Annual Review of South Asian Languages and Linguistics
2007

Trends in Linguistics
Studies and Monographs 190

Editors

Walter Bisang
Hans Henrich Hock
(main editor for this volume)
Werner Winter

Mouton de Gruyter
Berlin · New York

Annual Review of South Asian Languages and Linguistics

2007

edited by
Rajendra Singh

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

♾ Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

ISBN 978-3-11-019583-5
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.

© Copyright 2007 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin
All rights reserved, including those of translation into foreign languages. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording or any information storage and retrieval system, without permission in writing from the publisher.
Cover design: Christopher Schneider, Berlin.
Printed in Germany.

Contents

Editiorial Preface vii

General Contributions

The Nominal Left Periphery in Bangla and Asamiya 3
Probal Dasgupta and Rajat Ghosh

Special Contributions: *Indian English / English in India*

The Nature, Structure, and Status of Indian English 33
Rajendra Singh

Seeking the Holy Grail of Nativeness. 47
Ad Backus

On the Native/Non-native Distinction 55
Rakesh M. Bhatt

The Athletics of English in India 73
Probal Dasgupta

Let's Face the Music: The Multilingual Challenge 87
Claudia Lange

The Notion 'Native Speaker': A Philosophical Response 97
Syed A. Sayeed

Regional Reports

Europe 113
John Peterson

Reviews

Rama Kant Agnihotri – Hindi: An Essential Grammar 149
by *Ayesha Kidwai*

Yamuna Kachru – Hindi 156
by *Pramod Pandey*

P. G. Patel, Pramod Pandey and Dilip Rajgor (eds.) –
The Indic Scripts: Paleographic and Linguistic Perspective 162
by *Tej K. Bhatia*

Kashi Wali – Marathi: A Study in Comparative South Asian Languages . 165
by *Alice Davison*

Dialogue

Whole Word Morphology and its Indian Critics 173
Rajendra Singh

Appendices

Announcements: The Gyandeep Prize / Housekeeping 193

Notes on Contributors 195

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 I think the time has come to go a bit further and incorporate a slightly modified form of such a forum into *Current Trends in Linguistics*. *ARSALL* does exactly that.

Each volume of this annual will have four major sections:

i. **General Contributions** consisting of selected open submissions that focus on important themes and provide various viewpoints.

ii. **Special Contributions** consisting of inter-related or easily relatable, invited contributions on important issues, ranging from the narrowly grammatical to the wide-scope socio-linguistic/socio-political. This 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, Reviews and Abstracts** consisting of reports from around the world, reviews of important books and monographs and abstracts of doctoral theses.

iv. **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 first, issue of *ARSALL* contains Probal Dasgupta and Rajat Ghosh's *The Nominal Left Periphery in Bangla and Asamiya*.

The Special Contributions section is dedicated to 'Indian English/English in India' this year. English has become an important and salient issue in South Asia since the decision of several federal/central and provincial governments of that region to introduce it very early in schools. The contributions here, built around a position paper I was invited to present at the Symposium on 'Indian English/English in India', held at the Central Institute of Indian Languages, Mysore, Jan. 4, 2007. Our contributors come from India (Dasgupta and Sayeed), Europe (Backus and Lange) , and North-America (Bhatt and Singh). Although I am tempted to respond to my interlocutors, I shall refrain from doing so because I do not want to misuse my editorial privilege. I do, however, hope that other interested scholars, particularly scholars from Bangladesh, Nepal, Pakistan, and Sri Lanka, will respond to these contributions by submitting short replies and comments for the Dialogue section of the next issue of *ARSALL*.

As Europe has been a major center for the study of South Asian languages for over two centuries now, we thought it was appropriate to have our Regional Reports section publish a report on European research on South Asian languages in this first issue of *ARSLL*.

The Dialogue section contains an invitation to open a debate regarding the nature and structure of morphology.

The Review section of this issue contains reviews of Pandey et al's significant book on South Asian writing systems, Y. Kachru's recent important book on Modern Hindi, and of the collection of Wali's decade-long insightful explorations of Marathi. These have been written by T. K. Bhatia of Syracuse University, Alice Davison of the University of Iowa, and Promod Pandey of Jawaharlal Nehru University.

I am grateful to Dr. Ursula Kleinhenz of Mouton de Gruyter and Stephen Moran of the Université de Montréal for help far beyond the call of duty in the preparation of this issue.

Rajendra Singh

General Contributions

The Nominal Left Periphery in Bangla and Asamiya

Probal Dasgupta and Rajat Ghosh

This study of left peripheral positions in the nominal construction in Bangla and Asamiya reexamines the way particular configurations of nominal and classifier material correlate with definite and/or specific readings, considers the interaction of these phenomena with nominal specifiers and adjuncts, and presents a range of cases where these adjuncts move from the nominal periphery to clausal topic and focus sites.

1. Introduction

Generative work on the contrast between Bangla/Asamiya indefinite nominals such as (1i, ii) and their definite variants (2i, ii) has given rise to several successive accounts – we return to the details of who said what in our brief survey of the literature in section 2:

(1) i. Bangla: *duTo chele*
two.Cla boy
'two boys'

ii. Asamiya: *duta lOra*
two.Cla boy
'two boys'

(2) i. Bangla: *chele-duTo*
boy-two.Cla
'the two boys'

ii. Asamiya: *lOra-duta*
boy-two.Cla
'the two boys'

Early generative proposals had the quantifier-classifier sequence move from a base position as in (1) to a postnominal position as in (2). Later proposals involved moving some material to the left. There has been disagreement about just what moves and how this movement is triggered. But it is generally agreed that a position within the nominal phrase, more or less in the left section of its area, becomes active in such definite nominal constructions as (2); and that, whatever precise characterization it may turn out to require, this active position follows Dem[onstrative]:

(3) i. Bangla: *Sei chele-duTo*
that boy-two.Cla
'those two boys'

ii. Asamiya: *xei lOra-duta*
that boy-two.Cla
'those two boys'

Either of the two nominal definitizing strategies just illustrated – the presence of an overt Dem and the preposing of nominal material to the active post-Dem position – suffices to mark a particular nominal phrase as definite. However, neither of these is a necessary condition. Certain nominals, call them canonical genitive constructions, do their definitizing work even further to the left:

(4) i. Bangla: [[*ramer* [*boi*]] 'Ram's book'
ii. Asamiya: [[*ramOr* [*kitap*]] 'Ram's book'

We take the stand that the nominal construction (4) is definite by virtue of its structure, and build our analysis around this "CGD" view, "Canonical Genitives are Definite". The point of the bracketing shown in (4) is to refer non-committally to a minimal structure, where the genitive is attached as low as possible in the nominal tree. A genitive construction with such a structure is here called a canonical genitive.

One important goal of the present study is to defend CGD against the natural alternative CGID, the view that Canonical Genitives are Indefinite or Definite. CGID would say of structure (4) that it is ambiguous between the preferred definite reading 'Ram's book' and an also available indefinite interpretation 'some book/s of Ram's'. On that view, (5) and (6), which are unambiguous, would correspond to those readings for (4):

(5) i. Bangla: *ramer boi-dukhana*
Ram's book-two.Cla
'Ram's two books'

ii. Asamiya *ramOr kitap-dukhOn*
Ram's book-two.Cla
'Ram's two books'

(6) i. Bangla: *ramer dukhana boi*
Ram's two.Cla book
'two books of Ram's'

ii. Asamiya *ramOr dukhOn kitap*
Ram's two.Cla book
'two books of Ram's'

The plausibility of CGID rests on evidence from strings like (7), which are clearly ambiguous:

(7) *amra ramer boi poRechi*
we Ram's book have.read
'We have read the book / some book/s by Ram'

Examples such as (7) are presented only in Bangla, for expository convenience. Wherever we make no explicit statement to the contrary, our argumentation is intended to cover Asamiya as well. But we present Asamiya material mainly in the context of the detailed study of the nominal construction, where the specific properties of the Asamiya classifiers are at stake.

To return to string (7), we take the stand that the string is structurally ambiguous. The definite reading is associated, we argue, with a nominal phrase of the (4) format that means 'Ram's book'. In contrast, the indefinite reading arises when the genitive 'Ram's' occupies what classical parametric syntax terminology used to call a non-argument position. If string (7) is given an indefinite construal, this position is either a clause-structural niche outside the nominal phrase or an adjunct position at the left periphery of the nominal phrase. Even in the variant that places 'Ram's' in a clausal niche, it nonetheless binds a trace in the nominal left periphery adjunct position. Part of our discussion will of course focus on that adjunct position, not to be confused with the specifier position that 'Ram's' occupies in structure (4).

The claim that […Ram's [t book]…] is one parse for string (7) enables us to make sense of cases where the displacement to a clausal niche is obvious, such as (8):

(8) *ramer to amra* [*t boi*] *poRechi*
Ram's Prt we t book have.read
[bad English:] 'By Ram, of course, we've read some book/s'

Making this claim also compels us to articulate our views on the nominal left periphery; hence this paper.

We are using the term "left periphery", operationally, to refer to positions preceding Dem. The major nominal left periphery players include such possessors as 'Ram's' in (9). Since issues of the detailed structure of the nominal are at stake in (9), we revert to multilingual display, with Hindi added this time to make a comparative point – more is said on matters of typology in section 2:

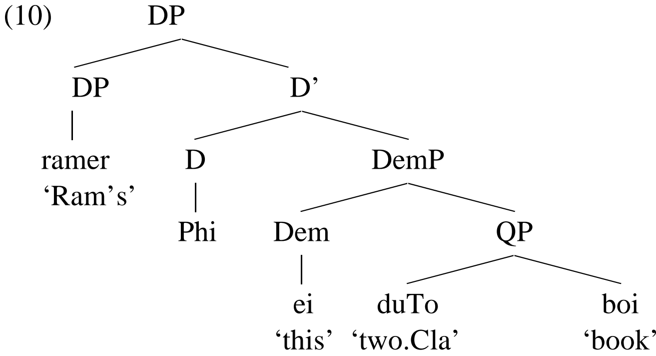
(9) i. Bangla: *ramer ei duTo boi*
Ram's this two.Cla book
'these two books of Ram's'

ii. Asamiya: *ramOr i dukhOn kitap*
Ram's this two.Cla book
'these two books of Ram's'

iii. Hindi: *raam kii ye do kitaabeM*
Ram GenFPl these two books
'these two books of Ram's'

The structure of possessives in the gender-number language Hindi, illustrated at (9iii), is minimally different from what goes on in the classifier languages Bangla and Asamiya, even though all three belong to the Indo-Aryan subfamily of Indo-European. The ways in which the gender-number workings of (9iii) differ from the mechanics of classifiers in (9i,ii) are indeed significant, as will become clear as our argument unfolds. However, at the level of broad parameters, Hindi exhibits much the same nominal construction as its eastern sisters. Provisionally, we formalize our understanding as follows, continuing to use Bangla as our main language of exemplification:

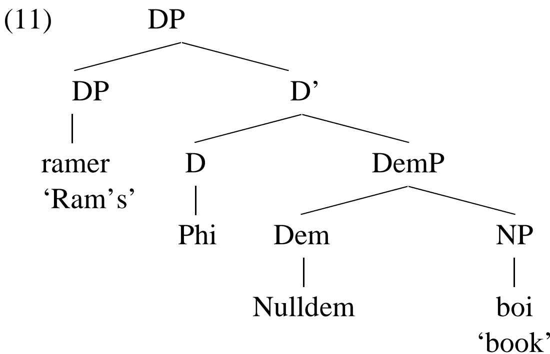
(10)



The phonologically null element Phi in (10) houses nominal identification features. In the gender-number language Hindi, Phi is activated by setting up an association with the immediately subordinate projection, here DemP, in terms of shared features of gender and number. For concreteness, we assume that Dem feature-unifies with D and in effect become a single hybrid head. In the classifier languages Bangla and Asamiya, Phi feature-unifies with subordinate heads to choose either deictic or quantificational referentialization for the DP. Formal devices for these manoeuvres are specified in section 2.

We return now to the canonical genitive construction. On these assumptions, it should have roughly the following structure:

(11)

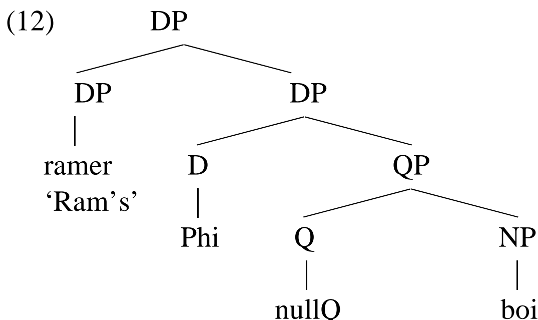


If the Dem node plays any role in ensuring the default definite interpretation, the phonologically null element written as Nulldem in (11) must be involved. We propose that, when a null demonstrative head-moves to D and the null complex at D asks its specifier DP for referential identification

help, the outcome is that the specifier DP obliges, and yields a definite reading. In effect, it is /ramer/ ‘Ram’s’ that is responsible, then, for definitizing /boi/ ‘book’ in such a construction.

We maintain that in certain non-canonical constructions, the genitive phrase is not in the specifier position shown in (10) and (11), but in the adjunct position shown in (12), giving rise to the possibility of the indefinite reading:

(12)



In such a construction the active nullQ plays a key role in fashioning a construal for the host DP to which /ramer/ ‘Ram’s’ is adjoined. Given the nature of nullQ, the reading comes out as indefinite.

Why is the host DP not able to choose between an indefinite and a definite construal depending on whether it contains an active nullQ or an active nulldem? Our stand is that the nulldem choice operates with a genitive only in the specifier position. A nulldem trying to make sense of an adjunction structure would expect the extrasentential context to provide guidance, and the genitive adjunct would fall between the two stools of proximate syntactic control and a remote pragmatic decision. However, little turns on whether this is the right account. If one were to propose instead that the definite reading arises either from (11) or from a variant of (12) where nulldem takes the place of nullQ, then this reading would be associated with two distinct structures.

Why, in the description we are in fact proposing, is (11) unable to resort to the nullQ option and come out with an indefinite reading? Because on our view (11) involves consulting the specifier /ramer/ to settle the issue, and such consultation is compatible only with a nulldem seeking a definite interpretation, not with a nullQ seeking a quantificationally oriented indefinite construal.

An adjunct so loosely connected to the core of the nominal cannot expect help from the core for its own licensing. We propose that the genitive /ramer/ in (12), in adjunct position, is licensed by mechanisms analogous to whatever allows the adjectival phrase double-bracketed in (13) to serve as an adjunct:

(13) [[*taj mOholer ceyeo purono*]] *ei praSad*
Taj Mahal.Gen than.Emph old this palace
'this palace (which is) even older than the Taj Mahal'

Observationally, the overt adjunct is permitted to be an adjectival or participial or genitive structure. The unavailability of other categories indicates that the relevant mechanisms cannot cope with the full range of predicate diversity. These observations constrain the class of mechanisms available; just which descriptive device best fits the facts is an issue we leave open.

We are thus able to cleave to the adjunction structure (12), avoiding the postulation of a Topic projection in the nominal left periphery. This series of decisions gives us a vantage point from which we can open up the issue of how nominal left periphery positions interact with clausal non-argument niches such as the topic position in (8), repeated below, the focus position in (14), and the ambiguous position in (15) – throughout, t is the trace left by 'Ram's':

(8) *ramer to amra* [*t boi*] *poRechi*
Ram's Prt we t book have.read
[bad English:] 'By Ram, of course, we've read some book/s'

(14) *ramer-i amra* [*t boi*] *poRechi*
Ram's-Prt we t book have.read
[bad English:] 'It is by Ram, of course, that we've read some book/s'

(15) *ramer amra* [*t boi*] *poRechi*
Ram's we t book have.read

String (15) can be read either with focal stress on 'Ram's', yielding a focal reading of the (14) type, or without such stress, resulting in a topic interpretation as in (8).

What lends support to CGD rather than to CGID is the fact that all three sentences robustly provide only an indefinite reading for /boi/ ‘some book/s’. This is unexpected under CGID assumptions. To make any further headway, we need to put some cards on the table.

2. The comparative and theoretical context

Historically Bangla and Asamiya (a.k.a. Bengali and Assamese) belong to the Eastern group within the Indo-Aryan subfamily of the Indo-European language family. Bangla is spoken in West Bengal and in the Republic of Bangladesh; Asamiya is spoken in Assam. A typological overview helps locate the salient phenomena of these languages in a global context. In both languages, a numeral normally appears in a “classification” format marking the semantic class of the nominal phrase. We repeat some examples from the beginning of section 1 to highlight the phenomenon:

(1) i. Bangla: *duTo chele*
two.Cla boy
‘two boys’

ii. Asamiya: *duta lOra*
two.Cla boy
‘two boys’

(2) i. Bangla: *chele-duTo*
boy-two.Cla
‘the two boys’

ii. Asamiya: *lOra-duta*
boy-two.Cla
‘the two boys’

(3) i. Bangla: *ramer boi-dukhana*
Ram’s book-two.Cla
‘Ram’s two books’

ii. Asamiya *ramOr kitap-dukhOn*
Ram’s book-two.Cla
‘Ram’s two books’

(4) i. Bangla: *ramer dukhana boi*
Ram's two.Cla book
'two books of Ram's'

ii. Asamiya *ramOr dukhOn kitap*
Ram's two.Cla book
'two books of Ram's'

Notice that the classification formats in Bangla, neutral /XTo/ and objectual /Xkhana/, closely match those in Asamiya, /Xta, XkhOn/, in that they place the same semantic constraints on nominals and are phonologically similar. Most studies describe these formats in terms of "the classifiers /To, khana, ta, khOn/". The theoretical framework we adopt here, as becomes apparent in the course of our argument, does not recognize classifiers as distinct morphosyntactic entities; nonetheless, the term Classifier, and the terms Gender and Number, continue to provide helpful abstract characterizations of categories manifested as morphological formats corresponding to syntactic features.

Both in Bangla and in Asamiya, classifiers in this sense play important roles in quantification and definiteness marking, in collaboration with the deixis system. To see all this in a larger comparative context, we need to place these languages in Allan's (1977) grid of languages that use the classifier device:

(5) a. Numeral classifier languages
b. Concordial classifier languages
c. Predicate classifier languages
d. Intra-locative classifier languages

Bangla and Asamiya belong to Allan's type (5a), but they also exhibit other classifier types whose syntax requires detailed investigation; see S. Ghosh and Dasgupta (2005) for some discussion. Most of the nominal structures in the languages in this type have three distinguishable constituents: a numeral or quantifier (Q), a classifier (Cl) and a noun (N). Within this group, Allan introduces a further subgrouping on the basis of the linear order of these constituents:

(6) a. Q-Cl-N: Amerindian, Eastern Indic, Chinese, Semitic, Vietnamese
b. N-Cl-Q: Burmese, Japanese, Thai
c. Cl-Q-N: Kiriwina (Oceanic)
d. N-Cl-Q: Louisiade Archipelago (Oceanic)

In terms of this subgrouping, Bangla and Asamiya belong to (6a); the sequencing within DP fits the template Q-Cl-N. However, the framework we adopt takes seriously the fact that the Cl never occurs as an independent word in Bangla and Asamiya, and therefore shows the classification-formatted numeral as a single Q item. This is clearly an inappropriate strategy for the enclave of Malayalam structure that exhibits what appear to be classifier words; see Hany Babu (1997) for (pioneering) discussion; for us to venture into Malayalam in this paper would take us too far afield.

We turn now to characteristics that Bangla and Asamiya share and do not share with Hindi. All three languages belong to the Indo-Aryan subfamily and work without an overt definite determiner. They have broadly similar nominal constructions. As we have seen in section 1 at (9) and (10), it makes sense to leave D technically empty in all three languages. However, the "Phi" place-holder that we indicated under D at structure (10) in section 1 works differently depending on the language type. In the agreement language Hindi, Phi's job is to make sure that gender and number are configured consistently throughout the nominal. In classifier languages, which lack grammatical number or gender, phi deals with interpretable features. Classifier languages handle semantic non-singularity in terms of interpretable features such as Individual and Collective expressed in the classifier feature matrix, while gender disappears entirely from the grammar. In response to this particular instance of complementary distribution, Dasgupta & Bhattacharya (1994) proposed an overarching "Badge" category, a device not adopted here. In our present account of classifier languages, Phi works with Dem and/or with Q in ways that are worked out in the course of this exposition.

The earliest work on these issues focused on the alternation in the classifier language Bangla between indefinite (1), with the (classifier-bearing) quantifier preceding the noun, and definite (2), with the noun preceding the quantifier. The first response to this alternation was to postulate a transformation postposing the quantifier. Early publications implementing this idea (Dasgupta 1981; Azad 1983; Dasgupta 1983) reflected a period of relatively

unconstrained syntax, but even as late as 1994, Dasgupta & Bhattacharya were still postulating quantifier postposing. Ghosh (1995) was the first to suggest preposing the nominal instead.

Bhattacharya's (1999, 2000) influential minimalist version (along the lines of Chomsky 1995) of this proposal, still accepted by many authors as the default analysis, has a NP carrying a feature [+*Specific*] move to Spec-QP, where the feature is checked. In the present study, we adopt the core of the idea – that the relevant constituent is specific and moves to Spec-QP. But if the constituent is taken to be an NP that does not include the quantifier, the mechanics of moving it cannot be made to work, and we are forced to reopen the issue of just what it is that moves. Consider the contrast between the facts at (1) and (2), above, and the pattern observed in (6) and (7):

(6) i. Bangla: *dOSTa chele*
ten.Cla boy
'ten boys'

ii. Asamiya: *dOxta lOra*
ten.Cla boy
'ten boys'

(7) i. Bangla: **chele-dOSTa*
boy-ten.Cla
'the ten boys'

ii. Asamiya: *lOra-dOxta*
boy-ten.Cla
'the ten boys'

The ill-formedness of (7i) in Bangla contrasts with the routine availability of its Asamiya counterpart (7ii). This contrast reflects the fact that the association between noun-quantifier order and definiteness is – in Bangla but not in Asamiya – subject to a small numbers constraint. There is no way to embed this fact in a Bhattacharya-type account where a feature in the NP drives the mechanics of NP preposing regardless of the content of the quantifier.

A second problem with the decision to move NP without reference to Q has to do with the impossibility of registering, within such a system, the known lexical incompatibility patterns involving nouns and quantifiers. In Bangla (but not in Asamiya), (8) is ill-formed for many speakers and (9) is

excluded for all speakers, whereas there is no such constraint if the /Xjon/ classifier format is not involved, as we see at (10) and (11):

(8) **chele-dujon*
boy-two.Cla
'the two boys'

(9) **chelejon*
boy.Cla
'the boy'

(10) *boi-dukhana*
book-two.Cla
'the two books'

(11) *boikhana*
book.Cla
'the book'

Such facts have been part of the crucial data set since Azad (1983) and Dasgupta (1983). The question is how to square the desirable features of Bhattacharya's account with the way these observations constrain the class of acceptable solutions.

One promising line of inquiry is opened up by Dhanwar (2004), who proposes a combination of lexical and syntactic devices. His account suggests that Word Formation Strategies in the sense of Ford, Singh & Martohardjono (1997) are responsible for extending the noun word and forming, say, /boi-dukhana/ at (10), and that the lexical nature of this process can accommodate the unavailability of (8) and (9). Likewise, (6) and (7) can be handled by constraining the Strategies. What then moves is an NP whose head N has been extended so that the N, and through percolation the NP, carries quantificational features. If we add that the relevant Strategies have the effect of mapping non-specific onto specific N, this account converges with the core insight of Bhattacharya (2000).

In the present paper, we adopt the substantivist perspective in linguistics in general (see Dasgupta, Ford & Singh 2000 for one exposition) and Ford, Singh & Martohardjono's Whole Word Morphology in particular, differing from Dhanwar's implementation only to the extent of revising the mechanism slightly. We propose, using a device introduced in Dasgupta (2005),

that forms like /boi-dukhana/ 'the two books' at (10) do not reflect a bidirectional Word Formation Strategy mapping from the lexicon to the lexicon. The mechanism in this case is instead a unidirectional Word Extension Strategy mapping from the lexicon onto syntactic subtrees, attaching /dukhana/ as a Q clitic to the N host in the morphology and allowing for relevant syntactic niching of word extensions to ensure that the subtree is well-formed at the syntactic level. In the present instance this means that, when the quantificational feature endowed NP moves to Spec-QP with its head N carrying a Q clitic along with it, the syntactic instructions associated with the Word Extension Strategy ensure that the N remains head of the NP, but that the Q lands in the Q head to whose specifier the NP has moved. In the syntax, Q is outside the NP constituent; it is an extension that has been niched appropriately. In the extended morphology, this Q is a clitic hosted by the NP's N head, preserving Bhattacharya's tree but radically altering his analysis.

Bhattacharya's proposal that /boi-dukhana/ is specific in contrast to non-specific /dukhana boi/ 'two books' works particularly well when these forms follow a demonstrative:

(12) *ei boi-dukhana*
this book-two.Cla
'these two books (specific)'

(13) *ei dukhana boi*
this two.Cla book
'these two books (non-specific)'

In such instances, the demonstrative ensures that the overall nominal is definite, and the contrast between (12) and (13) deals with specificity. However, the fact that (10) itself is definite remains unaddressed if one merely postulates a specificity feature driving a preposing to Spec-QP. We propose that a nominal such as (10), where there is no overt Dem, obtains its definiteness by moving the entire QP to Spec-DemP in order to check a [+definite] feature against a nulldem head attracting it. For concreteness, we assume that the noun /boi/ 'book' optionally picks up a definiteness feature, which, if chosen, must be checked against a nulldem. On this proposal, definiteness driven movement is to Spec-DemP whereas Bhattacharya's specificity driven movement is to Spec-QP, and (10) involves both.

This account of definiteness requires a second look at (13), where nothing has been done at the featural level to mark the nominal as definite. The overt Dem /ei/ 'this' itself in classifier languages does not carry enough features to do the work of referring on its own: sentences like (14), where the bare Dem is alone in its DP, are ill-formed:

(14) **ei bhalo nOY*
this good isn't, for
'This isn't good'

To make (14) work, we need to fortify the Dem with a classifier format or by cliticizing a full, classifier-formatted quantifier:

(15) *eTa bhalo nOY*
this.Cla good isn't
'This isn't good'

(16) *e-duTo bhalo nOY*
this-two.Cla good aren't
'These two aren't good'

On morphological grounds, we propose that /eTa/ 'this' in (15) reflects the work of a Word Formation Strategy whereas, in contrast, /e-duTo/ 'these two' in (16) arises through the application a Word Extension Strategy. For /duTo/ in such forms is a clitic, unlike the /Ta/ fragment of /eTa/.

This machinery enables us to revisit (13), which we now propose should in fact be represented as (17), with /dukhana/ cliticized to /ei/:

(17) *ei-dukhana boi*
this-two.Cla book
'these two books'

In other words, a Word Extension Strategy applies to /dukhana/ 'two.Cla' and extends it to /ei-dukhana/ 'this-two.Cla', thereby endowing it with the feature [+definite]. It is worth noticing that, for reasons that have to do with the way morphological strategies are formulated, the strategy maps from Q to Dem+Q, not from Dem to Dem+Q, since there are fewer Dems than Qs and thus treating the Dem as a constant in the morphological formalism makes for better strategy formulations. However, it is Q, not Dem, that

counts as a clitic. As far as a WES is concerned, each word is an extension of the other; the notion "clitic" is not formally defined in the theory.

To return to our concrete examples, the extended entity Dem+Q is merged at the Q head of QP, but the Dem extension continues to await proper syntactic niching. Definiteness driven movement of this definite QP to Spec-DemP provides the opportunity required. With QP landing at Spec-DemP, the extension /ei/, a definite Dem, is niched in the D immediately to the left. If D is associated with a definite Dem in this case, we must assume that this is uniformly so. Accordingly, we slightly revise the analysis of bare /boi-dukhana/ provided earlier. Definiteness checking, we have said, forces the definite QP to move to Spec-DemP, where it checks the feature against the nulldem head. We now add that nulldem subsequently moves to definite D. If we make reasonable assumptions about proper nouns, it follows that the D of every definite DP even in a classifier language contains a syntactic entity marked as definite.

The manoeuvre we just agreed to involves head movement, which many authors have abandoned. Our proposal is that, given the existence of morphological amalgams of contiguous heads in the case of functional categories (French *au, du, aux, des*, German *im, am, beim*, and so on), head movement should be retained as an option for functional heads, whereas lexical heads are only permitted to move as part of the phrases that they head, following what by now counts as the standard Kaynean procedure. To fine-tune the analysis on that front, along the lines indicated by Dhanwar (2004) but again substituting an extension process for Dhanwar's Word Formation Strategy, we propose for cases like Bangla /boiTa/ (Asamiya /kitapkhOn/) 'the book' that a Word Extension Strategy maps from the noun X to the form /X-Ta, X-khOn/ with a cliticized classifier. In our analysis, such a form is merged at N, endowing the NP with the features Specific and Definite. The specificity feature triggers NP preposing to Spec-QP (with niching of the clitic /Ta, khOn/ in Q). Later, if there is a definite nulldem attracting it, the QP, being definite, moves to Spec-DemP to satisfy the requirements of the definiteness feature. Notice that since N is a lexical head, our system does not allow it to undergo head movement.

The analysis developed here has the advantage of being able to accommodate a phenomenon called 'excapsulation' in earlier work by Ghosh (1999, 2001). Bangla exhibits it only in the case of one classifier format, /Tuku/ – which is best regarded as a massifier in the sense of Cheng & Sybesma (1999) – as exemplified below:

(18) *ei dudhTuku*
this milk.Ma
'this little bit of milk'

(19) *eiTuku dudh*
this.Ma milk
'so little milk'

We use the term Excapsulation to refer to a demonstrative morphologically fortified by classifier or (as in this case) massifier material, as in (19). In Asamiya, unlike Bangla, demonstratives can be fortified by classifier material not only when a bare Dem heads a DP, as in (15) above where the proximal Dem is fortified by the classifier format /XTa/ and appears as /eTa/ 'this', but also within a DP containing a lexical noun. Thus, compare the Asamiya excapsulated forms (20), (21) with their Bangla counterparts (22), (23):

Asamiya:

(20) *xei-to lOra*
that-Cla boy
'that boy'

(21) *xei-khOn kitap*
that-Cla book
'that book'

Bangla:

(22) *Sei cheleTa* (**SeiTa chele*)
that boy.Cla (*that.Cla boy)
'that boy'

(23) *Sei boikhana* (**Seikhana boi*)
that book.Cla (*that.Cla book)
'that book'

Within the analysis developed here, the description runs as follows. In Bangla, a Word Formation Strategy maps a demonstrative /X/ onto a demonstrativized mass quantifier /XTuku/ 'X little', which is endowed with a definite feature. Such an /XTuku/ word is merged at Q. The definiteness

feature drives movement of the QP to Spec-DemP. The word /XTuku/ carries the right feature composition to enable it to head-move to definite D.

In Asamiya, a host of Word Extension Strategies map demonstrative /X/ onto demonstrativized classifiers /X-to, X-khOn/ and so on, along the lines of the analysis of /ei-dukhana/ 'the two books' in Bangla given above. The parametric difference (rooted in factors identified in Ghosh 2001a, b) is that Asamiya classifiers are more distinctly meaning-bearing and have greater independence than the morphological material of classifier formats in Bangla. Thus Asamiya attaches bare classifier clitics /-to, -khOn, -zOn, -gOraki/ etc to demonstratives across the board, whereas Bangla only has a few particular Word Extension Strategies doing this – it has WESs for /Ta, khana/ but not for /jon/. Example (20) in Asamiya is thus unqualifiedly parallel to (24):

(24) Asamiya: *xei-duta lOra*
that-two.Cla boy
'those two boys'

Both (20) and (24) involve a Word Extension Strategy creating a definite demonstrativized Q. Even though there is a significant morphological difference – (20) is based on the Dem while (24) maps from the Q, since basing (20) on Q would have only been possible if the bare classifier Q could be shown to have independent existence as a word – we can nonetheless assume that a definite demonstrativized Q is consistently merged at Q and triggers QP movement to Spec-DemP, with niching of the definite Dem at D. Note that this contrasts with the treatment of the Bangla example (19), where the entire demonstrativized quantifier /eiTuku/ 'this small quantity' head-moves to D.

Any attempt to replicate our results in a formalistic framework will have to deal with the fact that, while Asamiya can combine any demonstrative with any classifier clitic, Bangla demonstratives are not uniformly hospitable to classifier format material:

(25) Bangla:

	i.	ii.	iii.
a. default:	eTa 'this'	oTa 'that (distal)'	SeTa 'that (sequent)'
b. objectual:	?ekhana 'this'	?okhana 'that (d)'	?Sekhana 'that (s)'

c. human:	i.*ejon 'this one'	ii. *ojon 'that one (d)'	iii. *Sejon 'that one (s)'

When we try to extend these variably acceptable series into interrogative and relative columns, we discover an overall "wh amnesty" papering over the heterogeneity of pattern (25), but then what hits us is an inexplicable gap at (26c-v):

(26) Bangla:

a. default:	iv.	konTa	'which'	v.	jeTa	'the one which'
b. objectual:	iv.	konkhana	'which one'	v.	jekhana	'the one which'
c. human:	iv.	konjon	'which one'	v.	*jejon	'the one who'

Substantivism, with its strategies, can manage what is manageable and recognize true regularities where they appear, without over- or understatement. Formalism – precisely because it tries in the name of science to push rules as far as they go and assumes far too hastily that the first rules it postulates are principled – is skilled in making sense of what is very smooth and at dismissing as rough what is obviously rough, but has few resources for walking on the semi-rough surfaces that language typically offers.

We are now ready ask exactly what work the Phi device does in our scheme of things. This involves putting Hindi in the picture. In Hindi, none of these movements take place; demonstratives and most quantifiers carry gender and number features; Phi, sitting at D, has to Agree with the phi-feature content of each of these agreement sites. In classifier languages, we have seen what happens to definite nominals, some of which are furthermore specific. Indefinite non-specific nominals carry out the grammatical integration of the construction by having the Phi device work with the overt or null Quantifier. For uniformity, we assume that Q head-moves to D. It is now possible to formulate a uniform statement. Phi in languages devoid of definite articles – classifier languages as well as agreement languages – is a null D merged as a place-holder oriented to the nature of integration devices in the nominal construction. In an agreement language, Phi presides over the agreement patterns. In a classifier language, Phi is marked either quantificational or definite. If Phi is quantificational, Q head-moves to D. If Phi is definite, an definite element moves to D in one of the ways indicated in our detailed account of Bangla and Asamiya. It is possible that Hindi

replicates some of these processes in non-obvious ways; this, if true, is for future research to uncover.

