the historical evolution of the past and perfect system in Indo-Aryan, especially in Hindi/Urdu, which she links to an interpretation of the ergative pattern as a stative predication of localization, while Montaut (2009b), in French, deals with ergativity and the classification of Indo-Aryan languages, touching upon a number of issues, most notably that of internal evolution versus contact with other families.

The past three years have also seen an upsurge in the number of studies on the Munda languages of central and eastern-central India at European universities. A number of these are grammatical sketches on Munda languages found in the extensive overview of this family edited by Gregory D.S. Anderson (Anderson, 2008), such as the chapters giving a grammatical overview of the three South Munda languages Gorum (Anderson & Rau, 2008), Gutob (Griffiths, 2008) and Kharia (Peterson, 2008a). Other studies on Munda languages include two studies on Kharia by myself, one in conjunction with Utz Maas. These include a study dealing with reduplication, especially with respect to the masdar (Peterson & Maas, 2009), as well as an extensive Kharia-English lexicon (Peterson, 2009).

There are also a number of projects within the framework of the DoBeS<sup>1</sup> program of the Volkswagen Foundation being conducted by, or in collaboration with, scholars at a number of European research institutes, and a number of descriptive studies have been published by researchers in some of these projects in the past three years. For example, the **Chintang / Puma** project, led by Balthasar Bickel, and the project on **Sri Lanka Malay**, led by Umberto Ansaldo, both of which have now been completed, and the new project on the **Kurumba Languages of the Nilgiris in South India**, led by Christiane Pilot-Raichoor and Frank Heidemann.

Among those studies by scholars involved in the first project just mentioned, a number of works have appeared within the past three years which should be mentioned here, such as Bickel, Banjade et al. (2007), dealing with prefix ordering in Chintang, and Bickel, Gaenszle et al. (2007), which discusses object agreement in Puma. Finally, Rai, Rai et al. (2009) is a Puma-English-Nepali dictionary and grammar of Puma, written in Nepali.

A number of studies on Sri Lanka Malay have also appeared in the past few years by scholars residing in Europe, many of which are directly linked the DoBeS project mentioned above. These include first and foremost the

extensive, three-volume (811 pages) overview of this language by Sebastian Nordhoff (Nordhoff, 2009), the author's Ph.D. dissertation, covering all aspects of the grammar (including history of the language, methodology, phonology, morphology, syntax, valency, grammatical relations and alignment, pragmatics and much more) and a number of texts. In a further study, Ansaldo & Nordhoff (2009) investigate the question of complexity in language creation, using data from Sri Lanka Malay to test various generally accepted assumptions. Finally, although not related to the DoBeS project, Slomanson (2008) should also be mentioned in this respect, which investigates the perfect construction in Sri Lanka Malay, using evidence from negation to argue that this construction is biclausal.

Finally, four further descriptive/typological studies should be mentioned here: Bickel (2008) presents a typological comparison of Tibetan epistemic categories with agreement systems to determine how they are similar and how they differ from one another. Khokhlova (2009a), in Russian, discusses control of reflexivity in Hindi, while Khokhlova (2009b) deals with the position of Hindi with respect to the typology of verbs of motion as being verb-framed or satellite-framed. Two further studies by Anju Saxena deal with Kinnauri, a language of the West Himalayish group of the Tibeto-Burman family: Saxena (2007) discusses contrastive focus in narratives in this language. More specifically, she deals here with the discourse functions of the ergative marker, which she argues is used to mark a change in perspective. Saxena (2008b) then deals with the spatial and temporal extensions of two morphemes in this language. The final study to be mentioned in this category is unique in that it is arguably a "South Asian" language, although it is not spoken in South Asia: Bizri (2009) presents a discussion of what the author refers to as "Pidgin Madam", a contact language which has arisen between female speakers of Sinhalese working as domestics in Lebanon and their Arabic-speaking employers.

#### 4. Computational linguistics

A large number of studies in this report come from the field of computational linguistics, many of them stemming from two multi-authored studies which have recently appeared, co-edited by Gérard Huet – Huet, Kulkarni & Scharf (2009) and Kulkarni & Huet (2009). These two collections contain revised selected and invited papers from the first, second and third international symposia in Rocquencourt, France, Providence, Rhode Island, USA and Hyderabad, India, respectively, dealing with various aspects of Sanskrit computational linguistics, many by authors residing in Europe.

We can begin here with a number of studies by Pawan Goyal and his associates: Goyal, Arora & Behera (2009) presents work towards building a dependency parser for Sanskrit, while Goyal, Kulkarni & Behera (2009) discusses programming concepts, techniques and paradigms used by Pāṇini. Finally, Goyal & Sinha (2009) deals with translation patterns between English and Sanskrit and Hindi and Sanskrit. Other studies contained in these two volumes by scholars working in Europe (and their associates) include Hellwig (2009), which describes a hybrid dependency tree parser for Sanskrit sentences; Houben (2009), which argues that confronting the Pāṇinian system of grammar with theories of constructional grammar and cognitive linguistics yields new perspectives on this ancient grammar; Huet (2009), which discusses "the mathematical structure of various levels of representation of Sanskrit text in order to guide the design of computer aids aiming at useful processing of the digitalized Sanskrit corpus." (Huet, 2009: 162); a work by Malcom Hyman (Hyman, 2009), summarizing Pāṇini's treatment of sandhi, his notational conventions, as well as other formal aspects of his theory; two studies by Anand Mishra – Mishra (2009a), which proposes a Pāṇinian system of Sanskrit grammar for computer representation, and Mishra (2009b), which provides a discussion of an extended version of the author's previously developed model for a computer representation of Pāṇinian grammar; finally, Petersen (2009) provides a formalization of Pāṇini's technique in the *Śivasūtra* for denoting sound classes.

Other works which should be mentioned here include various works by Anju Saxena and her colleagues, such as Saxena & Lind (2008), which describes the experiences made with the use of linguistic corpora in grammar teaching and learning at Uppsala University, Saxena et al. (2008), which presents a brief description of on-going work on the Uppsala Hindi Corpus, and Saxena et al. (2009), which describes the authors' work on two parallel treebanks (Swedish-Hindi-English and Swedish-Turkish-English) and their value for teaching and research.

There are also a considerable number of computational-linguistic studies by Andrew Hardie and his associates which deserve mentioning here. Hardie (2007a) deals with the problems of converting texts in the so-called "legacy" scripts into Unicode. A number of other studies deal with the status of adpositions in Nepali (and elsewhere) based on electronic corpora. These include Hardie (2007b), which deals with collocational properties of adpositions in English and Nepali, Hardie (2008), a collocation-based approach to Nepali postpositions, and Hardie & Mudraya (2009), a contrastive analysis of adpositions in English, Nepali and Russian. Finally, Hardie and his associates have produced a number of studies dealing with questions of corpus linguistics and South Asian languages, such as Hardie (2009), which presents a survey of work to date in the field as well as a brief exploratory study of the lexical differences between Hindi and Urdu, Hardie et al. (2009), which deals with questions of tagging for the Nepali National Corpus (NNC), and Yadava et al. (2008), which again deals with the NNC and questions of tagging.

Three further works which will prove useful to those interested in Sanskrit should also be mentioned here: Gérard Huet's *Sanskrit Heritage Dictionary*<sup>2</sup> and his massive (547pp.) Sanskrit-French dictionary, the *Héritage du Sanskrit. Dictionnaire sanskrit-français*, which is regularly updated.<sup>3</sup> Finally, Oliver Hellwig's highly useful *Digital Corpus of Sanskrit* (DCS), a searchable collection of lemmatized Sanskrit texts, hosted by the Heidelberg Research Architecture (HRA) at the University of Heidelberg, deserves mention here.<sup>4</sup>

## 5. Educational works

A considerable number of works designed for language learners have also appeared in the past three years. Indira Adilyevna Gazieva has published a number of Russian-language educational works in the past three years dealing with various aspects of Hindi and its usage in different contexts, such as business writings (Gazieva, 2007a) and geography (Gazieva, 2007b), in addition to a textbook for self-study for language learners (Gazieva, 2009b). Other works include Gazieva (2007d), dealing with the use of conventional and electronic dictionaries in teaching Hindi, Gazieva (2007e), dealing with finding appropriate teaching materials for pupils learning Hindi in secondary schools, and finally, together with four of her colleagues (Gazieva, Grishina, Kononova, Senichkina & Pereverzeva, 2007) a collec-

tion of authentic texts on various topics related to Indian culture and politics for students of Hindi.

Further works in this category include various works written in German by Rainer Krack and Daniel Krasa, such as works on Hindi for Bollywood fans (Krack & Krasa, 2007; Krasa, 2007a), Hindi slang (Krack & Krasa, 2008), two introductory works on Hindi for beginners (Krasa, 2009a, b), and an introductory textbook on Marathi (Krasa, 2007b). Finally, we should mention here Christophe Vielle's Malayalam manual (Vielle, 2008), written in French, which contains a brief overview of the Malayalam script and its history as well as five introductory chapters to the language.

## 6. Other studies

There are also a number of studies which either do not fit well into the categories in the preceding pages or which are contained in multi-authored works dealing with a number of different topics. We will now take a closer look at these.

A considerable number of studies by European researchers appear in Masica (2007), a collection of papers growing out of the Fifth International Conference on South Asian Linguistics (ICOSAL-5), held in Moscow in July, 2003. These are, in order of their appearance there, Oranskaia (2007), dealing with lexical terms in Indo-Aryan languages denoting 'fear', Sigorskiy (2007) on case marking in Hindi from a historical perspective; Khokhlova & Singh (2007a), dealing with resultative constructions with an overt agent / possessor in western NIA languages, concentrating on Punjabi, Gujarati and Marwari, in addition to Hindi-Urdu; Liperovsky (2007) on the marking of actants in Braj Bhasa in comparison with Modern Standard Hindi; Chevillard (2007b) on "syntactic duality" in poems in Classical Tamil, with two contrasting poles, "marked" or "M-syntax" and "unmarked" or "U-syntax"; Peterson (2007a), which provides an alternative analysis of Kharia (South Munda) which does not make use of traditional parts of speech such as "noun", "adjective" and "verb"; Sitchinava (2007), dealing with markers of past time in Santali (North Munda); and Zakharyin (2007) on Pāṇini's *lopa* and "zeroing" in Modern Standard Hindi.

Other titles which fall in this category include Asher (2008), the first chapter in Kachru et al. (2008), the extensive volume dealing with the languages of South Asia. Asher's study presents an overview of the linguistic situation in South Asia within a historical context; Bickel et al. (2009) is a probabilistic typology for the distribution of phonological word domains,

based on a sample of 70 typologically diverse languages, including a number of South Asian languages of Indo-European, Sino-Tibetan and Austro-Asiatic stock, and Rai, Bickel et al. (2009), dealing with the *mundum*, the oral tradition among the Kiratis in eastern Nepal, and Chintang ritual language.

A large number of studies by Russian scholars fall under this category as well, such as various studies by Indira Adilyevna Gazieva and her associates; Gazieva (2007c), in Russian, deals with onomatopoeic words in Hindi, while Gazieva (ed., 2008) contains a number of articles from participants at the conference "Oriental languages and cultures", held at the Russian State University for the Humanities in Moscow in November, 2007, two of which, Khokhlova (2008) and Sigorskiy (2008), deal with South Asian languages. Gazieva (2009a), in Russian, on the other hand, deals with problems of translation between Hindi and Russian.

The next two articles by this author and her colleagues are more sociolinguistic in nature, such as Gazieva (2008), which deals with the frequency of "Hinglish" (Hindi and English used together) and "Indlish" (Indian English) in a number of domains, such as movies, literature, and elsewhere, while Gazieva, Grishina, Kononova & Senichkina (2007), in Russian, deals with the issue of gender discourse in Indian society.

There are also a number of studies by Liudmila Khokhlova and her associates in this category: Khokhlova (2009c, in Russian, and 2009d) discuss the issue of Punjabi ethnicity, dealing with the role of religion and language in the ethnic self-identity of Punjabis; Khokhlova & Kozlova (2009), in Russian, takes a detailed look at various metaphors denoting *pain* in Hindi; and Khokhlova & Singh (2007b), also in Russian, is a comparative description of direct and metaphorical meanings of a number of verbs of motion in liquid from Hindi, Punjabi, Rajasthani and Gujarati.

Finally, Montaut (2009d) is a sociolinguistic article which takes a closer look at the use of "regional" or "vernacular" languages by Indian authors and the bewildered reactions such writers often evoke; Peterson (2008b) is a brief, German-language introduction for non-specialists to the linguistic richness and variation of Jharkhand; and Saxena (2008a) presents an overview of a number of languages of the South Asian subcontinent for non-specialists.



## 7. Concluding remarks

As the preceding pages have shown, research on South Asian languages is taking place throughout much of Europe, and there are few aspects of these languages and few language families that are not currently receiving attention in Europe. Also, at least three major research projects funded by the Volkswagen Foundation have either left their mark on the linguistic scene or are now setting out to do so.

The large number of individual homepages dedicated to these languages and their study is a further indication of the vibrancy of this field. Although there are far too many such pages to mention here, a few deserve mention, as they provide an important starting point for those who are trying to orient themselves in a particular field. Some of the more important sites run by individuals or research units in Europe include the following: My own *Bibliography of Seldom Studied and Endangered South Asian Languages*,<sup>5</sup> *Resources for the Study of South and Southeast Asian Languages and Cultures* at the University of Göttingen,<sup>6</sup> which also includes the highly useful page *GRETIL – Göttingen Register of Electronic Texts in Indian Languages and related Indological materials from Central and Southeast Asia*,<sup>7</sup> *Savifa, the Virtual Library South Asia* at the South Asia Institute of the University of Heidelberg,<sup>8</sup> and of course the indispensable *Indology* homepage,<sup>9</sup> in addition to the sites by Gérard Huet and Oliver Hellwig, mentioned above in section 4.

And finally, a new series came into existence in 2009, edited by myself and Anju Saxena, which is dedicated entirely to the languages of South and Southwest Asia, *Brill's Studies in South and Southwest Asian Languages*.

All these facts suggest that research on South Asian languages in Europe is in good condition, and it cannot be denied that the field is currently extremely vibrant and multifaceted. However, in past reports, e.g., Peterson (2007b), I struck a rather pessimistic note, as at that time a disturbingly large number of linguistic, Indological and other related departments had either already been closed or were set for closure. Unfortunately, little has changed in this respect since then, other than that more of these departments have indeed been closed. As the cuts have tended to hit the smallest departments the hardest, this usually means that Indological departments take the brunt of these financially motivated cuts. With that, work on South Asian languages in Europe is gradually becoming an area which is more and more dependent on the research interests of individual linguists and less a matter of institutionalized research programs. It remains to be seen

whether the current high level of studies can be maintained in the coming years, especially given the current difficult economic situation.

## Acknowledgements

A report such as the present one cannot possibly be compiled without the help of many others, and I would like to take this opportunity to thank all who have helped me in my work by sending me details on both their own works as well as that of their colleagues, and often by answering many further questions. I apologize in advance if I have overlooked anyone here: Ashok Aklujkar, Emilie Aussant, Jean-Luc Chevillard, Anne Daladier, Rahul Peter Das, Elisa Freschi, Andrew Hardie, Oliver Hellwig, Hans Heinrich Hock, Gérard Huet, Annie Montaut, Sebastian Nordhoff, Tiziana Pontillo, Anju Saxena, René Schiering, Hanna Thompson, Herman Tieken and Christophe Vielle. Special thanks also to Indira Adilyevna Gazieva, Liudmila Khokhlova and Ludmila Vasilyeva, not only for sending me the details of their works but also for all their patience while I struggled with the Cyrillic alphabet, and to Helena Olfert and Alena Witzlack-Makarevich for checking to make sure that my attempts at Russian orthography were successful. Thanks to all of you!

Of course, it goes without saying that I alone am responsible for any and all inaccuracies contained in this report.

## Notes

1. DoBeS – "Dokumentation bedrohter Sprachen" or "Documentation of endangered languages". For more information on these and other projects, see <http://www.mpi.nl/DOBES/>
2. Available under <http://sanskrit.inria.fr/DICO/index.html>
3. Available under <http://sanskrit.inria.fr/Dico.pdf>
4. Available under <http://kjc-fs-cluster.kjc.uni-heidelberg.de/dcs/index.php>
5. <http://www.SouthAsiaBibliography.de>
6. [http://www.sub.uni-goettingen.de/ebene\\_1/fiindolo/fiindole.htm](http://www.sub.uni-goettingen.de/ebene_1/fiindolo/fiindole.htm)
7. [http://www.sub.uni-goettingen.de/ebene\\_1/fiindolo/gretil.htm](http://www.sub.uni-goettingen.de/ebene_1/fiindolo/gretil.htm)
8. [http://www.savifa.uni-hd.de/home\\_en.html](http://www.savifa.uni-hd.de/home_en.html)
9. <http://indology.info/>

## Literature

- Anderson, Gregory D.S.  
 2008 *The Munda Languages*. (Routledge Language Families). London: Routledge.
- Anderson, Gregory D.S. and Felix Rau.  
 2008 Gorum. In Gregory D.S. Anderson (ed.), 381-433.
- Ansaldò, Umberto and Sebastian Nordhoff  
 2009 Complexity and the age of language. In Enoch O. Aboh & N. Smith (eds.), *Complex processes in new languages*. 345-363. Amsterdam: Benjamins.
- Asher, R.E.  
 2008 Language in historical context. In Braj B. Kachru, Yamuna Kachru & S.N. Sridhar (eds.), 31-48.
- Aussant, Émilie  
 2009 *Le nom propre en Inde. Considérations sur le mécanisme référentiel*. Lyons: ENS Editions. [*The proper noun in India. Considerations of the referential mechanism*, in French.]
- Bickel, Balthasar  
 2008 Verb agreement and epistemic marking: a typological journey from the Himalayas to the Caucasus. In Brigitte Huber, Marianne Volkart, Paul Widmer & Peter Schwieger (eds.), *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*. 1-14. (Beiträge zur Zentralasienforschung 12). Halle (Saale): International Institute for Tibetan and Buddhist Studies.
- Bickel, Balthasar, Goma Banjade, Martin Gaenszle, Elena Lieven, Netra Paudyal, Ichchha Purna Rai, Manoj Rai, Novel Kishor Rai and Sabine Stoll  
 2007 Free prefix ordering in Chintang. *Language* 83: 1-31.
- Bickel, Balthasar, Martin Gaenszle, Arjun Rai, Prem D. Rai, Shree K. Rai, Vishnu S. Rai and Narayan P. Sharma (Gautam)  
 2007 Two ways of suspending object agreement in Puma: between incorporation, antipassivization, and optional agreement. *Himalayan Linguistics* 7: 1-18.
- Bickel, Balthasar, Kristine A. Hildebrandt and René Schiering  
 2009. The Distribution of Phonological Word Domains: A Probabilistic Typology. In Janet Grijzenhout & Baris Kabak (eds.), *Phonological Domains: Universals and Deviations*. 47-75. (Interface Explorations, 16). Berlin: Mouton de Gruyter.