We return at last to the CGD vs CGID issue formulated in section 1. The main point of CGD as stated there was that Phi could host a nulldem moving to it and negotiating referential anchorage with a genitive in Spec-DP. Stated baldly, that hypothesis naturally left the reader wondering what Phi would do if there was no genitive in Spec-DP. Part of the point of section 2 was to answer this question as fully as necessary. When the genitive is an adjunct to a fully equipped host DP, then all the equipment displayed in section 2 become available. This realization leads to a new worry. Certain examples in section 1, repeated below as (27)-(29), gave the impression that the fully equipped DP host of an adjunct genitive would have to be indefinite. But we have seen that the various options open to a fully equipped DP include various forms of definiteness, and it is therefore unexpected that an adjunct genitive should force indefiniteness on its host:

(27) *ramer to amra* [*t boi*] *poRechi*
Ram's Prt we t book have.read
[bad English:] 'By Ram, of course, we've read some book/s'

(28) *ramer-i amra* [*t boi*] *poRechi*
Ram's-Prt we t book have.read
[bad English:] 'It is by Ram, of course, that we've read some book/s'

(29) *ramer amra* [*t boi*] *poRechi*
Ram's we t book have.read '(27)'/ '(28)'

The answer to that worry is straightforward. The point is that an adjunct genitive (as distinct from a specifier genitive) does not itself induce definiteness, and that therefore a bare nominal such as /boi/ 'book' will come out indefinite if the genitive (or its trace) is an adjunct rather than a specifier. When the nominal itself chooses definiteness, as in (30), of course the adjunct, as in (30i), or its trace, as in (30ii), does not prevent this:

(30) i. *amra [ramer* [*ei-duTo boi*]] *poRechi*
we Ram's this-two.Cla book have.read
'We have read these two books by Ram'

ii. *ramer to amra* [*t* [*ei-duTo boi*]] *poRechi*
Ram's Prt we t this-two.Cla book have.read
[bad English:] 'By Ram, of course, we have read these two books'

Genitive adjuncts to a nominal host can move out of this adjunct site into a clausal non-argument position. There are other adjuncts, however, that have to move out, for reasons we examine with some care in section 3.

3. Non-Genitive Adjunct Nominals

To the best of our knowledge, observations in Dasgupta (1982) concerning what is here called genitive adjunct movement to a clausal non-argument position and in Dasgupta (1983) for the case of non-genitive adjunct movement represent the empirical beginning for this inquiry in the case of Eastern Indic. The general point of departure is provided by Szabolcsi (1984) and Abney (1986), whose theoretical resources we all depend on to gather our threads and connect our descriptions. We take these points of departure for granted. Our concern in this section is with non-genitive adjuncts such as /botol/ 'bottles' in (1) or /jontu-janoar/ in (2):

(1) *botol to tumi parlpeT ar sripeT kinle*
bottles Prt you Pearlpet and Sripet bought
'As for bottles, you bought Pearlpet and Sripet [brand names]'

(2) *jontu-janoar-o tumi horin pabe na, Sudhu khOrgoS pabe*
animals-Emph you deer get won't, only rabbits will.get
'As for animals, you won't get deer, only rabbits'

We have chosen examples where the topic marker /to/ and the focus marker /o/ make it obvious that particular non-argument positions in the clause are at stake. When these markers are missing, the non-argument can be read either as a topic or a focus by modulating the intonation. For brevity, consider only (3), where the lexical choices make the topic reading natural:

(3) *botol tumi parlpeT ar sripeT kinle*
bottles you Pearlpet and Sripet bought '(1)'

Such a topic can occupy a clause-internal position, either with the topic marker /to/ as in (4) or without it as in (5) – to save space, we refrain from labouring this point in the case of focus positions:

(4) *tumi botol to parlpeT ar sripeT kinle*
you bottles Prt Pearlpet and Sripet bought '(1)'

(5) *tumi botol parlpeT ar sripeT kinle*
you bottles Pearlpet and Sripet bought '(1)'

We are agnostic about the clausal configurations involved in (1)–(5). Jayaseelan's (2000) account of intraclausal non-argument positions involves revising standard parametric theories quite radically. Belletti (2004) assimilates Jayaseelan's work to standard views by locating his internal non-argument positions at the edge of vP, echoing the familiar positions on the left periphery of the CP. These matters are orthogonal to what is at stake in the present study; Dasgupta (2005) provides further reasons for our agnostic stand.

We are concerned here with two crucial properties of (5). One is that /botol/ in (5) cannot be interpreted as occupying an adjunct position on the left periphery of a nominal [botol [parlpeT ar sripeT]], but must be taken to have moved (string-vacuously) into a non-argument niche in the clausal structure. We can ascertain this by manipulating the structure. (5) can be expanded into (6):

(6) *tumi botol parlpeT ar flask igl kinle*
you bottles Pearlpet and flasks Eagle bought
'On the bottles front, you bought Pearlpet; in the flasks category, you bought Eagle'

Now, if /botol parlpeT/ and /flask igl/ were constituents, it would be possible to say (7):

(7) **botol parlpeTer ceye flask igler dam beSi*
bottles Pearlpet than flasks Eagle.Gen price greater
'Eagle, as in flasks, costs more than Pearlpet, as in bottles'

But (7) is crashingly ungrammatical. The pseudo-cleft construction also crashes:

(8) **tumi ja kinle ta holo botol parlpeT*
you what bought that is bottles Pearlpet
'What you bought is Pearlpet, as in bottles'

The second crucial property of (5), which dramatizes the first property by choosing particular examples, emerges when we replace /parlpeT ar sripeT/ in (5) with /duTo/ 'two' or /ei-duTo/ 'these two', yielding:

(9) *tumi botol duTo kinle*
you bottles two.Cla bought
'As for bottles, you bought two'

(10) *tumi botol ei-duTo kinle*
you bottles these-two.Cla bought
'As for bottles, you bought these two'

In these examples also what may look like a sequence of the form Noun Quantifier in (9) or Noun Demonstrative Quantifier in (10) is in fact not a constituent, as can be demonstrated by using the same diagnostics. The existence of sentences like (9) and (10) presents a dangerous trap for investigators taking up the study of these constructions, for examples like these get mixed up with the first set of primary data that a linguist looks at, and of course throw the analysis completely out of kilter if the linguist jumps to the conclusion that /botol duTo/ in (9) and /botol ei-duTo/ in (10) are DPs of some sort.

Now that we know they are not, given that our analysis says /botol/ in these sentences begins its career as a DP adjoined to a DP and then moves from that site to a non-argument position in the clause, it is natural to ask why they are compelled to move – in other words, just why a DP constituent such as /botol duTo/ in (9) or /botol ei-duTo/ in (10) cannot stay intact. After all, it can be shown, by applying our diagnostics, that a genitive DP originating in a DP adjunct position has the right to stay in situ:

(11) *ramer boiyer kaTti jodur boiyer ceye beSi*
Ram's books' sales Jodu's books than greater
'Books by Ram sell / Ram's book is selling better than books by Jodu / Jodu's book'

(12) *ami ja cai ta holo ramer boi*
I what want that is Ram's books
'What I want is books by Ram / Ram's book'

Recall from section 1 that the availability of an indefinite reading of /boi/ 'book/s' – in all its occurrences in (11) and (12) – can be taken to demonstrate that the relevant structure contains a DP /ramer boi/ with /ramer/ in adjunct position within that DP. The 'Ram's book' reading, it will be recalled, betokens a structure with /ramer/ in specifier position.

This paper seeks to open up this question for investigation from various viewpoints. Our own hunch is that the semantics of these constructions plays a role. 'Bottle' and the brand name 'Pearlpet' cannot, for instance, change places:

(13) *parlpeT to tumi botol kinle
Pearlpet Prt you bottles bought
'As for Pearlpet, you bought some bottles'

We seem to be observing a construction where the DP's non-genitive adjunct 'exported' to a non-argument niche in the clause structure typically designates a superordinate term and the host DP designates a hyponym. It seems to us that this construction, though both its terms are full nominals, is broadly reminiscent of *There is a man in the room,* on the (frequently made) assumption that the element *there* is merged as part of a complex nominal *there – a man* in which *there* serves as a formal introducer (an ad hoc term for its role within the complex). If 'bottles' is a substantive introducer for 'Pearlpet' in the constructions at stake, the presumption must be that the 'bottles-Pearlpet' A-bar-chain has properties parallelling those of the *there-a-man* A-chain, modulo the systematic (Case-theoretic, theta-theoretic and other) differences between the two chain types.

One class of examples of the non-genitive adjunct DP construction that deserves careful study is a variant of the measure word structure. Measure words are elements such as 'cup' in 'two cups of tea':

(14) *du kap ca*
wo cup tea
'two cups of tea'

(15) *tin bOsta gOm*
three sack wheat
'three sackfuls of wheat'

(16) *Ek camoc cini*
one spoon sugar
'a spoonful of sugar'

It has long been known that measure words are analogous in some respects to classifiers and massifiers. In the literature on Eastern Indic, there seem to be no formal proposals about their categorial identity. We propose that the structure is as shown in (17), with QP recursion (based perhaps on the fact that the measure word is a Q that also carries a noun feature – this conjecture plays no role in the argument):

(17) $[_{QP}$ $[_{Q}$ du][$_{QP}$ $[_{Q}$ kap] $[_{NP}$ ca]]]

The machinery set up earlier in this paper implies that, unless there are word formation/ extension strategies targeting measure words, they cannot participate in the syntax of specificity or definiteness, and indeed they do not; there is no (18):

(18) **ca-dukap*
tea-two.cup, for
'the two cups of tea'

What requires our attention is the fact that the string does exist, within (19):

(19) *amra ca du kap khete pari*
we tea two cup drink can
'We can drink two cups of tea'

However, /ca/ 'tea' is here a moved constituent occupying a clausal A-bar position and binding an adjunct trace in the nominal [t [du kap]] 't two cup'. Standard diagnostic testing shows that sequences such as 'tea two cup' are non-constituents: (20) is ill-formed, and the variant (21) must be used instead:

(20) **gOm tin bOstar ceye cini tin bOstar dam kOm*
wheat three sack than sugar three sack's price less
'Wheat, three sackfuls, will cost you less than sugar, three sackfuls'

(21) *tin bOsta gOmer ceye tin bOsta cinir dam kOm*
three sack wheat than three sack sugar's price less
'Three sackfuls of wheat will cost you less than three sackfuls of sugar'

The theory of classifiers, massifiers and measure words must some day formulate the obvious question – why do measure words not take part in the specificity and definiteness system in Bangla and Asamiya – at a level that enables us to offer non-speculative answers. We are compelled to close these remarks without such a formulation, noting only that, in examples like (22) in Bangla, a classifier formatted numeral can be used with a measure word, which may have something to do with the matter. No doubt progress in the study of the nominal left periphery in Chinese, Indonesian or Vietnamese will help make sense of these and related facts.

(22) *ora ei-duTo kap ca-o dite deri korlo*
they this-two.Cla cup tea-Emph to.give delay did
'They took forever even to provide these two cups of tea'

Acknowledgements

We would like to acknowledge the role of interlocutors whose participation has made it possible for inquiry on these matters to move forward, especially the readership of *Baromas* in 1981, the workshop audience at CIEFL's memorable ISIS (International Summer Institute on Syntax) of 1994, and the audience of the 2006 Utrecht conference on DP-internal topic and focus positions. By the same token, we are grateful to Asok Sen, editor of *Baromas*, to K.A. Jayaseelan and R. Amritavalli, who organized ISIS, to Enoch Aboh, who organized the Utrecht workshop, and to Jacqueline Guéron, Jean-Yves Pollock, Alice Davison, Henk van Riemsdijk, Richie Kayne, and Noam Chomsky, whose response to this thread of inquiry has consistently gone beyond the call of professional responsibility. The usual disclaimers apply.

References

Abney, Stephen Paul
1987 The English noun phrase in its sentential aspect. Doctoral dissertation, MIT.

Allan, Keith
1977 Classifiers. *Language* 53 (2): 285–311.

Azad, Humayun
1983 *Pronominalization in Bengali*. Dhaka: Dhaka University.

Belletti Adriana
2004 Aspects of the Low IP Area. In *The Structure of CP and IP: The Cartography of Syntactic Structures*, Vol. 2, Luigi Rizzi (ed.), 16–51. London: OUP.

Bhattacharya, Tanmoy
1999 Specificity in the Bangla DP. In *The Yearbook of South Asian Languages and Linguistics 1999*, Rajendra Singh (ed.), 71–79. New Delhi: Sage.
2000 Structure of the Bangla DP. Doctoral diss., University College, London.

Cheng, Lisa Lai-Shen & Rint Sybesma
1998 Classifiers and Massifiers. *Tsing Hua Journal of Chinese Studies* 28 (3): 385–412.

Chomsky, Noam
1995 *The Minimalist Program*. Cambridge, MA: MIT Press.

Dasgupta, Probal
1981 Txaaye txoye. *Baaromaas* 3 (5): 53–8.
1982 On conceiving of South Asia as a linguistic area. *Indian Linguistics* 43 (3): 4.37–48.
1983a Bangla equatives, complementizers, final foci, and roots. *Linguistic Analysis* 11 (2): 103–137.
1983b The Bangla classifier /Ta/, its penumbra and definiteness. *Indian Linguistics*. 44: 11–26.
2005 *Q-baa and Bangla clause structure. In The Yearbook of South Asian Languages and Linguistics 2005*, Rajendra Singh & Tanmoy Bhattacharya (eds.), 45–81. Berlin/New York: Mouton de Gruyter.

Dasgupta, Probal & Tanmoy Bhattacharya
1994 Classifiers and the Bangla DP. In *SALA XV: Papers from the 15th South Asian Language Analysis Roundtable Conference 1993*, Alice Davison & Frederick R. Smith (eds.), 59–69. Iowa: University of Iowa South Asian Studies Program.

Dasgupta, Probal, Alan Ford & Rajendra Singh
2000 *After Etymology: Towards a Substantivist Linguistics*. München: Lincom Europa.

Dhanwar, Vijay Louis
2004 The syntax of classifier constructions in Assamese. M.Phil dissertation, University of Hyderabad.

Ford, Alan, Rajendra Singh & Gita Martohardjono
1997 Pace *Panini: Towards a Word-Based Theory of Morphology*. New York: Peter Lang.

Ghosh. Rajat
1995 The DP-analysis of English and Bangla Noun Phrases. M.Phil dissertation. CIEFL, Hyderabad.
1999 Excapsulation in Asamiya and Bangla DPs. Paper presented at the International Conference on South Asian Languages (ICOSAL-II), Patiala, January 1999.
2001a The role of light nouns in definiteness marking in Asamiya. Paper presented at the South Asian Language Analysis Roundtable, Konstanz, October 2001.
2001b Some Aspects of Determiner Phrase in Bangla and Asamiya. Ph.D. dissertation, Tezpur University.

Ghosh, Sanjukta & Probal Dasgupta
2005 Honorificity in Bangla and the movement description of classifier constructions. Paper presented at the annual conference of the Linguistic Society of India, Hyderabad, December 2005.

Hany Babu, Musaliyarveettil Tharayil
1997 The syntax of functional categories. Ph.D. dissertation. CIEFL. Hyderabad.

Jayaseelan, Karattuparambil Achutan
2000 IP-internal Topic and Focus Phrases. *Studia Linguistica* 55(1): 39–75.

Szabolcsi, Anna
1984 The possessor that ran away from home. *The Linguistic Review* 3: 89–102.

Special Contributions:
Indian English / English in India

The Nature, Structure, and Status of Indian English[1]

Rajendra Singh

This paper attempts to characterize the nature, structure, and status of Indian English and, by implication, of other so-called 'non-native varieties' of English. In order to accomplish the task, I examine three theses: (1) that IE is a substratum-laden deviant variety and a member of a class called 'non-native varieties', (2) that IE does not quite belong to the class of genuinely Indian languages, and (3) that an account of IE cannot be provided without looking at the sociolinguistics of English in India and the political economy of the contemporary world and India's place in it. As for the first thesis, I argue that even well-meaning (synchronic) analyses of IE are, unfortunately, anchored in a tradition that is guided by pedagogical concerns and is responsible for encouraging analysts to (1) compare performance with competence, (2) turn their analyses into error/contrastive analyses and (3) assume substratum-ladenness, ignoring the fact that historical influences on IE are just that and no different from comparable historical influences on any other variety of English or any other language for that matter. I point out that it is ironic that the autonomy conferred on all linguistic systems by the naive linguist is taken away by some in the name of sociolinguistic responsibility. There are, I claim, no linguistic reasons for classifying systems such as IE as 'non-native varieties'. I also argue against the position that sees IE as 'non-native' for putative reasons of local language ecology for it is clear that IE is an integral part of the language ecology of contemporary India. As for the claim that the linguistic dimensions of IE cannot be separated or isolated from the social and politico-economic context in which it is embedded, I argue that while there is no need to deny the role that political economy and ideology can and do play in encouraging some linguists to treat varieties such as IE as somewhat deviant, these factors have little bearing on who is to count as a (native) speaker of English, and should be discussed and negotiated elsewhere.

As the study of Indian English is deeply embedded in the study of what is generally referred to as 'the non-native phenomena of English', I must be-

gin by opening that box. The expression in question invites two related but distinct interpretations: the pedagogical one, according to which Indian English is a deviant variety, and the ecological one, according to which English is an outsider in India. Both of them go back at least to Manu (2.21–22), who characterized the growing Aryavrata exactly the way B. Kachru characterizes, without any mention of Manu, the expansion of English (cf. Singh 2003). In the recent past, there are, of course, Iranian discussions of the nature and quality of the Persian of the upstart Indians. As for English, it is simply the new kid on the block, with the difference that its leader speaks a variety of English which is a bit of a challenge for everyone!

As the very notion of non-native phenomena in English is predicated on the legitimacy of the error-driven, pedagogical enterprise that studies what it calls non-native varieties of English, perhaps I should begin with my very early reactions to Error Analysis in SLA, expressed in my review of Nickel (1971). In it I had expressed some discomfort with the procedures followed by this field of inquiry and with the very notion 'non-native errors'. The review in fact is a plea for abandoning the comparison of native competence with non-native performance, something which seemed to me to be the empirical bedrock of these enterprises, and an invitation to colleagues to undertake systematic investigation of rigorous hypotheses of the following sort (from Singh 1974: 76):

1. The class of non-native errors always includes a subclass that will never be included in the class of native performance errors.
2. The class of performance errors in Lj+1 given Lj as the mother tongue will always include a subclass which will never be found in a comparable body of Lj+1 given Lk as the mother tongue.

As these reactions were based on some actual experience and some painfully collected data, let me briefly summarize the experience and the work these comments were based on. In 1972, when I was involved with a Freshman Composition Programme at a US university, I noticed that some of my freshmen were doing things natives were not supposed to do or only non-natives were supposed to do. I tried to make sense of this in Singh (1972), where I point out that conclusions drawn from a comparison of native competence with non-native performance were flawed in fundamental ways. The standard procedure of collecting performance data from L2 learners of English and having native speakers of English evaluate the grammaticality

of the structures found in that data was like mixing oranges and apples. I argued that we needed to compare native performance with non-native performance and native competence with non-native competence for the latter, as we have known at least since Corder (1967), do develop their own competence, not always directly reflected in their performance.

Although the point may seem obvious now, it is interesting to note that my intervention of the early 1970's did not quite have the effect I thought it would. In order to drive the point home, some colleagues and I conducted a series of judgmental and operational acceptability tests in the late 1970's and published the results in 1983 in *Language Learning* (cf. Singh, D'Anglejan & Carroll 1983). Our results established, once for all, that interlanguage speakers do not necessarily accept the structures they produce.

Although all the three interventions mentioned above use data from Indian speakers, the emerging endo-normative nature of Indian English is not introduced as an issue in these interventions. To keep things nice and easy, the way some scholars like them, I simply use the traditional, pedagogically motivated dichotomy native/non-native speaker. The real questions clearly go beyond the deliberately limited mandates of these interventions: who counts as a native speaker and why? What does it mean to talk about non-native phenomena? These questions acquire particular significance when one looks at the contrast between countries like China, Bulgaria, Denmark, The Netherlands, and Sweden, where one CAN speak of mistakes and errors because adults in these places don't interact with each other in English, and countries like India, Nigeria, Pakistan, and Singapore, where one CANNOT, at least not so easily, because adults in these places do interact with each other in English. Moreover, there is the problem that people all over Florida and South Carolina regularly say things like *He might could do it*. Such facts are carefully hidden by the Non-native English enterprise from their Asian readers.

Let me first take up the issue of who counts as a native speaker. In Singh (in press) I attempt to answer the question somewhat as follows. Although the rise of the idea of a (naive) native speaker can be, as pointed out by Dasgupta (1998), traced as far back as the anti-urbanist impulse released by German Romanticism of the 19th century, the use of the expression *native speaker* has become prevalent in modern linguistics particularly since the Chomskyan intervention in linguistics, an intervention which in perhaps its least appreciated aspect brings linguistics truly back 'home', after several detours to various peripheries. Although some would argue that Chomsky's

ideal native speaker does not look very different from the native speaker of those who hold a more prescriptivist position, he takes the theoretical position that we are all native speakers of the steady state grammar we develop on the basis of innately specified language capacity.

Scholars concerned with what Chomsky calls E-language are, of course, preoccupied with the question 'of which language?' This, as Muysken (1998) points out, is not a straightforward matter for the I(nternal) L(anguage)/E(xternal) L(anguage) mapping is more often than not asymmetrical (Cf. Hindi and Urdu in South Asia, which arguably represent the same IL, and 'Patois', 'Dialect' and 'Quechua' in South America, which presumably manifest different IL's). Nor is the relationship between language 'competence' and language 'use' a straightforward one for the former seems crucially to depend on the latter, as can be clearly seen in language attrition (cf. Seliger & Vago 1991). It is these considerations – of asymmetrical mapping and of use – that bring social parameters into the picture. A further complication is added by the so-called indigenized varieties of some European languages, particularly English (for obvious reasons behind its international spread). The debates regarding the status of these varieties, at least some of which are demonstrably fully rule-governed linguistic systems, have made it increasingly clear that multilingualism must be taken into account in providing a more viable characterization of the notion of *native speaker*.

Earlier accounts of the notion 'native speaker', such as the ones collected in Paikeday (1985) and Coulmas (1981), try to come to terms with some of the complications summarized above. However, they do so, with only a couple of exceptions, with what must be seen as a monolingual bias, remarkably clearly spelled out by Crystal (1985) and Quine (1985), and with an almost complete unawareness of questions thrown up by the existence of varieties such as Indian and Singaporean English. The old, monolingualist characterization of the concept of *native speaker* in terms of mother tongue or first language may no longer be sufficient (cf. Pattanayak 1981 and Harris & McGhee 1992, amongst others). Token homage to multilingualism or minor adjustments to such a characterization can't solve the problem because they tend to take a rather simplistic view of multilingualism.

The functionally determined distribution of the use of particular languages and the concomitant acquisition and competence in them in multilingual societies makes such accounts inadequate because neither the proficiency nor the competence of a multilingual speaker can be described in

simple, additive terms – bilingual speaker is NOT a simple, additive union of two monolingual speakers. The existence of indigenized varieties of English (and some other European languages) makes these accounts look even worse. It is one thing to de-emphasize or question the role of introspection in linguistics, as many of the contributors to both Paikeday and Coulmas do, but quite another to come to terms with what Coulmas himself calls "the common reference point" for all of linguistics. Although the question of the relationship between use and the acquisition and sustenance of competence needs to be researched more thoroughly than it has hitherto been, no harm is done if the expression *native speaker* is understood as *native speaker/user* as no one will deny that to become and to remain a native speaker of a language one must use it.

The question, as Kandiah (1998: 90) puts it, is NOT whether the native speaker/user exists – Paikeday's dismissal of her is much too cavalier – but "what we mean when we say that people know, use and view a language in a manner that allows them to see themselves as and to be recognized and accepted as native speakers/users of it". "To be recognized as" and "to be accepted as" add the dimension of ownership or proprietorship to an already complex set of parameters that must be taken into account in defining the native speaker/user. The native speaker/user is, in other words, not dead but has been, as the very titles of the collections edited by Coulmas and Paikeday suggest, somewhat prematurely buried by some. Kandiah rightly insists that the fact that large numbers of ordinary people consciously or unconsciously assume the notion in their ordinary interactions guarantees that the notion *native speaker/user* captures something real. Although it is instructive to deconstruct certain construals of *native speaker/user*, not much is to be gained by throwing the baby out with the bath water. Even if the assumption that one is naturally proficient in one's mother tongue is rejected, as it must be, Paikeday's suggestion that we use "proficient user of a specified language" instead of *native speaker/user* is a non-starter because we need to know how to measure it and who determines the norms against which such measuring will take place. We cannot, in other words, cut off the nose we must in the final analysis count on.

As the considerations that preoccupy most of the contributors to Paikeday – mother tongue, the age at which the acquisition of the language in question began, and the order of acquisition, for example – are rendered problematic by the functionally distributed use of and competence in several languages in multilingual societies, the only way to avoid being sidetracked

by them is to attempt a characterization grounded squarely in the reality and psycholinguistics of mutilingualism. Singh (1994: 370) offers the sort of characterization I believe is needed :"Linguistically speaking, a native speaker of a language is a person who has relatively stable and consistent grammaticality judgements, which he shares with some other speakers, regarding structures alleged to be from his language." A native speaker/user is, in other words, a speaker/user whose well-formedness judgements on utterances said to be from her language are shared by the community she can be said to be a member of. Only a definition of this sort can, it seems, preserve the innocent grain of truth in structuralist and generativist conceptions of the native speaker and acknowledge the sort of considerations Kandiah rightly brings to our attention. It also exposes the oxymoronic nature of labels such as 'non-native varieties of X' by making it clear that whereas one can legitimately say that native speakers of Texan English are not native speakers of Heartland Canadian English, one cannot legitimately say that native speakers of Texan English are not native speakers of English (because they do not speak Standard British or Standard Mid-Western American English). As for the alleged differences between the psycho--neurolinguistic causes/cosequences of simultaneous and sequential bilingualism, it is perhaps enough to point out that Paradis (1994), who looks at the question, reports finding none.

Although some of the questions thrown up by the emergence of varieties of English such as Indian and Singapore English are very important for characterizing the notion of 'native speaker/user', the debates regarding the status of such varieties tend to be, unfortunately, almost journalistic. Consider the easily understood matter of lexical innovation and morphology, for example. The preoccupation with pedagogy and an almost complete neglect of grammar in the contemporary sense reduce most discussions of lexical innovation in such varieties to journalistic reports on exotica. It is true that Indian English (IE), for example, has words that are peculiarly its own, but all varieties of English have words that are peculiarly their own. This sort of peculiarity is, in other words, nothing to write home about. Although the delight of discovering words that are unknown in other, particularly standard, varieties of English is understandable, the unfortunate conclusions that are drawn from such excursions into exotica ARE unwarranted. These conclusions seem to me to stem from an absence of attempts to understand the morphology of IE. It is important to look carefully at the subset of morphologically complex words in IE because they result from an interaction

between what the material landscape requires and what grammar permits, and here varieties such as IE do not offer much to write home about.

The fact that *goonda* 'gangster' or *lathi* 'stick' exist only in IE is no more interesting than the fact that *toque* exists only in Canadian English. There would be something to write home about if the peculiarity of the lexicon of IE could be attributed to a distinct and peculiar morphology. Such an attribution, however, seems unwarranted. Like the peculiarity of the lexica of all other varieties of English, the peculiarity of IE seems limited to simplexes. As an example consider Hosali's (1998) example of *lathi-charge* 'an attack with *lathis*' intended to show a distinctive morphological pattern in IE. She notes that the distinctive feature of this morphologically complex item is " the use of a *lathi* or ' a long heavy stick made of bamboo and bound with iron'". This explanation shows, contra her own suggestion, that the item is not a result of substratum-influenced morphology or of an unlicensed extension of English rules of word-formation. Rather it reflects the fact that the simplex *lathi* is a word of IE, a fact which is of no particular relevance to the morphology of IE. Other complex words also suggest that no such substratum influence or illegal extension is involved in the morphology of IE. There is, as I argue in Singh (2002), little in IE morphology that cannot be seen as an entirely natural extension of patterns or rules of word-formation used or exploited in other, so-called 'native' varieties of English. IE certainly has (simple and complex) words that don't exist in these other varieties, but, then, each one of them has (simple and complex) words that don't exist in the other varieties. Lexical differences are, in other words, nothing to write home about. Morphologically complex words in IE are, in other words, fully licensed by word-formation rules of English morphology. *Batch-mate* exists in IE because *class-mate* and *room-mate* exist in all varieties of English and *collectorate* exists in IE because *directorate* exists throughout the English speaking world. The rules that can and do generate *room-mate* and d*irectorate* will also generate *batch-mate* and *collectorate.*

Comparable illustrations from syntax and phonology are easy to find, but perhaps it is sufficient to point out here that there are no structural features, at any level of grammatical description, that characterize all "non-native" varieties of English to the exclusion of all "native" varieties. Given that most linguists who have made serious efforts to find such features acknowledge/concede that there aren't any (cf. Trudgill 1995), we are fully justified

in concluding that the dichotomy native variety/non-native variety cannot be structurally or grammatically sustained. And if it indeed cannot be sustained, speakers of at least the varieties that can be shown to have their own norms, such as Indian English and Singapore English, must be classified as native speakers of English by virtue of the fact that they are native speakers of their respective varieties – the fact that they are not native speakers of some other variety is irrelevant. And perhaps so is the fact that what is being transmitted today may well have been coloured yesterday by the mother tongues of those who leant it as a second language before transmitting it as a first language to the next generation This is, of course, consistent with the definition in Singh (1994), cited above. Although I fully recognize the importance of acceptance, recognition, and ownership, the definition itself does not have anything to say directly about them. It does not because I believe, and have argued extensively, that these are clearly politico-economic matters, and are better discussed and negotiated elsewhere.

It is at least mildly ironic that whereas the asocial tradition of linguistic or grammatical inquiry sees and characterizes the speakers of the sorts of varieties mentioned above as native speakers of these varieties, the allegedly socially responsible tradition of sociolinguistics is responsible for creating the expression *non-native variety.* The former honours its commitment to treat all viable, rule-governed systems of linguistic communication at par, but the latter seems more than willing to sacrifice the grain of innocence contained in the impulse released more than a century ago. It is the sociolinguist's intervention that adds to the understandable pedagogical dichotomy native/non-native speaker the unlicensed dichotomy native/non-native variety. Why some native-speakers of English want to treat some other native-speakers of English as non-native speakers is an important question the answer to which is to be found in the political-economy of the contemporary world, though socio-linguists are welcome to try to answer it. Why some English-speaking sociolinguists also want to do that is perhaps an even more important question, at least for theorizing about language and society. And in what is perhaps the final irony, the only sustainable interpretation of 'non-native variety' may well be the interpretation 'not of the land' or 'still retaining its otherness' – that is why the linguistic argument that IE, for example, is just as self-contained a system as RP, for example, sounds like a threat to speakers of other Indian languages in India. That it also sounds like a threat to speakers of RP is easy to explain – such a status is seen as a demand for a share in the cultural and political power wielded by the native

speakers of English in the so-called "inner circle" inhabited by RP speakers, who do not amount to more than 2% of the population of U.K. This interpretation is, at any rate, not the one that the creators of the expression *non-native variety* have in mind. It is not available to them because the non-nativeness they see in or want to confer on varieties such as Indian English and Singaporean English resides in their view, as they make repeatedly clear, in the Indianness or Singaporeannes of these varieties. It can be invoked only by those who, like Dasgupta (1993), believe that the non-nativeness of these varieties resides in their Englishness instead.

Having referred or perhaps deferred to political economists, I must, to complete the story, now turn to what some of them actually say or might say. As Lele is the only political economist who has written on the subject (cf. Lele 2005), I shall illustrate that point of view with reference to his paper. They may tell you that the sociofunctional approach of the Kachrus, which gave birth to and still nurtures the strange creature called Non-native English, the almost post-modernist sociolinguistic approach of Dasgupta, and my analysis of the nature and status of systems of communication such as Indian English are EQUALLY INADEQUATE. Given that the Kachrus actually go out of their way to show that what political economists prefer to call the narrow linguistic criteria are violated only in such varieties, I would urge you to be careful. In order to make their claim that both form-driven and sociofunctional analyses of such varieties as Indian English are equally inadequate, they are using too broad a brush.

They might also tell you that Dasgupta and I do not quite face the larger issues involved in a proper characterization of English and its place in India. Although I am flattered by an attribution of concern regarding the place "English occupies in the contemporary political-economy of India and with what consequences for the people of India", I must insist that I have never said anything about the place of English in India. Lele is, of course, right in saying that all three of us, Dasgupta, Kachru, and me "posit, at least implicitly, contemporary India as a relatively autonomous community, as a post-colonial nation; and hence, presumably, coterminous with the territorial boundaries of the Indian nation-state", but given that Indians have never subscribed to two-thirds of the project of nation-state, expressed somewhat dramatically by the author of *Mein Kampf*, I am not sure what to make of that objection. I agree that "the rapidly changing broader, global context" "must alert a linguist to questions about the current world order and the place of both India and of English in it", but I would also insist that unless

political economists are willing to sensitize themselves to matters of linguistic form they can, to play with a famous Canadian name a bit, only Cohen around language. I am afraid that they show their unwillingness to see that challenging scientism taken to be science according to its own evaluation metric is not necessarily subscribing to it. Although I have no difficulty with the critique, which I in fact share, of the superficial nature of the sociolinguistic solution to the problem at hand, I am not sure if I fully understand what is to be gained by dismissing the characterization that the allegedly narrow linguistic criteria in fact provide. Minimally, it directly points to the fact that the distinction such as old/new or native/non-native, between Englishes, relate directly to the central-peripheral status people occupy in the political economy. It applies to the differences, not only between those who live in the metropolis as against those living in its international periphery but to the people at the centre and on the periphery within the metropolis itself. This I see as an important step in the direction of making it clear that the peculiarity of English in India" is not unrelated to what happens in Anglo-America and the rest of the world. It is appropriate, of course, to raise questions about attempts such as Dasgupta's "to put the users of Indian English in an 'outer circle' by positing an 'inner circle' constituted by the (native?) speakers of Indian languages" and in outlining the path that those claiming to study the sociolinguistics of English in India must follow for those who wish to understand and transform the place of English in India cannot, they might rightly add, afford to ignore the questions a true sociolinguistics must ask. Please note that they are talking about the place of English in India and not about Indian English.