- Bizri, Finda  
2009 Sinhala in Contact with Arabic: the birth of a new pidgin in the Middle East. In Rajendra Singh (ed.), *Annual Review of South Asian Languages and Linguistics, 2009*. 135-148. (Trends in Linguistics, Studies and Monographs 190). Berlin / New York: Mouton de Gruyter.
- Candotti, Maria Piera and Tiziana Pontillo  
2007 The (in)separable parts of a plant in the *Mahābhāṣya* imagery i.e. how nature may inspire a grammarian. *Pandanus '07*: 43-63.
- Chevillard, Jean-Luc  
2007a Le rôle de l'exemple dans l'argumentation: le cas de la littérature tamoule savante. *Langages* 166: 32-46. [The role of the example in argumentation: the case of learned literary Tamil, in French.]  
2007b Syntactic Duality in Classical Tamil Poems. In Colin P. Masica (ed.), 177-210.  
2008a *Companion Volume to the Cēṇāvaraiyam on Tamil Morphology and Syntax (Le commentaire de Cēṇāvaraiyar sur le Collatikāram du Tolkāppiyam : Volume 2)*. (Collection Indologie 84.2). Pondichery: Institut Français de Pondichéry / Ecole française d'Extrême-Orient.  
2008b The concept of *ticai-c-col* in Tamil grammatical literature and the regional diversity of Tamil classical literature. In M. Kannan (ed.), *Streams of Language: Dialects in Tamil*. 21-50. Pondichery: French Institute of Pondicherry.  
2009a The Pantheon of Tamil grammarians : a short history of the myth of Agastya's twelve disciples." Colas Gérard & Gerschheimer Gerdi, (eds), *Écrire et transmettre en Inde classique*. (Études thématiques N°23). 243-268. Paris: École Française d'Extrême-Orient.  
2009b The Metagrammatical vocabulary inside the Lists of 32 Tantrayukti-s and its Adaptation to Tamil: Towards a Sanskrit-Tamil Dictionary." Eva Wilden (ed.), *Between Preservation and Recreation: Proceedings of a workshop in honour of T.V. Gopal Iyer*. 71-132. (Collection Indologie, 109). Pondichery: Institut Français de Pondichéry / Ecole française d'Extrême-Orient.
- Chevillard, Jean-Luc, B. Colombat, J.-M. Fournier, J.-P. Guillaume and J. Lallot  
2007 L'exemple dans quelques traditions grammaticales (formes, fonctionnement, types). *Langages* 166: 5-31. [The example in select grammatical traditions (forms, functioning, types), in French.]
- Freschi, Elisa  
2008 Structuring the chaos: Bhāṭṭa Mīmāṃsā hermeneutics as depicted in Rāmānujācārya's *Śāstraprameyapariccheda*. *East and West* 58/1-4: 157-184.

Gazieva, Indira Adilyevna

- 2007a *Хинди. Деловое письмо: материалы для практических занятий по написанию деловых писем.* Moscow: Sapport ST. [Hindi. Business letters: materials for training Hindi on writing business letters, in Russian.]
- 2007b Проектная методика в обучении языку хинди в Российском Государственном Гуманитарном Университете. In L.K. Geuxman, E.V. Demidova, I.I. Zuyev, D.A. Kotkovjес, M.A. Mosina & A.L. Nazarenko (eds.), *Иностранные языки в дистанционном обучении.* Perm: Perm State Technical University. 283-287. [The Project method in teaching Hindi at Russian State University for the Humanities, in Russian.]
- 2007c Характеристика звукоподражательных слов в хинди. In E.V. Ryabceva, M.N. Makeyeva & A. A. Arestova (eds.), *Альманах современной науки и образования*, volume 1. Tambov: Gramota. 69-71. [Characteristics of onomatopoeic words in Hindi, in Russian.]
- 2007d Позиционирование обычных и электронных учебных словарей в парадигме обучения языку хинди. In A.A. Atabekova, *Профессионально-ориентированное обучение иностранному языку и переводу в ВУЗе.* Moscow: Издательство Московского Государственного Открытого Университета (МГОУ). 75-78. [The positioning of conventional and electronic dictionaries in the educational paradigm of teaching Hindi, in Russian.]
- 2007e О трудностях составления учебных пособий по языку хинди для учащихся младших классов российской школы с углубленным изучением языка хинди. In N.V. Kremļjeva (ed.), *Информационно-Образовательные Ресурсы Обучения Языку и Культуре: Традиции и Иновации.* Moscow: Russian State University for the Humanities. 29-30. [The difficulties of compiling of teaching materials on Hindi for pupils of Russian school where Hindi is taught, in Russian.]
- 2008 "Contact languages in modern India. In D. Susniene (ed.), *Nation and Language: Modern aspects of socio-linguistic development. October 2008. Proceedings of the 3rd International Conference.* Kaunas, Lithuania: Kaunas University of Technology (KTU), Panevėžys Institute Centre of Languages. 29-32.
- 2009a Особенности синонимии юридической лексики в языке хинди. In A.A. Atabekova (ed.), *Профессионально-ориентированное обучение иностранному языку и переводу в вузе.* Moscow: Российский Университет Друзвы Надоров (RUDN). 80-83. [The features of the synonymy of legal vocabulary in Hindi, in Russian.]

- 2009b *Язык хинди за один месяц. Самоучитель разговорного языка.* Moscow: AST, Vostok-Zapad. [Hindi per month. Teach yourself Hindi. For beginners, in Russian.]
- Gazieva, Indira Adilyevna (ed.)  
2008 *Oriental languages and cultures.* Cambridge: Cambridge Scholars Publishing.
- Gazieva, Indira Adilyevna, A. Grishina, T. Kononova, A. Senichkina and S. Pereverzeva  
2007 *Индия: культура и политика. Сборник учебно-методических материалов по языку хинди.* Moscow: Sapport ST. [India: Culture and Politics. The collection of teaching materials for Hindi language, in Russian.]
- Gazieva, Indira Adilyevna, A. Grishina, T. Kononova and A. Senichkina  
2007 Гендерный дискурс в индийском измерении. In O.B. Smirnova (ed.), *Язык и межкультурная коммуникация.* Astrakhan: Издательский Дом "Astrakhan University". 94-97. [Gender Discourse in the Indian dimension, in Russian.]
- Goyal, Pawan, Vipul Arora and Laxmidhar Behera  
2009 Analysis of Sanskrit text: Parsing and semantic relations. In Gérard Huet, Amba Kulkarni & Peter Scharf (eds.), 200-218.
- Goyal, Pawan, Amba Kulkarni and Laxmidhar Behera.  
2009 Computer simulation of *Aṣṭādhyāyī*: Some insights. In Gérard Huet, Amba Kulkarni & Peter Scharf (eds.), 139-161.
- Goyal, Pawan and Mahesh K. Sinha  
2009 Translation divergence in English-Sanskrit-Hindi language pairs. In Amba Kulkarni & Gérard Huet (eds.), 134-143.
- Griffiths, Arlo  
2008 Gutob. In Gregory D.S. Anderson (ed.), 633-681.
- Hardie, Andrew  
2007a From legacy encodings to Unicode: the graphical and logical principles in the scripts of South Asia. *Language Resources and Evaluation* 41(1): 1-25.  
2007b Collocational properties of adpositions in Nepali and English. Proceedings of the Corpus Linguistics 2007 conference. [http://ucrel.lancs.ac.uk/publications/CL2007/paper/88\\_Paper.pdf](http://ucrel.lancs.ac.uk/publications/CL2007/paper/88_Paper.pdf)  
2008 A collocation-based approach to Nepali postpositions. *Corpus Linguistics and Linguistic Theory* 4(1): 19-62.  
2009 Corpus linguistics and the languages of South Asia: some current research directions. P. Baker (ed.), *Contemporary Approaches to Corpus Linguistics.* London: Continuum. 262-288.

- Hardie, Andrew, R.R. Lohani, B.R. Regmi, and Y.P. Yadava  
2009 A morphosyntactic categorisation scheme for the automated analysis of Nepali. Rajendra Singh (ed.), *Annual Review of South Asian Languages and Linguistics 2009*. 171-198. (Trends in Linguistics, Studies and Monographs 190). Berlin / New York: Mouton de Gruyter.
- Hardie, Andrew and O. Mudraya  
2009 Collocational patterning in cross-linguistic perspective: adpositions in English, Nepali, and Russian. *Arena Romanistica* 4: 138-149.
- Hellwig, Oliver  
2009 Extracting dependency trees from Sanskrit texts. In Amba Kulkarni & Gérard Huet (eds.), 106-115.  
no date *Digital Corpus of Sanskrit (DCS)*. Hosted by the Heidelberg Research Association (HRA), University of Heidelberg.  
<http://kjc-fs-cluster.kjc.uni-heidelberg.de/dcs/index.php>
- Houben, Jan E.M.  
2009 Pāṇini's grammar and its computerization: A construction grammar approach. In Amba Kulkarni & Gérard Huet (eds.), 6-25.
- Huet, Gérard  
2009 Formal structure of Sanskrit text: requirements analysis for a mechanical Sanskrit processor. In Gérard Huet, Amba Kulkarni & Peter Scharf (eds.), 162-199.  
no date *Héritage du Sanskrit. Dictionnaire sanskrit-français*. [Heritage of Sanskrit. Sanskrit-French dictionary, in French.]  
<http://sanskrit.inria.fr/Dico.pdf>  
no date *Sanskrit Heritage Dictionary*.  
<http://sanskrit.inria.fr/DICO/index.html>
- Hyman, Malcolm D.  
2009 From Pāṇinian sandhi to finite state calculus. In Gérard Huet, Amba Kulkarni & Peter Scharf (eds.), 253-265.
- Huet, Gérard, Amba Kulkarni and Peter Scharf (eds.)  
2009 *Sanskrit computational linguistics. First and Second International Symposia Rocquencourt, France, October 29-31, 2007, Providence, RI, USA, May 15-17, 2008. Revised selected and invited papers*. (Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence, 5402). Berlin et al.: Springer.
- Kachru, Braj B., Yamuna Kachru and S.N. Sridhar (eds.)  
2008 *Language in South Asia*. Cambridge et al.: Cambridge University Press.

Khokhlova, Liudmila V.

- 2007 Синтаксическая эволюция западных новоиндийских языков в 15-20вв. In I.S. Smirnov (ed.), *Orientalia et Classica: Труды Института восточных культур и античности. Выпуск XI. Аспекты компаративистики*. 2nd edition. Moscow: RGGU, Institut vostochnyx kul'tur i antichnosti. 151-186. [Syntactic evolution of western New Indo-Aryan languages (Hindi-Urdu, Punjabi, Gujarati, Rajasthani) in the 15th-20th centuries, in Russian.]
- 2008 Evolution of case marking and verbal agreement systems in [the] course of ergative-to-accusative changes in western New Indo-Aryan languages. In Gazieva (ed., 2008), 11-13.
- 2009a Проблемы рефлексивизации в хинди в типологическом аспекте. *Vestnik MGU* 13/1: 72-85. [Problems of reflexivization in Hindi in a typological perspective, in Russian.]
- 2009b Opposition of verb-framed versus satellite-framed languages applied to Hindi Data. In A.R. Fatihi (ed.), *Language Vitality in South Asia*. Aligarh: Aligarh Muslim University. 249-260.
- 2009c Исторические причины приоритета религиозных факторов над языковыми в самосознании этнических панджабцев. *Vestnik MGU* 13/4: 157-171. [Historical reasons for the priority of religious over language factors in self-identification of ethnic Punjabis, in Russian.]
- 2009d The role of Punjabi language in self-identification of Punjabi community. *Journal of Pakistan Vision*, 10/1: 1-11.

Khokhlova, Liudmila V. & E. Kozlova

- 2009 Метафоры боли и неприятных ощущений в языке хинди. In V.M. Bricyn, E.V. Raxilina, T.I. Reznikova & G.M. Yavorskaya (eds.), *Концепт боль в типологическом освещении*. Kiev: Nacional'na Akademiya Nauk Urkajny, Instytut Movoznavstrva Im. O.O. Potebni. 229-280. [Metaphors of *pain* and unpleasant bodily sensations in Hindi, in Russian.]

Khokhlova, Liudmila V. and Chandrajit Singh

- 2007a Resultative constructions with overt agent/possessor in Hindi-Urdu and western NIA languages. Colin P. Masica (ed.), 91-108.
- 2007b Глаголы перемещения в жидкой среде и движения жидкости в западных индоарийских языках. In T.A. Maisak, E.V. Rakhilina (eds.), *Aquatotion: Глаголы движения в воде: лексическая типология*. Moscow: Indrik. 513-549. [Predicates denoting motion of objects in a liquid medium and motion of liquid in western New Indo-Aryan languages, in Russian.]



- Krack, Rainer and Daniel Krasa  
 2007 *Hindi für Bollywoodfans, Hindi Slang*. Bielefeld: Reise Know-How Verlag Rump. [*Hindi for Bollywood fans, Hindi slang*, in German.]  
 2008 *Hindi Slang – das andere Hindi*. Bielefeld: Reise Know-How Verlag Rump. [*Hindi slang – the other Hindi*, in German.]
- Krasa, Daniel  
 2007a *Hindi für Bollywoodfans Wort für Wort*. Bielefeld: Reise Know-How Verlag Rump. [*Hindi for Bollywood fans word for word*, in German.]  
 2007b *Lehrbuch des Marathi*. Hamburg: Buske. [*A Textbook of Marathi*, in German.]  
 2009a *Einstieg Hindi für Kurzentschlossene. Audio-Sprachkurs*. München: Hueber Verlag. [*Introduction to Hindi for the spontaneous*, in German.]  
 2009b *Hindi 3 in 1 – Kauderwelsch Jubiläumsband 6: Kauderwelsch Hindi*. Bielefeld: Reise Know-How Verlag Rump. [*Hindi 3-in-1 – Kauderwelsch Anniversary Volume 6: Scraping by in Hindi*, in German.]
- Kulkarni, Amba and Gérard Huet (eds.)  
 2009 *Sanskrit computational linguistics. Third International Symposium, Hyderabad, India, January 15-17, 2009. Proceedings*. [Lecture Notes in Computer Science; Lecture Notes in Artificial Intelligence, 5406]. Berlin et al.: Springer.
- Liperovsky, V.P.  
 2007 Notes on the marking of actants in Braj (in comparison with Modern Standard Hindi). In Colin P. Masica (ed.), 144-152.
- Masica, Colin P. (ed.)  
 2007 *Grammatical and semantic perspectives on South Asian and other languages. Papers growing out of the Fifth International Conference on South Asian Linguistics (ICOSAL-5), held at Moscow, Russia in July 2003*. Delhi: Motilal Banarsidass.
- Mishra, Anand  
 2009a Simulating the Pāṇinian system of Sanskrit grammar. In Gérard Huet, Amba Kulkarni & Peter Scharf (eds.), 127-138.  
 2009b Modelling the grammatical circle of the Pāṇinian system of Sanskrit grammar. In Amba Kulkarni & Gérard Huet (eds.), 40-55.
- Montaut, Annie  
 2007a Le morphème *-kar* en hindi/ourdou: coordination ou subordination? In Isabelle Bril & Georges Rebuschi (eds.), *Coordination et subordination : typologie et modélisation*. 181-192. (Faits de Langues, 28). Paris: Ophrys. [The morpheme *-kar* in Hindi/Urdu: coordination or subordination?, in French.]

- 2007b Réfléchi, transitif et intransitif en hindi/ourdou. In André Rousseau, Daniel Roulland & Didier Bottineau (eds.), *L'Enoncé Réfléchi*. 229-248. (Rivages Linguistiques). Rennes: Presses Universitaire de Rennes. [Reflexive, transitive and intransitive in Hindi/Urdu, in French.]
- 2008 Reduplication and 'echo words' in Hindi/Urdu. In Rajendra Singh (ed.), *Annual Review of South Asian Languages and Linguistics, 2008*. 22-62. (Trends in Linguistics, Studies and Monographs 190). Berlin / New York: Mouton.
- 2009a Ergative and pre-ergative patterns in Indo-Aryan as predications of localization. A diachronic view of past and future systems. In Ali R. Fathi (ed.), *Language Vitality in South Asia*. 295-325. Aligarh: Aligarh Muslim University.
- 2009b Ergativité et classification des langues indo-aryennes : évolutions internes ou contacts externes. In J.-M. Durand & A. Jacquet (eds.), *Centre et périphérie : approches nouvelles des Orientalistes*. Actes du colloque organisé par l'Institut du Proche-Orient Ancien du Collège de France, la Société Asiatique et le CNRS (UMR 7192), les 31 mai et 1er juin 2006. 347-361. (Cahiers de l'Institut du Proche-Orient ancien du Collège de France (CIPOA), 1). Paris: Librairie d'Amérique et d'Orient, J. Maisonneuve. [Ergativity and the classification of the Indo-Aryan languages: internal evolutions or external contacts?, in French.]
- 2009c Réduplication et constructions en écho en hindi/ourdou. *Faits de Langue. Revue de Linguistique*. 9-44. (Cahiers de l'Inalco, 1). [Reduplication and echo constructions in Hindi/Urdu, in French.]
- 2009d But why do you write in Hindi? *Etudes Anglaises (Ecrire l'Inde, romanciers et poètes contemporains)* 62/3: 332-344.
- Nordhoff, Sebastian  
2009 *A Grammar of Upcountry Sri Lanka Malay*. Utrecht: LOT (Ph.D. dissertation, University of Amsterdam). Three volumes.
- Oranskaia, Tatiana  
2007 Words with the meaning 'fear' in Indo-Aryan languages. In Colin P. Masica (ed.), 24-33.
- Petersen, Wiebke  
2009 On the construction of Śivasūtra-alphabets. In Amba Kulkarni & Gérard Huet (eds.), 78-97.
- Peterson, John  
2007a Languages without nouns and verbs? An alternative to lexical classes in Kharia. Colin P. Masica (ed.), 274-303.  
2007b Europe. In Rajendra Singh (ed.), *Annual Review of South Asian Languages and Linguistics 2007*. 113-146. (Trends in Linguistics, Studies and Monographs 190). Berlin / New York: Mouton de Gruyter.  
2008a Kharia. In Gregory D.S. Anderson (ed.), 434-507.

- 2008b Die Sprachen Jharkhands. Ein Beispiel für die sprachliche Vielfalt Südasiens und was wir daraus lernen können. *Südasiens*. Theme of this issue: *Unbehaust? Adivasis im 21. Jahrhundert*. 28/3:88-91. [The languages of Jharkhand. An example of the linguistic diversity of South Asia and what we can learn from it, in German.]
- 2009 A Kharia-English Lexicon. *Himalayan Linguistics Archive* 5. <http://www.linguistics.ucsb.edu/HimalayanLinguistics/>
- no date *Bibliography for seldom studied and endangered South Asian Languages*. <http://www.SouthAsiaBibliography.de>
- Peterson, John and Utz Maas  
2009 Reduplication in Kharia. *Morphology* 19/2: 207-237.
- Pollock, Sheldon  
2006 *The language of the gods in the world of men: Sanskrit, culture, and power in premodern India* (A Philip E. Lilienthal Book in Asian Studies). Berkeley and Los Angeles: University of California Press.
- Pontillo, Tiziana  
2007 The "attributes" of the sea as a "theme" in the *Rāmāyaṇa* and in the *Raghuvamśa*: a lexical analysis. I. *Indologica Taurinensia* 33: 291-317.
- 2008a The edible part of the rice in the *Mahābhāṣya* imagery: what are the husks of rules? What is *a-tantram*? *Pandanus* '08, 1: 79-96.
- 2008b The "attributes" of the sea as a "theme" in the *Rāmāyaṇa* and in the *Raghuvamśa*: a lexical analysis. II. *Indologica Taurinensia* 34: 283-310.
- 2009a La "Libertà del parlante" nell'orizzonte dei commentari pāṇiniani: dal *Mahābhāṣya* alla *Kāśikāvṛtti*. *Atti S.G.M.* III/2 n.s. 51-62. [The "freedom of speech" in the horizon of the Pāṇinian commentaries: from the *Mahābhāṣya* to the *Kāśikāvṛtti*, in Italian.]
- 2009b Late Vedic *rūpakas* based on nature imagery: Ritual identifications as a sort of *alamkāra*-pattern. *Pandanus* '09. 9-24.
- Rai, I.P., B. Bickel, E. Lieven, G. Banjade, M. Gaenszle, M. Rai, N.P. Paudyal, N.K. Rai, S. Stoll and T.N. Bhatta  
2009 *Mundum*: A case study of Chintang ritual language. In R. Mukherjee & M.N. Rajesh (eds.), *Locality, history, memory: the making of the citizen in South Asia*. 20-33. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Rai, V.S., S.P. Rai, N.P. Gautam (Sharma), A. Rai, G.P. Cintan, S. Günter, D. Schackow, M. Gaenszle, J. Pettigrew and B. Bickel.  
2009 *Pumā śabdakoś tathā vyākaraṇ*. Lalitpur: Kirāt Pumā Rāi Tupkhābāṅkhālā. [Puma dictionary and grammar, in Nepali.]