The difference between Dasgupta (1993, 2005) and me, apart from the fact that he is a very good story-teller and I am incapable of reading fiction and telling good stories, is that whereas he is interested in retooling sociolinguistics, I have simply been unable to find it anywhere (cf. Singh 1996), but perhaps it is a difference only of style. At any rate, I restrict myself to the linguistics of the situation because I am convinced that there really can't be a sociolinguistics of English in India or elsewhere – there can only be a form-related linguistics and a political economy of English, and the latter can determine only when the linguistic argument would be heard and NOT what it would look like. I withhold the third cheer from political economists because they systematically confuse language with language-institutions. Political economy can plant, transplant, or kill speakers or language institutions creating the illusion that it has played havoc with the ar-

chitecture of human languages. The fact that pre-Revolution Russian is not very different from post-Revolution Russian or that neither is very different from post-liberalization Russian is in my view a serious problem for those who avoid facing issues of linguistic form (for a parallel argument regarding "language-death", see Dressler 1996). As for the sociolinguistics of English in India, I have, unfortunately, no cheers at all because what passes for it is bad linguistics combined with a total disregard for society and political economy.

I AM painfully aware of the fact that being a native speaker is no protection against discrimination, but I draw some consolation from the fact that my modest demonstration that speakers(=/learners) of Indian English are native speakers of Indian English just as speakers of Midwestern American English are native speakers of Midwestern American English not only establishes what needs to be established but also makes it clear that Indian English needs to be studied with all the opposites of standard British and American English, Englishes that are treated as marginal by speakers and sociolinguists of Standard British or American English. When the currently deaf ears would, to twist Wittgenstein's famous phrase a bit, begin to hear this demonstration would, of course, be determined by political economy, but neither the demonstration nor the allegedly narrow linguistic criteria it is based on would need to change. Anything other than an analysis, according to the narrow criteria of lingustics, of linguistic form or an analysis of the political economy of English is, of course, bound to exhibit only horizontal depth, though it will, of course, vary from the very thin veneer of slogans of the sort the Kachruvian enterprise adds on to its analyses to the somewhat more absorbent, though not quite water-tight, buffer provided by some linguists by embedding their linguistically informed investigations in language ecology.

Be that as it may. Given the considerations I have attempted to share with you, I am inclined to conclude that the justification for talking about non-native phenomena of English under either interpretation is very meager indeed, actually non-existent, at least in the context of countries like India.

The only thing to remember is that we are talking about speakers and NOT learners. If we were talking about learners, there would be some justification, BUT then we couldn't possibly restrict ourselves to this or that country for the rather simple reason that just as there are universals of first language acquisition, there are universals of second language acquisition, and to concentrate on or study problems in only one language or even one

language family in isolation is to practice Afghanistanism. It does, fortunately, make sense to study learners' errors in the context of India because trivially more than a dozen languages are spoken here and, importantly, those languages represent very rich typological diversity. This is a matter of particular importance for the study of syntactic and morphological errors. It is, after all, the typological difference between Indo-Aryan and Dravidian that is responsible for the fact that whereas the latter has penetrated relatively deeper layers of the grammar of the former, the former has had to content itself with just landing words to the latter. To get the maximum benefit, we need to contrast NOT English and Marathi or English and Hindi, as is often done in Indian universities, but to contrast Marathi Inter-English with Hindi Inter-English on the one hand and Malayalam Inter-English on the other, never forgetting, of course, that this is being done in an increasingly endo-normative context. If we do this right, we might actually find out, for example, precisely what expressions like "honorary Dravidian", often applied to Marathi, actually mean.

Notes

1. An earlier version of this paper was delivered as the key-note address at the Symposium on "Indian English/English in India", The Central Institute of Indian Languages, Mysore, January 3, 2007.

References

Afendras, A. Evangelos et al.
1995 Discussion: On 'new/non-native' Englishes: A *gamelan*. *Journal of Pragmatics* 24: 283–294.

Coulmas, F. (ed.)
1981 *A Festschrift for Native Speaker*. The Hague: Mouton.

Crystal, D.
1985 I could not make love in English, said one man to Professor Crystal. In Thomas M. Paikeday, 65–71.

Dasgupta, Probal
1993 *The Otherness of English*: *India's Auntie Tongue Syndrome*. New Delhi: Sage Publications.

1998 The Native Speaker: A short history. In *The Native Speaker: Multilingual Perspectives*, Rajendra Singh (ed.), 182–193. Thousand Oaks: Sage.

2005 Retooling Sociolinguistics. Paper presented at the workshop On Rethinking the Agenda for Sociolinguistics in India. C.I.I.L., Mysore. Dec. 27, 2004 – Jan. 4, 2005.

Dressler, W. U.

1996 Language Death. In *Towards a Critical Sociolinguistics*, Rajendra Singh (ed.). Amsterdam: Benjamins.

Harris, R. & M. N. McGhee

1992) Bilingualism: Not the exception anymore. In *Cognitive Processes in Bilinguals*, R. J. Harris (ed.), 3–14. Amsterdam: North-Holland.

Hosali, P.

1998 Word formation in an Indian Variant of English. In *Word Formation in Indian Languages,* B. Vijayanarayana & C. Ramarao (eds.), 135–148. Hyderabad: Book Links.

Kachru Braj B.

1998 English as an Asian Language. *Links and Letters* 5 (89–108).

Kandiah, T.

1998 Epiphanies of the Deathless Native User's Manifold Avatars. In *The Native Speaker: Multilingual Perspectives*, Rajendra Singh (ed.), 79–110. Thousand Oaks: Sage.

Lele, Jayant

1980 The Two faces of nationalism: on the revolutionary potential of tradition. In *National and Ethnic Movements*, J. Dofny & A. Akiwow (eds.). Thousand Oaks: Sage.

1981 Community, Discourse and Critique in Jnanessvar. In *Tradition and Modernity in Bhakti Movements*, Lele J. (ed). E.J.Brill.

2005 The Missing Public: Searching Vaguely for a Community in the Indian English Debate. Paper presented at the Emeneau Centenary Seminar, Mysore, India, January 1–4, 2005.

Lele Jayant K. & Rajendra Singh

1991 And never the twain shall meet or language and politics *chez* Chomsky: A review of Noam Chomsky's *Language and problems of knowledge* and *On power and ideology. Journal of Pragmatics* 15: 175–194.

Muysken, Pieter

1998 We are all Native Speakers, but of which Language? In *The Native Speaker: Multilingual Perspectives,* Singh (ed.), 193–204. Thousand Oaks: Sage.

Paikeday, Thomas M.

1985 *The Native Speaker is Dead!* Toronto: Paikeday.

Paradis, Michel
1994 Neurolinguistic aspects of implicit and explicit memory: implications for bilingualism and SLA. In *Implicit and Explicit Learning of Language*, N. C. Ellis (ed.), 393–420. New York: Academic Press.

Pattanayak, D. P.
1981 *Multilingualism and Mother Tongue Education*. Delhi: Oxford University Press.

Quine, W. V. O.
1985 What Quirk and Quine think. In Thomas M. Paikeday: 6–9.

Seliger, Herbert W. & Robert M. Vago, (eds.)
1991 *First Language Atrrition*. New York: Cambridge University Press.

Singh, Rajendra
1972 Syntactic interference: a case study. In *From Sound-stream to Discourse*, Hays & Lance (eds.), 214–223. Columbia: University of Mis–souri Press.
1974 Review of Nickel, Papers in Contrastive Linguistics. *Lingua* 33: 69–77.
1994 Afterword (Indian English: Some Conceptual Issues). In *Second Language Acquisition: Socio-cultural and Linguistic Aspects of English in India,* R. K. Agnihotri & A. L. Khanna (eds.), 369–381. New Delhi: Sage.
1995 Coda: 'New/non-native' Englishes revisited: A reply to my colleagues. *Journal of Pragmatics* 24: 323–333.
1997 *Lectures against Sociolinguistics.* New York.
1998 *The Native Speaker: Multilingual Perspectives*. New Delhi: Sage.
2001 Looking for Sociolinguistics. Paper presented at PLM-30. Poznan.
2002 Against Afghanistanism: A Note on the Morphology of Indian English. In *The Yearbook of South Asian Languages and Linguistics* 2002: 269–273.
2003 The languages of India: A Bird's eye view. In *The Yearbook of South Asian Languages and Linguistics* 2003: 173–182.
in press Native Speaker. *Elsevier Encyclopaedia of Linguistics*. 2nd ed.

Singh Rajendra & others
1995 Discussion note: On' new/non-native' Englishes: A Quartet. *Journal of Pragmatics* 24: 283–294.

Singh, Rajendra, A. D'anglejan & S. Carroll
1983 The Elicitation of Inter-English. *Language Learning* 33 (2): 271–288.

Trudgill, Peter
1995 Contribution to "On 'new/non-native' Englishes: A *gamelan.*" Evangelos A. Afendras et al. *Journal of Pragmatics* 24: 295–321.

Seeking the Holy Grail of Nativeness

Ad Backus

I write this response in general sympathy with Rajendra Singh's attack on the dichotomy 'native' versus 'non-native', certainly given how it is commonly used. I wish to add two dimensions to the discussion. First, I will discuss the issue from the perspective of cognitive linguistics; second, I will discuss a situation that has hitherto been largely overlooked and can be seen as the mirror image of the situation in which Indian English finds itself.

As a cognitive linguist, I am generally suspicious of dichotomies, since better descriptive adequacy can often be reached if phenomena are described as gradient categories. Dichotomous classification is superior only if the categories are indeed 'categorically' distinct, or if we tend to focus on the extreme ends of the continuum rather than on the zone in the middle. Whether the distinction between 'native' and 'non-native' makes much sense depends on whether one of these two extenuating circumstances holds for this domain. This question, which I think should be answered negatively, will feature between the lines of what I'm going to say in this response, and I will come back to it explicitly in the conclusion.

Singh tackles two interpretations of the phenomenon of 'non-native speech'. The first assumes that people born into a speech community are native speakers and that everyone else who speaks its language is a learner, with proficiencies ranging from poor to very good. In fact, 'learners' can be so good that their speech is perhaps indistinguishable from that of native speakers: such speech tends to be labelled 'near-native', hinting at the problem at hand. Why not call them 'native'? Why withhold the Holy Grail of native speakerhood if they are apparently indistinguishable from native speakers linguistically? The term 'near-native' suggests that there must still be a reason to make the distinction between native and non-native, irrespective of someone's command of the language. If that reason is not linguistic in nature, it must be socio-political. Most of this response will be about this interpretation.

The second interpretation Singh discusses in his key-note is another issue familiar from the discussions on World Englishes that have been going on

for some time now: is there any linguistically relevant reason to distinguish the English from monolingual speakers in places like the US, Britain and Australia from that of bilingual speakers in places like India, Singapore and Nigeria? Are there linguistic reasons to call only the former 'native speakers'?

Native speakerhood is often defined in terms of mother tongue: someone is a native speaker of the language that is his or her mother tongue. Ignoring for the moment the thorny question of how long after birth exposure to a language may begin to still call it one's mother tongue – one day, a week, a year, three years?), it seems obvious that the distinction between native and non-native makes sense only if there is enough evidence that native speakers and non-native speakers do different things, for example that non-native speakers make types of errors not found in native speech. Singh draws attention to the fact that we still don't have much evidence that can be brought to bear on this issue, and that what evidence we have does not suggest that speakers of Indian English are not native speakers of English.

Interesting as that may be, I want to draw attention here to a different kind of multilingual setting that is, on the one hand, the mirror image of the postcolonial Indian English-type setting, and yet, on the other hand, shows the same mechanisms at work. Just like groups of speakers may have trouble attaining the status of native speakers, it is also possible to lose it. This is what happens to many immigrant groups. Ever since Thomason & Kaufman (1988), we have a clear global picture of what happens to languages in contact. Languages may be maintained in the face of contact, but they will usually change in the process, for instance by borrowing features from the other language. At the same time, the other language may become a normal vehicle of communication for the speech community, with or without maintenance of the ancestral tongue, and in the process it too may become altered, for instance by conventionalizing L1-inspired phonological and syntactic substrate features. Indian English is an example of the latter scenario; Immigrant Turkish, to be discussed below, exemplifies the first type of outcome. The point of my comparison will be that immigrant Turks lose the status of native speakers (of Turkish) for the same reasons that it seems to be withheld from speakers of Indian English.

In our work on the language of Turkish immigrants in Western Europe, we have noted that the discourse about the status of the immigrant variety is very similar to the World Englishes discourse. Turks in Turkey tend to label the Turkish of the second and third generation descendants of guest

workers in Germany, Holland, France, and other countries 'non-native', because they perceive it to be different from their own 'native' Turkish. Actually, they perceive it as 'wrong Turkish', an assessment shared by the immigrant speakers themselves. The sources of the differences are, of course, to be found in the universal workings of language contact. The communities abroad are bilingual and at the same time cut off from the norm-enforcing mechanisms of life in Turkey. The results are influence from European languages ('borrowing'), and deviations from the norms that hold in Turkey ('attrition'), respectively. When tested for their command of these norms, the immigrant community dutifully confirms the picture: they don't have command of those norms as well as 'real' native speakers do. Whether those norms are relevant for the issue of their competence is a question rarely asked, but I think it should be.

The upshot is that these speakers are generally regarded as semi-speakers, though the term is politically corrected out these days. But the general picture in Holland at least is that the Dutch see them as second language learners of Dutch, because they have Turkish as their mother tongue, and that Turks in Turkey see them as non-native speakers of Turkish, because they don't follow the norms of Turkey, a negative assessment the subjects themselves tend to share, albeit without using the epithet 'non-native'. This status quo produces a no-win situation, since they can only become native speakers of Dutch if they give up Turkish completely, and native speakers of Turkish if they move back to Turkey.

If the variety the Turkish immigrants speak is indeed different, that is because languages in contact undergo contact-induced changes. What we do to detect 'changes' in Immigrant Turkish is to compare the speech of bilinguals in Holland with that of monolinguals in Turkey, identify the deviations in Dutch Turkish, and see whether we can attribute them to language contact. Ostensibly, that procedure obeys the guideline Singh suggests for SLA, to stop comparing native competence and non-native performance, since we compare performance in both varieties and even refrain from calling only one of them native. However, in practice we still use Turkey-Turkish as a yardstick. That may be legitimate given our goal of finding instances of contact-induced change (cf. Heine & Kuteva 2003), but we still invite judgments from panels of 'native speakers' from Turkey on what their cousins in Holland produce.

A tricky complication is that the 'errors' we observe in Immigrant Turkish are not always absent from the speech of monolinguals, just like the so-

called deviations in Indian English can sometimes be found in British or American dialects. In Example (1a), accusative case is missing on the object noun. Initially we thought this to be a typical case of contact-induced change, since Dutch doesn't have an accusative marker. In addition, every grammar book of Turkish will tell you that definite direct objects get accusative case, period (compare the expected 1b). Imagine our surprise, however, when we found the same structure occasionally in our monolingual data from Turkey (cf. Example 1c). Apparently, there is more variation in Turkish than we thought, and accusative is perhaps marked less consistently in contexts of lesser transitivity. We don't really know very well, since the spoken vernacular is under-researched in Turkish linguistics. Note that grammaticality judgments, Singh's suggested remedy, probably won't help here: every Turkish speaker, certainly in Turkey (having been exposed to school teachers), will identify (1a) and (1c) as ungrammatical.

(1) a. Dutch Turkish:

Türkçe iyi konuş-uyor-lar mı?
Turkish good speak-PROG-3PL Q.
'Do they speak Turkish well?'

b. Expected:

Türkçe-yi iyi konuş-uyor-lar mı?
Turkish-ACC good speak-PROG-3PL Q

c. TR-Turkish:

Ben Kırşehir yemek-leri bil-ir-im.
I Kırşehir dish-POSS.3PL know-PRES-1SG.
'I know [how to make] Kırşehir dishes.'

d. Expected:

Ben Kırşehir yemek-leri-ni bil-ir-im.
I Kırşehir dish-POSS.3PL-ACC know-PRES-1SG.

Be that as it may, we find quite a bit of evidence for the changed nature of Turkish in the contact setting, that evidence mainly being found linguistically in 'unconventional' collocations and turns of phrase, rather than in a changing syntactic character, and sociolinguistically in the clear impression in both Holland and Turkey that Dutch Turkish is different. Such changes entail that new norms are developing in Immigrant Turkish. Like Indian

English (a case of Thomason & Kaufman's 'shift-induced interference'), Immigrant Turkish (a case of 'borrowing') has emerging endo-normative norms. And both varieties have in common that they are subjected to exo-normative standards. It is perhaps typical of the difference between post-colonial and immigrant settings that speakers of Immigrant Turkish readily accept these 'foreign' standards, while speakers of Indian English reject them. But that difference is of a socio-political nature, not a linguistic one. It seems a worthwhile sociolinguistic endeavour to look for generalizations about endo-normativity: under what circumstances does it come about, and under what circumstances are these norms not recognized?

The central point seems to be that endo-normativity automatically ensues when the variety it pertains to is the normal vehicle of communication in the speech community. If speakers of English in India were only using it in communication with, say, foreign tourists, or perhaps in international work settings, like many Europeans do, orientation on international norms (or American or British norms) would be logical, but if they use it to converse among themselves, adherence to those outside norms loses its relevance. Similarly, if speakers of Turkish in Holland would stop using Turkish among themselves, and hence become Dutch speakers, some of them may wish to learn Turkish and become what is these days commonly called 'heritage language speakers'. Since their express goal would be communication with people in Turkey, not among themselves, orientation to the language as spoken in Turkey, would, again, be a straightforward choice.

One comment is in order. I have talked about norms as if we all know what we are talking about. That is not the case, however. Since linguistically speaking it makes little sense to define norms prescriptively, norms can only be descriptive generalizations of how people in a given community speak. However, especially for the types of communities we may wish to empower with endo-normative norms, we don't know all that much about how they speak. We know some features of their speech, mostly the ones that are eye-catching, such as codeswitching and deviations from the outside norms, but we don't know how systematic their codeswitching is, how systematic the 'deviations' (that is, their innovative or changed features) and how systematic their use of the structures shared with the outside norm. Whether we want to know all this is a different matter, but to help bury the wrongful assumption that speakers of these varieties are non-native speakers of the languages their hyphenated varieties are said to be varieties of (Dutch Turkish is a variety of Turkish; Indian English is a variety of English), it would help

to have a few reasonably accurate descriptive grammars of such varieties. The real desideratum, though, is a culture change on the parts of linguistic conservatives, be they language guardians in state-sponsored language bodies or the dispossessed speakers themselves, that allows them to see the undesirability and negative societal consequences of imposing outside norms on one's native speech. In short, to start seeing speakers as native speakers of the variety in which they communicate with other members of their speech communities.

Dutch Turkish speakers themselves orient to the norms of Turkey and therefore feel insecure about their Turkish. I, as a linguist, may feel that they have the right to accept their own norms, and that that is the only linguistically sensible way to behave, but they, as sociolinguistic beings, feel that they fall short in the skill that matters more in life: competence in the Turkish of Turkey. Changing these attitudes may be desirable, it may also prove a quixotic task.

Most of this is political work, however, and certainly outside my field of expertise. What I can make some, hopefully helpful, comments on, though, is how to identify the endo-normative norms to which the speakers of hyphenated varieties implicitly orient, even if they don't so so explicitly. I depart here from Singh's position. As a cognitive linguist, I put less faith in grammaticality judgments, since they tend to reflect the sociolinguistic tendency to orient to outside standards. Instead, I rely more on corpus evidence. Cognitive linguistics tends to emphasize usage rather than abstract rules, since usage is assumed to directly determine psycholinguistic representation. Under that assumption, a reasonably sized corpus of someone's speech gives you a fairly good idea of that person's internal grammar. Corpus linguistics has made great strides forward in the last few decades, thanks to the ever greater computational possibilities of computers, though obviously most efforts have been geared to building corpora of the same written standards we witness playing their exo-normative roles in our multilingual settings, such as British English, American English, Dutch, etc. It should be only a matter of time, though, before we have some reasonably adequate representations of spoken and multilingual varieties of these languages, including Texan English, Heartland Canadian English, Dutch Turkish, and Indian English. At Tilburg University, we're building such a corpus of spoken Turkish in Turkey and Holland. Comparing the two should give us a reasonably sound view of the differences and similarities between the varieties. The next question will be, of course, how much structural dif-

ference in phonology, lexicon, morphology, syntax, and discourse structure we want to see before calling the varieties 'different'. We should see this as an empirical question, though, rather than as an ideological one.

Conclusions

The only use of 'non-native' that makes sense is when it refers to the speech of learners who are still actively learning the language and do not use it for everyday interaction within their speech community. In such cases, a dichotomy between native speakers and learners is certainly defensible and useful. However, languages that are used for actual communication inside the community (Turkish in the case of immigrants in Holland, or Dutch in case they shift their main language of interaction among each other; Indian English in case of India, alongside other native languages), do not lend themselves to such categorization. The road from native to non-native (in case of a shift away from Turkish in the immigrant community), and from non-native to native (in case of the shift towards Indian English that has taken place in India – despite the maintenance of other languages its speakers speak natively), is gradual, so the purported categories 'native' and 'non-native' cannot be completely separate categories. They are, instead, regions on the opposite ends of a continuum of development. These are relevant for the technical study of second language acquisition; they are not for the branch of sociolinguistics that deals with the issues tackled in Singh's address. The failure of sociolinguistics to make this view commonplace is indeed 'mildly ironic', given the general emancipatory stance of the field.

Acknowledgements

Although I could not attend it, I am grateful to the organizers of the C.I.I.L. Symposium on Indian English/English in India for their invitation to participate in it.

References

Heine, Bernd & Tania Kuteva
2003 On contact-induced grammaticalization, *Studies in Language* 27 (3): 529–572.

Thomason, Sarah & Terence Kaufman
1988 *Language Contact, Creolization, and Genetic Linguistics.* Berkeley: University of California Press.

On the Native/Non-native Distinction

Rakesh M. Bhatt

Singh's paper is clearly a very well argued paper problematizing the term "non-native English speakers"; especially as it is used to identify speakers of English in countries such as India, Singapore, etc. where speakers frequently interact with each other in English. I tend to agree with most of the arguments in his paper, and would like to present some more arguments to bolster his claims. I would, however, argue that in order to fully understand the "non-native phenomena of English" we need both the political-economy model as well as the "narrow" linguistic model. The political-economy analysis is necessary as it provides a macro-discursive understanding of the phenomena – the linguistic ideologies that legitimate the native/non-native dichotomies – whereas the linguistic analysis presents the structural nuances of individual varieties that are "native" to the speakers who use it. I discuss both of these analyses in that order.

Native/Non-native English: Availability, Accessibility, and Economy

It is undoubtedly a truism that English has served, and still serves, as a prime cultural legitimation of the world division of labor into core and periphery. The dominant core ('center') has, over the last few decades now – in an effort to bolster and express its control over a greater linguistic diversity in the 'periphery' – attempted to "style" its English-use as definitive for the English-using community as a whole, often, as I have argued elsewhere (Bhatt 1995, 2001a,b, 2002, 2005), in the form of an overarching grammaticology that made the norms of the core the presumptive standard for all English linguistic behavior. This is precisely where the contemporary distinctions of native/non-native English appear – the core speakers are the standard bearers, but only those who have (had) the power to control the norm and successfully impose it on "others." This presumption provides an interpretation of linguistic markets where only the dominant system ("na-

tive" English) is able to establish a global standard; such a standard is simply internal to itself but is successfully portrayed as universal. The standard serves only to reproduce socio-economic inequalities as it privileges only those who have access to its possession, leaving "others" disenfranchised. From this economic-theoretic perspective, the struggle between Cockney and Standard English, between African American and General American English or between English and Hindi in India, or between English and Filipino in the Philippines can indeed be interpreted as a struggle between competing economic interests: Standard English serving the elite and native languages serving mainly other classes.

The story of English is, however, more complex than a simple core-periphery issue. The issue is complicated by the role that semi-periphery plays in the contexts of English around the world (cf. Ramanathan 1999). Semi-periphery, represented by the 'middle class', presents the possibility of upward mobility. The quest for inclusion in the core – the "inner circle" of native English speakers – makes the semi-peripheral members ideal agents for the core to control the periphery, the lower classes studied by Ramanathan (1999). The conflicts that arise between classes, those that benefit from the acquisition of the standard-native variety and those that do not, while political in form and cultural in expression, are invariably economic in origin (cf. Bourdieu 1991).

Within the political economy perspective, the role of the educational system in the legitimation of the native-nonnative distinction needs to be fully explored. In the context of India, both in colonial times and now, educational institutions are, and have been, the most important instruments of the reproduction of English symbolic capital since schools[1] have the monopoly over the reproduction of the market on which the value of linguistic competence depends (Bourdieu 1977). In India, where education is/was the only source for the acquisition of cultural capital[2] and apprenticeship into the "fellowships of discourse" (à la Foucault 1972),[3] the principal medium of that initiation is English. Even though English is, in principle, available, the accessibility of the perceived "native" model of English is, however, distributionally restricted, following an economy logic: most restricted in government (poorly funded) schools and least restricted in the elite schools (e.g., Doon School), in terms of the models (teachers) and the local "linguistic culture" in the sense of Schiffman (1996). The role of the school in legitimizing the correlation between economy and access is most profoundly pointed out by Woolard (1985: 740–741), who writes: "The family

initially endows children with linguistic and cultural capital, but the school establishes the authority and legitimacy of the scarcest, and therefore most highly valued, linguistic and cultural forms and secures universal recognition of this legitimacy." The scarcest linguistic form in India is English, that too, the "native" form, which appears in different local linguistic markets as the most highly valued commodity. The value of the "native" model becomes most visible when (many, if not all, of) those endowed with the cultural capital – school teachers, university professors, scholars, writers, media personalities – engage in mimetic acts of "native" model of English, to the extent they can, in their public appearances, in their writings, in all "formal" contexts. It is in the public spaces that "native" speech/speakers get their accreditation, as the only standard. Even where local models of English are used, they are perceived, in the best-case scenario, as linguistic acts of resistance, which serves, unfortunately, only to reinforce, make visible, the native/non-native distinction.

In sum, from a political-economy perspective, the perceived linguistic differences ("native" and "non-native") become indices of social positions of speakers and reflections of the quantities of linguistic, cultural, and/or symbolic capital they possess. The more linguistic capital speakers possess, the more they are able to exploit the system of differences to their advantage (native-standard/nonnative-nonstandard) and thereby secure a profit of distinction. The forms of linguistic expression that receive the greatest value and secure the greatest profit are those that are most unequally distributed, both in the sense that the conditions of the capacity to produce them are restricted and in the sense that the expressions themselves are relatively rare on the markets where they appear.

Expert Discourse and the Production of the Native/Non-native Distinction

Let me now turn to the sociology and politics of knowledge production, a productive field of inquiry that exposes the network of integrated systems – of knowledge legitimation, of exclusion, of ideological management, and of normalization – that coordinate in complex ways to produce discourses that legitimize the native/non-native distinction (cf. Bhatt 2002). In Bhatt (2002), I have argued that the native/nonnative distinction is a theoretical tool used to maintain the autonomy and privilege of agents and agencies invested in

selling English world-wide. The distinction – native/non-native – gets validated by the kind of intellectual imperialism whereby an idealized model of language (à la Chomsky 1986) assumes a paradigmatic status in the linguistic sciences as a whole. This idealization produces 'the illusion of linguistic communism' (Bourdieu 1991) and ignores and trivializes the socio-historical and economic conditions that have established a particular set of linguistic practices as dominant and legitimate. Consider, for instance, Chomsky's (1997) view on acquisition, particularly (presumably, adult) second language acquisition:

> Learning a first language is like growing up. You just do it …But learning a second language is like learning, you know, gymnastics – getting to be a pole-vaulter in the Olympics or something. It's ***not* a *normal*** human activity. (emphasis added)

The methodology used in this dominant paradigm to discover the knowledge of language – the linguistic competence – requires assembling the set of all and only grammatical sentences of *'the ideal native speaker-hearer in a completely homogenous speech community'*. Even with such monotheistic views on language and language acquisition, especially second language acquisition, the academic community keeps producing second language acquisition studies that validate intellectual practices that follow the dominant discourse. This continues in spite of the fact that there is considerable circularity, complexity, and ambiguity within the Chomskyan model itself (O'Grady 1997). What keeps the momentum going on the reproduction of Chomskyan ideas may partially be explained in terms of the non-discursive requirements of academic success, leading to the familiar paradigm trap. Dixon (1996: 841–842) observes the following :

> The majority of publishers specializing in linguistic books (including university presses) are most concerned with theoretical novelties, which generally have pleasing sales and do not feel any obligation to support sound scholarly studies which may not have substantial initial sales. … We get the vicious circle of few buyers → small print run → highish prices → fewer buyers still → even smaller print run → further price rises.

This insight into the political economy of knowledge production shows the structure of the dominant discourse that creates conditions – for getting jobs, tenure and promotion – where only certain forms of knowledge are considered legitimate. This process of intellectual imperialism produces an

academic culture where rival forms of thought are excluded and divergent ideas are either completely ignored or denigrated as uninteresting (cf. Lantolf 1996; Corson 1997).

Although studies in second language acquisition and teaching have predominantly used native/non-native interactions to demonstrate the logic of second language development, there have been voices of dissent: Bley-Vroman (1983) and others (Singh 1974; Anderson 1984; White 1989, 1996; Klein 1994; Schwartz 1995; and Schwartz & Sprouse 1996; see also Singh 2006: p. 1) have shown that such methodologies suffer from "comparative fallacy": referring to the researcher imposing the structure of the native/target language onto interlanguage. These scholars have argued that the structure of the interlanguage at various stages should be considered on its own terms, not from the structural perspective of the target language. As Schwartz (1995: 8) puts it: "If there's one thing we often know about developing interlanguages, it's that they **don't** have the structure of the target grammar – so why such a fuss about the syntax of the target language." Yet the native/non-native distinctions are used as mathematical axioms, beyond debate, by the "experts" in the profession, creating sets of understandings that continue to legitimize attitudes and methodological practices of linguistic monotheism (cf. Sridhar 1994; Y. Kachru 1994; Cook 1999; Bhatt 2002). These distinctions have thus come to belong to what Michael Polanyi (1967) calls the "tacit dimension" of scholarly understanding.

Such monotheistic practices, unfortunately, leave no account of the regularity and systematicity with which such linguistic forms are used even by monolingual English native speakers, in Mid-West America, as shown in (1) below (cf. Bhatt 2002):

(1) a. You should have went to your commencement.
  b. Them guys only play for money anymore.
  c. I never would have wrote that song.
  d. He wouldn't hardly eat nothing.
  e. You wanna have your picture took?
  f. Everything we took out on a picnic had to be throwed.

It is possible to interpret the data in (1) as fossilized expressions of a monolingual interlanguage, fossilized for some unknown reasons in the grammar of some native speakers of "native" English. The alternative is to accept the

data in (1) as examplars of a monolingual grammar that follows a computational logic that is different, not deviant, from other "native" varieties of American English. This alternative interpretation of data in (1) forces a view of grammar that neutralizes the native/non-native distinction in favor of native grammarS – grammar of Mid-west American English, grammar of Southern American English, African American English, etc. This is precisely what the professional "experts" of ELT are unwilling to accept (cf. Davies 1989, 1991; Quirk 1996).

In India, too, one observes such a difference in native varieties of Indian English: one shares its grammar with English varieties spoken outside India, and the other has certain grammatical properties that are not shared by the grammars of English varieties spoken outside India. In the next section, I present the grammar of these two Indian varieties of English to demonstrate the DIFFERENCE between them, and to claim that these differences are just as (in)significant as the differences between two varieties of American English. The analogical argument here is that if mid-western American English and "Standard"[4] American English are two "native" varieties of American English, then the Vernacular Indian English and "Standard" Indian English must be two "native" varieties of Indian English. In other words, I claim that there is NO "non-native" variety of English that has any independent linguistic status.

The Linguistic Argument: The Syntax of Vernacular Indian English

In this section, I will briefly present a snapshot of the arguments presented in Bhatt (1997, 2000), to demonstrate the logic of English language use in India. I will show below the logic of language use in at least two varieties of English spoken in India: the Vernacular Indian English used mainly in casual conversation and all informal contexts, and Standard Indian English used in most formal and written contexts. Most English-educated speakers control, I argue, the grammar of both of these varieties and use them in contextually appropriate situations. Although eventually a restrictive theory of language use is obligated to declare the precise nature of the "context of situation", which presumably yields observed realization of linguistic expressions of a certain communicative act, I can only conjecture here that some articulated theory of diglossia, along the lines of Ferguson (1959) – where certain (High/Low) forms are indexed to certain (High/Low) func-

tional domains – may account for the observed choices among the competing candidates of linguistic expressions.

The term "non-native English" in the literature is used to refer to expressions that belong to the set of admissible expressions in what I call the Vernacular Indian English grammar (cf. Bhatt 2000). I will show that the underlying grammatical logic that governs the use of this variety is DIFFERENT from the logic that governs the use of the other, e.g., Standard Indian English, variety, in a manner similar to the difference between Standard American and African American Vernacular English (cf. Labov 1969; Sells, Wasow & Rickford 1996). I will use the discussion of questions, direct and indirect *wh*-questions, as an illustration of difference: for details of other syntactic differences, see Bhatt (1997, 2000, 2004). The rationale for this discussion is twofold: to show on the one hand that both varieties are rule-governed – a necessary exercise to motivate the assumption that the native speakers of the two varieties of Indian English in fact "know" the underlying logics that govern their use – and on the other hand to show that the difference between the two varieties of Indian English lies in their grammatical design – how, for example, the two varieties prioritize their grammatical constraints, in an optimality-theoretic (Prince & Smolensky 2004) sense. Once this is established, the linguistic argument for the native/non-native distinction can no longer be maintained, neither in the domain of language acquisition nor in the domain of language use.

Let me, in the remainder of this paper, present the linguistic argument with respect to the grammar of questions in two varieties of Indian English: Standard Indian English (SIE) and Vernacular Indian English (VIE). In SIE, direct (root) questions are formed by moving the wh-phrase to the left-edge of the clause (Spec-CP) followed by the auxiliary verb (in Comp) – the inversion operation, in those questions where the wh-phrase is not a subject. A couple of examples are given in (2) below.

(2) a. What$_i$ has$_j$ he t$_j$ eaten t$_i$?
b. Where$_i$ has$_j$ he t$_j$ gone t$_i$ now?

Embedded indirect questions on the other hand involve movement of the wh-phrase to the left-periphery (Spec-CP) of the embedded clause, without, however, any auxiliary verb following it (in Comp) – the adjunction operation, as shown in (3) below:

(3) a. They know who_i Vijay has invited t_i tonight.
b. I wonder $where_i$ he works t_i.

The well-known empirical generalization about data such as (2) and (3) is that the rule of subject-auxiliary inversion is restricted to matrix sentences; it does not apply in embedded contexts.

In Vernacular Indian English, on the other hand, direct questions are formed also by moving the wh-phrase to the left-periphery (Spec-CP) of the clause but without the Aux-to-Comp movement – the adjunction operation, as shown in (4) below:[5]

(4) a. $What_i$ he has eaten t_i?
b. $Where_i$ he has gone t_i now?

Embedded (indirect) questions in Vernacular Indian English involve wh-movement to the left-periphery (Spec-CP) of the embedded clause. The wh-phrase, however, is followed by the auxiliary verb, i.e., wh-movement in embedded contexts is accompanied by auxiliary verb movement – the inversion operation – to Comp. The relevant data is given in (5) below:

(5) a. They know who_i has_j Vijay t_j invited t_i tonight.
b. I wonder $where_i$ does he work t_i.

The simple empirical generalization that emerges from the data in (4) and (5) is that in Vernacular Indian English, inversion is restricted to embedded questions; it does not apply in matrix questions. The generalization of question formation strategy in Vernacular Indian English is just the mirror opposite of the generalization of question formation strategy in Standard Indian English where inversion is restricted to matrix contexts; it does not apply in embedded questions.

These generalizations can be easily encoded in a grammatical theory, such as Optimality Theory (OT) (Prince & Smolensky 2004), that assumes that *knowledge of language* consists of a universal set of candidate structural descriptions, a universal set of well-formedness constraints of these structural descriptions, and a language particular ranking of these constraints from strongest to weakest. In other words, grammars are assumed to contain ranked constraints – arranged in a strict domination hierarchy –

on the well-formedness of linguistic structure.[6] Given these assumptions, it becomes possible to argue that:

(6) a. VIE is just as systematic and logical as SIE;
b. The grammars of VIE and SIE are constrained by the same set of grammatical (universal) constraints;
c. The differences in the two varieties are a function of how each grammar prioritizes these constraints.

In the context of matrix questions in SIE and VIE, we need to address the problem of Inversion vs. Adjunction, i.e., we need to explain the fact that SIE allows subject-verb inversion whereas VIE does not. The universal constraints that need to be recruited to yield direct questions are: OP-SPEC and STAY.[7] The interaction of these two constraints in the order given in (7) yields the categorical prediction of direct questions in SIE: the wh-phrase in specifier position of CP is followed by an aux in Comp.

(7) SIE: OP-SPEC>>STAY

The tableau in (8) shows a competition between two candidates, an adjunction structure and an inversion structure. Both candidates violate the low ranking constraint STAY; however, the inversion candidate structure gets two asterisks for violating STAY twice: once when the *wh*-phrase moves to the specifier position of CP and once when the modal auxiliary moves from Infl to Comp. However, STAY remains inactive on the candidate set since the adjunction structure violates OP-SPEC, a higher ranked constraint, while the inversion structure does not. The grammar of SIE therefore chooses inversion over adjunction as more harmonic, optimal, thus grammatical.

(8) Tableau: SIE

Candidates	OP-SPEC	STAY
adj [$_{IP}$ What [$_{IP}$ you would like to eat *t*]	*!	*
inv ⇒ [$_{CP}$ What would [$_{IP}$ you *t* like to eat *t*]]		**

Turning to VIE, recall that direct questions in VIE involve an adjunction structure; the wh-phrase adjoins to IP-Spec instead of moving to CP-Spec as it does in SIE. It turns out that both OP-SPEC and STAY yield the adjunction structure too, albeit with a different ranking. The VIE grammar ranks STAY over OP-SPEC (as in (9)), which gives the desired results in (10).

(9) VIE: STAY >> OP-SPEC

(10) Tableau: VIE

Candidates	STAY	OP-SPEC
adj ⇒ [$_{IP}$ What [$_{IP}$ you would like to eat *t*]	*	*
inv [$_{CP}$ What would [$_{IP}$ you *t* like to eat *t*]]	**!	

The tableau in (10) shows two competing candidates, both violating the highest ranking constraint STAY. Notice however, the inversion structure incurs two violations of STAY – one by moving the wh-phrase and the other by moving the Infl/Aux to Comp – as opposed to only one violation of STAY, moving the wh-phrase in the adjunction structure. In this competition, inversion loses because it incurs more violations (of STAY) than adjunction.

The difference between the grammars of SIE and VIE, with respect to direct question formation, reduces to different rankings of the same constraints, which is expected in OT.

The generalization about indirect questions is the following: SIE does not permit inversion in indirect questions (=Noninversion) whereas VIE allows inversion in indirect questions (=Inversion). This grammatical distribution of inversion in the two varieties of English under consideration can be accounted for by the interaction of three constraints, two previously recruited to account for direct questions, viz., OP-SPEC, STAY, and OB-HEAD.

Consider first SIE. Since SIE does not permit inversion in indirect questions, OB-HEAD must have a lower prominence vis-a-vis OP-SPEC and STAY. We have already established that the grammar of SIE ranks OP-SPEC over STAY (7, above); the relevant ranking is given in (11).

The tableau (12) shows two competing candidates, both deferential to OP-SPEC. Since OP-SPEC cannot distinguish between the two candidates, the evaluation is passed on to the next important constraint, STAY. Again

both violate STAY, but it is the inversion structure that incurs two violations of STAY as opposed to non-inversion structure which violates STAY only once. In this competition, then, non-inverted structure is harmonic, and wins.

(11) SIE: OP-SPEC >> STAY >> OB-HD

(12) Tableau: SIE

Candidates	OP-SPEC	STAY	OB-HD
no-inv ⇒ I wonder [$_{CP}$ what ***e*** he is eating *t*]		*	*
inv I wonder [$_{CP}$ what is he *t* eating *t*]		**!	

Turning to indirect questions in VIE, recall that these require inversion with wh-movement, i.e., the fact that wh-phrase is followed by a head suggests that the movement of the wh-phrase is to a specifier position. The inversion facts in indirect questions in VIE follow straightforwardly from a constraint hierarchy where OB-HEAD outranks STAY and OP-SPEC, as shown in (13).

Once again, the tableau in (14) shows two competing candidates. The optimal, grammatical, output, given the dominance hierarchy in (13), is the inverted structure because the non-inverted structure violates OB-HEAD.

(13) VIE: OB-HD >> STAY >> OP-SPEC

(14) Tableau: VIE

Candidates	OB-HD	STAY	OP-SPEC
no-inv I wonder [$_{CP}$ what ***e*** he is eating *t*]	*!	*	
inv ⇒ I wonder [$_{CP}$ what is he *t* eating *t*]		**!	

With respect to indirect question formation, the difference between the grammars of SIE and VIE reduces, again, to different rankings of the same constraints, which is only expected given that OT appeals to variation in ranking to provide different grammars.

In sum, I have shown that the differences between the observed patterns of generalization of question formation in SIE and VIE are straightforwardly accounted for in a conceptualization of grammar that is based on a

general notion of priority. This OT-theoretic conceptualization allows us to capture the important generalization that the grammatical constraints that govern the syntactic behavior of VIE are *not* unique to it. Specifically, as I have shown that the difference between the grammars of SIE and VIE is reducible to different rankings of the same constraints, which is only expected given that OT appeals to variation in ranking to provide different grammars.

Concluding Remarks

In this paper, I have argued that to a (socio-)linguist the use of the term "non-native" English in the context of India makes no sense, just as it makes no sense to a biologist calling Penguins fossilized, non-native, birds. Having said that, I must admit that I have also unfortunately used the term "non-native English" in my own works on world Englishes (Bhatt 2001a), merely as a convenient term, without indulging of course in any critical assessment of that use. I am, needless to say, guilty of being part of the disciplinary culture that promotes the use of these ideologically loaded terms.

We must exercise caution against laying all the blame of "non-native" phenomena on the coiners of the term, non-native English, without a full consideration of the socio-historical and politico-academic conditions under which the term was first proposed and discussed. Let us not forget Kachru's response to Prator's "Linguistic heresy" paper in 1976, which was the first of its kind that established, with the help of subsequent works that followed, a counter-discourse – granting legitimacy to various "non-native" Englishes – that contested the dominant monolithic paradigm of "native" English; its acquisition and use world-wide. Since Prator's (1968) paper and Kachru's (1976) response, the field of inquiry has evolved, inviting more penetrating questions about the spread, variation, change and use of English world-wide, and, most importantly, a need to re-evaluate our own theoretical assumptions, our methodological approaches, and the terms used to provide theoretical analyses of English in the global context: its structure and practice.

Perhaps the question that needs to be addressed is: What, if any, is the need for us to discuss prescriptive prejudices of prescriptive grammarians? Clearly, one could argue that the prescriptive-linguists of the Quirk kind have obviously either missed the works of Saussure (1916), Chomsky

(1965) and Labov (1972) or failed to understand the implications of those works; what, for instance, do we mean by knowledge of language, what do dialect differences index, and how do we/can we account for difference? So, why bother? I will argue that unless we do not question their – and our own – assumptions, methods, and frameworks of analyses as an important part of an academic exercise, we will not be able to keep ourselves, or our profession, intellectually honest, or critically engaged with "disciplinary institutions" (TESOL, British Council) that manufacture "régimes of truth" (native/non-native, standard/non-standard) that serve only those in power, keeping the rest disenfranchised. The "govern-mentality" of those in power can only be exposed by questioning their analytical terms and by examining and challenging their theoretical-conceptual and empirical claims. In his contribution, Singh has opened up a discussion that will hopefully bring clarity to the choice of terms we use in producing and understanding our disciplinary discourses.

Acknowledgements

I am grateful to the organizers of the C.I.I.L. Symposium on Indian English/ English in India for the invitation to participate in this discussion.

Notes

1. It is in schools, argues Giroux (1981: 24), that the production of hegemonic ideologies 'hides' behind a number of legitimating forms. Some of the most obvious include: "(1) the claim by dominant classes that their interests represent the entire interests of the community; (2) the claim that conflict only occurs outside of the sphere of the political, i.e., economic conflict is viewed as non-political; (3) the presentation of specific forms of consciousness, beliefs, attitudes, values and practices as natural, universal, or even eternal."
2. Cultural capital here refers to the 'system of meanings, abilities, language forms, and tastes that are directly or indirectly defined by dominant groups as socially legitimate' (Apple 1978: 496).
3. The function of "the fellowships of discourse" is, according to Foucault (1972: 225–226), "to preserve or to reproduce discourse, but in order that it should circulate within a closed community, according to strict regulations, without those in possession being dispossessed by this very distribution. An archaic model of

this would be those groups of Rhapsodists, possessing knowledge of poems to recite or, even, upon which to work variations and transformations. But though the ultimate object of this knowledge was ritual recitation, it was protected and preserved within a determinate group, by the, often extremely complex, exercises of memory implied by such a process. Apprenticeship gained access both to a group and to a secret which recitation made manifest, but did not divulge. The roles of speaking and listening were not interchangeable"

4. By "Standard" I mean the variety that is assumed to carry prestige in local-national contexts of use, the variety that is used, in the American context for example, by evening news anchors on major networks (ABC, NBC, CBS, etc).
5. The non-inversion facts of Vernacular Indian English direct *wh*-questions are not mysterious; Standard Indian/British/American English questions with the question phrase *how come*, as in (a) below, also do not involve inversion.

   (a)How come this is grammatical?
6. It is not possible in this "response paper" to go over the details of optimality theory. Interested readers are referred to Grimashaw (1997) and Bhatt (2000).
7. The constraints used here have been formally analyzed/discussed in Grimshaw (1997):

   OP-SPEC: Operators must be in Specifier position.
   STAY: No movement (=trace) is allowed.
   OB-HD: Heads of selected projections must be filled (either by trace or overt material)

References

Anderson, Roger
  1984 What's gender good for, anyway? In *Second languages: a cross-linguistic perspective*, Roger Anderson, 77–90. Rowley, MA: Newbury House.

Apple, Michael W.
  1978 The new sociology of education: Analyzing cultural and economic reproduction. *Harvard Educational Review* 48: 495–503

Bhatt, Rakesh M.
  1995 The uprooted, the indentured, and the segregated: South African Indian English. *Journal of Pidgin and Creole Languages* 10: 381–396.
  1997 Optimality and the syntax of lectal variation. *Penn Working Papers in Linguistics* 4: 271–297.
  2000 Optimal expressions in Indian English. *English Language and Linguistics* 4: 69–95.
  2001a World Englishes. *Annual Review of Anthropology* 30: 527–550.

2001b Language economy, standardization, and world Englishes. In *The Three Circles of English*, E. Thumboo (ed.), 401–422. Singapore: UniPress.

2002 Experts, dialects, and discourse. *International Journal of Applied Linguistics* 12 (1): 74–109.

2004 The syntax of India English. In *Mouton Handbook of Varieties of English*, Elizabeth Traugott et al. (eds.), 213–227. Berlin/New York: Mouton de Gruyter.

2005 Expert discourses, local practices, and hybridity: The case of Indian Englishes. In *Language Practices and Identity: Negotiating the Global and Local*, S. Canagarajah (ed.), 25–54. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Bley-Vroman, Robert

1983 The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning* 33: 1–17

Bourdieu, Pierre

1977 *Outline of a Theory of Practice.* Cambridge: Cambridge University Press.

1991 *Language and Symbolic Power.* Cambridge: Harvard University Press.

Chomsky, Noam

1965 *Aspects of the Theory of Syntax.* Cambridge: MIT Press

1997 Questions – Belém. *D.E.L.T.A.* 13 (Special issue: Chomsky no Brasil): 39–48.

Cook, Vivian

1999 Going beyond the native speaker in language teaching. *TESOL Quarterly* 33: 185–209.

Corson, David

1997 Critical realism: An emancipatory philosophy for applied linguistics? *Applied Linguistics* 18: 166–188.

Davies, Alan

1991 *The native speaker in applied linguistics.* Edinburgh: Edinburgh University Press.

Dixon, R. M. W.

1996 Review of Francesca Merlan's book 'A grammar of Wardman, a language of the northern territory of Australia'. *Language* 72: 839–842.

Ferguson, Charles

1959 Diglossia. *Word* 15: 325–340.

Foucault, Michel

1972 *The Archaeology of Knowledge and the Discourse on Language.* Transl. A. M. Sheridan Smith. New York: Pantheon Books

Giroux, Henry
1981 Ideology, Culture and the Process of Schooling. Philadelphia: Temple University Press.
Grimshaw, Jane
1997 Projections, heads, and optimality. *Linguistic Inquiry* 28: 373–422.
Kachru, Braj
1976 Models of English for the third world: White man's linguistic burden or language pragmatics? *TESOL Quarterly* 10: 221–239.
Klein, E.
1994 Evidence for a "wild" L2 grammar: when PPs rear their empty heads. *Applied Linguistics* 16: 87–117.
Labov, William
1972 *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.
Lantolf, James P.
1996 SLA theory building: letting all the flowers bloom! *Language Learning* 46: 713–49.
O'Grady, William
1997 Second language acquisition theories of mind: A general nativist perspective. Paper presented at the 1997 Second Language Research Forum, Michigan State University, East Lansing, MI.
Polanyi, Michael
1967 *The Tacit Dimension*. New York: Anchor Books.
Prator, Clifford
1968 The British heresy in TESOL. In *Language Problems of Developing Nations*, Joshua Fishman, Charles Ferguson & Jyotindra Das Gupta (eds.), 459–476. New York: Wiley.
Prince, Alan & Paul Smolensky
2004 Optimality Theory: Constraint Interaction in Generative Grammar. Oxford: Blackwell.
Quirk, Randolph
1996 *Grammatical and Lexical Variance in English.* London: Longman.
Ramanathan, Vaidehi
1999 "English is here to stay": A critical look at institutional and educational practices in India. *TESOL Quarterly* 33: 211–231.
Saussure, Ferdinand de
1916 *Cours de linguistique générale*.
Schiffman, Harold
1996 *Linguistic Culture and Language Policy*. New York and London: Routledge.
Schwartz, Bonnie
1995 Transfer and L2A: Where are we now? Plenary paper delivered at the Second Language Research Forum, Ithaca, NY.

Schwartz, Bonnie & Sprouse, Rex
1996 L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12: 40–72.
Sells, Peter, John Rickford & T. Wasow
1996 An Optimality theoretic approach to variation in negative inversion in AAVE. *Natural Language and Linguistic Theory* 14: 591–627.
Singh, Rajendra
1974 Review of Nickel, *Papers in Contrastive Linguistics*. *Lingua* 33: 69–77.
2006 On Non-native Phenomena of English. Ms.
Singh, Rajendra, et al.
1995 On new/non-native Englishes. *Journal of Pragmatics* 24: 283–333.
Sridhar, S. N.
1994 A reality check for SLA theories. *TESOL Quarterly* 28: 800–805.
White, Lydia
1989 *Universal Grammar and Second Language Acquisition.* Amsterdam: John Benjamins.
1996 Universal grammar and second language acquisition: current trends and new directions. In *Handbook of Second Language Acquisition*, William Ritchie & Tej Bhatia (eds.), 85–120. New York: Academic Press.
Woolard, Kathryn
1985 Language variation and cultural hegemony: Toward an integration of sociolinguistic and social theory. *American Ethnologist* 12: 738–748.

The Athletics of English in India

Probal Dasgupta

This intervention seeks to respond to the reiteration by Rajendra Singh of his position that Indians who speak English proficiently are to be considered as native speakers of their variety of English and that correspondingly the term "non-native speaker" is inapplicable to them. His position is one of the important articulations of efforts by Indian users of English to contest the ownership claims over English by certain subcommunities of speakers. Those subcommunities are in fact hard to characterize. The term metropolitan, which is sometimes pressed into service for this purpose, has no formal means of zeroing in precisely on the American and British nations, and within those nations on the standard-owning geographical and social centres, with reference to which these discourses function. Nonetheless, for the lack of a better term, we may speak of the metropolitan speakers with reference to whom many others, including in particular such non-white speakers as Indians who use English, are routinely disenfranchised.

My own stand on these matters – formulated in The Otherness of English – needs to be reformulated in a way that may help certain readers to see more fully just how they propose to take on board the point that Rajendra Singh is making. I hope to provide in this text a reformulation that is sensitive to the closeness between the tools I employ and certain theoretical initiatives by Rajendra Singh himself (working in a different context and in collaboration with Alan Ford). At the same time I need to continue to stress certain aspects of the matter that the claim that proficient Indian users of English are to be categorized simply as native speakers of certain varieties of English might lead us to miss. Specifically, it seems to me that the default portrayal about the relation between a speech community and its language that the term "native speaker" (left unsupplemented) keeps in place is one that does not raise questions concerning the pedagogy of that language in the process whereby that community's children make it part of their lives. I have maintained, and will reiterate here, the position that questions of pedagogy and effort arise every time the use of English by Indians is

considered at a practical or theoretical level. Indians who are proficient in English exert themselves to cultivate and maintain this proficiency in a way that is written into their relation with the language. In the present text I stress the athletic dimension to point this up.

It has been a very long time since Indians began to learn English in a big way. Many missed opportunities in the colonial period made sense in that context. The opportunities missed after 1947 should elicit more concern and corrective action. Academics and other thinkers who look at English and hail from India or contextualize their work in relation to this country frequently miss the opportunities rendered accessible by the Chomsky revolution in linguistics. In some of my work I have pointed to certain paths that open up if we try and interconnect clusters of ideas that have been niched in mutually distant sectors in the linguistic, textual and social sciences. The purpose of the present intervention is to bring out some methodological coordinates of substantivist research on the English language in India. I propose to show here in some detail what role substantivist ideas play in the argument of The Otherness of English (Dasgupta 1993), a matter not explicitly worked out in that text.

Substantivism as a linguistic concept, introduced in Dasgupta 1989, was first shared with others in Abel 1998 (a paper based on earlier work done by Abel with me) and Dasgupta, Ford and Singh 2000. I myself use it to denote a set of interconnected ways of investigating matters of language that keep in view both the sentence grammar questions highlighted in several approaches and various exits from these formal questions into discourse.

The variousness of the locations of these exits remains a puzzle in a linguistic theory that seeks to model the individual speaker's performance and views a listener as only reconstructing what a speaker has done. Substantivism, on my take, seeks to consider these and related issues from the viewpoint of a listener in a state of dialogue (Ghosh 2003) who is sensitive to the potentially separate specificities of the speaker's planning, the listener's individual vantage point, the unfolding of the microdialogue in which the sentences are being uttered and received, and its niching within various macrodialogues invoked in the contextual coordinates a serious listener brings into play.

In seeking to attain such a multi-coordinated viewpoint, substantivism relies on the method of requiring that each level of formal representation count as co-specified with respect to two or more structural modules. This co-specification requirement ensures that the substance or body of a repre-

sentation does not remain uniquely imprisoned in any single system. Simultaneous multiple accountability keeps the body of a representation, be it a word or a phrase or a sentence, autonomous to the extent that it prevents unique attachment to any privileged specification.

This prevention of unique attachment is a postmodern trait. Dasgupta (1993) does bear markers of the postmodern persuasion. However, the formally focused theoretical articulation of substantivism in Dasgupta, Ford & Singh (2000) remains an important formulation of what this research programme shares with Ford & Singh's. Their emphatically modernist contributions in morphology (Whole Word Morphology) and in phonology (Generative Phonotactics) find, in Dasgupta, Ford & Singh (2000), a reformulation that associates substantivist linguistics with the programme of characterizing the patterns of what strikes particular speakers as easy (more natural) or as strenuous (less natural).

The morphological and phonological methods pioneered by Ford & Singh (classically formulated in Ford, Singh & Martohardjono 1997) contribute to the enterprise of identifying strenuous and easy stretches of interpretable sound at a microlinguistic level. The fuller development of a substantivist programme also requires the study of ease and strain in the macro domain. From Dasgupta (1993) to Dasgupta, Misra & Datta (2002), the substantivist study of English in India has steadily focused on the fact that speaking it in this country is a strenuous activity. As our understanding of the matter advances, all inquirers can be expected to see more clearly just how the modules of substantivism apply to this and related cases.

Indians who live in India and use English proficiently do so athletically, with competitive exertion written into the practice. It has to do with keeping up with certain Joneses, without focusing much on the faces of these competitors. The shift from a linguistic science rhetoric to a formulation mentioning economy and implying politics may lead some readers to misread me as rooting for this or that team in the political agora. Speaking at the level of Indian political team sports, though, my main point in The Otherness of English and predecessor texts was that the Congress, the BJP, the CPI(M) and other non-minor players uniformly sponsor Indian athletics in English in the same configurational relation to significant others. That point, and the pattern that it responds to, will remain stable for the foreseeable future.

Serious politics, in contrast, is more directly connected to scholarly inquiry. For it must formulate a praxis reconstituting the polis in keeping

with the growth of the rational in the logic of social formations. It is standard to distinguish the economy, one of the specific (and major) subsystems, from the general dimension of the economical, a dimension informing the way we run all subsystems of human effort. Given this terminological decision, what ultimately directs major social formation change belongs to the dimension of the economical and not to the economy.

Politics must then formulate a praxis geared to that dimension. The theoretical wing of that praxis must correspondingly seek to characterize what is often termed seriousness, as in the notion of serious speech acts in interpretive analytics. If this reasoning holds, it follows that serious politics needs to do business with a theory of language that provides substantive distinctions between speech and writing, between lightness and heaviness, between work and play, all over the range of language use even in a cybernetic age that redraws some boundaries. In that case, what is needed at the interface between the study of language and that of political economy is not just increased accountability on the part of the linguist to questions of politics as currently comprehended, but also some willingness on the part of political thinkers to recomprehend politics in tune with a substantivist linguistics.

The substantivist requirement of co-specification means at the macrolinguistic level that the discursive circulation in the speech community is to be anchored not only in an English that some Indians seek to wrest from hegemonic ownership by certain metropolitan users, but also in Indian languages whose sense of transparent and homogeneous self-presence does not survive the critique by Annamalai (2001). The specifics of Indophony (our use of our languages) get rearticulated as we expand and deepen our communities in concrete opposition to long entrenched practices of upper caste patriarchal hegemony masquerading as (and sometimes, complexly, doing double duty as) cultural self-defence against alien forces. The way our terminology makers keep thrusting truly alien Sanskrit "technical terms" on us does thinly veil a tendency (if not always a conscious desire) to postpone forever the moment of fully equipping our languages for technical and intellectual versatility. Once we see this, we begin to insist on our right to borrow certain English terms legitimately and naturally. For the point is to find ways of alleviating misery and of sustainably maximizing ease and naturalness in all appropriate contexts. This of course is not to say that all contexts can be entirely pain-free.

The inevitable pain of pedagogy is minimized when we face it with 'grace under pressure', Hemingway's definition of courage. Some readers tend to misunderstand courage as merely a moral matter of no academic significance. To preempt that misprision, and by way of emphasizing that courage encapsulates what considerations of methodology in the philosophy of science have helped us all to learn about empirical and scientific content, we shall use the term 'epistemic investment'. A statement, a theory, a move will only count as contributing to the advancement of sustainable knowledge, often called science, to the extent that putting forward that statement or theory or move means that the proponent is sticking his or her neck out and running the risk of being refuted, a matter of epistemic courage – or of epistemic investment, as we shall call it.

At the level of pedagogy, it represents a major epistemic investment when one proposes, as Anita Ravanam does, that we should move out of the collective mirage of the ideal dictionary and adopt the postmodernist lexical matrix, an implementation of Ford & Singh's Whole Word Morphology perspective (Ford, Singh & Martohardjono 1997) inflected in terms of the differentiated entry theory of the lexicon that Ravanam (2002) takes from Dasgupta (2002). This is perhaps as far as the postmodern transition can take us.

But we need to aim higher, for a transmodern transition enabled by substantivism. At the level of providing tools for context-sensitive sustainable pedagogy, taking the next step of lexico-phrasal kernel archiving (Dasgupta 2005) will require an even larger epistemic investment, and will involve articulating a politically vibrant level of metapedagogy as part of public discourse. An informed public discourse in a period marked by social activism, in order to aim for appropriately high standards of intellectual and political accountability, will have to balance the modern need for playing by transparent rules with the postmodern need to provide equitable and therefore differentiated playing fields for differently equipped constituencies.

One way to work towards this balance, whose terms and coordinates obviously cannot be specified a priori, is to put in place certain protocols of translation as a constitutive presence in the new public space. Sarukkai (2002) provides an multi-disciplinary and translation-focused theoretical bridge between the modernist foundations of scientific inquiry and the postmodern cultural discourse of the new public space. It is the affinity between his transmodern tools (which rigorously underwrite the freedom of inter-notational translation in science) and his views on the sphota-theoretic foundations of substantivism that make his work relevant to our concerns

here (see especially Sarukkai 2005: 240–250). My work on the transition to transmodernity (Dasgupta 2005) is in large part a response to Sarukkai's take on these issues.

The particular formulation of the programme of inquiry and archival action provided in Dasgupta 2005 takes the issue of notation as one of its points of departure. A linguistics capable of underwriting a serious politics cannot afford to leave the notational infrastructure of linguistic theory in its current semi-articulation. That phonologists have the international phonetic alphabet, syntacticians have trees, and semanticists have logico-mathematical symbolism means that three out of four components possess adequate equipment. But lexical workers have no formal means to explore the cross-linguistic and intra-linguistic lexico-grammatical correspondences between more compact and more diffuse word structures to which Tesniere (1959) drew the attention of all linguists.

Once we see the importance of lexical difficulty in the sense of Ravanam (2002) and the need for some type of lexico-phrasal kernel archiving as part of a public pedagogic apparatus enabling non-proficient users of a written language to access all its resources, it follows that interlexical inquiry must provide a lexical characterization of Tesnierean correspondences prior to any syntactic or semantic redescription. While the Hale & Keyser (2002) mode of lexico-syntactic research has rightly attracted attention, it tendentiously subsumes lexical work under syntactic inquiry, and it even more tendentiously builds the controversial incorporation account of the lexicon-syntax interface into its tree notation. What we need is something that combines the impartial independence of an IPA with specific fitness for the task of interlexical translation.

One obvious way to not reinvent the wheel in this domain is to use, as an interlexical archival glossing apparatus, some parts of the vocabulary and lexicological devices of the constructed language Esperanto, whose construction is sensitive precisely to interlexical realities just as IPA was built for the purposes for which it has been employed. It is shown in Dasgupta (2005) that such a move – which must of course not be confused with an adoption of Esperanto as a language, just as using IPA for phonetic transcription must be distinguished from underwriting an enterprise of orthographic reform that proposes the use of IPA based spelling systems – makes different predictions from the Hale & Keyser apparatus at crucial points. This tactical use of Esperanto devices performs demonstrably better on the descriptive adequacy front than the Hale & Keyser alternative does.

To effect that breakthrough at the appropriate epistemic investment level will involve facing the primordial and unexamined fears that currently deter many linguists and social scientists from having anything to do with the image of artificial forced globality that they associate with the idea of Esperanto. This fear of artifice – which takes the form of scholars feeling free to laugh at a colleague who takes Esperanto seriously – needs of course to be squarely faced without failing to notice the consequences of the fact that many of us have bought into the American model of international achievement so far as the form of our intellectual and political endeavours is concerned. I have argued (Dasgupta 2002) that English today is globally an "Olympian" language associated with projects of aggregating the "best" minds and their "best" products in certain central locations, condemning all peripheries to a permanent brain drain. The presentation in Dasgupta (2002) of Esperanto's alternative global vision took a default nationalism as its counterpoint and did not link up with the theme of India as "an area of listening" and of inclusion – or with transparency rooted in epistemic investment; hence the present intervention, which includes an attempt to spell out how Esperanto encodes a anti-centralizing, anti-Olympian approach to the dimension of the economical, and how this touches base with substantivism.

The identification of India, in Dasgupta (1993), as an area of listening and inclusion is to be read in the context of that text's reiteration of the Dasgupta (1988) notion of the society function of a particularized human language anchored in its literature, a visualization reaffirmed and expanded in Dasgupta (2005). Contemporary proposals to reinvent India as an American orbit dependent superpower compete with India's potential for articulating and advancing its listening-based enterprise of maximally open social inclusion of marginalized or disenfranchised subcommunities. When Ghosh (2003) reformulates sphota theory as a linguistics of listening, she thereby brings back into currency an essential ingredient for this sector of a cuisine that cannot afford to remain just Indian. It has long been independently clear that the Ford & Singh project in phonology and morphology as well as the Dasgupta thread of substantivism agree with and extend Bhartrihari's sphota architecture (Pillai 1971). The way Ghosh locates sphota theory in the thematics of listening is her own distinctive contribution to the substantivist sequel without with formal grammars miss the generative bus. Contemporary formal syntaxes, both those based on and those counterpointed to abstract derivational approaches, need the sphota oriented concrete sequel just as crucially as the ancient Indian version of formal deriva-

tional grammar did. Substantivism had been launched (Dasgupta 1989) precisely to work out the details of how this sequel sets things right.

Just where does the thematics of substantivism converge with that of Esperanto? Just how does the resulting vision differ from the Olympian vision that notions of world English stand for and that many of my interlocutors have bought into?

World English encodes a rationality that takes it that rational people and non-rational people are distinct individuals; that rational people are supposed to dominate non-rationals; that English encodes rationality to the extent that it has taken over the scientific enterprise and the marketplace; that various others have to fall in line in order to get a bit part in the comedy of emergent world civilization that will put the bitter legacies of colonialisms behind us; and that there is no constitutive reason for rational people to work for the conservation of the languages of the world. This vision involves erasing everybody's histories if that is what the relevant individuals, reshaped by an Americanoid education, "freely choose" to do, and making all discourse anaphoric to the Anglo-American canon in the sense of discursive anaphora invoked throughout Dasgupta (1993).

The vision of rationality that has been cultivated in the community of Esperanto speakers and writers (see Janton (1993) for a point of entry and Montagut (1994) for a specific articulation) opposes the assumptions outlined above point for point.

Rationality, a set of interconnected (and interconnective) projects on the plane of the economical in the sense explicated above, is not predicable of individuals per se. Individuals who strive for maximization of interconnective transparency do not set themselves the task of dominating others who do not yet see the point of such maximization. English has become a monopoly system for the prize-focused, championship-focused aspect of the scientific and industrial game, and has been associated with a massive disenfranchisement of the right of non-Anglophone scientists to enunciate scientific claims in non-English languages and to expect these claims to be tested and validated in those languages. This disenfranchisement, to the extent that scientists have permitted this to happen, is not only an atrocity at the level of the practice of science, but is demonstrably a process undermining rationality itself and the principle of maximizing the scientific task of seeking relevant criticism from as many potential critics as possible. To the extent that scientists have worked to strengthen the monopoly of English, they have permitted themselves to surrender to a commercial process that

undermines the very foundations of open scientific traffic. Not only does the defence of freedom require opposition to this increasingly monopolistic use of English in such domains as science. Furthermore, the defence of the legacy of anti-colonial and other struggles for equity, which amounts to a defence of the public space and the open society, requires the explicit and rigorously argued-for promotion of other languages and rolling back the domain growth of English.

It is the responsibility of those who write, who work in the public space, to remain not only accountable to the public today, in all its multilinguality, but also capable of keeping this public plugged into the legacy of the publics of the shared past. Since the publics of that past were also given to sharing and in that fundamental sense rational (though all rationality is punctuated by pathology), and since they expressed their rationality in the crucibles of various languages that are still spoken, it follows that articulators of consciousness who express the way we, the various "we"s, are able to anchor ourselves in these pasts, have the responsibility for cultivating full and uncurtailed self-expression in all the languages we find ourselves initially anchored in, in all the languages children first acquire (we can afford to call them "mother tongues" only if this term does not impose the image of a unique, monolingual, nationalism-nurtured symbol of blood and soil).