- Saxena, Anju  
 2007 Contrastive focus in Kinnauri narratives. In Peter K. Austin & Andrew Simpson (eds.), *Linguistische Berichte*. Sonderheft 14. 191-206.
- 2008a South Asian Languages. In Peter K. Austin (ed.), *1000 Languages. The Worldwide History of Living and Lost Tongues*. London: Thames & Hudson. 112-135.
- 2008b On *ñum* and *əm* in Kinnauri. In Brigitte Huber, Marianne Volkart, Paul Widmer & Peter Schwieger (eds.), *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*. 153-163. (Beiträge zur Zentralasienforschung 12). Halle (Saale): International Institute for Tibetan and Buddhist Studies.
- Saxena, Anju and Mikaëla Lind  
 2008 Corpora in grammar learning. Evaluation of ITG. In Joakim Nivre, Mats Dahllöf & Beáta Megyesi (eds.), *Resourceful Language Technology: Festschrift in Honor of Anna Sägvall Hein*. 149-158. <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8933>
- Saxena, Anju, Beáta Megyesi, Éva Á. Csató and Bengt Dahlqvist  
 2009 Using parallel corpora in teaching & research: The Swedish-Hindi & Swedish-Turkish parallel corpora. In Anju Saxena & Åke Viberg (eds.), *Multilingualism. Proceedings of the 23rd Scandinavian Conference of Linguistics Uppsala University. 1-3 October 2008*. 93-101. (Studia Linguistica Upsaliensia 8). Uppsala: Acta Universitatis Upsaliensis.
- Saxena, Anju, Pranava Swaroop Madhyasta and Joakim Nivre  
 2008 Building the Uppsala Hindi corpus. *Proceedings of the Second Swedish Language Technology Conference. SLTC 2008*. Stockholm. 11-12. <http://www.speech.kth.se/sltc2008/>
- Sigorskiy, Alexander A.  
 2007 Case, split nominativity, split ergativity, and split accusativity in Hindi: A historical perspective. In Colin P. Masica (ed.), 34-61.
- 2008 Formation of mixed languages in the history of the Hindi languages: Urdu and Hinglish. In Gazieva (ed., 2008). 20.
- Sitchinava, Dmitri V.  
 2007 Past markers in Santali: a typology-oriented approach. In Colin P. Masica (ed.), 304-322.
- Slomanson, Peter  
 2008 The perfect construction and complexity drift in Sri Lankan Malay. *Lingua*, 118/10: 1640-1655.
- Tieken, Herman  
 2008a The process of vernacularization in South Asia. *Journal of the Social and Economic History of the Orient* 5/12: 338-383.

- 2008b Metaphony, or height harmony, in written Tamil and Malayalam. In M. Kannan, *Streams of Language: Dialects in Tamil*. 53-63. (IFP-Publications hors série 6). Pondichéry: Institut Français de Pondichéry.
- 2009 Two peculiar usages of the participle *kira/kiri* in Apabhraṃśa. *Sambodhi* 32: 33-36.
- Vasilyeva, Ludmila
- 2008 От «шахского двора» до языка урду. In I.P. Gluškova (ed.), *Язык до Индии доведет: памяти А.Т. Аксёнова*. Moscow: Oriental Literature Publishing Co. 2008. 275-306. [From "royal court" to Urdu language, in Russian.]
- Vielle, Christophe
- 2008 *Manuel de Malayālam [leçons 1-5]*. Bruxelles: Societas Belgica Indiae Orientalis / Louvain-la-Neuve: Institut orientaliste de Louvain. [S.B.I.O. E-Library, 2] [*Manual of Malayālam [lessons 1-5]*, in French.]  
<http://belgianindology.lalibreblogs.be/list/s-b-i-o-library/manuel-de-malay%C4%81%E1%B8%B7am.html>
- Yadava, Yogendra P., A. Hardie, R.R. Lohani, B.N. Regmi, S. Gurung, A. Gurung, T. McEnery, J. Allwood, & P. Hall
- 2008 Construction and annotation of a corpus of contemporary Nepali. *Corpora* 3(2): 213-225.
- Zakharyin, Boris A.
- 2007 PaaNini's 'LOPA' and zeroing in Modern Standard Hindi. In Colin P. Masica (ed.), 337-348.

# Reviews



**R. K. Agnihotri and H. K. Dewan (Eds.)**  
**Knowledge, Language and Learning.**

2010. Delhi: Macmillan. 276 p., ISBN 023 032 881 4

Reviewed by *Anjani K. Sinha*

The book contains a revised version of papers read at an international seminar on “The Construction of Knowledge” organized by Vidya Bhawan Society, Udaipur, India, in 2004. Nine of them are on aspects of interrelation between mind, cognition, language and knowledge, and thirteen discuss practical experiences of teachers of language and social, mathematical and physical sciences. The common thread that binds these scholars is their interest in and commitment to the growth of knowledge.

Knowledge is a powerful tool for human beings not only because we know but also because we articulate what we know. Language plays a crucial role in giving structured manifestation to our conceptual knowledge. It not only improves our consciousness but also makes the relation between concepts and language problematic. Tista Bagchi argues that it is essential to know how knowledge is constructed, if we want to enable children to acquire it effectively. Rajendra Singh argues that the most effective way to understand the construction of knowledge is to *deconstruct*. He recapitulates the controversy between Internalized Knowledge and Externalized Knowledge and summarizes Carstairs-McCarthy’s (1999) argument regarding the evolution of language. Chomsky (2002: 148) insists that knowledge of language is unique to humans; it is internal to human mind/brain and it does not share any interesting property with “nonhuman primate calls” or “with gestural systems”. Belletti and Rizzi (2002: 151) speculate that “language could have come to existence suddenly through a single mutation” but Singh (pp. 28-29) is not ready to accept “a radical discontinuity between ourselves and other primates”. He is not satisfied with *what* the knowledge of language looks like at this point but wants to know *how* we got to this point. He is fairly convinced by Carstairs-McCarthy’s argument that the most distinguishing features of language--- large vocabulary, duality of patterning, distinction between an NP and a sentence and recursion-- could be accounted for in terms of the modification of the vocal tract, plasticity and secondary increase in the size of the human brain. Though I do not find this argument particularly convincing, it does open up fresh areas of enquiry. Though Chomsky assumes (2002: 146) that language is “the



result of biological evolution”, he also suggests that language faculty is biologically isolated (p.144). He is interested in showing how the language faculty meets *certain conditions* and he wants to discover them “in the course of the process of asking how the language faculty satisfies them” (p.160). Chomsky (1988:157) considers language faculty as a central element of the human mind which “operates quickly, in a deterministic fashion, unconsciously and beyond the limits of awareness and in a manner that is common to the species, yielding a rich and complex system of knowledge, a particular language”. He concludes that “for problem-solving and theory construction, there is nothing so specific”. However, it does not bother him. If a module of human mind is designed to be able to do *X* (e.g., to acquire language), it may *ipso facto* be incapable of doing *Y*. For instance, “bats can echolocate with wonderful facility but understandably the physics of echolocation is beyond them. We can manage the latter but not the former”, he argues (Smith 1999: 161-162).

Chomsky claims that human mind is richly structured in the domain of social reality, in scientific reasoning... and in aesthetic and moral judgments. Modularity, whether of language, vision or morality, is the inner properties of the system, i.e., it is genetically determined. It constitutes a part of what it is to be human. There may not be any direct, verifiable evidence but if a proposal is falsifiable, it is on the correct path. Chomsky accepts that every proposition is questionable. That is why he (1975: 5) observes that a significant notion of language “as an object of rational enquiry can be developed only on the basis of far-reaching abstraction”.

While discussing the relation between language and cognition, Singh considers construction of knowledge and ‘cognition’ broadly as synonyms. Language may not necessarily be ‘thought’ but it allows thought to be expressed. However, as Vygotsky (1962) suggests, there is no thought process without language process, which does not mean that they are identical. Paintings and sculpture may also express thought but they need to be explained through language. Devitt (2006) believes in “intentional realism” and thinks that thought is prior to language which suggests that an organism could logically have the conceptual competence to think certain thoughts without having linguistic competence to express them. It is difficult to decide whether language had priority over thoughts or vice-versa but, undoubtedly, a human being is a biological organism and a thinking organism which acts because it thinks.

According to Bhattacharya (p.39), Chomsky thinks that the faculty of language “is inserted into already existing ‘external’ systems: the sensory motor system and some kind of system of thought both of which are somewhat independent of language” but Chomsky would consider it erroneous.

When he (1988: 161) says that the mind is “modular”, he clearly means that it consists of “separate systems with their own properties”. Of course, the systems interact though we do not know how. Chomsky (2000: 117-118) does not use ‘modularity’ in the sense of Fodor (1983) which keeps to ‘input’ and ‘output’ systems. Though psychological mechanisms are “composed of independent and autonomous faculties like perception of faces and of languages”, they “do not appear to fit within the framework of modularity so narrowly construed”, he asserts. Jayaseelan’s concept of modular view of mind and the place of Universal Grammar with reference to its innate endowment is close to Chomsky’s.

Bhattacharya (pp.41-42) discusses Chomsky’s notion of *phase* (as developed in Chomsky, 2001), differently from Chomsky’s. Chomsky argues that syntactic structures are built up in phases (e.g. a complimentizer phrase or a VP). Once a phase has been produced, the complement/head of the phase undergoes transfer to the PF and LF components and becomes impenetrable to further operations in syntax. Given this concept, it is surprising that Bhattacharya considers phases as “chunks of thoughts” (p.42). If a phase is a somewhat crude representation of thought itself, one wonders what chunks of thought are represented by the italicized phrases in the following lines from Lewis Carroll’s *Jabberwocky*:

*’T was brillig, and the slithy toves*  
*Did gyre and gimble in the wabe:*

Though they do not express chunks of thought, they show that the patterns (e.g. VP, NP) have a degree of life, irrespective of the words that make them up.