This self-expression must, in order to keep faith with the fundamental mandate of the human enterprise, include the full intellectual range of writing, including scientific work. If there are languages in which nobody writes science articles, it becomes the responsibility of the scientific and social systems to ensure that science writing is produced or translated into those languages and is made available through such channels that the primary speech communities of those languages are able to receive and criticize the proposals made by the science and engineering elite that affect the lives of those speech communities. If the scientific and social systems do not meet this criterion, they are working against science, not for science, and this point has to be hammered into the heads of the directors of institutes of science and technology.

Esperanto is a medium of sharing and self-expression whose textual history encodes the will of people who interconnect rationalities across obvious cultural boundaries, who thereby seek to express the human as a distinctive element and to downgrade the fascist authority of monoglot cultural "histories" (which in fact rest on the falsification of the historical record in order

to bend the general light into special ethnic gravitational fields), and who have worked long, hard and articulately to mediate between the notion of universal language as "language for everybody, as easy as possible" and that of universal language as "language for everything, as complex and far-reaching as possible". That mediation has played out in the cultural history and self-consciousness of the Esperanto literary and intellectual landscape over more than a century, in texts (by such polar opposite figures as Kalocsay Kalman and Baghy Gyula, for instance) whose value will become obvious to participants in this debate when they learn the language and look up its standard bibliographies. Janton is only a starting point.

To return to the task of working out just where the vision of Esperanto converges with that of India as an area of listening, we now see that Esperanto, correspondingly, is a medium marked by individuals who have made it a point to listen, and by philosophy in the sense of an enterprise that "considers all the evidence" (Whitehead) and "gives no prizes" (Wittgenstein).

My interlocutors who wish to make serious claims for English in the present context and who believe that an English-based future for humanity can achieve equity will perhaps find themselves trying to reinvent these wheels in English and turning English into a version of Esperanto. It might make better sense for them to adopt the original design of the wheel in the first place, and to make the adjustments there rather than first force the irrational orthography and prepositions of English down everybody's throat, then apologize for this atrocity, and finally take corrective action that could have been taken much earlier. English is known to produce high degrees of dyslexia because of its pathological spelling system, and a world struggling to achieve widespread literacy can ill afford the consequences of continuing to ride the wave of the English language and the hegemonies complicit with it.

Substantivism in linguistics is built around a theoretical characterization of economy that avoids extractive generalizations. This is as clear in the Dasgupta & Ghosh components of substantivism as it is in the Ford & Singh corpus. It is characteristic of Ford & Singh morphology (Ford, Singh & Martohardjono 1997) that one does not extract an affixation process and call it an affix morpheme reified as a lexical entry; one keeps the affixations in place and beholds the generalizing pattern as a totality that leaves each token intact and in place. On the one hand, substantivist linguistics is rigorous and accountable at the level of the best contributions of modernist science; at that level, it keeps faith with the scientific impulse and strives to articu-

late equations of the kind that drive the translation machine in Sarukkai 2002. On the other hand, substantivist linguistics is non-extractive and therefore respects the specificity of each precinct or locality, and in this sense touches base with the central insight of postmodern inquiry and politics, namely, that every region has distinctive patterns that cannot be exhaustively explained by universalistic scientific postulates and axioms. It has been argued elsewhere that the transmodern combination of modernist and postmodernist imperatives is the pattern that emerges when we pursue a substantivist enterprise in microlinguistics on the basis of the full range of morphological and syntactic evidence available (Dasgupta 2006). One goal of the present intervention is to give some access to the threads of reasoning that show that the same transmodern combination becomes crucial also when we take a substantivist path in macrolinguistic inquiry. The particular case of the English language in the Indian context provides food for this type of thought especially because English, as an area of special and difficult learning for Indians, is in substantive articulation with the Indian languages, which are the easy base from which this learning is achieved.

I am sure that scholars who disagree with this account, and who believe that it is epistemically acceptable on their part to not face the evidence (much of it in Esperanto) that this account is based on, will produce their own response to the fact that the English performance by Indians is an athletic performance. If they do, then there will be a common ground on which I can meet those critics who do not wish to make the epistemic investment of meeting me on the much older common ground of the debates in Esperanto about appropriate forms of universality in language and literature. If they do not, then the ways of the abridged enlightenment will remain as they are.

Acknowledgements

I am grateful to the organizers of the C.I.I.L. Symposium on Indian English/ English in India for their invitation to participate in it.

References

Abel, P. Rekha
1998 Diglossia as a linguistic reality. In *The Yearbook of South Asian Languages and Linguistics 1998*, Rajendra Singh (ed.), 83–103. New Delhi / Thousand Oaks / London: Sage.

Annamalai, E.
2001 *Managing Multilingualism in India: Political and Linguistic Manifestations.* New Delhi / Thousand Oaks / London: Sage.

Dasgupta, Probal
1988 The external reality of linguistic descriptions. *Canadian Journal of Linguistics* 33 (4): 345–365.
1989 Outgrowing Quine: towards a substantivist theory of translation. *International Journal of Translation* 1 (2): 13–41.
1993 *The Otherness of English: India's Auntie Tongue Syndrome*. New Delhi / Thousand Oaks / London: Sage.
2005 The south and Esperanto in the postorientalist public space. Invited talk, Institute for General Linguistics, University of Helsinki, 26 October 2005. [Forthcoming in Indian Journal of Linguistics 33 [2006].]
2006 Transmodern linguistics and its puurvapakshas. Paper presented at seminar on 'Foundations of Sciences', National Institute of Advanced Studies, Bangalore.

Dasgupta, Probal, Alan Ford & Rajendra Singh
2000 *After Etymology: Towards a Substantivist Linguistics.* Muenchen: Lincom Europa.

Dasgupta, Probal, Udayon Misra & Amaresh Datta
2002 *English at Crossroads: The Post-colonial Context*. Guwahati: Students' Stores.

Ford, Alan, Rajendra Singh & Gita Martohardjono
1997 *Pace Panini: Towards a Word-based Theory of Morphology.* New York: Peter Lang.

Ghosh, Sanjukta
2003 The syntax-pragmatics interface in Bangla. University of Hyderabad Ph.D. dissertation.

Hale, Ken & Samuel Jay Keyser
2002 *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.

Janton, Pierre
1993 *Esperanto: Language, Literature and Community*. Albany: State University of New York Press.

Montagut, Abel
1994 *Poemo de Utnoa*. Wien: Pro Esperanto.

Pillai, K. Raghavan
1971 *The Vakyapadiya: Critical Text of Cantos I and II (with English translation, summary of ideas and notes).* Delhi: Motilal Banarsidass.

Ravanam, Anita
2002 Three levels of lexical codification. In *The Yearbook of South Asian Languages and Linguistics 2002*, Rajendra Singh (ed.), 135–155. New Delhi/ Thousand Oaks/ London: Sage.

Sarukkai, Sundar
2002 *Translating the World: Science and Language*. New York: University Press of America.
2005 Indian philosophy and philosophy of science. *History of Science, Philosophy and Culture in Indian Civilization* 15. New Delhi: Centre for Studies in Civilizations/Motilal Banarsidass.

Tesnière, Lucien
1959 *Éléments de syntaxe structurale*. Paris: Klincksieck.

Let's Face the Music: The Multilingual Challenge*

Claudia Lange

Let me start with a quote that may ring a bell for those who are familiar with Rajendra Singh's work:

> […] What all this means is that the monolingual approach is neither appropriate nor adequate for the investigation of language use in a society where multilingualism was endemic and where, for the educated at least, monolingualism was the exception and not the norm.

This quote might have been lifted directly from Rajendra Singh's paper, but it actually comes from David A. Trotter (Trotter 2000: 2f.), a medievalist and scholar of Anglo-Norman, the language spoken by the French conquerors of England who were so successful in and after 1066. Along with others (e.g. Rothwell 2000; Wright 2000), Trotter has relentlessly campaigned for a re-appraisal of Middle English as a contact language.

The "Old Englishes", as we might call them in analogy to the "New Englishes", can only be adequately conceptualized in their multilingual setting. After the Norman Conquest, Old English almost completely ceased to be the language of record, to be replaced by the pan-European prestige language Latin and later French, or rather Anglo-Norman, the language of the rulers. When English, still the language of the vast majority of the population, came to be used as a written language again to a larger extent, it had been developing alongside French and Latin, and there are many later medieval texts where the boundaries between the three languages are very difficult to draw indeed. This kind of evidence for a multilingual scenario in medieval England has prompted Trotter's stance against "the monolingual approach".

But whereas multilingualism in medieval England would probably be classified as "elite bilingualism" (Skutnabb-Kangas 1984: 75) because it was tied to literacy, this is not the case in countries like India, where multilingualism is very much a fact of everybody's life, be they literate or not. Leaving this difference aside, accepting the multilingual challenge necessar-

ily entails a less categorical, more flexible approach to the concept of "native speaker" for scholars of "Old Englishes" as well as "New Englishes". Practitioners of "Old Englishes" obviously have a harder time here, since they tend to come from Western countries where – leaving Latin as the language of the educated aside – monolingualism rather than multilingualism is endemic. Nowadays, as linguists from monolingual countries, we are in principle aware of the fact that monolingualism is the exception and multilingualism the norm world-wide, but to truly let go of the naïve idea about native language as the sole language of an individual and his or her speech community is a different matter. The equation "one nation – one language" was only established in Europe in the last centuries (e.g. in France in the name of Enlightenment, cf. Bourhis 1997) and not without extensive conflict, but it has deeply penetrated our politics and our theories.

Singh, Trotter and the many other scholars of "New" and "Old" Englishes who have exposed the monolingual bias in contemporary linguistic theory as well as in more general attitudes towards language(s) are thus perfectly justified in doing so. I would, however, maintain that this myopia did not arise because the West is intrinsically evil or upholding the image of its supremacy by deliberately withholding information about variation in the "native" varieties from the "non-natives". Theories and models are not context-free, they are embedded in contexts of discovery and contexts of justification. The Chomskyan idealized native speaker in a completely homogeneous speech community has loomed large in linguistics for decades. Now, in the 21st century, there may well be a paradigm shift under way. The linguistic evidence is overwhelming: there is a growing interest in World Englishes and increasing recognition of variation even in fields like historical linguistics and formal syntax.[1] Further, scholars like Sarah Grey Thomason (Thomason & Kaufman 1991; Thomason 2001), Rajend Mesthrie (Mesthrie 1992, 2003) and Salikoko Mufwene (e.g. Mufwene 2001) have paved the way for a new and better understanding of language contact. And there are several approaches to the study of standardization in language and the mechanisms by which one variety/language gets selected over others to become the national standard language.[2]

Linguistic variation is also asserted and acknowledged at the political level. More and more individuals and communities within Europe realize that monolingualism is neither natural nor desirable. The Council of Europe has passed the European Charter for Regional or Minority Languages in 1992, in response to demands from linguistic minorities within European

countries.[3] Regionalism has developed into a powerful force in Europe over the last years, and it may be a historical irony that this process is particularly evident in the United Kingdom, where Scotland and Wales now have their own national assemblies. The "one nation – one language" formula is clearly losing its grip on Europe and Europeans.

How, then, is this paradigm shift apparent in the study of Indian English? Obviously, not all users of English in India are native speakers of English in the traditional sense that English is acquired during the period of primary socialisation. This is only to be expected in a multilingual society where there never has been a one-to-one relationship between "nation" and "national language". Mufwene points out that English is a native language in India for only a negligible proportion of the population, "however, regardless of the proportion of native speakers, each new variety of a European language exported to (former) colonies has developed its own autonomous norm" (Mufwene 1998: 112). I fully agree with Mufwene when he continues:

> […] the umbrella term for the arbiter of well-formedness and appropriateness in any community is the proficient speaker, one who is fully competent in a language variety according to the established norm of the community using it. Whether or not such a speaker it typically native will vary from one community to another. (Mufwene 1998: 117)

This is of course entirely consistent with Singh's definition of a native speaker/user as "a speaker/user whose well-formedness judgements on utterances said to be from her language are shared by the community she can be said to be a member of." Keeping this in mind, there is no further obstacle to treating Indian English as an established variety worthy of description in its own right. Many scholars have already made significant contributions towards an assessment of the "range and depth in Indian English" (D'Souza 2001: 145). S. V. Parasher, for example, has carried out a survey of "Educated Indian English (EIE)" (Parasher 1991: 53), "to study the use of English by educated bilinguals at a fairly high point on the cline of bilingualism […] those bilingual users of English who were engaged in learned professions and white-collar jobs" (1991: 66). The other end of the cline of bilingualism is covered, for example, by Priya Hosali's studies on "Butler English" (e.g. Hosali 2005). Reliable descriptions of different varieties of Indian English are indispensable for eventually arriving at a reference grammar of contemporary Indian English, as S. N. Sridhar states:

> A grammar of an IVE [indigenized variety of English] can be written only if it is possible to determine reliably whether a given structure occurs in a specific variety, whether it has a regional or panregional distribution and acceptability, and so forth. (Sridhar 1996: 56)

It is frequently argued that there is no such thing as Indian English, and that there never will be in a country where the range of variation, or simply the access to English is determined by so many different factors, among them caste, the urban-rural divide etc. But here we need to distinguish between the teacher's and educator's point of view and that of the linguist. The former two are naturally more concerned with learner varieties of Indian English. These should not be excluded from consideration; in fact, Sridhar's paper quoted above is a description of such a learner variety, one which he classifies as "lower mesolectal" relative to a "standard or acrolect that (1) is not too strongly marked by varietal features of one particular region, (2) is free from stigmatized features such as gross agreement violations, and (3) enjoys pan-South Asian distribution, intelligibility, and positive evaluation" (Sridhar 1996: 67). The former two, and Sridhar among them, thus take for granted that there is a common ground for proficient speakers/users of English in India and that there is some supralocal norm either already in place or emerging right now.

Let me call research that simply treats Indian English as one variety among others "New School" research. "Old School" work that considers Indian English in terms of its "deviance" has fortunately become rarer, but is by no means extinct. Raja Ram Mehrotra's contributions to the study of Indian English, for example, are unabashedly Old School. The title of his article "A British response to some Indian English usages: Is Indian English significantly opaque to outsiders?" (2003) already indicates that it is a prime example of what Rajendra Singh in his paper refers to as the "error-driven, pedagogical enterprise that studies what it calls non-native varieties of English". Mehrotra presented a list of purportedly Indian English expressions (such as *chaste, bed-tea, beer-bottle, to do the needful*) to "native speakers of English", that is, British English speakers, and asked them to give an intelligibility rating. Even if he concludes his paper by asking "Why should then the correctness or acceptability of IE be all the time required to look upon BrE for norms?" (Mehrotra 2003: 25), the whole setup of his study just serves to perpetuate the native/non-native distinction it is supposed to dispense with. In earlier papers (Mehrotra 2001, 2002), Mehrotra again is

primarily concerned with value judgments about "Indianisms". To quote Rajendra Singh once more, these papers represent "journalistic reports on exotica" rather than serious contributions to a grammar of Indian English.

Coming back to "New School" research: I would like to take a short look at the new research possibilities opened up by Corpus Linguistics, particularly the *ICE* (*International Corpus of English*) project. In doing so, we shall also see how the paradigm shift I referred to above has found its way even into the corpus design.

The *ICE*-project as initiated by the late Sidney Greenbaum is designed to provide parallel corpora of international varieties of English, so that comparative studies of World Englishes can be undertaken.[4] The ICE-India corpus was released in 2002; it follows the general ICE text selection and coding principles in including one million words of spoken and written *standard* Indian English, where *standard* is defined according to speaker:

> The authors and speakers of the texts are aged 18 or over, were educated through the medium of English, and were either born in the country in whose corpus they are included, or moved there at an early age and received their education through the medium of English in the country concerned.
> (http://www.ucl.ac.uk/english-usage/ice/design.htm)

This definition is not a very practical one for a country like India, where English-medium education is only available (or affordable) for a fraction of the population. The compilers of *ICE*-India have unobstrusively adapted this requirement to the Indian context in their selection of speakers: "The category of 'conversations' are drawn largely from the trained ELT teachers, though they have not been educated in the medium of English at all levels" (Shastri 2002). S. V. Shastri and his team have just gone along with the Indian reality with generally

> [...] more than one mother tongue for a person, as in a multilingual context the socialisation of a child may involve multiple languages. For instance, an Oriya boy marrying a Tamil girl, speaking mostly English at home and employing a Hindustani 'Ayah' could bring up a child who would be using four languages before formal training. (Pattanayak 1998: 127–128)

Thus, what might have been left of a monolingual bias in the original definition of "standard speaker" is gone. The "trained ELT teachers" referred to above certainly are among the "arbiter[s] of well-formedness and appropri-

ateness" (Mufwene 1998: 117) who provide the norms within their speech community, whether they went to English-medium schools or not. The proficient speaker/user of English has now taken the place of the Old School "native speaker", and he or she is likely to stay.

I still think that a term like "New Englishes" makes sense – even if the native/non-native distinction does not – because it reminds us of the multilingual settings in which English is just one option in speakers' repertoires. From this perspective, the study of New Englishes is a particularly rewarding pursuit: It will deepen our understanding of the role of contact and the interplay of internal and external factors in language change, and it will offer new insights into the process of standardization. It will, eventually, lead to better linguistic theory.

Notes

* An earlier version of this paper was presented at the Mysore Symposium on Indian English/English in India, the Central Institute of Indian Languages, Mysore, January 4, 2007. Singh will, I hope, forgive me for exploiting the metaphor he and his colleagues used when they published their article "On new/non-native Englishes: A quartet" (Singh et al. 1995); my title refers to his quartet as well as the *gamelan* to which it gave rise. Afendras et al. (1995: 295) labelled their responses to the original quartet a *gamelan*, "'the indigenous orchestra of Java and Bali', characterized by its 'highly developed polyphony and heterophony'".

My first ever visit to India was in 2004 when I took part in an exchange programme funded by the German Academic Exchange Service (DAAD), to which I am grateful, and spent three months as a guest lecturer in the English department of the University of Pune. After three further visits to India in the following years, I am still only beginning to understand the complexities of the Indian multilingual context, but the experience of living and working in a multilingual society clearly has had a profound impact on my thinking about my profession and its main pursuit, namely linguistics and the investigation of language(s).

1. Besides Trotter's work (2000) on multilingualism, there are further the relatively new fields Historical Sociolinguistics (for example Milroy (1992), Nevalainen (1996)) and Historical Pragmatics (cf. Jucker (1995)) which explicitly deal with variation rather than abstracting away from it. Barbiers, Cornips & van der Kleij (2001) treat "Syntactic Microvariation" from a generative perspective.
2. For standardization processes relating to the Germanic languages see Deumert & Vandenbussche (2004). An ongoing research project at the University of Leiden, Netherlands is concerned with "The codifiers and the English language: tracing the norms of Standard English" (www.lucl.leidenuniv.nl/index.php3?m=13&c=126).

Mazzon (2000) examines the ideological underpinnings of standardization processes with reference to New Englishes.
3. http://conventions.coe.int/Treaty/EN/Treaties/Html/148.htm.
4. For a comprehensive overview of the corpus design, sampling techniques and research possibilities see Greenbaum (1996). Some recent studies that are based on several *ICE*-subcorpora, among them *ICE*-India, are Sand (2004) and Schneider (2004).

References

Afendras, Evangelos A. et al.
1995 Discussion: On 'new/non-native' Englishes: A gamelan. *Journal of Pragmatics.* 24: 283–294.
Barbiers, Sjef, Leonie Cornips & Susanne van der Kleij (eds.)
2001 *Syntactic Microvariation.* Amsterdam: Meertens Institute. Online publication: http://www.meertens.knaw.nl/books/synmic/.
Bourhis, Richard Y.
1997 Language policies and Language Attitudes: le Monde de la Francophonie. In *Sociolinguistics: A Reader and Coursebook*, Nikolas Coupland & Adam Jaworski (eds.), 306–322. Houndmills: Macmillan.
D'Souza, Jean
2001 Contextualizing range and depth in Indian English. *World Englishes* 20 (2): 145–159.
Deumert, Ana & Wim Vandenbussche (eds.)
2004 *Germanic Standardizations: Past to Present.* Amsterdam: Benjamins.
Greenbaum, Sidney (ed.)
1996 *Comparing English Worldwide: The International Corpus of English.* Oxford: Clarendon.
Hosali, Priya
2005 Butler English: An account of a highly distinctive variety of English in India. *English Today* 81 (21, 1): 34–39.
Jucker, Andreas H. (ed.)
1995 *Historical Pragmatics: Pragmatic Developments in the History of English.* Amsterdam: Benjamins.
Lange, Claudia
2007 Focus Marking in Indian English. *English World-Wide* 28 (1): 89–118.
Mazzon, Gabriella
2000 The ideology of the standard and the development of extraterritorial Englishes. In *The Development of Standard English 1300–1800. Theories, Descriptions, Conflicts*, Laura Wright (ed.), 73–92. Cambridge: Cambridge University Press.

Mehrotra, Raja Ram
2001 English in bereavement advertising in India. *English Today* 67 (17, 3): 31–35.
2002 English in private letters in India. *English Today* 72 (18, 4): 39–44.
2003 A British response to some Indian English Usages. *English Today* 75 (19, 3): 19–25.
Mesthrie, Rajend
1992 *English in Language Shift: The History, Structure and Sociolinguistics of South African Indian English.* Cambridge: Cambridge University Press.
2003 The World Englishes paradigm and contact linguistics: refurbishing the foundations. *World Englishes* 22 (4): 449–461.
Milroy, James
1992 *Linguistic Variation and Change: on the Historical Sociolinguistics of English*. Oxford: Blackwell.
Mufwene, Salikoko
1998 Native speaker, proficient speaker and norms. In Rajendra Singh (ed.), 111–123.
2001 *The Ecology of Language Evolution.* Cambridge: Cambridge University Press.
Nevalainen, Terttu (ed.)
1996 *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence.* Amsterdam: Rodopi.
Oesterreicher, Wulf
2001 Historizität – Sprachvariation, Sprachverschiedenheit, Sprachwandel. In *Language Typology and Language Universals: An International Handbook*, Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), 1554–1595. Berlin/New York: Mouton de Gruyter. (Handbücher zur Sprach- und Kommunikationswissenschaft).
Parasher, S. V.
1991 *Indian English: Functions and Form.* New Delhi: Bahri Publications.
Pattanayak, D. P.
1998 Mother Tongue: An Indian Context. In Rajendra Singh (ed.), 127–147.
Rothwell, William
2000 Aspects of Morphosyntactic Mixing in the Languages of Medieval England. In *Multilingualism in Later Medieval Britain*, David A. Trotter (ed.), 213–232. Suffolk: D. S. Brewer.
Sand, Andrea
2004 Shared morpho-syntactic features in contact varieties of English: article use. *World Englishes* 23 (2): 281–298.
Schneider, Edgar W.
2004 How to trace structural nativization: particle verbs in world Englishes. *World Englishes* 23 (2): 227–249.

Shastri, S. V.
2002 *Overview of the Indian Component of the International Corpus of English (ICE-India)*. Distributed with the *ICE*-India corpus. http://www.ucl.ac.uk/english-usage/ice/download/index.htm.

Singh, Rajendra (ed.)
1998 *The Native Speaker: Multilingual Perspectives*. New Delhi: Sage.

Singh, Rajendra et al.
1995 On new/non-native Englishes: A quartet. *Journal of Pragmatics* 24: 283–294.

Skutnabb-Kangas, Tove
1984 *Bilingualism or Not: The Education of Minorities.* Clevedon, Avon: Multilingual Matters.

Sridhar, S.N.
1996 Toward a syntax of South Asian English: Defining the lectal range. In *South Asian English: Structure, Use, and Users*, Robert J. Baumgardner (ed.), 56–69. Urbana/Chicago: University of Illinois Press.

Thomason, Sarah Grey
2001 *Language Contact*. Washington, DC: Georgetown University Press.

Thomason, Sarah Grey & Terrence Kaufman
1991 *Language Contact, Creolization, and Genetic Linguistics*. Berkeley/Los Angeles/Oxford: University of California Press.

Trotter, David A.
2000 Multilingualism in Later Medieval Britain. In *Multilingualism in Later Medieval Britain*, David A. Trotter (ed.), 1–5. Suffolk: D. S. Brewer.

Wright, Laura
2000 Bills, accounts, inventories: Everyday trilingual activities in the business world of later medieval England. In *Multilingualism in Later Medieval Britain*, David A. Trotter (ed.), 149–156. Suffolk: D. S. Brewer.

The Notion 'Native Speaker': A Philosophical Response*

Syed A. Sayeed

In this response, I will not try to engage with all or even most of the several complex issues Singh discusses in his paper. I shall confine myself to a response to what I see as the central contention he makes, which is as follows: the notion of a native speaker represents a legitimate linguistic category, even though it was motivated by weak pedagogic concerns and is somewhat outdated and of not much consequence in the context of the current state of the discipline. But the idea of 'non-native varieties of a language' is an incoherent 'socio-linguistic' concept, derived through an illegitimate extension of the notion of native speaker, and existing as part of a dubious discursive enterprise that belongs neither to rigorous linguistics nor to critical social science.

My response will be largely sympathetic to Singh's position but with certain reservations. Whether the reservations I articulate here are very pertinent or not, I hope they will contribute to moving the debate further in a meaningful direction.

One of the main points with which Singh is concerned in his paper is related to the question of what constitutes a linguistic unit for certain purposes. His own position seems to me to consist of an application of a straightforward Wittgensteinian point according to which if there is a consistent, rule-governed structure, it must be treated as a system, and error should be defined only in terms of deviation from the rules that constitute that structure. It makes no sense to treat that structure itself as a deviant, abnormal form of another, more primary structure and treat its rules as distortions of the rules of the 'original', 'normal' structure. And if, for some taxonomical purposes we choose to so designate the second structure in terms of deviation from the 'original' structure, we must avoid drawing normative consequences from it. That is to say, we must not interpret compliance with the rules of the 'deviant' system as imperfect or incomplete obedience of the rules of the 'original' system. In other words, any coherent

set of language games is a language and the relation of that set to any other set in terms of structural resemblance, overlap of constitutive subsets, or history of emergence, is totally irrelevant. This also means that there are no a priori limits on the size or scope of any such set. That is, there is no minimum required richness without which such a set cannot be treated as an independent language. A fortiori, there is no minimum number of speakers required to qualify such a system to the status of a language. In principle there can be a language in this sense with only one speaker. (We must avoid private-language anxieties here. The issue of private language concerns *privacy* in the sense of theoretical inaccessibility: A one-man language that contingently happens to remain that way is not at all problematic. What is contested is a one-man language that can imperviously remain that way.). I must confess that I find Singh's position on this point unexceptionable. I agree that while it is harmless to talk in terms of varieties of English ordered and labelled in a certain way, we must not attach normative significance to that taxonomy. The notion of 'non-native' involves precisely such significance. And it is also true that such a move would have not only discursive consequences in the sense that it would introduce methodological irritants into the study of the languages in question, but also political consequences in the sense that it would create tensions that can be resolved only by hierarchising the entire domain comprising the concerned speech communities. That in turn would generate pressures on some of them to accept a secondary status and face the choice either to live the guilt of 'illegality' or try to gravitate towards the originary language. However, while his linguistic and political points are well taken, I think Singh does not pay sufficient attention to the *social* dimension of the issue. What I mean is this: In principle, there can be a language spoken by just half a dozen people and there is no problem as long as they live in relative isolation or in distinct social spaces. But in practice that rarely happens. In the case of the English languages (I use the term advisedly) it almost never happens. Structural normativity becomes inevitable. By this I only mean that a set of rules become the de fault norm and the pressure to conform to them is spontaneously generated. This pressure may have political consequences but may not be politically motivated. It can just be a reaction to the frustration attendant on such a situation. There can be two factors to this frustration: first, the rule of my variant generates a meaning different from that of my interlocutor and the constant reinterpretation can be a strain. Second, even if there is no distinct meaning variation, the sense of arbitrariness may be quite annoying and obstructive

to smooth communication. That is to say, in practice, there may be problems in granting the status of a unit to every coherent set of language games. Of course it can be argued that this is to do with the issue of standardisation. But I think the notion of 'nativity' involves the compulsions of standardisation at some level.

As far as Singh's reservations about the notion of 'native speaker' are concerned, once again I find them on the whole acceptable. But I feel that in his anxiety to be sensitive to certain aspects of the issue, he tends to be rather short with other equally legitimate aspects of the question. I have no quarrel with his point that the phenomenon of bilingualism complicates the notion of native speaker understood in terms of first language or mother tongue, but it is by no means obvious that the notion of native speaker implies that one can be a native speaker of only one language. Further, I find myself in agreement with Singh when he says that Thomas Paikeday's definition of a native speaker as a 'proficient user of a specific language' raises more problems than it solves, but I cannot help complaining that he is not willing to draw out one important implication of his own criticism of Paikeday which is that there is a historical dimension to the notion of native speaker which cannot be reduced to functional criteria. His own definition, which is characteristically Wittgensteinian, makes 'relatively stable and consistent grammatical judgements' central to the notion in question. However, there is a circularity to this definition. It is not a logical circularity as much as a sort of explanatory redundancy. Let me explain what I mean.

Singh takes meta-assertions by a language community about its language as primary. A native speaker, on this approach, is one who is taken to be a member by that community. This intra-community consensus approach leaves out any way of recognizing a native speaker on the basis of the relation she has to the language. That relation cannot be, as pointed out earlier, only functional. It will have many strands – structural as well historical. But bracketing them all and concentrating on shared 'well-formed judgements' on utterances from that language makes the notion of native speaker virtually redundant. Singh himself would of course not be particularly upset by that prospect. But there is a contradiction inasmuch he concedes the relevance of the notion in pedagogic as well as other contexts.

Regarding the question of lexicon, I find Singh's position a little ambiguous. I would like to know how he would like to see his own assertion that 'morphologically complex words in IE are…fully licensed by word-formation rules of English morphology'. My point here is that Singh him-

self is positing a criterional relation between English morphology and the morphology of IE and he does not seem to be aware that such a move involves him in a normative relation between the two languages which provides ground for such expressions as 'non-native varieties of English'.

Extending the same point further, we can say that Singh is guilty of contradiction insofar as he argues that speakers of Indian English, for instance, should be classified as native speakers of English. Let me explain: If we accept his paradigm, English is no longer a language; it is a family of languages. So, there can be no such thing as a native speaker of English. There can only be native speakers of only this or that variety of English. From among them one might feel justified in referring to the original variety of English as native English. And by stressing on the historical relation a little bit one might be able to extract a certain amount of normativity. It is no help at all to say that his definition of native speaker is adequate linguistically although it has nothing to say about acceptance, recognition and ownership. Granted that they are politico-economic questions but Singh tells us neither how to dissociate nor how to integrate them into a cohesive paradigm. On top of that, he permits himself to be scathing about those who are trying their best to study this complex phenomenon. My own view is that contexts such as these are proper objects of what I call regulative conceptual analysis and should be approached in a much broader framework than Singh would be sympathetic to. For what they are worth, I will briefly give my general views on this question at the end of this response. But to anticipate myself, what I would like to say with reference to this specific issue is that while Singh's criticism that the extension of the (already problematic) dichotomy of native/non-native speaker to the dichotomy of native/non-native varieties (of English etc) is unwarranted and illegitimate, is substantially correct, his criticism is not a helpful one. In the first instance, Singh's claim that the second dichotomy is an extension – illegitimate or otherwise – of the first dichotomy is itself open to question. The latter involves a relation between a language and a speaker whereas the former is concerned with the relation between two languages. How we define a native speaker or for that matter whether we recognize the concept of native speaker may have no bearing on the question whether it would be legitimate to speak of varieties of a particular language, and further whether it would be legitimate (and/or politically desirable) to designate the older, original version as the native one and the other versions as non-native varieties and impose certain regulative requirements of conformity on the latter. The entire cluster of arguments

that Singh brings to bear is concerned only with the last issue. The arguments doubtless have considerable theoretical force. However, the issue has a quantitative dimension which in practice complicates the matter considerably. Every variation does not merit the status of a 'variety'. A few idiosyncratic deviations from normal usage practised by a handful of people do not make for a variety. The deviations and the number of people should be substantial. But again, if the deviations are too many, it may no longer make sense to speak of a 'variety'. It would be a separate language.

While it would be naively positivistic to assert that the social sciences are stuck in some pre-paradigmatic stage, there is substance to the view that unlike the physical sciences, the human sciences are not practiced within the boundaries of a definite paradigm – if by paradigm we mean definite, stable axiomatic frameworks comprising broad metaphysical and epistemological assumptions which facilitate articulation of reality. If by paradigm we mean just a conceptual framework of some sort, we will still have to concede that the paradigms of the natural sciences are considerably different from those of the social sciences. The axiomatic structures that constitute the platform for the articulation of social reality exhibit neither the nomological rigidity nor the structural stability that characterizes the paradigms of the natural sciences. The reason for this is not, I think, methodological but ontological and consequently epistemological. It is to do with the fact that the social sciences chiefly deal with the rational, decisional aspects of human behaviour which cannot be adequately captured by rigid conceptual structures.

There would be many qualifications to be made to this general dichotomy drawn here with reference to the paradigms of the natural and the social sciences and one can certainly show several domains in the social sciences which successfully capture certain aspects of social reality through structural and even formalistic modes of investigation. But I am ignoring them here because I am concerned with only one consequence of that basic difference, and that consequence is not affected by any qualifications a more nuanced characterization of this dichotomy might require. My reference is to the fact that *regulative analysis* is a very necessary part of the practice of the social sciences whereas the natural sciences have no need of any such activity. Though this activity is a fairly common one in the social sciences and is engaged in all the time, for some reason it has not been an explicit object of attention and consequently has not received a definite name or description.

So, let me begin by briefly explaining what exactly I mean by regulative analysis.

Concepts and theories stand in what one might call an intersectional relation to each other. While theories represent the structure of relations between certain concepts, concepts themselves represent, in the final analysis, points of intersection of fundamental theories. Now, depending on how loose-textured the relation is between different theories concerning related phenomena (which is determined not by the putative rigour or otherwise of the methodologies in question but by the dynamics of the phenomena the theories seek to describe or explain), the concepts retain a certain flexibility, a certain fuzziness around the edges. In other words, the fact that the social sciences are essentially historical, the fact that social reality does not consist of changing configurations of unchanging elements, but is a dynamic flux of changing elements, from which the social scientist extracts an ontology that is stable only in a limited, provisional way, sufficient for the immediate purposes, results in the requirement that if the new theories are to capture the changed structures of social reality, these new theories must also undertake some amount of conceptual revision since the elements of social reality denoted by those concepts have themselves changed.

The main difference here between the physical and the social sciences is that the relation between concepts and basic theories in the physical sciences is quite rigid. You cannot retain a concept from an older paradigm and use it as a component in the theories that are part of the new paradigm. You cannot, for instance, properly articulate the quantum theory with Aristotelian or for that matter Newtonian concepts. We must remember that what is retained from the older paradigms in these sciences is only a scattered bunch of signifiers. The corresponding concepts in each succeeding paradigm are quite distinct and are deeply entrenched in the basic theories that constitute the paradigm. There is hardly anything in common between the atoms of Democritus, Dalton and Neils Bohr except the word. Therefore, in the course of what Kuhn has called ordinary science, a physical scientist need not worry about the structure of concepts. He can take their stability for granted. That in fact is what is meant by the stability of the paradigm. When the paradigm becomes inadequate, due to the discovery of new and anomalous facts etc., the scientist discards it – concepts and all – and opts for a new paradigm.

In the social sciences, on the other hand, the relation between basic theories and concepts is loose. This is not, as I emphasised above, due to

the laxity of the methodology of the social sciences but because of the historical as well as synchronically diverse character of the elements of social reality represented by the concepts. Therefore, there are neither spectacular, revolutionary replacements of paradigms, nor corresponding, radical replacement of theories; what we find are evolutionary adjustments and adaptations. Obviously, these adjustments are not ad hoc or arbitrary but follow a trajectory of what Karl Popper called versimilitude. However, the important point is that the direction of the changing theories is not a matter of increasingly efficient theories alone. It is not just that a more comprehensive theory with greater explanatory potential or greater hermeneutic scope replaces one with lesser potential and scope. The additional – significantly additional – fact is that the older theories have become inadequate because the social reality that they had sought to describe and interpret has itself changed significantly. Moreover, the change in the social reality is not morphologically similar to the change in physical reality, namely, changed configurations of essentially unchanged elements. In this case, the very elements that constitute the reality are subject to change. In such a situation, the social scientist has no option but to resort to conceptual reform. He cannot, unlike the natural scientist, overthrow the entire paradigm along with the older concepts and basic theories: there is no warrant for this from the ontological situation. Nor can he continue with the older concepts and try to forge new theories that could adequately capture the changed reality. He must retain the concepts and make structural adjustments to them and then weave new and adequate theories. I am not suggesting that the social scientist is constantly doing this. But I think that often – more often than is realised or acknowledged – he engages in this activity. In fact I think that a great deal of theoretical activity in the social sciences consists of this effort to achieve a conceptual equilibrium of the paradigm. Most of the times it is not conspicuous since it is a gentle manoeuvre, involving small nudges and pushes not at all resembling the upheavals that characterise the revolutions in the physical sciences. However, sometimes, largely due to the cultural and political significance of the concepts or the direction in which they are altered, this activity becomes an object of – occasionally heated – attention. The concept of the native speaker I think is one such concept. Concepts such as secularism, fundamentalism, terrorism, democracy, etc represent other, in some ways more vividly illustrative, instances of this phenomenon.

It is this activity of conceptual revision that I call regulative analysis, in contrast to the scrupulously detached, descriptive analysis that is favoured

by the philosophers of the analytic tradition. (While the latter kind of analysis has its undeniable value, I think the analytic philsophers' inability to realise the importance of regulative analysis for the practice of the social sciences has contributed to the sense of sterility that has come to surround routine activity in that tradition.).

Now, I do not think there is any definite pattern, any definite procedure for regulative analysis understood in this sense. There would be, naturally, time-tested guidelines that can be usefully followed, and similarly there would be an inventory of some of the characteristic ways in which this kind of analysis can go awry. Such guidelines would be specific to each discipline and they must emerge from a self-conscious, meta-level reflection on the concepts in question from within the discipline. But one can delineate the broad orientation that any regulative analysis would exhibit.

A regulative analysis will obviously consist of two components: an analysis of the structure of the concept and an examination of the possible directions in which the concept can be moved. This is not to imply that there will be two distinct activities sequentially performed, first a descriptive analysis and then a prescriptive exploration, since the whole point of regulative analysis is to avoid that sort of a descriptive-prescriptive dichotomy. However, of necessity, a regulative analysis must take an examination of the mutual relations of the constituent elements of the concept under consideration as a point of departure.

As for the direction of the possible shift of the concept, the social scientist has two options available to him (I am talking mainly about those cases where the region or aspect of social reality under investigation has undergone a change): in terms of sense and in terms of reference. The first option is to consider the sense of the concept, its structure, non-negotiable and suggest that hence forth, we better use that concept to denote some other phenomenon, or some other combination of phenomena. This can be very roughly described as a rationalist option. The other – broadly empiricist – option is to stay focused on the phenomenon to which the concept has been referring to and see how one can make internal adjustments in it and bring it closer to the current form of that phenomenon.

However, what makes both these options problematic is the issue of evaluating the change. Should one take some nodal points within the structure of the phenomenon as its essential features and treat the phenomenon as persisting with some incidental changes? But what is one to do if the changes in the phenomenon in question include changes in those nodal

points, while the phenomenon in some ways still retains a certain stability with reference to other related phenomena? This would present the dilemma whether to try to use the old concepts in new ways to describe the phenomenon or discard the old concepts and coin new concepts. This may amount, in some cases, to a dilemma whether to continue with a revised version of the old theory or look for a new theory. However, in practice, the matter turns out to be more complicated than this since, given the constructed character of a great deal of social reality, the heuristic luxury of looking at the phenomenon on the one side and the conceptual structure that purports to explain it on the other may just not be available. In other words, if there is sufficient space between the descriptive concepts and the explanatory categories, and provided that the former are relatively theory-free, it may be easier to take a position as to whether the phenomenon in question is the 'same'. But such a situation is much less frequent than imagined in the social sciences. The descriptive concepts through which one tries get a purchase on the phenomena in order to explain them are themselves frequently subject of theoretical conflict. As a result, those according to whom a phenomenon has not changed, those for whom the phenomenon has changed but only to a certain extent and those who believe that the old phenomenon has been replaced by an altogether new phenomenon will be found, while not exactly inhabiting the notorious 'different worlds' of Thomas Kuhn, to a considerable extent talking past each other.

Now, I would like to make a few observations in the form of questions with reference to the concept of the native speaker against the backdrop of what I have said above.

1. Given that native speaker means 'native speaker of a language', an examination of the concept of native speaker must obviously involve an analysis of the notion of nativity (in the context of language), the notion of language and the kind of relation that binds the two terms. The question is: can we define these constituent notions independent of each other? Or does the assumption that the concept of native speaker is shorthand for native speaker of language, imply that there is an analytic relation between the two terms?

2. If the answer to the second question is in the negative, how exactly would we characterise the relation between native speaker and language that can be indicated by the assertion that it is not possible that a normal human (understood as an individual with normally functioning faculties) is

not a native speaker of any language – except as an anomaly? What follows form the apparent fact that native speaker implies a language while the converse is not obtained (as evidenced by the existence of dead or artificial languages)? Given the fact that what we are discussing is conceptual dependence, does the fact that those are special cases make a difference? Is this the point to be gleaned from the above question: that the notion of native speaker is tied to the notion of a language *as a unit*, whether that unit is, strictly speaking, a language, a dialect or something else being a different matter?

3. Does this fact mean that the notion of a native speaker is tied to the notion of a linguistic community – at least in the narrower sense of the term?

4. In this context, is there reason to fear that the relation between native speaker and linguistic community, might bring the notion of native speaker close to that of 'mother tongue' to the point of their collapse into one another? Or can we maintain that such a fear is unfounded since, although both are socio-linguistic terms, the notion of native language is primarily a linguistic concept while the notion of mother tongue is primarily a sociological concept?

5. If the points mentioned till now are conceded as relevant, can we attach significance to the fact that the ongoing debate regarding the native speaker seems to centre largely on the English language and its variants? The point that the question of native speaker has great political significance in the context of the English language is well taken. But what about the fact that centring the debate on the English language has tended to warp the framework of the discussion to a considerable extent?

6. It cannot be denied that the notion of native speaker has tended to be seen a little too much in terms of power and cultural hegemony on the one hand and linguistic competence in the sense of achievement on the other. This being so, is there not the danger that the former move might tend to carry with it, subconsciously so to speak, a desire to deflate the notion, to dilute it and rob it of its hegemonic potential? While the democratic motives of such a move are admirable, can it escape the blame of not doing justice to the essence of the notion? To raise this is not to imply that there is an ahistorical essence of concepts – particularly our social concepts. But is not relevant to point out that it would be wrong to characterise a concept altogether in terms of its features as obtained in one particular cultural/his-

torical setting? To put the question in rather simple terms, is not the analysis that is performed in the context of the English language likely to yield results and draw us towards conclusions that may look far from balanced in the context of the language of some vanishing tribe?

7. On the other hand, there may be this fact to be considered: the move of looking at the notion of native speaker predominantly in terms of competence must be conceded to be a plausible, natural move since such an idea was integral to the initial positing of this concept as a useful – indeed necessary – construct as a terminating point of appeal to settle questions of linguistic correctness. However, is there not need to distinguish the notion of competence from that of a more spontaneous, intuitive, a more poetic, a more erotic relation to language? A relation, no matter how difficult to define, which captures the capacity to take delight in the play a language, the capacity to 'appreciate' its beauty? Is there not a need to look closely at the proposition 'a native speaker is necessarily a competent speaker'? Would it be correct to say that the model of man-language relation that employs the notion of competence has already incorporated a notion of a standard language with objectified, ossified rules with corresponding ability to correctly apply those rules with different degrees of efficiency? That this model tends to first posit a language and its rules and then look at the native speaker in terms of his competence in that language? How do we deal with the objection that this model makes sense only up to a point and doesn't take sufficiently into account the ageny-character of the native speaker, his initiative in the diachronic direction of his language?

8. To raise this question need not necessarily imply a rejection of the notion of a standard language. It may be conceded that obviously, there is no way that notion can be dispensed with. As soon as you posit some definite language, you are required to treat it in terms of standard usage etc. One does in pedagogic contexts say about a particular expression, 'the native speaker won't say that'. This can't be dispensed with "The choice to delimit the domain of a language is after all always open". One might say for instance that 'a native US, or Australian (or whatever variant) English speaker won't say that'. But the question might still remain as to how far can you go in the direction of delimiting the domain of a language.

9. Let us expand the above point in this way: having settled the question as to what 'native speaker' means, we would need to look at what the

boundaries of the linguistic units would be in relation to a native speaker. In this context there would be questions as regards the size as well as the heterogeneity of the linguistic unit. Specifically, how much variation can one language *as a unit* tolerate? At what point can one say that it ceases to be that language and begins to be another language? Similarly, in the opposite direction, how small can a linguistic unit be? Supposing I insist that my own rather peculiar variant of the English language must be treated as a unit and I am the native speaker of that unit whose linguistic community consists of only one member? It may be preposterous but does it raise problems at the theoretical level? If it does, what light do those problems shed on the question of the notion of the native speaker?

10. Can one say that the changed situation of certain languages, for instance the proliferation of their variants, has only exposed the inherent contradictions of the concept of native speaker? Or would one prefer to say that the concept was once useful and appropriate to a past situation, but the situation having changed, it has outlived its usefulness and should be erased from the discursive lexicon? Or would one take a somewhat more radical position and say that it is not a question of one particular concept but a matter of restructuring the entire discourse in the light of our awareness of the implications – linguistic as well as non-linguistic – of certain theoretical assumptions, and that the problem with a particular concept, in this case the concept of the native speaker, represents only an indication of, and a point of entry into, what is problematic with the entire discourse of language as a cultural phenomenon? To put the matter in deliberately simplistic terms: what precisely is at stake? – A term (with its troubling connotations)? A concept? A theory? An entire discourse? And what is the perceived objective of the exercise? To alter the meaning of a word? To change its referent? At another level – To achieve a greater degree of verisimilitude in description? To bring cohesiveness to or generate greater potential in explanatory categories? At still another level – To retrieve some lost cultural ground? To render some sort of political justice?

Needless to say it is not for philosophers to advise how the social scientist should execute that difficult manoeuvres involved in answering these questions in such a way as to keep the concept useful as a theoretical category while not doing violence to its denotative relation to a changing reality, while being aware of his obligation as a critical social scientist not to be-

come unmindful of the normative dimension of his practice. The philosopher can only stand in the wings and occasionally prompt him. How helpful or superfluous that assistance is, it is for the social scientist to decide.

Acknowledgements

I am grateful to the editor of this volume for his invitation to respond to his intervention.

Regional Reports

Europe

John Peterson

1. Introduction

For over two centuries now Europe has been a major center for the study of South Asian languages. Although the Indo-Aryan languages were undoubtedly the first to receive serious attention by scholars in Europe, resulting in the birth of Indo-European linguistics, and ultimately of modern linguistics itself, attention soon also turned to the Dravidian, Austro-Asiatic and Tibeto-Burman languages of the region, culminating in the last century in Grierson's landmark survey of the languages of South Asia. This tradition continues to the present day, with a large number of descriptive and historical studies devoted to these languages, as well as an increasing number of theoretical works based on data from them. Although the emphasis may have shifted somewhat, with the often still highly under-documented Tibeto-Burman languages of the northern subcontinent now probably receiving the most attention in Europe, work on historical and modern Indo-Aryan languages is still flourishing, while work on the Dravidian and Austro-Asiatic languages would seem to be experiencing somewhat of a comeback.

In fact, there is currently so much work now being conducted on South Asian languages in Europe that, when I first sent out calls to colleagues asking for information on their works on these languages, I was literally inundated with data and it soon became clear that my original plan for a nearly exhaustive bibliography with a discussion of the topics covered in all of the works had to be abandoned simply for reasons of space in favor of the present general discussion, with only a few words on each title, and all claims of exhaustiveness have been abandoned. Nevertheless, the present report should provide the reader with a good overview of the breadth of current work on South Asian languages in Europe, as it discusses work by scholars from ten different European countries.

One final note: Although the present report is dedicted to works on South Asian languages by European scholars and despite considerations of space,

in the case of collections of studies such as Saxena (2004), Saxena & Borin (2006), Tikkanen & Hettrich (2006), and Chevillard & Wilden (2004) which have been (co-)edited by scholars residing in Europe, I have chosen not to "sort out" those works written by scholars who are not resident in Europe but rather to include all works contained therein, keeping my comments here to a bare minimum for reasons of space. Also for reasons of space, these articles will not be discussed again in the relevant sections following the general section (§ 2).

2. General studies / more than one language family / English in South Asia

We begin our discussion here with two recent volumes, one edited by Anju Saxena and the other co-edited by her and Lars Borin, which contain a number of studies on various individual languages and other topics.

Saxena (ed.) (2004) contains a selection of the papers on the Indo-Aryan and Tibeto-Burman languages of the Himalayas which were presented at the Himalayan Languages Symposium at Uppsala University, Sweden, in August, 2001. The volume begins with a contribution by the editor herself (Saxena, 2004a) presenting an overview of the volume and of the linguistic situation in the Himalayas. This is followed by 15 individual contributions: Schmidt (2004b), a grammatical comparison of tenses and noun inflections in various Shina (Dardic) dialects; Heegård & Mørch (2004), a discussion of the phonology of Kalasha, also Dardic; Grunow-Hårsta (2004), a discussion of direction and differential dative case marking in Magar, a West Central Himalayish language of the Bodic subgroup, spoken in central and western Nepal; Turin (2004a), a description of the largely undocumented kinship terminology of the Tibeto-Burman language Thangmi of central eastern Nepal; Bickel (2004a), dealing with the notion of subject in the Kiranti language Belhare, with special emphasis on pivots; Oetke (2004), a review of the notion of "sentence" in Classical Tibetan; Saxena (2004b), which examines the structure and functions of finite verbs in narratives in Kinnauri (West Himalayish, Himachal Pradesh, India); Winter (2004), a discussion of the complex system of preverbal modifiers in the Kiranti language Sunwar, spoken in Nepal; Hargreaves (2004), which is primarily devoted to the system of directionals in Kathmandu Newar; Honda (2004), which takes a look at the grammaticalization of deictic motion verbs in Seke, a Tamangic lan-

guage spoken in the district of Mustang, Nepal; Andvik (2004), a study of the use of the morpheme 'do' in Tshangla (Bodic, eastern Bhutan, Arunachal Pradesh and the Tibetan Autonomous Region) in subordinating constructions; Gvozdanović (2004), which focuses on the "agentive" suffixes *-a* and *-u* in the Kiranti language Bantawa, spoken in eastern Nepal; Matisoff (2004), dealing with the question of areal semantics, not only in the Himalayas but including data from a number of languages from around the globe; Bielmeier (2004a), which takes a closer look at Shafer's proto-West Bodish hypothesis and the formation of verb paradigms in Tibetan; and finally, van Driem (2004a), a brief discussion in which the author critically reviews his own earlier Mahakiranti hypothesis.

The second volume referred to above, Saxena & Borin (eds.) (2006), aims to "discuss the status of the lesser-known languages in South Asia and to disucss how modern technology can be a tool in documenting these languages and in spreading awareness about them." (Saxena 2006: 5). It begins with a general introduction to the topic and to the individual contributions by Anju Saxena (Saxena 2006), followed by three sections containing individual contributions.

The first section, entitled "Language situation and language policies in South Asia", contains three studies dealing primarily with the status of lesser-known/minority languages in India (Singh 2006), Nepal (Turin 2006a) and Pakistan (Rahman 2006) where, among others, the legal status of the many languages of the region is discussed, as well as problems in implementing strategies to promote these languages.

The second section, "Lesser-known language communities of South Asia: Linguistic and sociolinguistic case studies", contains the following five studies: Abbi (2006), which presents preliminary data on Great Andamanese, based on a pilot study; Bradley (2006), dealing with the various orthographical systems of Lisu, spoken in China, Burma, Thailand and India, especially with regards to email; Kohistani & Schmidt (2006), a discussion of the status of the Shina language in contemporary Pakistan; Noonan (2006), dealing with "the state of play between the rise of ethnic consciousness, attitudes toward language, and the state of language endangerment of some Tibeto-Burman speaking peoples of west-central Nepal." (Noonan 2006: 162); and finally Zeisler (2006), which describes the strong opposition in Ladakh to the development of literacy and literature in Ladakhi (Tibeto-Burman), due primarily to the influence of the conservative Buddhist elites who favor the teaching of Classical Tibetan only.

The third and last section, "Information and communication technologies and languages of South Asia" contains nine studies: Annamalai (2006), dealing with the effect of technology on indigenous languages through disruption of the environment; Renganathan & Schiffman (2006), dealing with the impact technological advances are having on Tamil language use, attitudes and planning; Hardie et al. (2006), which discusses some of the issues related to the construction of corpora of South Asian languages in the EMILLE (Enabling Minority Language Engineering) Project at the Universities of Lancaster and Sheffield (UK); Michailovsky (2006b), which describes a number of current applications of information technology for the languages of Nepal, concentrating on linguistic documentation; Nathan & Csató (2006), a disscussion of how information and communication techology can be used in communities where lesser-known languages are spoken as part of what they refer to as the "deliver *to*" framework of language documentation; Allwood (2006), which discusses how language technology can be used in "language survival kits"; Trosterud (2006), which looks at ways of building language technology for minority languages; Borin (2006), a brief introduction to language technology, focusing on issues of corpus collection and the creation of tools for the annotation of corpora; and Grinevald (2006), which deals with a number of ethical issues relating to fieldwork on endangered languages, especially with respect to the issue of "informed consent".

In addition to these two volumes, a number of other, article-length studies dealing with more than one language have appeared in the past three years. For example, Bickel (2004b) deals with the syntax of experiencers in Himalayan languages, both Indo-Aryan and Tibeto-Burman, while Ebert (2005) is a brief, encyclopedic overview of the data and literature on the topic of South Asia as a linguistic area, including a discussion of possible sub-areas on the subcontinent. Kulikov & Manevskaia (2004) then deal with systematic correspondences in the translation of Sanskrit Buddhist texts into Tibetan, above all with respect to the various syntactic strategies employed when translating active and passive constructions from Sanskrit into Tibetan. Montaut (2004a, in French) presents a discussion of mirativity and other "evidential" categories in Hindi and Nepali, in addition to data from the Dravidian languages Telugu and Tamil. In contrast, Montaut (2005, in English) is more sociolinguistic in nature, dealing with the often contentious issue of language movements and multilingualism in India.

Further studies under this heading include Sheiderman & Turin (2006), who deal with issues relating to the “run” for tribalness in Darjeeling and Sikkim, especially with respect to the ethnic groups whose ancestors migrated to there from Nepal. Turin (2005) is a brief discussion of language endangerment and linguistic rights in the Himalayas, focusing on Nepal and covering a number of topics. Verma (2005) then discusses the status of English in India, focusing on the economic advantages of English’s increasing use in education, commerce and information technology.

Finally, I would like to call the reader’s attention to my own online “Bibliography of Endangered and Seldom Studied South Asian languages” (Peterson, no date), which has been online since January, 2001 and which is updated at regular intervals. The site documents literature on all seldom-studied languages of the subconitinent in order to faciliatate their study, and contributions from other researchers are always welcome:

http://www.SouthAsiaBibliography.de/

3. Old and Middle Indo-Aryan / Historical Indo-Aryan linguistics

The past few years have seen quite a large number of studies appear on historicial aspects of Indo-Aryan, of which the following have come to my attention. We begin here with the works contained in the 154-page Russian-language study edited by Elizarenkova et al. (2004) from the well-known “Languages of the World” (*Языки Мира*) series. This study contains an introduction, followed by eight overviews (Elizarenkova 2004a–h) of the following languages / language periods: Old Indo-Aryan, Vedic, Sanskrit, Middle Indo-Aryan, Pāli, Prākrit, ApabhraÞÐa and Buddhist Hybrid Sanskrit, and finally an overview of Indian scripts of the Old and Middle periods (Koryakov, 2004).

Another volume which should be mentioned here is Tikkanen & Hettrich (eds.) (2006), which contains studies based on 12 of the papers presented at the 12th World Sanskrit Conference held in Helsinki, Finland, in July, 2003: Kobayashi (2006), which takes a fresh look at the development of Proto-Indo-Iranian **sć* into Sanskrit /(c)c^h/; Hock (2006), dealing with reflexivization in the Rig-Veda (and beyond); Kulikov (2006a), a discussion of a group of athematic middle participles marked by the suffix *-āna-* in early Vedic; Bubenik (2006), which discusses changes in the Old and Middle Indo-Aryan case and adpositional systems; Oguibénine (2006), dealing

with the case marking of the subject/agent in different syntactic constructions in Buddhist Sanskrit; Seldeslachts (2006), which examines Prakrit-like phonological developments in Sanskrit, applying the "Kölver-principle" to a wide range of Old Indo-Aryan forms; Baghbidi (2006), a discussion of Iranian elements in Sanskrit; Pinault (2006), dealing with links between the Indo-Iranian substratum and the "Bactria-Margiana Archaeological Complex" (BMAC) or "Oxus culture" of the Bronze Age; Scharfe (2006), which traces two developments showing Dravidian influence on Old Indo-Aryan; Sheldon (2006), a discussion of the Sanskrit translation of the Avestan Haoma liturgy in the light of recent research; Glass (2006), which addresses various issues involved in creating a dictionary for Gāndhārī; and finally, Huet (2006), a discussion of the lexicon-directed methodolgy for computer processing of Sanskrit.

There are also two collections of studies by Italian scholars which have been brought to my attention. The first such work, Ronzitti & Borghi (eds.) (2004), is a selection of five studies from two meetings of the *Associazione Genovese di Studî Vedici e Pāṇiniani* held in July and October, 2003. The volume is divided into two sections: The first section is dedicated to Vedic and Pāṇinian linguistics and contains two studies: Candotti & Pontillo (2004), an English-language study dealing with substitution in Pāṇini's grammar, especially with rule A 1 1 56. This is followed by Ronzitti (2004), an Italian-language study which deals with a number of derivations in **-mo-* in Old Indo-Aryan, a topic on which this author has also published a number of more recent studies as well (see below). The second section is devoted to Indo-European and Indo-Aryan and contains the following three studies: Busetto (2004), a discussion of phonations and articulations in Old Indo-Aryan, written in Italian; Milizia (2004), also in Italian, dealing with the Indo-Aryan "conspiracy" against the Indo-European voiced fricatives; and finally, Scala (2004), a brief discussion in Italian of the phoneme /f/ in Romanī, especially with regard to its alternation with /k^h-/ in initial position.

Another collection of studies in this category by Italian scholars – all of which are written in Italian – is Ronzitti et al. (eds.) (2006), the result of three meetings of the *Associazione Genovese di Studî Vedici e Pāṇiniani* held in Genoa between 2004 and 2006. This collection contains eight studies altogether, of which the following five deal specifically with Indo-Aryan. The first study is Ronzitti (2006a), which deals with simplex primary derivatives ending in *-ma- / -mā-* (< **-mo- / *-meh*$_{2/4}$) in post-Vedic Indo-Aryan. This topic is also dealt with by this same author in greater detail in her

book-length study (119 pp.), also written in Italian (Ronzitti 2006b), which consists of two chapters, "etymological analysis" and "functional analysis". The remaining studies contained in Ronzitti et al. (eds.) (2006) are, in order of appearance: Viti (2006), dealing with subordination in Vedic; Borghi (2006), a critical discussion of the comparison of Old Indo-Aryan *bʱūkă-* and Italian *buco*, both with the meaning 'hole', a topic which he has also discussed in a somewhat shorter study together with Rosa Ronzitti (Borghi & Ronzitti 2005); Busetto (2006), which discusses issues relating to Devanagiri, particularly with respect to Unicode; and finally, Fortuna (2006), dealing with words denoting "doctor", "medical science" and "medicine" in a number of South Asian languages, but also in many other languages, such as Romani, Vietnamese, Mon, Thai, Khmer, Indonesian, Chinese and Japanese.

We now turn to a number of studies on historical aspects of Indo-Aryan by individual authors, presented here in alphabetical order. François Heenen has published a book-length (267 pp.) study of the desiderative in Vedic (Heenen 2006), an enlarged and modified version of his doctoral thesis in Vienna. The study consists of four principal parts: a general overview of the respective category, a chapter dealing with its formation, another dedicated to its semantic and pragmatic functions, as well as a complete inventory of the attested forms for each root.

Korn (2005) is a ten-page study of the Iranian language Balochi, spoken in Pakistan, dealing with earlier contacts between Balochi and other Iranian languages. Kulikov (2005a), in the same volume, then takes a closer look at the length vacillation *-īy-//-iy-* and related phenomena in Vedic. Other works by this same author include Kulikov (2005b), which discusses reduplication in the Vedic verb, Kulikov (2006b), dealing with diathesis in Vedic, primarily with various functions of the passive and middle in this language, and finally, Kulikov (2006c), a typological study of case systems from a diachronic perspective, containing data from many language families, including a discussion of new case markers from postpositional phrases and nominal compounds in New Indo-Aryan languages.

Two works by Thomas Oberlies have also come to my attention, both of which deal with historical aspects of Hindi: Oberlies (2005a), a brief, two--Aryan is the result of haplological shortening, providing examples s page article in which the author shows that "shortened" case marking in Indo howing that this phenomenon is found in all stages of the Indo-Aryan languages, beginning with the Rigveda. This same topic is covered – along

with a great many others – in Oberlies (2005b), a discussion in book format (vii, 70 pp.) of Hindi morphology from a historical perspective, covering virtually all of the morphology of modern Hindi.

There are also a number of works on Old Indo-Aryan by Carlotta Viti which have appeared since 2004. Viti (2004a), written in Italian, deals with the relative clause in Vedic; Viti (2004b), also in Italian, is a study of possession in Old Indo-Aryan; Viti (2005), again in Italian, then deals with the suffix *-vat* in Vedic. Finally Viti (2007), a revision of the author's doctoral dissertation, is a book-length (302 pp.) study in English on subordination in Vedic from a functional-typological perspective. It contains a general introduction, a chapter presenting the heterogeneous structures of clause linkage in Vedic, a chapter on relative clauses, five chapters devoted to various adverbial constructions (temporals, conditionals, causals, purposives, and concessives), a lengthier chapter devoted to completive relations, and finally the author's conclusions.

Finally, three Russian-language works by Boris Zakharyin deserve mention here: Zakharyin (2004), in which it is shown that as early as the MahābhāÒya commentary of Patañjali there already existed a subcategorization of Pāṇini's *karmaṇ* into *nirvartana* 'production' and *vikāra* 'modification', and that further subcategorizations of this concept were continued by Katyāna, Bhartªhari and KaiyāÔa; Zakharyin (2007a), which deals with verbs of *movement in* and the *movement of* liquids in the history of Indo-Aryan, similar to the studies by Kolotova (2007) and Khukhlova & Singh (2007a), discussed below for modern Indo-Aryan languages; and finally, Zakharyin (2007b), an 800-page translation of Varadarāja's *Laghu-siddhānta-kaumudī* into Russian, supplemented by an introductory article, commentaries, and notes by the translator.

4. Modern Indo-Aryan

We now turn to studies on the modern Indo-Aryan languages which have appeared within the past two years. These include descriptive studies – both book-length as well as shorter studies – covering issues such as phonology, morphology and syntax, but also works of a more theoretical nature.

It is in this last category that the following study belongs: In their 25-page study, de Hoop & Narasimhan (2005) propose a model of case-marking the takes not only functions of case such as identification and dis-

tinguishability into account but also the notion of argument strength, which can vary from one language to another, to account for the data in a language with differential case-marking such as Hindi.

Singh & Khokhlova (2004) and Khokhlova & Singh (2007b) then deal with resultative constructions in the western New Indo-Aryan languages Hindi-Urdu, Punjabi, Gujarati and Marwari, pointing out similarities to constructions found in Slavic languages such as Russian. Also, similar to Zakharyin (2007a), discussed above, these same authors then deal with verbs of motion in liquid, such as Hindi *tairnā/tarnā, bahnā, andūbnā* and their cognates in Punjabi, Marwari and Gujarati, in a general typological framework (Khokhlova & Singh, 2007a, in Russian). Kolotova (2007, in Russian) then deals with this class of verbs in Bengali.

Annie Montaut has also published a number of studies devoted to Hindi/Urdu in the past few years. For example, Montaut (2004b) is a detailed grammar of modern Hindi (xii, 319, v pages), covering virtually all aspects of this language. Two further studies by this same author which deal with various topics in Hindi/Urdu are Montaut (2006a, in French; 2006b, in English), both of which deal primarily with the "mirative" uses of the aorist in this language, including surprise, argumentative and polemic use, and saliency.

Finally, we mention here works by three authors on Indo-Aryan languages of the western Himalayas: Heegård Petersen (2006) is an English-language doctoral dissertation which deals with local case-marking in the Dardic language Kalasha. This study consists of two sections: The first section (xiv, 295 pp.) is an extensive discussion of the semantic and morphosyntactic aspects of local case-marking in this language and also contains much information on the Kalasha people and their language, previous linguistic research on Kalasha, a sketch of Kalasha grammar, as well as a discussion of the term "Dardic" itself. The second section (135 pp.) contains four maps and 36 appendices for reference. Schmidt (2004) then deals with compound verbs in Shina, i.e., complex predicates formed by a lexical verb in its "conjunctive participle" or converbal form plus one of a number of "vectors", in the case of Shina forms deriving from verbs meaning *go, sit, fall, give, leave* and *release*, which have lost most or all of their original semantics. Finally, the dictionary by Claus-Peter Zoller (Zoller 2005) of the Indus Kohistani language, referred to in my last report, has since appeared. It contains around 8,000 entries along with etymological information, as well as English – Indus Kohistani and Old Indo-Aryan – Indus Kohistani indices.

5. Dravidian

In my last report (Peterson 2004: 131), I regretted to note that I had not come across even one single study devoted at least primarily to Dravidian languages published by a scholar residing in Europe. Fortunately, things are different this time around and I am happy to be able to include in this issue a number of such works.

The large majority of these studies – all in English – stem from the section entitled “Studies in Language and History of Language Description” in the volume edited by Jean-Luc Chevillard and Eva Wilden (Chevillard & Wilden 2004). These include the following studies, given here in their order of appearance: Zvelebil (2004), a prolegomena to the author’s proposed etymological dictionary of the Ir̠ula language of the Nilgiri Mountains in southern India; Subramonian (2004), a brief study dealing with the suffix *-āre* in finite verbs in Hermann Gunder’s *Kēraºa pašama* with respect to person and number marking; Schiffman (2004), which deals with the case system of Tamil; Lehmann (2004), dealing with pronoun incorporation in Old Tamil; Agesthialingom (2004), dealing with the generation of numerals in Tamil; Pollock (2004), which reflects on a number of principal texts and persons of early Kannada philology, a topic which, as the author notes, “is virtually a blank slate for western readers” (Pollock 2004: 389); Vacek (2004), which provides correspondences for words denoting *sheep*, *deer* and *cattle* in Dravidian and Altaic languages; Schalk (2004), which takes a critical look at Robert Caldwell’s highly influential derivation of the Tamil toponym *īšam* for the island of Sri Lanka from the Sanskrit / Pāli forms *siÞhala / sīhaºa*; Tieken (2004), dealing with the question of the whether various characteristics of the language of CaÉkam poetry which have often been assumed to be especially archaic features are in fact proof of the antiquity of these poems; and finally Chevillard (2004), which deals with ideophones in various periods in Tamil, in what he refers to as the *X-en̠al* template, i.e. some element “X” followed by a form of the quotative verb *en̠al*.

In addition to the studies in the volume just discussed, I have come across two further studies – both in French – which should be mentioned here: Murugaiyan & Pilot-Raichoor (2004), who deal with “non-differentiated” predications in Dravidian, above all in Tamil, and also argue on the basis of their data for a re-evaluation of many assumptions of historical Dravidian linguistics; and finally, Pilot-Raichoor (2006), a discussion of

time and space in an oral tradition of the Badaga of the Nilgiri, with special reference to the use of grammatical resources of the language to this end.

6. Tibeto-Burman

As I have noted in previous reports, Europe has established itself in the past few decades as a major center for the study of the many Tibeto-Burman languages of Tibet and the Himalayas, and studies on both diachronic and synchronic aspects of these languages now far outnumber those on Austro-Asiatic or Dravidian languages and even rival in terms of number those on Indo-Aryan languages, for which Europe has traditionally been known.