Bhattacharya’s view that a word “cannot be decomposed into meaningful sub-units” (p.43) is equally surprising. No proponents of new morphology believe in mindless ‘chopping’ of words. If *marwaanaa* ‘caused to be killed’ can be morphologically derived from *marnaa* ‘to die’ and *maarnaa* ‘kill’ in Hindi, kill can also be conceptually decomposed as {CAUSE [TO DIE]}. Bhattacharya’s plea for “a shared space in the community” of the speakers of a language is welcome and so can be a plea for “shared” knowledge. Of course, they cannot be shared “with eyes wide shut” if one speech community is not ready to accept something simply because it belongs to another community. This point has been made amply clear when Satyanath discusses instances of sharing of knowledge in a pluralistic society. It is through the complex interaction of continuity and change that a knowledge system is constructed and enriched. Such a process has little scope for

counter-posing ‘indigenous knowledge system’ to a so-called alien knowledge system even if it is a colonial legacy.

Amritavalli argues for pragmatic transparency by making idealized form of knowledge interact with the real world knowledge. It is the latter which prevents the speaker from saying:

\*The boy was eaten by the apple.

though he will have no difficulty in saying:

The boy was eaten by monsters.

The prototype theory (Rosch 1975) enables us to name an object on the basis of *apparent*, rather than *real*, transparency. It is because of fuzziness that the theory has come up as a better alternative to the theory of absolute categorization. A *grapefruit* is a citric fruit like an orange or a tangerine, yet it is called a *grapefruit* because it grows in a bunch like grapes. The Arabic word *tamar-al-hind* ‘dates from India’ (from which *tamarind* is derived) is called so even though there is minimal resemblance between *tamar* (*dates*) and *tamarind*.

Mukherji’s semantic and linguistic explanation is a collection of loosely organized thoughts on semantics. One wonders how the syntactic ambiguity of *flying planes* becomes the “ambiguity of a single sound” (p.102) and how it is connected with the next point he discusses, namely, the verbs (such as *believe*) which take a clausal complement. He fails to distinguish the *believe* type verbs from *seem* type verbs though they have been discussed by several generative linguists and philosophers of language. Many issues Mukherji raises in regard to semantic decomposition were raised by generative semanticists such as McCawley (1973, 1979), George Lakoff (1970, 1971) and Fillmore (1971) but he does not refer to them. Had he done so, he would have discovered the difficulty that Fillmore (1972) faced with regard to the decomposition of *kill*, *murder* and *assassinate*. Though all the three verbs have a common point (e.g. X CAUSE Y to DIE), there are basic differences involving the presuppositions attached to their arguments. *Murder* presupposes that the murderer has a malice towards the victim while *assassinate* presupposes that the victim is a celebrity. The question is: Where and how do we incorporate them while positing their semantic primes?

Jean Aitchison’s discussion on the role of *nature* and *nurture* is lucid and well-known (cf. Aitchison 1994, 2007, 2008). Sharma discusses the same issue, presumably from a behaviourist point of view (pp. 132-133). Pandey accepts the significance of both nature and nurture but focuses mainly on the development of phonological theories. He posits the phonological inventories of Angami, Bhumij, Kashmiri and Toda, though he does not discuss them.

Vasishth *et al.* talk about noun modifier preferences with precision and visual clarity (e.g. on p. 170). The results are based on informants' judgements on sentences of relative complexity.

The rest of the papers are reports on down-to-earth practical teaching of various subjects in challenging situations. While Patnaik and Banerjee discuss how to teach "summarization", Jangid reports on a project on Whole Language Approach to the teaching of English to young children in an English medium school in Hyderabad. Her graphs indicating the growth of their literacy in English are self-explanatory. Devaki presents a comparative study of tribal children inside and outside the classroom. Inside, there is one-sided transmission of relevant information from the teacher; outside, the same students narrate their experiences excitedly presumably with some variations. Joshi talks about variations in the composition of his students even though they responded to the same pictures. Bhaduri talks about interpersonal thinking with emphasis on the relationship between their concepts and language. Vandana Puri focuses on the crystallization of thought in children with a view to testing their intelligence whereas Saxena and Mahendroo explain how they developed an interactive and innovative pedagogy for teaching Science at Hoshangabad. Subramaniam and Paliwal analyze their own experiences of teaching history to children in a remote area. They suggest that the matter be presented in the form of tangible images. It is as much true of history as of other social sciences. All these are narrations of felt experience. One wonders if they were ever exposed to what the proponents of basic/Gandhian system of education (such as E.W. Aryanayakam) experienced during 1930-1950. The basic approach was similar, i.e., to teach every subject, including science and mathematics, using whatever tangible was available in the area concerned so that children were gradually exposed to abstraction. The Gandhian approach to education did not succeed because it ignored the urge for change in Indian society, and not because its pedagogy was not innovative.

The two papers which demand special attention are the one on learning mathematics by H.K. Dewan and P. Ashok and the other on facilitating the understanding of physics. Both subjects are dreaded by children almost everywhere because they are taught in a dreadful manner. Perhaps one could pay attention to how uneducated villagers and tribals add, subtract and multiply while conducting business transactions in village markets. Vijaya Varma deserves special appreciation for removing our misunderstanding about some common phenomena of nature. Though students receive correct information in the classroom, it is not realized as a *felt* experience. The papers in this section remind us that pre-school as well as elementary education needs a fresh approach to get the children genuinely in-

volved in what they are learning. As Jean Piaget (1971: 18) observes, mathematics teachers should learn about “the natural psychogenetic development of the logico-mathematical operations” in children. He (1975: 266-345) discusses three forms of knowledge: (a) Innate (of *a priori* categories), (b) logico-mathematical structures and their biological significance, and (c) acquired knowledge of physical environment. It is heartening that the volunteer-teachers involved in the Hoshangabad project have come up with such generalizations.

In brief, reading this book is a rewarding experience if one considers the theory of acquisition of knowledge in the light of experiences of actual practitioners. The book is of immense value to not only them but also to those who are interested in the philosophical and psychological aspects of construction of knowledge and their communication through language.

## References

Aitchison, Jean

1994 *Words in the Mind*. Oxford: Blackwell.

2007 *The Word Weavers*. Cambridge: Cambridge University Press.

2008 *The Articulate Mammal*. London: Routledge.

Bulletti, Adriana and L. Rizzi

2002 In Chomsky, Noam. 2002, 151.

Carstairs-McCarthy, A.

1999 *The Origins of Complex Language*. Oxford: Oxford University Press.

Chomsky, Noam

1975 Knowledge of Language. In Keith Gunderson (ed.) 1975. *Language, Mind and Knowledge*. Minneapolis: University of Minnesota Press, 299- 320.

1988 *Language and Problems of Knowledge*. Cambridge, MA: MIT Press.

2000 *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

2001 Derivation by Phase, in M. Kenstowicz (ed.) 2001. *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 1-52.

2002 *On Nature and Language*. Cambridge: Cambridge University Press.

Devitt, Michael

2006 *Ignorance of Language*. Oxford: Clarendon Press.

Fillmore, Charles

- 1971 Types of Lexical Information. In Danny D. Steinberg and Leon A. Jakobovitz (eds.) 1971. *Semantics*. Cambridge: Cambridge University Press. 370-392.
- 1972 Subject, Speakers, Roles. In Donald Davidson and Gilbert Harman(eds.) 1972. *Semantics of Natural Language*. Dordrecht: Reidel, 1-24.

Fodor, Jerrold J.

- 1983 *The Modularity of Mind*. Cambridge, MA: MIT Press.

Lakoff, George

- 1970 *Irregularity: in Syntax*. New York: Halt, Rinehart and Winston.
- 1971 On Generative Semantics. In Danny D. Steinberg and Leon A. Jakobovitz (eds.) 1971. *Semantics*. Cambridge: Cambridge University Press, 232-296.

McCawley, James D.

- 1973 *Grammar and Meaning*. Tokyo: Taishukan Pub. Com.
- 1979 *Adverbs, Vowels and Other Objects of Wonder*. Chicago: University of Chicago Press.

Piaget, Jean

- 1971 *To Understand is to Invent*. New York: Penguin.
- 1975 *Biology and Knowledge*. Chicago: Chicago University Press.

Rosch, Eleanor

- 1975 Cognitive Representations of Semantic Categories. *Journal of Experimental Psychology* 104:192-233.

Smith, Neil

- 1999 *Chomsky: Ideas and Ideals*. Cambridge: Cambridge University Press.

Vygotsky, Lev S.

- 1962 *Language and Thought*. Cambridge, MA: MIT Press.



***Shreesh Chaudhary***  
**Foreigners & Foreign Languages in India:  
A Sociolinguistic History**

2009. New Delhi: Cambridge University Press. *xiii* + 586 pp.  
ISBN: 978-81-7596-628-4.

Reviewed by *Rajend Mesthrie*

This work (*FFLI*) focuses on the histories of largely standard and official languages which were taken to India: Greek, Hebrew and Sanskrit (chap. 2), Arabic, Persian and Turkish (chap. 3), Armenian, Portuguese, Dutch and French (chap. 4), and English – the latter broken up as English in relation to Indian languages (chap. 5) and English as an Indian language (chap. 6). Gathering material from a number of different traditions and archives, the author provides an impressive account of the impact of these languages upon the development of several of India’s languages from their MIA (Middle Indo-Aryan) to MIL (Modern Indian Language) stage. The focus is mostly on vocabulary development; syntax and morphology – considered the backbone of a language – are not discussed. As such this work is accessible to a non-specialist with an interest in language and the languages of the area, since it falls to some extent into the genre of “external history” rather than strict linguistic history in the classical sense of the term. It is an engaging and detailed account from which the linguist interested in more technical areas like phonology and syntax can glean a great deal from. As Rajendra Singh notes in his preface (p. ix) what is interesting in this account is not the structural linguist’s laws and tenets but “how emperors and ambassadors, harlots and compradore elites, translators and interpreters helped and hindered new linguistic arrivals in India”. This work therefore falls into a new area of interest increasingly called colonial linguistics and having such illustrious precursors as Frances Karttunen’s study of translators and interpreters in the New World (*Between Worlds*, 1994).