In the following, these will be divided into several sub-sections for convenience, beginning with general studies and then dealing separately with studies on the Kiranti languages of Nepal, Tibetan and finally "Other", i.e., Tibeto-Burman languages of the region whose classification is either uncertain or which lie outside of these two other groups, such as Lepcha, Newar and Thangmi.

Tibeto-Burman, general

We begin our discussion of general works on Tibeto-Burman with a study by Roland Bielmeier (Bielmeier 2005) which provides a general overview of the Sino-Tibetan languages. This is followed by a number of studies by George van Driem, dealing with more general topics with respect to Tibeto-Burman. For example, van Driem (2004b) presents an overview of three highly endangered languages of Bhutan which are unique to this country, namely Lhokpu, Black Mountain and Gongduk, including much information on the geography and history of the region as well. Van Driem (2004c) then provides a critical discussion of three competing theories about the genetic relationships of the Tibeto-Burman languages during Brian Houghton Hodgson's lifetime, namely Tibeto-Burman, especially as Hodgson understood and used this term, Turanian and Indo-Chinese. The study concludes by highlighting the importance of Hodgson's contribution to Tibeto-Burman linguistics, noting that there is undoubtedly much still to be learned from a careful study of Hodgson's notes from over a century ago. In a similar vein, van Driem (2005) provides a discussion of Tibeto-Burman and Indo-Chinese

but also includes a discussion of the implications of linguistic classifications for interpreting the prehistory of a large portion of Asia.

Another article under this heading which has come to my attention is Michailovsky (2006a), a brief encyclopedic article presenting an overview of the langauges of Bhutan. Finally, Turin (2006b) deals with the topic of "linguistic identities" in the Himalayan periphery, arguing against the uncritical extension of linguistic classifications in categorizing ethnic communities, focusing on the Thakali and Thangmi communities of Nepal.

Kiranti

We begin our discussion of studies of the Kiranti languages of Nepal with a number of works which have been published by members of the Chintang and Puma Documentation Project, a DOBES-Program project by the University of Leipzig in conjunction with the Department of Linguistics at Tribhuvan University in Nepal, sponsored by the VW Foundation.[1]

Two of these studies deal specifically with the language Puma: Sharma et al. (2005), a discussion of the personal and possessive pronouns in this language, and Bickel et al. (2007a), which deals with two types of "object suspension" or detransitivizing strategies in this same language, including a typological comparison of the Puma data with those of other languages.

Three other works by members of this research team which should be mentioned here all deal with Chintang: Rai et al. (2005) is a discussion of triplication and ideophones in this language in which the authors show, among other things, that triplication in Chintang cannot be considered a case of recursive reduplication. Bickel et al. (2007b) then takes a look at the extremely rare and typologically highly interesting issue of free prefix ordering in Chintang. Among other things, the authors argue here that Chintang also possesses endoclitics, an equally rare phenomenon in the world's languages. The last article to be discussed here, Gaenszle et al. (2005), begins with a discussion of the Chintang deities Rajdeu and Budhahang. The authors then procede to discuss a number of interesting characteristics of Chintag ritual language, characterized among other things by parallelism in nominals (bi-, tri- and even multinominals) as well as a use of the subjunctive which differs considerably from that of the spoken language.

Finally, a further study by one of the authors involved in this project but on a different Kiranti language should be mentioned, Bickel (2006), a typo-

logical study dealing primarily with data from Belhare (and German) in which the author suggests a typological variable, the "PSA-LEVEL VARIABLE", according to which constructions differ as to whether their "PSA" or "privileged syntactic argument" is selected at the predicate or at the clause level.

Moving along to works on Kiranti languages by researchers associated with the Himalayan Languages Project in Leiden,[2] we begin with three works by Jean Robert Opgenort, starting with two reference grammars on the closely related languages Wambule and Jero. Opgenort (2004a) is an extensive (xxix, 900pp., 29 plates) overview of Wambule, also known under the names "Chouras'ya", "Chouraśya", "Chourase" and "Umbule", and is a revised and enlarged version of the author's dissertation. It contains an introductory chapter on the Wambule people and their culture and nine further chapters covering virtually all aspects of Wambule grammar, such as nominals, verbal simplicia, finite verbs, non-finite deverbatives, gerunds and complex sentences, and verbal constructions and complex verbs. This is then followed by five appendices containing 16 segmented, glossed and translated texts, Wambule-English and English-Wambule lexica, affirmative and imperative paradigms and a final appendix entitled "The 'Chouras'ya' Materials", with a comparison of earlier work on Wambule vocabulary conducted by Hodgson & Konow with the forms as analyzed by Opgenort, as well as a copy of the grammatical note on "Chouraśya" from Grierson (1909: 369f.).

Opgenort (2005) is a shorter (xxv, 404pp.) reference grammar of Jero and is similarly structured: It contains an introductory chapter to the Jero language and its relatives, with much historical discussion, followed by chapters on phonologoy and morphophonology, nominals and adverbials, finite verb forms, non-finite deverbatives, gerunds and verbal constructions, and complex verbs. Finally, there are four appendices: Jero-English and English-Jero lexica, affirmative and imperative paradigms, and a comparative Kiranti word list. Finally, Opgenort (2004b) is a 27-page study of implosives and preglottalized stops found in the Western Kiranti languages Bahing, Sunwar and Wambule from a historical perspective.

Tolsma (2006), a revised version of the author's dissertation, presents an overview (xiv, 286pp.) of the Kiranti language Kulung, or *kuluri3*, spoken by an estimated 15,000 people in eastern Nepal. It begins with a general overview of the language, the people and their culture, followed by a further 11 chapters covering virtually all aspects of the language: phonology, nomi-

nal categories, verb stem morphology, verbal affixes, compound verbs, subordination, verbal derivation, non-finite verbal constructions, subordinators, mood, and particles. Finally, there are four appendices, containing 13 segmented, glossed and translated sample texts, verbal paradims, a Kulung-English lexicon and information on Kulung kinship terms.

Tibetan

We begin our discussion of works on Tibetan with a German-language article by Roland Bielmeier (Bielmeier 2004b), dealing with lexical variation and change in Tibetan as illustrated by designations for various body parts. Two further works on Tibetan by Bielmeier's colleague, Felix Haller, should also be mentioned here. The first, Haller (2004, in German), provides an in-depth overview of the language of the Themchen district of North Amdo. This 442-page work contains three main sections: The first section provides an overview of the structure of this dialect, covering among other things phonetics and phonology, nominals, adverbs and verbs as well as a brief chapter on syntax. The second section contains three narratives given first in phonological transcription, followed by a German translation, and finally a phonetic transcription. The third section then contains a Tibetan-German glossary, an index of place names, and a German-Tibetan index. Haller (2006), by the same author and written in English, then discusses verbal valence in Shigatse and Themchen Tibetan with respect to control, the case marking of the arguments of these verbs, and the semantic roles they express.

Finally, Zeisler (2004), a revision of her 1999 doctoral dissertation, is an extensive (xxv, 986pp.) study of temporal and aspectual concepts of verbal expressions in Tibetan languages belonging to four different dialectal groups: West Tibetan (including Balti, Purik, Ladakhi), "Lhasa" Tibetan, Kham and Amdo. The study consists of four main sections, dealing with the concepts of tense, aspect and mood from a more general perspective (Part I), the Tibetan system of relative tense and aspectual values (Part II), West Tibetan (Part III) and closes with a comparative discussion (Part IV).

Other

We now turn to the remaining Tibeto-Burman languages to be discussed here: Newār, Lepcha and Thangmi / Thami, proceding in alphabetical order of the authors. Hale & Shrestha (2006), a comprehensive grammar (xviii, 252 pp.) of Newār (also known as Newārī, Nepāl Bhāsā or Newāh), concentrates on the Kathmandu-Patan dialect of this language, covering virtually all areas of Newār grammar, including phonology, nominal and verbal morphology, and also contains an extensive discussion of syntax (covering four chapters), topic marking and a short Newār text.

Two works by Heleen Plaisier have also come to my attention, both dealing with Lepcha or *róngríng*, spoken in Sikkim, the Darjeeling district of West Bengal, the Ilām district of Nepal, and in a few villages of the Samtsi district in southwestern Bhutan. The first, Plaisier (2005), is an introduction to Lepcha orthography and literature. This study contains not only information on the indigenous Lepcha script and the literature in this language but also much information on the Lepcha people, clans, religion and a brief history of previous work on this language. The other work by this author of which I am aware, Plaisier (2007), is a reference grammar (xiii, 254 pp.) of Lepcha. It consists of two main parts: In the first part, a comprehensive overview of the language is given, with an introduction to the people, their customs and their region, as well as a discussion of previous work on Lepcha, and chapters on phonology and orthography, parts of speech, nominal and verbal morphology, clause-final particles, coordination and subordination. This is then followed by six texts given first in the Lepcha script, followed by a segmented and glossed text, and finally by an English translation. A Lepcha-English glossary rounds off the volume.

Finally, we come to three works authored or co-authored by Mark Turin on the Thangmi language of Nepal. Turin (2004b) is a 41-page overview of the phonology of Thangmi. This detailed study covers virtually all aspects of phonology, focusing on the Dolakhā dialect of this language. Turin (2004c) is then an analysis of linguistic evidence both for and against a genetic link between Thangmi and Newar. Finally, Turin & Thami (2004) is a 166-page Nepali-Thami (Thangmi)-English dictionary.

7. Austro-Asiatic

We now turn to the Austro-Asiatic languages of South Asia, the last family to be dealt with here separately. Although work on this family has traditionally lagged behind that on the other three major families of the subcontinent, the past ten years have seen an increase in research on these languages, both of the Mon-Khmer as well as of the Munda branches.

We begin with a brief discussion of Daladier (2007), a 43-page study in French dealing with the intricate nature of mutual Indo-Aryan/Austro-Asiatic borrowings with respect to mythical figures in the War-Khasi (Mon-Khmer, Meghalaya) oral traditions and in the War religion in which the author makes use of both comparative grammar and mythology as a means of recovering common features of Austro-Asiatic cosmogony, at the same time providing information on the period when Indo-Aryan and Austro-Asiatic cultures came into contact with one another. The article also contains much information on the linguistic diversity of the Mon-Khmer groups of Meghalaya.

Finally, there are three further studies which have appeared in the past two years dealing with the Munda languages of central eastern India. Two of these studies are in the form of responses to Evans & Osada (2005) on parts of speech in Mundari (North Munda) in an issue of *Linguistic Typology* dedicated primarily to this topic. The first of these, in their order of appearance there, is one of my own studies, Peterson (2005), in which I argue that Kharia (South Munda) as well as Mundari can best be considered languages where parts of speech in the traditional sense do not exist, as what is at issue in these languages are not lexical parts of speech (i.e., noun, verb, adjective, etc.) but rather "phrasal" units, in which enclitic markers for tense, aspect, mood and person attach to units which are syntactic in nature. Hengeveld & Rijkhoff (2005), in the same issue, then also argue that Mundari is in fact a language without distinct verb and noun classes, at least as far as the basic, non-derived vocabulary is concerned, although they concede that derived forms do exist in the language which cannot be used predicatively and must therefore be classified as non-verbs.

One further recent work on a Munda language is Peterson (2006), my *Habilitationsschrift* or "professorial dissertation", a three-volume description of the South Munda language Kharia. Volume I is a grammatical analysis (v, 375 pp.) covering virtually all aspects of this language. Volume II (v, 181 pp.) is a collection of Kharia texts, segmented, glossed, translated

into English and annotated, while Volume III (ii, 304 pp.) is a Kharia-English lexicon.

Finally, it should be mentioned here that an extensive volume on the Munda languages, Anderson (in press), is scheduled to appear later this year and will contain, among other things, a number of studies on aspects of Munda linguistics by scholars residing in Europe. As this volume has not yet appeared, these studies will be dealt with in the next regional report.

8. Language acquisition

In the past two years a number of studies on language acquisition have also appeared which are based on data from South Asian languages by Bhuvana Narasimhan and her associates. Narasimhan (2005) deals with agentivity and its acquisition in early childhood by examining the use of the ergative marker *ne* by Hindi-speaking children. Narasimhan & Gullberg (2006) then investigate children's sensitivies to patterns of perspective-taking in adult language on the basis of verbs of placement in colloquial Tamil, such as *veyyii* 'put' vs. the more fine-grained *nikka veyyii* 'make stand' and *paDka veyyii* 'make lie', based on work with 23 children (aged 3;11 to 6;7) and a control group of ten adult native speakers. The authors also take a closer look at the role animacy plays in shifting to a more fine-grained perspective. Finally, Narasimhan et al. (2005) deals primarily with argument ellipsis in Hindi and its implications for child language acquisition.

9. Computational linguistics

Finally, a number of works by European scholars dealing with computational aspects of South Asian languages have appeared over the past few years, which will be the topic of the present section. We begin with a discussion of work carried out by Andrew Hardie and his associates at Lancaster and Sheffield Universities, UK, in cooperation with the Central Institute of Indian Languages (CIIL) in Mysore, the EMILLE (Enabling Minority Language Engineering) Project.[3]

Baker et al. (2004) provides an overview of the now completed EMILLE Project, with information on the EMILLE Corpus, containing data on 14 South Asian Languages,[4] the parts-of-speech tagger for the Urdu data, and other topics such as the difficulties encountered in obtaining the diffierent

types of data. Issues involved in an automated parts-of-speech analysis for Urdu are then also dealt with in Hardie (2005) in more detail. Finally, Xiao et al. (2004) is a general discussion of the standards required for developing Asian-language corpora in general so as to facilitate international data exchange. The study also includes a discussion of two corpora which have been developed in Lancaster, namely the Lancaster Corpus of Mandarin Chinese and the EMILLE Corpus described above.

We conclude this section with a brief overview of works authored or co-authored by Gérard Huet. Sanskrit poses a number of special problems for computational linguistics which are not encountered (or, at least not to the same degree) in written texts in other languages, most notably (internal and external) sandhi and compounds, especially *bahuvrīhi* compounds, where gender issues complicate the analysis. These and other related issues are dealt with in Huet's studies (e.g., 2003a; 2005). Of special interest here are the "phantom phonemes" **e* and **o*, with which Huet is able to cope with peculiarities of certain preverbs such as *ā* with respect to sandhi rules. Huet (2004) then presents the architectural design rationale of a Sanskrit computational linguistics platform, one in which the lexical database plays a central role. Finally, Huet (2003b) and Huet & Razet (2006) are more general computational linguistic studies dealing with methodolgy related to the achievement of the above-mentioned goals.

10. Concluding remarks

As the preceding pages show, the study of virtually all aspects of South Asian languages is currently thriving in Europe. Nevertheless, despite the present situation, the last version of this report (Peterson 2004) struck a rather pessimistic tone, and unfortunately, the not-so-distant future does in fact look somewhat bleak. Conversations with colleagues from various European countries, as well as a look at internet sites for various departments, often give cause for concern as to the future of such work, as several Indological and linguistic departments have already been targetted for closure, with a number of other departments having only narrowly escaped closure, at least for the present round of cutbacks. This topic has also figured repeatedly in electronic discussion groups such as INDOLOGY, and it would seem to be only a matter of time before reports such as the present one become rather brief.

Interestingly, although discussions on this topic regularly take place on the INDOLOGY list, as well as discussions of constructive counter-measures, such discussions seem to be entirely lacking in linguistic discussion groups dedicated to these languages, such as VYAKARAN, although to my knowledge the number of linguistic departments now set to close in the coming years is probably higher than the number of Indological departments set to close.

However, discussion groups such as INDOLOGY also show that this situation is not necessarily inevitable, as a number of decisions to close a particular department have in fact been reversed in the face of international pressure. It is of course debatable as to how much of a role such protests from the international community have played in these decisions, nevertheless, as most of Europe's universities are state-run, with elected politicians ultimately pulling the strings, such protests can certainly serve to draw "unwanted" attention to a particular decision, perhaps forcing those responsible to at least reconsider their position.

It is perhaps time to follow the Indologists' lead in this issue: What is needed now is a kind of "Agenda for South Asian Studies", summarizing the many benefits of such studies for modern society, for example with respect to language technology, literacy programs, the development of writing systems for as yet unwritten languages, cultural heritage programs, and much, much more. If we wish to save what is left of linguistic and Indological departments, and with that a large portion of Europe's future contribution to the study of South Asian languages, a concerted effort will necessary, one which will, however, require much coordinated work before effective action can be taken.

Acknowledgements

A report such as the present one cannot be compiled without the help of many people, and I would like to take this opportunity to thank all those who aided me in this work, apologizing in advance should I have overlooked anyone: Balthasar Bickel, Roland Bielmeier, Jean-Luc Chevillard, Anne Daladier, George van Driem, Karen Ebert, Felix Haller, Andrew Hardie, Jan Heegård, Gérard Huet, Yana Kolotova, Daniel Krasa, Leonid Kulikov, Thomas Lehmann, Annie Montaut, Bhuvana Narasimhan, Thomas Oberlies, Ruth Laila Schmidt, Bertil Tikkanen, Will Tuladhar-Douglas, Mark

Turin, Mahendra Verma, Christophe Vielle and Marius Zemp. Thanks to all of you for sending me information on your own works as well as on those of your colleagues, and for answering my many questions!

Special thanks go to Liudmilla Khokhlova for all her assistance as well as her patience with my problems with the Russian language, and to Igor Fortuna and Carlotta Viti, who not only sent me information on their own works and that of their colleagues but even sent me the works themselves!

Of course, it goes without saying that I alone am responsible for any errors which may have found their way into this report.

Notes

1. For more information on this project, see http://www.uni-leipzig.de/~ff/cpdp.
2. For more details on this project and its members, see http://www.iias.nl/himalaya.
3. For a more detailed discussion of work in this area, the reader is referred here to the following website: http://www.emille.lancs.ac.uk/
4. These are: Assamese, Bengali, Gujarati, Hindi, Kannada, Kashmiri, Malayalam, Marathi, Oriya, Punjabi, Sinhala, Tamil, Telugu and Urdu.

References

Abbi, Anvita
2006 Vanishing voices: A typological sketch of Great Andamanese. In Anju Saxena & Lars Borin (eds.), 107–123.

Agesthialingom, S.
2004 Numeral system in Tamil: Generation. In Jean-Luc Chevillard & Eva Wilden (eds.), 323–332.

Allwood, Jens
2006 Language survival kits. In Anju Saxena & Lars Borin (eds.), 279–292.

Anderson, Gregory D. S. (ed.)
in press *The Munda Languages.* (Routledge Language Family Series.) London/ New York: Routledge.

Andvik, Erik
2004 'Do' as subordinator in Tshangla. In Anju Saxena (ed.), 311–40.

Annamalai, E.
2006 The impact of technology on language diversity and multilingualism. In Anju Saxena & Lars Borin (eds.), 195–201.

Baghbidi, Hassan Rezai

2006 Iranian elements in Sanskrit. In Bertil Tikkanen & Heinrich Hettrich (eds.), 143–166.

Baker, Paul, Andrew Hardie, Tony McEnery, Richard Xiao, Kalina Bontcheva, Hamish Cunningham, Robert Gaizauskas, Oana Hamza, Diana Maynard, Valentin Tablan, Christian Ursu, B. D. Jayaram & Mark Leisher

2004 Corpus linguistics and South Asian languages: corpus creation and tool development. *Literary and Linguistic Computing* 19 (4): 509–524.

Bickel, Balthasar

2004a Hidden syntax in Belhare. Anju Saxena (ed.), 141–90.

2004b The syntax of experiencers in the Himalayas. In *Nonnominative Subjects,* Peri Bhaskararao & Karamuri Venkata Subbarao (eds.), 25–59. Amsterdam: Benjamins.

2006 Clause-level vs. predicate-level linking. In *Semantic Role Universals and Argument Linking*, Ina Bornkessel, Matthias Schlesewsky, Bernard Comrie & Angela D. Friederici (eds.), 155–190. Berlin/New York: Mouton de Gruyter.

Bickel, Balthasar, Goma Banjade, Martin Gaenszle, Elena Lieven, Netra Paudyal, Ichchha Purna Rai, Manoj Rai, Novel Kishor Rai & 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 & Narayan P. Sharma (Gautam)

2007 Two ways of suspending object agreement in Puma: between incorporation, antipassivization, and optional agreement. *Himalayan Linguistics* 7: 1–18.

Bielmeier, Roland

2004a Shafer's proto-West Bodish hypothesis and the formation of the Tibetan verb paradigms. In Anju Saxena (ed.), 395–412.

2004b Lexikalische Variation und lexikalischer Wandel im Tibetischen am Beispiel einiger Körperteilbezeichnungen. [In German: "Lexical variation and lexical change in Tibetan as illustrated by designations for certain body parts".] In *Lexical Data and Universals of Semantic Change*. (Stauffenburg Linguistik 35.) Wiltrud Mihatsch & Reinhild Steinberg (eds.), 167–202. Tübingen: Stauffenburg.

2005 Sino-tibetische Sprachen. [In German: "Sino-Tibetan languages".] In *Die Sprache ist die Seele eines Volkes. Die großen Sprachfamilien* (Brennpunkt Mensch). Wilhelm R. Baier & Christian Zinko (eds.), 59–71. Graz: Leykam.

Borghi, Guido

2006 Note aggiuntive sulla comparazione di antico indiano *bʱūkă-* 'buco' e italiano *buco* 'ĭd.' [In Italian: Additional notes on the comparison of Old India *bʱūkă-* 'hole' and Italian *buco* 'ĭd.'.]. In Rosa Ronzitti, Guido Borghi & Luca Busetto (eds.), 53–116.

Borghi, Guido & Rosa Ronzitti
2005 Sanscrito *bhūka-* ‘buco’ : italian *buco* ‘id.’ (con un’Appendice su latino *făucēs*). [In Italian: “Sanskrit *bhūka-* ‘hole’ : Italian *buco* ‘id.’ (with an appendix on Latin *făucēs*)”.] *Quaderni di Semantica* 26 (1): 153–180.

Borin, Lars
2006 Supporting lesser-known languages: The promise of language technology. In Anju Saxena & Lars Borin (eds.), 317–337.

Bradley, David
2006 Lisu orthograhies and email. In Anju Saxena & Lars Borin (eds.), 125–135.

Bubenik, Vit
2006 On the evolutionary changes in the Old and Middle Indo-Aryan systems of case and adpositions. In Bertil Tikkanen & Heinrich Hettrich (eds.), 65–88.

Busetto, Luca
2004 Fonazioni e articolazioni in antico indoario. [In Italian: “Phonations and articulations in Old Indo-Aryan”.] In Ronzitti & Borghi (eds.), 67–80.
2006 La codifica della scrittura devanagarica. [In Italian: “The encoding of the Devanagari writing system”.] In Ronzitti, Rosa, Guido Borghi & Luca Busetto (eds.), 117–144.

Candotti, Maria Piera & Tiziana Pontillo
2004 Substitution as a descriptive model in Pāṇini’s grammar: toward an opposition between phonological and morphological levels? In Ronzitti & Borghi (eds.), 1–45.

Chevillard, Jean-Luc
2004 Ideophones in Tamil: a historical perspective on the *X-eṉal* expressives (*OlikkuŠippu ĀŠŠuppaṯai*). In Jean-Luc Chevillard & Eva Wilden (eds.), 407–433.

Chevillard, Jean-Luc & Eva Wilden (eds.)
2004 *South Indian horizons: felicitation volume for François Gros on the occasion of his 70th birthday. with the collaboration of A. Murugaiyan, with a preface by R. E. Asher.* (Publications du Département d’Indologie 94.) Pondichéry: Institut Français de Pondichéry.

Daladier, Anne
2007 Éléments cosmogoniques et vocabulaire austroasiatique du groupe môn-khmer du Meghalaya. Interactions avec le monde indien. [In French: “Cosmogonal elements and Austro-Asiatic vocabulary of the Mon-Khmer group of Meghalaya. Interactions with the Indian world”]. *Bulletin d’Études Indiennes* 22. Paris: Collège de France.

Driem, George van
2004a Newaric and Mahakiranti. In Anju Saxena (ed.), 413–418.

2004b Bhutan's endangered languages programme under the Dzonghka Development Authority: The three rare gems. In *The Spider and the Piglet: Proceedings of the First International Seminar on Bhutan Studies*, Karma Ura & Sonam Kinga (eds.), 294–326. Thimphu: Centre for Bhutan Studies.

2004c Hodgson's Tibeto-Burman and Tibeto-Burman Today. In *The Origins of Himalayan Studies. Brian Houghton Hodgson in Nepal and Darjeeling 1820–1858*, David M. Waterhouse (ed.), 227–248. London/New York: Routledge

2005 Tibeto-Burman vs Indo-Chinese. Implications for population geneticists, archaeologists and prehistorians. In *The Peopling of East Asia. Putting Together Archaeology, Linguistics and Genetics*, Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds.), 81–106. London/New York: Routledge.

Ebert, Karen H.

2005 South Asia as a linguistic area. In *Encyclopedia of Language and Linguistics. 2nd Ed.*, Keith Brown (ed.), Vol. 11. Amsterdam et al.: Elsevier.

Elizarenkova, Tatjana Ya.

2004a Древнеиндийские языки. [In Russian: "Old Indo-Aryan".] In Tatjana Ya. Elizarenkova, Andrej A. Kibrik & Leonid I. Kulikov (eds.), 14–16.

2004b Ведийский язык. [In Russian: "Vedic".] In Tatjana Ya. Elizarenkova, Andrej A. Kibrik & Leonid I. Kulikov (eds.) 16–25.

2004c Санскрит. [In Russian: "Sanskrit".] In Tatjana Ya. Elizarenkova, Andrej A. Kibrik & Leonid I. Kulikov (eds.), 25–69.

2004d Среднеиндийские языки. [In Russian: "Middle Indo-Aryan languages".] In Tatjana Ya. Elizarenkova, Andrej A. Kibrik & Leonid I. Kulikov (eds.), 70–73.

2004e Пали. [In Russian: "Pali".] In Tatjana Ya. Elizarenkova, Andrej A. Kibrik & Leonid I. Kulikov (eds.), 73–106.

2004f Пракриты. [In Russian: "Prakrit".] In Elizarenkova, Tatjana Ya., Andrej A. Kibrik & Leonid I. Kulikov (eds.), 106–123.

2004g Апабхранша. [In Russian: ApabhraÞĐa.]. In Tatjana Ya. Elizarenkova, Andrej A. Kibrik & Leonid I. Kulikov (eds.), 123–136.

2004h Буддийский гибридный сансрит. [In Russian: "Buddhist Hybrid Sanskrit".] In Tatjana Ya. Elizarenkova, Andrej A. Kibrik & Leonid I. Kulikov (eds.), 136–141.

Elizarenkova, Tatjana Ya., Andrej A. Kibrik & Leonid I. Kulikov (eds.)

2004 *Языки мира. Индоарийские языки древнего и среднего периодов.* [In Russian: "Languages of the World: Indo-Aryan languages of the old and middle periods".] Moscow: Academia.

Evans, Nicholas & Toshiki Osada
2005 Mundari: the myth of a language without word classes. *Linguistic Typology* 9: 51–90.

Fortuna, Igor
2006 Parole significanti 'scienza medica', 'dottore' e 'farmaco' nelle lingue del subcontinente indiano e del sud-est asiatico. [In Italian: "Words meaning 'medical science', 'doctor' and 'medicine' in the languages of the Indian subcontinent and of Southeast Asia".] In Rosa Ronzitti, Guido Borghi & Luca Busetto (eds.), 247–320.

Gaenszle, Martin, Balthasar Bickel, Goma Banjade, Elena Lieven, Netra Paudyal, Ichchha Purna Rai, Manoj Rai, Novel Kishor Rai & Sabine Stoll
2005 Worshipping the King God: a preliminary analysis of Chintang ritual language in the invocation of Rajdeu. In *Contemporary Issues in Nepalese Linguistics*, Yogendra P. Yadava, G. Bhattarai, R. R. Lohani, B. Prasain & K. Parajuli (eds.), 33–48. Kathmandu: Linguistic Society of Nepal.

Glass, Andrew
2006 A preliminary study of Gāndhārī lexicography. In Bertil Tikkanen & Heinrich Hettrich (eds.), 273–303.

Grierson, George A.
1909 *Linguistic Survey of India. Vol. III. Tibeto-Burman Family. Part I. General Introduction, Specimens of the Tibetan Dialects, The Himalayan Dialects, and the North Assam Group.* Calcutta: Government of India.

Grinevald, Colette
2006 Worrying about ethics and wondering about 'informed consent': Fieldword from an Americanist perspective. In Anju Saxena & Lars Borin, 339–370.

Grunow-Hårsta, Karen
2004 Direnction and differential dative case marking in Magar. In Anju Saxena (ed.), 77–100.

Gvozdanović, Jadranka
2004 Morphosyntactic transparency in Bantawa. In Anju Saxena (ed.), 341–46.

Hale, Austin & Kedār P. Shrestha
2006 *Newār (Nepāl Bhāsā).* (Languages of the World/Materials 256.) München: Lincom Europa.

Haller, Felix
2004 *Dialekt und Erzählungen von Themchen. Sprachwissenschaftliche Beschreibung eines Nomadendialektes aus Nord-Amdo.* [In German: "Dialect and Narratives of Themchen. Lingusitic Description of a Nomadic Dialect from North Amdo".] (Beiträge zur tibetischen Erzählforschung, Band 14.) Bonn: VGH Wissenschaftsverlag.

2006 Verbal valence in Shigatse Tibetan and Themchen Tibetan. *Linguistics of the Tibeto-Burman Area* 29 (2): 63–77.

Hardie, Andrew
2005 Automated part-of-speech analysis of Urdu: conceptual and technical issues. In *Contemporary issues in Nepalese linguistics*, Yogendra P. Yadava, G. Bhattarai, R. R. Lohani, B. Prasain & K. Parajuli (eds.), 49–72. Kathmandu: Linguistic Society of Nepal.

Hardie, Andrew, P. Baker, A. McEnery & B. D. Jayaram
2006 Corpus-building for South Asian languages. In Anju Saxena & Lars Borin (eds.), 211–241.

Hargreaves, David
2004 Directional prefixes in Kathmandu Newar. In Anju Saxena (ed.), 273–284.

Heegård Petersen, Jan
2006 Case-marking in Kalasha. Unpublisched Ph.D. thesis. University of Copenhagen.

Heegård, Jan & Ida Elisabeth Mørch
2004 Retroflex vowels and other peculiarities in the Kalasha sound system. In Anju Saxena (ed.), 57–76.

Heenen, François
2006 *Le désidératif en védique.* [In French: "The desiderative in Vedic".] (Leiden Studies in Indo-European 13.) Leiden: Rodopi.

Hengeveld, Kees & Jan Rijkhoff
2005 Mundari as a flexible language. *Linguistic Typology* 9–3: 406–431.

Hock, Hans Henrich
2006 Reflexivization in the Rig-Veda (and beyond). In Bertil Tikkanen & Heinrich Hettrich (eds.), 19–44.

Honda, Isao
2004 Grammaticalization of deictic motion verbs in Seke. In Anju Saxena (ed.), 285–310.

de Hoop, Helen & Bhuvana Narasimhan
2005 Differential Case-Marking in Hindi. In *Competition and Variation in Natural Languages. The Case for Case*, Mengistu Amberber & Helen de Hoop (eds.), 321–345. (Perspectives on Cognitive Science). Amsterdam et al.: Elsevier.

Huet, Gérard
2003a Towards Computational Processing of Sanskrit. In *ICON-2003, Mysore, India, Dec. 2003. Proceedings*, Rajeev Sangal, S. M. Bendre & Udaya Narayana Singh (eds.), 40–48. Mysore: Central Institute of Indian Languages.

2003b Automata Mista. In *Verification: Theory and Practice: Essays Dedicated to Zohar Manna on the Occasion of his 64th Birthday*. Nachum Dershowitz (ed.), 359–372. Berlin et al.: Springer (LNCS vol. 2772).

2004 Design of a Lexical Database for Sanskrit. *COLING Workshop on Electronic Dictionaries*: 8–14. Geneva, August 29, 2004.

2005 A Functional Toolkit for Morphological and Phonological Processing, Application to a Sanskrit Tagger. *Journal of Functional Programming* 15 (4): 573–614.

2006 Lexicon-directed segmentation and tagging in Sanskrit. In Bertil Tikkanen & Heinrich Hettrich (eds.), 307–325.

Huet, Gérard & B. Razet

2006 The Reactive Engine for Modular Transducers. In *Algebra, Meaning and Computation, Essays Dedicated to Joseph A. Goguen on the Occasion of His 65th Birthday,* Kokichi Futatsugi, Jean-Pierre Jouannaud & José Meseguer (eds.), 355–374. Berlin et al.: Springer (LNCS 4060).

Khokhlova, L.V. & Charanjit Singh

2007a Глаголы перемещения в жидкой среде и движения жидкости в западных индоарийских языках. [In Russian: “Verbs of motion in liquid in Western Indo-Aryan Languages”.] In *Aqua Motion: Глаголы движения в воде: Лексическая типология*, T. A. Maisak, E.V. Rakhilina, M. Indrik (eds.), 509–545. Moscow: Indrik Press.

2007b Resultative Constructions with Overt Agent/Possessor in Western NIA Languages (Hindi-Urdu, Punjabi, Gujarati, Rajasthani). *Proceedings of the Moscow Conference ICOSALL 5,* July 2003. Delhi: Motilal Banarsidass.

Kobayashi, Masato

2006 The development of Proto-Indo-Iranian **sć* into Sanskrit /(c)c^h/. In Bertil Tikkanen & Heinrich Hettrich (eds.), 1–18.

Kohistani, Razwal & Ruth Laila Schmidt

2006 Shina in contemporary Pakistan. In Anju Saxena & Lars Borin (eds.), 137–160.

Kolotova, Yana V.

2007 Глаголы перемещения в жидкой среде в бенгальском языке. [In Russian: “Aquamotion in Bangla”.] In *Aqua Motion: Глаголы движения в воде: Лексическая типология*, T. A. Maisak, E.V. Rakhilina, M. Indrik (eds.), 546–567. Moscow: Indrik Press.

Korn, Agnes

2005 Das Balochi im Kontakt mit den anderen iranisichen Sprachen. [In German: “Balochi in contact with the other Iranian languages”.] In *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft,* September 17–23, 2000, Halle an der Saale, Gerhard Meiser & Olav Hackstein (eds.), 285–294. Wiesbaden: Reichter.