One has to admire the wealth of historical and sociolinguistic detail in the work, extending to the Hebrew links with the subcontinent (from 2000 BC onwards) and the Armenian presence (from c. 150 BC). But treating Sanskrit as akin to the other languages of foreigners and invaders is perhaps controversial; whilst most non-partisan accounts would see Sanskrit as a language that entered India from Central Asia, its longevity and evolution



into the modern Indo Aryan languages via Prakrit and subsequent developments give it a special place different from Hebrew and Greek, its co-lingual counterparts of chapter 2. This concern is strengthened by Chaudhary's acceptance of the modern geneticists' view that Dravidian also has its ultimate origins in south-eastern Europe. Moreover, as Chaudhary portrays it, Sanskrit can be considered both a mother language (in the historical linguist's sense of mother and daughter languages) as well as a father tongue (p. 95), insofar as it continues to give immensely to the modern vernaculars, and still has people speaking it as a second or even (putatively according to the censuses) a first language.

It is not possible to go into details of the contact histories of the major foreign languages with the MILS provided in FFLI. It would be better for me to whet the potential reader's appetite by citing some vivid characterizations and facts to savour along the way. The fragility of communication that relies on translators and interpreters and the interests of the latter within a communicative economy is evident in several places. Chaudhary (p.69) cites Lipsius (1974:192), a biographer of Alexander the Great, who quotes a seer's opinion that using interpreters 'is like asking pure water to flow through mud'. On p. 186 Chaudhary cites a well-preserved example of early Urdu from the mouth of an upstart parrot in Humayun's court – surely a unique piece of historical linguistic data. On pp. 251-5 one learns of the impact of the Inquisition upon Konkani in Goa, and of the contest between Franciscans and Jesuits in their attitude to Konkani: the former being against the use of local languages, the latter fully supportive, albeit with proselytising ambitions. We learn (on p. 260) of Brother Richard Stevens whose observations on the languages of India and their relation to Greek and Latin predated the more famous statements by Sir William Jones by almost two centuries. In connection with Sir William Jones we are reminded that he had translated Kalidasa's *Shakuntala* from Sanskrit into English (p. 325). We also hear of the Scotsman John Gilchrist's codifying of bazaar jargon into a grammar and lexicon of Hindustani around 1800, and of his efforts to eschew the words provided by his advisers who gave learned and poetic equivalents from the 'deserts of Arabia', 'the mountains of Persia' or from the 'dark intricate mines and caverns of Sanskrit lexicography' (p. 376) .

While FFLI is rich in detail, it is less certain on theory. Here it suffers in two regards: (a) its generalizations about reasons for vocabulary change, and (b) its recourse to pidgin and creole theory. I take the first point first. On p. 197 we are told that the masses spoke without regard to the origin of words. This is of course a truism, especially where non-religious or religious-cultural terms are concerned (Chaudhary's examples come from words in

Eastern Hindi for ‘within’, ‘no’, ‘enter’, ‘got and ‘began’). Etymology is the business of historical linguists, it plays little part in ordinary language processing and use. In summing up his narrative Chaudhary asks why it is that some borrowings from English and Portuguese survive better than other words from these languages in Hindi, and again why native words also have fluctuating fortunes. His conjectural response is not over-convincing: “In either of these cases, words might occasionally wear out under constant use, may gain or lose [*sic*] one or two features of grammar, vocabulary and pronunciation may also change a little in meaning, but, on the whole they endure”. The idea of wear-and-tear in vocabulary is not in fact endorsed in historical linguistics, with basic terms widely expected to survive over long periods of time: as in the famous Swadesh list: *mother, father, sister, sun, moon, stars, tree, fire, hand, walk* etc. which was constructed on the basis of resistance to wear-and-tear in language after language. One is instead reminded of Jaberg’s dictum (1908) that every word has its own history. We may have to accept that vocabulary replacement is relatively ad hoc apart from some classes of religious terms, taboo, and cultural and technological concepts characteristic of a specific culture. Colonialism and related kinds of conquest, of course, bring in new technologies, status and prestige, so that vocabulary replacement is heightened at such periods. Again, easy cross linguistic generalisations might not be feasible: some cultures borrow more than others. A linguistic ideology in some cultures favours calquing (loan translation) or internal compounding over direct borrowing.

The second shortcoming pertains to the overuse of the P-C-N cycle of pidginisation – creolisation – nativisation in explaining language change. This motif occurs early in the book and is referred to in almost all accounts of the contacts between Indian languages with the different incoming languages of the last 2000 years. Most sociolinguistic writing would not characterize these instances of contact-induced change in India as following a pidgin-creole cycle. If such a claim were to be made, it would have to be robustly motivated and illustrated. The only example of creolisation that is linguistically uncontroversial is the variety of Portuguese in Sri Lanka (see Jackson 1990). All other cases show other kinds of contact-induced change: via adult second language acquisition, bilingualism and language shift. Chaudhary’s discussions often point to the bilingualism of former times, the learning of foreign languages via education and coaching, and bilingual writing. None of these are even vaguely suggestive of pidginisation and creolisation, which rest on the learning of a new code under great time constraints and little access to the TL (target language). Early pidginisation may even involve the formation of a new code for

communication in multilingual societies under social duress, without there being a TL in mind (Baker 1995). It is rare to find an indigenous community in Asia and Africa which evolved a pidgin and later creole (in the strict sense of these terms) in response to an incoming prestige language and shifted to this creole. In the colonies and related dominated territories (like the former ‘protectorates’), the local languages survived well, bilingualism was relatively slow to develop and when language shift did take place, the new language was adopted in a form resembling an L2, rather than a pidgin or creole. Southworth once provided a detailed conjectural account in the 1970s of the evolution of Marathi using the tools of pidgins and creole linguistics. He soon retracted the theory, acknowledging that the kind of changes evident in Marathi compared to its antecedent Prakrit occurred gradually over generations, rather than in a telescoped pidgin-creole crucible.

Fortunately for FFLI, these two criticisms do not seriously detract from the pleasures of reading the work and the treasures of its external history and exemplification. I would recommend to the reader to simply ignore the discussions in the text of matters pertaining to PCN and vocabulary replacement theory. The work is generally well produced, though there was a rather large number of typos, exemplified in the misspelling of *lose* (as *loose*) in the citation above and several other places. In one place the use of plural for generic singular is unintentionally comic: “even Akbar sometimes enjoyed wearing Portuguese dresses” (p. 238). One assertion should be queried: the claim that Gandhi generally wrote in English seems unfair to this champion of vernacular languages (p526). The classics *Hind Swaraj*, *Satyagraha in South Africa* and his *Autobiography* were all originally written in Gujarati, and the newspaper *Indian Opinion* which he founded in South Africa and which survived for over fifty years used Gujarati as a major language, with some contributions in Hindi and Tamil in its earlier phase.

Notwithstanding my critique, FFLI is a work I shall be pleased to have on my shelves. For the future, it would be a wonderful complement to this work for someone to research in similar detail the impact of incoming languages upon the Austronesian and Dravidian languages of India, two major families that predated the Indo-Aryan presence. More especially, the give and take between Sanskrit and Proto-Dravidian would make a wonderful parallel study.

## References

- Baker, Phillip  
 1995 Motivation in Creole Genesis, in P. Baker (ed.) 1995. *From Contact to Creole and Beyond*. London: University of Westminster Press, 3-15
- Jaberg, Karl  
 1908 *Sprachgeographie*. Aarau: Sauerländer.
- Jackson, David  
 1990 *Sing Without Shame: Oral Traditions in Indo-Portuguese Creole Verse*. Amsterdam: Benjamins.
- Karttunen, Frances  
 1994 *Between Worlds: Interpreters, Guides and Survivors*. New Brunswick, NJ: Rutgers University Press.
- Lipsius, Frank  
 1974 *Alexander the Great*. London: Weidenfeld & Nicholson.
- Southworth, Franklin  
 1971 Detecting prior creolization: an analysis of the historical origins of Marathi, in D. Hymes (ed.) 1971. *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press, 255-76.



# Dialogue



# The Standard, (Non-) Rhoticity and Rhythm in Indian English: A Response to Lange

*Pingali Sailaja*

Indian English phonology has been studied more extensively than any other aspect of the variety. Yet, controversy surrounds this area. As will be seen below, contrary positions have been taken in the literature. The issues related to the standard, (non-) rhoticity and rhythm raised by Lange (2009) in her review of my book *Indian English* (Sailaja 2009) are addressed below.

On the issue of the standard, it is best to begin by considering how the standard may be conceived of. A good survey of the different approaches to the standard is presented in Crowley (1997). It could be uniform, excellent or common, or a combination of these, as Crowley argues. Usually, it is the model used in education even if it is used by a minority. The standard, therefore, wherever it is drawn from, becomes a prescriptive model.

The problem of defining standard Indian English applies equally to the definition of Educated Indian English as described by Nihalani et al (1979). First, as CIEFL (1972) says, English is used in India only by the educated. Next, there are several different types of educated Indian English and the bilingualism cline has been described by Kachru (Kachru 1965, 1982). The accent varies considerably across the groups. Even among those who belong to the highest class i.e., at the level of the acrolect, there is considerable variation in accent. Given the several different accents in India, which of these is referred to by Educated Indian English? In fact, after providing a brief description of the differences between British Received Pronunciation and Educated Indian English, Nihalani et al use another term—Indian Recommended Pronunciation—and lay down prescriptive rules. It seems as if EIE is deemed to be the standard since it is considered to be more commonly prevalent. The standard described in my book is an accent which is completely devoid of regional markers. It is an accent considered to be excellent by people and one that they wish to have.

Curiously, this standard cuts across the country and is usually free of regional features that mark the speech of most Indians. Whether the speaker is from Delhi, Calcutta, Bombay, Madras or Hyderabad, this is a speech variety that is Indian, but of a higher status than other varieties. (Sailaja 2009: 18)



This is not an accent that can be associated with any group that is homogenous in terms of its socio-economic, geographical or linguistic background. As stated in the book, this is a variety that people aspire to acquire. Moreover, when overt teaching of pronunciation happens, what I have identified as the standard in my book is used as the model. CIEFL (1972) also describes an accent that is neutral and one that people aspire to. McArthur (2002) illustrates how the aspiration manifests itself.

The proof of rhoticity or non-rhoticity in Indian English lies in the hearing. Several speech samples of speakers from various parts of India were collected for the work and some of them are hosted on the website <[www.lel.ed.ac.uk/dialects](http://www.lel.ed.ac.uk/dialects)> providing the necessary auditory evidence. There are samples that are close to the standard and some that are non-standard, and it will be observed that they are all completely non-rhotic, except one which is occasionally rhotic. To describe all (or most) of Indian English as rhotic is wrong.

McArthur (2002) clearly acknowledges the existence of a variety of English pronunciation that is perceived as standard, carries prestige and one that people aspire to acquire. He traces the source of this standard in a fairly detailed manner. However, he does caution that it is not spoken by a majority of the people, which point I concur with. At the same time, I do believe that there is scope for misinterpretation in his writing. The paragraph that describes the development of a standard accent is immediately followed by a list of Indian English features in which IE is described as rhotic. While I do not wish to second-guess what McArthur intended, it is doubtful whether he meant that the standard is rhotic. The features of the standard as identified by him, I believe, have not been provided by him. Finally, not everyone claims that IE is rhotic. Trudgill & Hannah (2002) state “[t]he English of most educated Indians is non-rhotic” (130).

The point that IE rhythm may not be syllable-timed is argued out in the book. In particular, I question the supposed non-existence of weak forms in connected speech. No speaker of standard Indian English articulates the full form of *and* in the phrase *bread and butter* in a neutral context. One sample on the website, the transcript of which is provided on page 156 begins with “You’ll probably notice that. . .” The speaker does not say “You will probably notice that. . .” Earlier works on this topic do have statements like this: “General Indian is *not* stress-timed” (CIEFL 1972: 8, emphasis as in original). However, this is followed by “[t]iming APPEARS to be mora- or syllable-weight based” (CIEFL 1972: 8, emphasis mine). Other works have simply removed the uncertainty and have stated this as a fact. Most works which claim that IE rhythm is syllable-timed (such as

CIEFL 1972, Nihalani et al 1979, McArthur 2002, Trudgill & Hannah 2002) do not give a single example to illustrate this point. Gargesh (2004) gives one example, which I refute in my work.

In fact, to the best of my knowledge, Prabhakar Babu's (1971a, 1971b) work is the only research that arrives at the conclusion that IE rhythm is neither stress-timed nor syllable-timed. And this conclusion is based on experimental work, small though it is which none of the other works has conducted.

To repeat, the standard in Indian English pronunciation is an aspirational model and not one that is necessarily spoken by the majority. Aspirational models find a place in the curriculum and many schools have given importance to pronunciation. The mere fact of non-rhoticity does not imply that the speaker has a standard accent. "Most non-standard varieties of IE are rhotic.... There are those whose speech would be somewhere in the middle of the cline but they still may have non-rhotic speech" (Sailaja 2009: 20). However, a speaker of standard IE will have a non-rhotic accent.

IE must be viewed more in terms of a set of features that may manifest themselves in the speech of individuals, rather than as a constant. Individual variation is quite considerable in IE. The extent to which Indian features of pronunciation will occur in the speech of an individual varies from person to person. (Sailaja 2009: 19)

The set of features described are "by and large, tendencies rather than absolutes" (Sailaja 2009: 19).

## References

- CIEFL  
1972 *The Sound System of Indian English*. Monograph No.7. Hyderabad: CIEFL.
- Crowley, Tony  
1997 Uniform, excellent, common: Reflections on standards in language. *Language Sciences* 19: 15-21.
- Gargesh, Ravinder  
2004 Indian English: Phonology. In *The Handbook of Varieties of English: vol.1, Phonology*, Edgar W Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie and Clive Upton (eds.), 993-1002. Berlin: Mouton de Gruyter.
- Kachru, B. B.  
1965 The Indianness of Indian English. *Word* 21: 391-410.

Kachru, B. B.

- 1982 Models for non-native Englishes. In *The Other Tongue: English across Cultures* B. B. Kachru (ed.), 31-57. Urbana: University of Illinois Press.

Lange, Claudia

- 2009 Review of Pingali Sailaja *Indian English*. *Annual Review of South Asian Languages and Linguistics*. 213-216.

McArthur, Tom

- 2002 *The Oxford Guide to World English*. Oxford: OUP.

Nihalani, Paroo, R.K. Tongue, Priya Hosali & Jonathan Crowther

- 1979 *Indian and British English: A Handbook of Usage and Pronunciation*. New Delhi: Oxford University Press.

Prabhakar Babu, B. A.

- 1971a *Prosodic Features in Indian English: Stress, Rhythm and Intonation*. Post-graduate research diploma project. Hyderabad: CIEFL.

Prabhakar Babu, B. A.

- 1971b Prosodic features in Indian English: Stress, rhythm and intonation. *CIEFL Bulletin* 8: 33-39.

Sailaja

- 2009 *Indian English*. Edinburgh: Edinburgh University Press.

Trudgill, Peter & Jean Hannah

- 2002 *International English: A Guide to Varieties of Standard English*. London: Arnold.

# Appendices



# Announcements

## The Gyandeeep Prize

We are happy to re-announce the continued availability of this annual prize. It will be awarded to the most outstanding student contribution to *ARSALL*.

## Housekeeping

As it is still our intention to bring out future issues earlier than November/December of each year, potential contributors to *ARSALL* should get in touch with the editor as soon as possible. Our new deadlines are:

*November 1*: initial submission

*March 1*: final versions of accepted papers.

Papers submitted after these deadlines will be processed, but only for a later issue. A paper initially submitted after November 1, 2010, for example, will be considered only for the 2011 issue of *ARSALL*.

Potential contributors are encouraged to send their initial submissions as word and pdf files (on the US letter-size template for contributions to edited volumes). The template is furnished by Mouton at their web-site: <http://www.degruyter.com/cont/imp/mouton/moutonAuthors.cfm>).

Authors of accepted contributions must, however, follow the Mouton style-sheet and furnish their articles as word files (on the appropriate template) and their PDF versions. They must also send a 50 word bio-blurb as a word file at the time of final submission.

We reserve the right not to process papers requiring unnecessary editorial work. We would also like potential contributors whose primary language is not English to have their initial submissions looked at by a competent writer of English



## Notes on Contributors

**Shishir Bhattacharja**, Ph.D. (Montréal) is Associate Professor of French at the University of Dhaka. He was trained in Indology and Linguistics at the University of Sorbonne, Paris and the University of Montreal. His most recent books are *Sanjanani Byakaran* (1998), a collection of articles, written in Bengali, on generative syntax, and *Word formation in Bengali: a Whole Word Morphological description and its theoretical implication* (2007), an exhaustive morphological description of Bengali. (shishir.bhattacharja@gmail.com)

**Probal Dasgupta**, Ph.D. (New York University) has taught in Kolkata, Pune, Hyderabad and is now professor of linguistics at the Indian Statistical Institute, Kolkata. His best known books include *Kathaar kriyaakarmo* (1987), *The otherness of English: India's auntie tongue syndrome* (1993), *Primico* (1977), and *Projective syntax: theory and applications* (1989). (Probal53@yahoo.com).

**Rajesh Kumar**, Ph.D. (Illinois) teaches linguistics at the Indian Institute of Technology Patna, India. He works in the areas of natural language syntax, sociolinguistics, language and education, and language and cognition. He has presented his research at many professional meetings, including annual meetings of the Linguistic Society of America and the Chicago Linguistic Society.

**Chandrika Maliyagoda**, M.Phil. (Delhi) works on Sinhala syntax and semantics.

**Rajend Mesthrie**, Ph.D. (Texas) is Professor of Linguistics at the University of Cape Town. His areas of interest include new varieties of English, pidgins and creoles as well as Indic dialects and Anglo-Saxon. He has published extensively on pidgins, Creoles, koines and language shift. His recent books include *Introducing Sociolinguistics* (Edinburgh University Press, 2000) and *Language in South Africa* (Cambridge University Press, 2002).



**Robert K Painter**, (Ph.D. candidate, State University of New York at Buffalo). Trained originally as a classicist, his research interests include Latin linguistics and philology; theories of sound change and historical experimental phonology; and processes of grammaticalization and lexicalization, particularly in English and Germanic.

**John Peterson**, Ph.D. (Kiel), currently visiting professor of Linguistics at the University of Leipzig, has published extensively on the Indo-Aryan and Munda languages of South Asia. He is founder and co-manager (with Tej Bhatia, Syracuse) of the email-discussion group *Vyākaran* and is co-editor (with Anju Saxena, Uppsala) of the series *Brill's Studies in South and Southwest Asian Languages*. He also manages the "Bibliography for Sel-dom Studied and Endangered South Asian Languages" (<http://www.southasiabibliography.de>).

**Shakthi Poornima**, (Ph.D. candidate, State University of New York at Buffalo) utilizes corpus research techniques in order to explore the syntax-semantics interface of aspect markers, primarily in South Asian languages.

**Pingali Sailaja**, Ph.D. (CIEFL, Hyderabad) is Professor of English at the University of Hyderabad. She is interested in the historical, educational and linguistic aspects of English in India, and in morphology. Her previous publications include *English Words: Structure, Formation and Literature* (2004, Pertinent Publishers, Mumbai) and *IndianEnglish* (2009, Edinburgh University Press). ([sailajapingali@yahoo.com](mailto:sailajapingali@yahoo.com))

**Anjani K Sinha**, Ph. D. (University of Chicago) retired as Professor of Linguistics from the University of Delhi. He has published several research articles in national and international journals in the areas of syntax, semantics, applied linguistics and sociolinguistics. His monograph *Lectures on Chomskyan Paradigm* was published by CIIL, Mysore in 2000 and his book on applied linguistics will soon be published by Cambridge University Press, India.

**Ghanshyam Sharma**, Ph.D. (Agra and Bologna) is on the faculty of Indian Studies at the University of Venice and has worked as adjunct at the University of Bologna. His research interests are in the fields of pragmatic theories of meaning, Hindi grammar, and forms of literary enunciation in Hindi.