Koryakov, Yu. B.
2004 Индийские письменности древнего среднего периодов. [In Russian: “Indian scripts of the Old and Middle periods”.] In Tatjana Ya. Elizarenkova et al. (eds.), 142–146.

Kulikov, Leonid
2005a Length vacillation *-iy-//-iy-* and related phenomena in Vedic. In *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft,* September 17–23, 2000, Halle an der Saale, Gerhard Meiser & Olav Hackstein (eds.), 295–317. Wiesbaden: Reichter.
2005b Reduplication in the Vedic verb: Indo-European inheritance, analogy and iconicity. In *Studies on reduplication.* (Empirical Approaches to Language Typology 28.) B. Hurch (ed.), 431–454. Berlin/New York: Mouton de Gruyter.
2006a The Vedic medio-passive aorists, statives and their participles: Reconsidering the paradigm. In Bertil Tikkanen & Heinrich Hettrich (eds.), 45–63.
2006b Passive and middle in Indo-European. Reconstructing the early Vedic passive paradigm. In *Passivization and typology: form and function.* (Typological studies in language 68.) W. Abraham & L. Leisiö (eds.), 62–81. Amsterdam: Benjamins.
2006c Case systems in a diachronic perspective: A typological sketch. In *Case, Valency and Transitivity.* (Studies in language companion series 77.) Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.), 23–47. Amsterdam: Benjamins.

Kulikov, Leonid & Ilona Manevskaia
2004 Text generation in the Indo-Tibetan tradition and its linguistic aspects: Nominative/accusative and ergative/absolutive patterning in Indo-Aryan and Tibetan syntactic systems (preliminary remarks). In *L’unité texte – Actes du colloque ‘Regards croisés sur l’unité texte / Conjoint Perspectives on Text’*. Sylvie Porhiel & Dominique Klingler (eds.), 153–170. Pleyben: Perspectives.

Lehmann, Thomas
2004 Pronoun Incorporation in Old Tamil. In Jean-Luc Chevillard & Eva Wilden (eds.), 307–322.

Matisoff, James A.
2004 Areal semantics – Is there such a thing? In Anju Saxena (ed.), 347–394.

Michailovsky, Boyd
2006a Bhutan. In *Encyclopedia of Language and Linguistics*, *2nd Edition*, Keith Brown (ed.), Vol. 1: 751–753. Amsterdam et al.: Elsevier.
2006b Digitized resources for languages of Nepal. In Anju Saxena & Lars Borin (eds.), 243–256.

Milizia, Paolo
2004 Sulla cospirazione indoaria contro le fricative sonore. [In Italian: "On the Indo-Aryan conspiracy against the voiced fricatives".] In Ronzitti & Borghi (eds.), 81–141.

Montaut, Annie
2004a Quotatif, médiatif et miratif dans les langues d'Asie du Sud: une ou plusieurs catégories?. [In French: "Quotative, mediative and mirative in the languages of South Asia: one or several categories?"] In *Les Médiations langagières, Vol. 1, Des faits de langue aux discours. Actes du colloque international «La médiation: marquages en langue et en discours»*, Régine Delamotta-Legrand (ed.), 111–122. Rouen: Publications de l'Université de Rouen.
2004b *A Grammar of Hindi.* (LINCOM Studies in Indo-European Linguistics 2.) München: Lincom Europa.
2005 Colonial language classification, postcolonial language movements and the grassroot multilingualism ethos in India. In *Living Together Separately. On the historicity of India's composite culture.*Mushirul Hasan & Asim Roy (eds.), 75–106. Amsterdam: Benjamins.
2006a Figures du sujet énonciateur: discontinue et continu en hindi/ourdou. [In French: "Figures of the subject of utterance: discontiuous and continuous in Hindi-Urdu".] In *Antoine Culioli, Un Homme dans le langage: Originalité, diversité, ouvertures: Actes du colloque de Cerisy-la-Salle*, June 2005. Claudine Normand & Dominique Ducard (eds.), 187–208. Paris: Ophrys.
2006b Mirative Meanings as Extensions of Aorist in Hindi/Urdu. In *The Yearbook of South Asian Languages and Linguistics*, Rajendra Singh (ed.), 49–70. Berlin/New York: Mouton de Gruyter.

Murugaiyan, Appasamy & Christiane Pilot-Raichoor
2004 Les prédications indifférenciées en dravidien: Témoins d'une évolution typologique archaïque. [In French: "Indifferentiated predications in Dravidian: Witnesses of an archaic typological evolution".] *Mémoires de la Société de Linguistique de Paris. Nouvelle Série, Tome XIV. Les Constituants prédicatifs et la diversité des langues.* 155–177.

Narasimhan, Bhuvana
2005 Splitting the notion of 'agent': case-marking in early child Hindi. *Journal of Child Language* 32: 787–803.

Narasimhan, Bhuvana, Nancy Budwig & Lalita Murty
2005 Argument realization in Hindi caregiver-child discourse. *Journal of Pragmatics* 37 (4): 461–495.

Narasimhan, Bhuvana & Marianne Gullberg
2006 Perspective-shifts in event descriptions in Tamil child language. *Journal of Child Langage* 33: 99–124.

Nathan, David & Éva Á. Csató
2006 Multimedia: A community-oriented information and communication technology. In Anju Saxena & Lars Borin (eds.), 257–277.

Noonan, Michael
2006 The rise of ethnic consciousness and the politicization of language in west-central Nepal. In Anju Saxena & Lars Borin (eds.), 161–174.

Oberlies, Thomas
2005a Language economy: 'Short(ened)' case-endings in Indo-Aryan. *Encyclopaedia of Indian Wisdom, Prof. Satya Vrat Shastri Felicitation Volume.* Vol. I., Delhi 2005: 513–515.
2005b *A Historical Grammar of Hindi.* (Grazer Vergleichende Arbeiten, Band 19). Graz: 2005.

Oetke, Claus
2004 On the notion of sentence in Classical Tibetan. In Anju Saxena (ed.), 191–212.

Oguibénine, Boris
2006 Notes on the instrumental case of the subject/agent vs. other cases in Buddhist Sanskrit. In Bertil Tikkanen & Heinrich Hettrich (eds.), 89–119.

Opgenort, Jean Robert
2004a *A Grammar of Wambule. Grammar, Lexicon, Texts and Cultural Survey of a Kiranti Tribe of Eastern Nepal.* (Brill's Tibetan Studies Library, Languages of the Greater Himalayan Region, Volume 5/2). Leiden/Boston: Brill.
2004b Implosive and preglottalized stops in Kiranti. *Linguistics of the Tibeto-Burman Area* 27 (1): 1–27.
2005 *A Grammar of Jero. With a Historical Comparative Study of the Kiranti Languages.* (Brill's Tibetan Studies Library, Languages of the Greater Himalayan Region, 5/3). Leiden/Boston: Brill.

Peterson, John
2004 Europe. In *The Yearbook of South Asian Languages and Linguistics – 2004,* Rajendra Singh (editor-in-chief), 131–144. Berlin/New York: Mouton de Gruyter.
2005 There's a grain of truth in every 'myth', or, Why the discussion of lexical classes in Mundari isn't quite over yet. *Linguistic Typology* 9–3: 391–405.
2006 *Kharia. A South Munda Language.* Vol. I: *Grammatical Analysis.* Vol. II: *Kharia Texts. Glossed, Translated and Annotated.* Vol. III: *Kharia-English Lexicon.* Unpublished Habilitationsschrift. Osnabrück: Universität Osnabrück.
no date Bibliography of Endangered and Seldom Studied South Asian Languages. *http://www.SouthAsiaBibliography.de/*

Pilot-Raichoor, Christiane
2006 Temps et espace dans la ballade de Giriji Mādi.]. (Monts Nilgiri, Inde du Sud). [In French: “Time and space in the ballade of Giriji Mādi”.] Samia Naïm (ed.), *La Rencontre du temps et de l’espace.* (Approches linguistique et anthropologique, numéro spécial 32.) Leuven/Paris/Dudley, MA: Peeters.

Pinault, George-Jean
2006 Further links between the Indo-Iranian substratum and the BMAC language. In Bertil Tikkanen & Heinrich Hettrich (eds.), 167–196.

Plaisier, Heleen
2005 A brief introduction to Lepcha orthography and literature. *Bulletin of Tibetology* 41 (1): 7–24.
2007 *A Grammar of Lepcha.* (Brill’s Tibetan Studies Library, Languages of the Greater Himalayan Region, Volume 5/5). Leiden/Boston: Brill.

Pollock, Sheldon
2004 A New Philology: From Norm-bound Practice to Practice-bound Norm in Kannada Intellectual History. In Jean-Luc Chevillard & Eva Wilden (eds.), 389–406.

Rahman, Tariq
2006 Language policy, multilingualism and language vitality in Pakistan. In Anju Saxena & Lars Borin (eds.), 73–104.

Rai, Novel Kishor, Balthasar Bickel, Goma Banjade, Martin Gaenszle, Elena Lieven, Netra Paudyal, Ichchha Purna Rai, Manoj Rai & Sabine Stoll
2005 Triplication and ideophones in Chintang. In *Contemporary issues in Nepalese linguistics,* Yogendra P. Yadava, G. Bhattarai, R. R. Lohani, B. Prasain & K. Parajuli (eds.), 205–210. Kathmandu: Linguistic Society of Nepal.

Renganathan, Vasu & Harold F. Schiffman
2006 The impact of technological advances on Tamil language use and planning. In Anju Saxena & Lars Borin (eds.), 203–210.

Ronzitti, Rosa
2004 Vecchie e nuove proposte etimologiche per alcuni derivati in **-mo-* dell’antico indiano. [In Italian: “Old and new etymological suggestions for some derivations in **-mo-* of Old Indo-Aryan”.] In Ronzitti & Borghi (eds.), 47–63.
2006a Derivati primari semplici in **-mo-/*meh*$_{2/4}$ della lingua postvedica (*A – DHYĀ-*). [In Italian: “Primary simple derivates in **-mo-/*meh*$_{2/4}$ of the post-Vedic language”.] In Ronzitti et al. (eds.), 3–20.
2006b *I derivati in* *-mo- *della lingua vedica (SaÞhitā e Brāhmaṇa).* [In Italian: “The Derivates in **-mo-* of the Vedic Language (SaÞhitā and Brāhmaṇa)”.] Perugia: Guerra Edizioni.

Ronzitti, Rosa & Guido Borghi (eds.)
2004 *Atti del secondo incontro genovese di Studî Vedici e Pāṇiniani. Genova, 23 luglio 2003; 15 ottobre 2003.* (Associazione Genovese di Studî Vedici e Pāṇiniani). Recco: Le Mani Microart's Edizioni.

Ronzitti, Rosa, Guido Borghi & Luca Busetto (eds.)
2006 *Atti del terzo, quarto e quinto incontro genovese di Studî Vedici e Pāṇiniani*, Geneva, July 26, 2004 / July27, 2005 / June 29, 2006. (Associazione Genovese di Studî Vedici e Pāṇiniani). Milano: Qu.A.S.A.R.

Saxena, Anju
2004a Linguistic synchrony and diachrony on the roof of the world – The study of Himalayan languages. In Anju Saxena (ed.), 3–29.
2004b On discourse functions of the finite verb in Kinnauri narratives. In Anju Saxena (ed.), 213–236.
2006 Introduction. In Anju Saxena & Lars Borin (eds.), 1–27.

Saxena, Anju (ed.)
2004 *Himalayan Languages. Past and Present.* (Trends in Linguistics, Studies and Monographs 149.) Berlin / New York: Mouton de Gruyter.

Saxena, Anju & Lars Borin (eds.)
2006 *Lesser-Known Languages of South Asia. Status and Policies, Case Studies and Aplications of Information Technology.* (Trends in Linguistics, Studies and Monographs 175). Berlin / New York: Mouton de Gruyter.

Scala, Andrea
2004 Il fonema /f/ nella romanī con particolare riguardo alla sua alternanza con /k^h-/ in posizione iniziale: considerazioni storice e applicazioni etimologiche. [In Italian: The phoneme /f/ in Romanī with particular regard to its alternation with /k^h-/ in initial position: historical considerations and etymological applications]. In Ronzitti & Borghi (eds.), 143–145.

Schalk, Peter
2004 Robert Caldwell's Derivation *īl̠am<sīhaºa*: A Critical Assessment. In Jean-Luc Chevillard & Eva Wilden (eds.), 347–364.

Scharfe, Hartmut
2006 Indo-Aryan and Dravidian convergence: gerunds and noun composition. In Bertil Tikkanen & Heinrich Hettrich (eds.), 197–253.

Schiffman, Harold F.
2004 The Tamil Case System. In Jean-Luc Chevillard & Eva Wilden (eds.), 293–305.

Schmidt, Ruth Laila
2004a Compound verbs in the Shina of Kohistan. *Acta Orientalia* 65: 19–31.
2004b A grammatical comparison of Shina dialects. In Anju Saxena (ed.), 33–55.

Seldeslachts, Erik
2006 Prākrit-like developments in Old Indo-Aryan: testing the 'Kölver-principle.' In Bertil Tikkanen & Heinrich Hettrich (eds.), 121–141.

Sharma (Gautam), Narayan P., Balthasar Bickel, Martin Gaenszle, Arjun Rai & Vishnu S. Rai
2005 Personal and possessive pronouns in Puma (Southern Kiranti). In *Contemporary Issues in Nepalese Linguistics*, Yogendra P. Yadava, G. Bhattarai, R. R. Lohani, B. Prasain & K. Parajuli (eds.), 225–232. Kathmandu: Linguistic Society of Nepal.

Sheldon, John S.
2006 The Sanskrit translation of the Avestan Haoma Liturgy in the light of recent research. In Bertil Tikkanen & Heinrich Hettrich (eds.), 255–272.

Shneiderman, Sara & Mark Turin
2006 Seeking the tribe. Ethno-politics in Darjeeling and Sikkim. *Himal Southasian* 18 (5): 54–58.

Singh, Charanjit & Liudmilla V. Khokhlova
2004 Resultative Structures in Russian, Hindi-Urdu and Punjabi. *Indian Journal of Russian Language, Literature and Culture* 5: 20–46. Delhi: JNU Publications.

Singh, Udaya Narayana
2006 Status of lesser-known languages in India. In Anju Saxena and Lars Borin (eds.), 31–59.

Subramonian, V.I.
2004 A Note on the *-āre* person number marking suffix in Gundert's writings. In Jean-Luc Chevillard & Eva Wilden (eds.), 291–292.

Tieken, Herman
2004 The Nature of the Language of CaÉkam Poetry. In Jean-Luc Chevillard & Eva Wilden (eds.), 365–387.

Tikkanen, Bertil & Heinrich Hettrich (eds.)
2006 *Themes and Tasks in Old and Middle Indo-Aryan Linguistics.* (Papers of the 12th World Sanskrit Conference, Vol. 5). Delhi: Motilal Banarsidass.

Tolsma, Gerard Jacobus
2006 *A Grammar of Kulung.* (Brill's Tibetan Studies Library, Languages of the Greater Himalayan Region 5/4.) Leiden/Boston: Brill.

Trosterud, Trond
2006 Grammatically based language technology for minority languages. In Anju Saxena & Lars Borin (eds.), 293–315.

Turin, Mark
2004a Thangmi kinship terminology in comparative perspective. In Anju Saxena (ed.), 101–39.

2004b The Phonology of Thangmi: A Tibeto-Burman Language of Nepal. *Journal of Asian and African Studies* 67: 63–103.

2004c Newar-Thangmi Lexical Correspondences and the Linguistic Classification of Thangmi. *Journal of Asian and African Studies* 68: 97–120.

2005 Language Endangerment and Linguistic Rights in the Himalayas. A Case Study from Nepal. *Mountain Research and Development* 25 (1): 4–9.

2006a Minority language policies and politics in Nepal. In Anju Saxena & Lars Borin (eds.), 61–71.

2006b Rethinking Tibeto-Burman: Linguistic identities and classifications in the Himalayan periphery. In *Tibetan Borderlands*, P. Christiaan Klieger (ed.), 35–48. Leiden/Boston: Brill.

Turin, Mark & Bir Bahadur Thami

2004 *Nepali-Thami-English Dictionary.* Kathmandu: Martin Chautari.

Vacek, Jaroslav

2004 Dravidian and Altaic: 'Sheep – Deer – Cattle.' In Jean-Luc Chevillard & Eva Wilden (eds.), 333–345.

Verma, Mahendra K.

2005 English as an Economic Investment: Who will Earn the Dividends? In *The Globalisation of English and the English Language Classroom*, Claus Gnutzmann & Franke Intemann (eds.), 41–54. Tübingen: Gunter Narr Verlag

Viti, Carolotta

2004a Osservazioni sulla frase relativa in vedico. [In Italian: "Observations on the relative phrase in Vedic".] In *Atti dell' undicesimo congresso nazionale di studi sanscriti*, Milan, Nov. 22, 2002, Oscar Botto (ed.), 241–256. Torino: Associazione Italiana di Studi Sanscriti.

2004b Funzioni semantiche e pragmatiche nelle strategie di possesso dell'antico indiano. [In Italian: "Semantic and pragmatic functions in the strategies of possession in Old Indo-Aryan".] *Archivio Glottologico Italiano* 89(1) : 41–83.

2005 Il suffisso vedico *-vat* fra comparazione e possesso. [In Italian: "The Vedic suffix *-vat* between comparison and possession".] Archivio *Glottologico Italiano* 90(1): 19–53.

2006 Lo sviluppo della subordinazione in vedico. [In Italian: "The development of subordination in Vedic".] In Rosa Ronzitti et al. (eds.), 21–31.

2007 *Strategies of subordination in Vedic*. Milano: FrancoAngeli.

Winter, Werner

2004 Preverbal modifiers in Sunwar. In Anju Saxena (ed.), 239–272.

Xiao, Z., A. McEnery, P. Baker & A. Hardie.
2004 Developing Asian language corpora: standards and practice. In *Proceedings of the 4th Workshop on Asian Language Resources. Sanya, China.* 1–8.

Zakharyin, Boris
2004 Категория ‘karmaṇ’ древнеиндийских грамматистов и её применимость к индоарийскому синтаксису. [In Russian: “The category ‘karmaṇ’ of the grammarians of ancient India and the possibilities of its application to Indo-Aryan syntax”.] In *Сравнительно-историческое и обсчее языкознание*. Moskow: Dobrosvet. 46–57.
2007a Глаголы перемесчения в жидкости и движения жидкости в историй индоарийского. [In Russian: “Verbs of movement in liquids and of movement of liquids in the history of Indo-Aryan”.] In *Aqua Motion: Глаголы движения в воде: Лексическая типология*. T. A. Maisak, E.V. Rakhilina, M. Indrik (eds.), 568–581. Moscow: Indrik Press.
2007b *Санскритская грамматика Варадараджи.* [In Russian: “The Sanskrit Grammar by Varadarāja”.] Moskow: Stepanenko.

Zeisler, Bettina
2004 *Relative Tense and Aspectual Values in Tibetan Languages. A Comparative Study.* (Trends in Linguistics, Studies and Monographs 150). Berlin/New York: Mouton de Gruyter.
2006 Why Ladakhi must not be written – Being part of the Great Tradition: Another kind of global thinking. In Anju Saxena & Lars Borin (eds.), 175–191.

Zoller, Claus Peter
2005 *A Grammar and Dictionary of Indus Kohistani.* Vol. 1, *Dictionary*. Berlin/New York: Mouton de Gruyter.

Zvelebil, Kamil
2004 Prolegomena to an Etymological Dictionary to the Iṟula Language. In Jean-Luc Chevillard & Eva Wilden (eds.), 281–290.

Reviews

Rama Kant Agnihotri – Hindi: An Essential Grammar

London: Routledge.
2007. xxii, 274 pp. US $ 120.00. ISBN 978-0-415-35670-1.

Reviewed by *Ayesha Kidwai*

Writing an essential grammar is a hazardous exercise. From decisions about what is essential in the grammar of a language, to its target audience, to questions about how prescriptivism and normativity are to be avoided, the terrain that such a grammarian must traverse is fraught with risk. For a writer of a grammar in the Indian context, the passage is even more precarious, because of the historical richness of the legacy available since ancient times. Often, the classical formulations of the grammars of ancient languages have been perceived as models that must simply be imitated, with the result that these exercises are doomed from their very inception. Only occasionally are such volumes worthy of any kind of review; happily, Rama Kant Agnihotri's *Hindi: An Essential Grammar* is an excellent example of both the do-ability and worthiness of such an exercise.

The most appealing aspect of Agnihotri's grammar is its clear conception of its own objectives and functions. For speakers of Hindi, it is an exposition of the systemacity and rule-governed nature of their language; for learners of Hindi, it is an instrument to further the learning of the language. In its jargon-free description of the patterns of Hindi grammar, the volume doubles up as an introduction to modern grammatical analysis for anyone trying their hand at grammar construction. In doing so, it produces an analytical learner/speaker who is not merely a *user* of language, but also its student.

Another important aspect of the book is the masterful definition of the object of study – Hindi; in simple, clear terms, and in just the five pages, of Chapter 1, Agnihotri distinguishes his project from official and community efforts towards defining Hindi. Rather than ascribing to a normative Hindi that is both homogenous and monolithic, Agnihotri points to the continuum between Hindi and its so-called dialects and its close relations to Urdu, in order to constitute an object of description that is both intersecting and in-

tersected. This understanding is well-reflected in the actual language data used in the book – just as in the mind of an ordinary speaker of Hindi, words and constructions from Perso-Arabic sources co-exist with words from indigenous and Sanskrit sources, they do so in the illustrations of grammar provided.

For a linguist, what is truly satisfying is that all this is achieved without any trace of naiveté or idealism in terms of a denial that people now *do* perceive of Hindi and Urdu as distinct languages. Instead, even while Agnihotri acknowledges that "what was hitherto only one language, often written in two scripts, is [now] perceived to be two different languages (p. 9)" – he can still continue with his plural description, because he takes the object of knowledge to be not some official or political construction of "Hindi", but rather the knowledge of her language that is an individual speaker's competence.

Agnihotri's examples quite naturally draw on as wide a range of lexical resources and contexts that an average Hindi speaker would be expected to have access to. The accompanying observations on the conditions of *use* of the examples, and in the Appendix on *Grammar in Context*, is also particularly worthy of commendation, as they not only relieve the work of the usual accusations of prescriptivism that grammars typically attract, they also reveal to the reader how grammatical analysis enriches our understanding of the social and the symbolic.

The book is divided into seven parts: Part 1 introduces the basic sentence structure of Hindi and elementary syntactic patterns; Parts II-IV discuss word-level structure; Parts V and VI return to the syntactic patterns; finally, Part VII discusses the sound system of Hindi and the mechanics of the mapping from sound to script. In this review, I will discuss each part in turn. Although, as it turns out, I will point out some omissions in Agnihotri's analyses in my comments, these should not serve to detract from an appreciation of the effort as a whole; rather, they are merely to be taken as suggestions for a future edition that the book richly deserves.

Part I – In this discussion of the basic syntax of Hindi, Agnihotri addresses questions of basic syntax and linear order (Chapter 2), negation (Chapter 3), interrogatives (Chapter 4) and exclamations in (Chapter 5). In general, the approach is very refreshing, as the objective here is not simply to demonstrate to the reader the rules governing these basic patterns; it is also to educate the reader to a level where she can begin to work with the author in

analysing the language. Chapters 4 and 5, in particular, display the advantages of Agnihotri's approach of specifying conditions of use for the examples employed, as what is significant about exclamations and imperatives is not a complex syntax, but when they may be used.

This appreciation notwithstanding, a few quibbles are still in order, particularly keeping in mind the intended non-linguist audience. In the Chapter 2 discussion of linear order (section 2.3), the free phrase order (scrambling) instantiated in Hindi is introduced, but unlike many other topics dealt with in the chapter, this is never revisited in any degree of detail. Yet, this flexibility of word order is used to great semantic and pragmatic effect by Hindi speakers, especially to encode topicality and focus (contrastive and presentational), and competence in the language necessarily entails an effective use of these information-packaging constructions. While it would be too much to expect an elementary grammar to delineate all the different ways in which word order variation packages information, a brief discussion of how movement to the left renders constituents presuppositional/old information and how preverbal positioning leads to focus/new information interpretations would be useful (particularly for Chapter 4, where the placement of question words is discussed).

Chapter 3, 'Negatives' is an informative exposition of the basic facts of Hindi negation; in particular, the discussion on the affective uses of *na* is novel and engaging. Unfortunately, however, there is also an important omission – constituent negation. In Hindi *nahĩĩ* can be used to negate just a phrase; for example *vah skuul nahĩĩ jaayegii* cannot only mean 'She will not go to school" but also 'It is not the school to which she will go tomorrow'. This narrow focus of negation is available for all constituents in a sentence, and is effected by placing *nahĩĩ* immediately after the phrase to be negated.

Chapter 4, 'Questions', does not discuss questions in complex sentences. It is important for the reader to know that Hindi does not allow question words left *in situ* in complex sentences to be answered – *vah jaantaa thaa ki kaun jaa rahaa hai* is not a question in Hindi, meaning the same as the English 'He knows who is going.' To transform this into a question analogous to 'Who does he know is going?', Hindi speakers may use either of three strategies:

(i) *kaun, vo jaantaa hai jaayegaa* (scrambling)
(ii) *vo kyaa jaantaa hai ki kaun jaayegaa* (scope-marking)
(iii) *vo kaun hai jo vo jaantaa hai ki jaayegaa* (clefting)

A mention of these strategies either here, or in Chapter 32 (on complex sentences) would make the account of interrogatives in Hindi complete. Although considerations of space and the specific nature of the target audience may well militate against such an elaboration, certainly the unavailability of wide scope interpretations for embedded questions must be noted.

Part II – Part II discusses the affixal word formation rules in Hindi. The focus on the category (noun, verb, adjective, and adverbs) turn-by-turn rather than the process (inflection, derivation) is learner-friendly. The exposition here is demanding however, as the simple presentation here takes a distinct linguistic turn here. Not that that is unwarranted, as the investigation of word-structure requires attention and commitment of a linguistic nature. My only query here is whether the use of bi-directional word-formation strategies (WFSs) does not demand too much of the learner, particularly as the bi-directionality part of the WFS is not exploited in the exposition, or presented as of special aid to the vocabulary building process. Furthermore, as WFSs are not ordered in terms of process (inflectional or derivational) or from the most productive to the least productive, the discussion can be a little difficult to internalize – perhaps a future edition could order WFS more explicitly in terms of productivity.

Part III – Part II discusses the Hindi WFSs of reduplication, compounding, and causativisation, and it must be said that here the energy of the book flags a little. In particular, reduplication is an extremely productive process in Hindi, and it has a number of conditions of grammatical and pragmatic use; but Agnihotri's discussion is a trifle lukewarm. Chapter 15, 'Reduplication', begins with the claim that any grammatical category can be reduplicated in Hindi (p. 112). What Agnihotri clearly has in mind a frame like 'What is this ___ that you keep on about?', into which any reduplicated category – article, auxiliary, non-participial verb can be inserted; however, it is nevertheless important to distinguish these cases (e.g., *thaa-thaa, har-har, dekhegaa-dekhegaa*) from the other more normal cases of reduplication. While the former instance of reduplication has chiefly a pragmatic function of conveying disapproval for the *form* of the language employed by an interlocutor, in the more normal sort of complete reduplication, distinct grammatical meanings are produced. It is also not the case, as Agnihotri suggests, that the meaning of complete reduplication always distributive: it can be

sometimes intensificatory, as in *laal-laal*, simultaneity, as in *calte-calte*, and the like. The conditions of grammatical use are also worthy of mention – compare *sitaa-ne raam-ko apne ghar-mẽ ghuste hue dekhaa*, 'Sita saw Ram entering her/his room' with *sitaa-ne raam-ko apne ghar-mẽ ghuste-ghuste dekhaa*, 'Sita saw Ram as she was entering her room'.

Another opportunity lost is in the discussion of partial reduplication of the echo-reduplication variety, as in *caay-vaay*. Echo reduplication is used primarily as a hedging device, in conditions in which the echoed word can be construed as shorthand for a list of metonymically associated items. Thus for example, *caay-vaay* is not only tea, but food items that are associated with tea – such as biscuits, *namkeen* and savouries, but unlikely to include *biryani*. Agnihotri lets the topic go, with a simple gloss of "etc".

In Chapter 16, 'Compounds', Agnihotri ascribes to the view that compounding must necessarily target two or more *words*, rather than the more dominant view of considering it to target two bases. As a consequence of this requirement of wordhood for compounding, Agnihotri's analysis presents any complex word that involves the concatenation of one or more bound bases as a WFS; for example, the analysis of *ghursvaar* and *hathkaRii* on p. 117. While this is an interesting analysis, it remains a puzzle as to why the section that follows continues to label the other WFSs listed (16.4) as 'Compounds' – if compounding is predictable by WFS rule, why is it not an instance of derivation? Perhaps this is simply a mistake; if so, it must be corrected.

Part IV – Part IV examines invariant words like pronominals, postpositions, particles, conjunctions, and other invariant words. The discussion on pronouns is particularly enjoyable; however, the chapter on postpositions is perhaps too much of a list. Given that Agnihotri's covert goal is to turn the learner into a language analyst, a more elaborate discussion of the conceptual axes on which relations between nouns are plotted by human cognition would be a welcome addition.

In addition, the discussion of emphatic particles in Chapter 21 is also in need of amendment, as Agnihotri, in my opinion, has only a partial analysis of both *-hii* and *-to*. He considers *-hii* to be the equivalent of English 'only'; but if that were indeed true, what should the analysis of *sirf/kewal* be? As Manindra K. Verma has observed in his *The Structure of the Noun Phrase in English and Hindi* (1971, Motilal Banarsidass), *-hii* marks the scope of *sirf* in Hindi. Thus, *raam* ***sirf*** **merii** ***hii*** *kitaab paRhegaa* means Ram will

read only the book that belongs to *me*, whereas *raam* ***sirf*** *merii kitaab* **paRhegaa** ***hii*** means that Ram will only *read* the book that belongs to me. In the examples that Agnihotri gives, it therefore seems reasonable to assume that *sirf/kewal* is elided.

Agnihotri does not also record the fact that *raam-hii seb khaata hE*, '***Ram*** eats apples' can also have a pure emphatic meaning. Suppose that this sentence is uttered as a response to a statement like "I don't know why Ram doesn't eat apples" – here the *-hii* marked noun will not have an 'only' interpretation, but rather will serve to emphasize that other interlocutor's presupposition is false.

The analysis of *-to* is also not sufficiently elaborate. There are two uses of the particle that need to be distinguished – the tag use, and the topic particle use (i.e. marking prior mention in the discourse or shared presupposition). Agnihotri mentions both uses, but does not elaborate. In the topic particle use, *-to* can have both a contrastive as well as a presentational interpretation – Agnihotri illustrates only the latter, glossing the interpretation of *mEhmaan to aa gaye*, as, 'as far as the guests are concerned, they have arrived'. However, a contrastive interpretation is also possible, and becomes clear when we embed an utterance like ***Raam to*** *aayegaa*, '*Ram* will come' in the context of a preceding utterance like "I don't know what Ram, Sita and Ramesh are up to, are they coming or not?"

Finally, Agnihotri does not discuss the fact that the *-hii* and *-to* particles can be used in combination, as in ***raam-hii-to*** *yeh baat phElaa rahaa hE,* 'Ram is the one who is spreading the rumour!" This is the only permissible combination of the two particles, and requires a context of the following sort: "I have learnt that there is a rumour being spread around that I am resigning. Many of my colleagues, and especially Ram, were outraged." Here, the *-to* particle sets up the contrast, and the *-hii* emphasizes it.

Part V – Part V marks a return to syntax, and the discussion here is pleasantly nuanced. The part begins with a comprehensive discussion of the habitual aspect in Hindi and its modal functions. This chapters of the use of the progressive and the subjunctive and the future in the chapters that follow are equally consummate and; however, the best chapter in this part is surely the description of the passive (Chapter 25). The only quibble I have here is that this chapter does not illustrate the passive of derived verbs like causatives (discussed earlier in Chapter 17) or compound verbs (discussed later in Chapter 31). The remainder of the part examines case phenomena in Hindi.

Part VI – Part VI moves to the syntactic patterns in complex clauses. The objective here is a broad sketch of the basic combinations beyond the simplex level, and in that respect, Agnihotri accomplishes what he has set out to do. Although the further detail that the subject merits would be unreasonable to expect, additional remarks on the Persian *ki* in Chapter 32, particularly with reference to its relatively recent incorporation into the grammar of Hindi, would be opportune. Mention of the fact that *ki*, along with the other subordinating conjunctions in Hindi, enforces an SOV order in complex clauses is also warranted; as would be of fact that Hindi ki-clauses may be used in construction with an object *yeh* or *yeh baat* or *aisaa* in the matrix clause (as in *mohan-ne yeh kahaa ki raam aayegaa*). This would complete the description of the basic patterns.

Part VII – Part VII reveals the systematicity and beauty of the Nagari orthographical system. In demonstrating how sounds are mapped to graphemes by the alphabet, Agnihotri has in effect prepared a handbook for all teachers of Hindi wrestling with the perplexing inability of students to master spelling in Hindi.

In conclusion, I would like to emphasize once again that the terms that have enabled the rather detailed discussion above have been set by the high standards of description that Agnihotri has himself set. If anything, these observations illustrate the difficulty of identifying what is to be considered *essential* in a grammar, as the adjective may well be motivated by the constraints imposed by publishers, expectations of reader knowledge and commitment, and the subjective selection that any linguist makes. These comments notwithstanding, I have no doubt that this work will be used by students of Hindi and linguists for many years to come – in fact, an abridged edition in Hindi translation could also be planned for use as a grammar book for school students of Hindi. Agnihotri's *Hindi: An Essential Grammar* will, from now on, count as essential reading for anyone interested in the language, irrespective of age or specialisation.

Yamuna Kachru – Hindi

London/Philadelphia: John Benjamins.
2007. xxi, 309 pp. Hc. US $ 169.00. ISBN 978-90-272-3812-2.

Reviewed by *Pramod Pandey*

Hindi is the official language of the Republic of India and is spoken by the fifth largest community of speakers in the world. The language has a long and continuous history that relates it to its neighboring sister Indo-Aryan languages such as Bangla, Gujarati and Punjabi. Having played a critical role in the socio-political history of India, especially in the past two centuries, it has come to acquire a character that languages functioning as efficient instruments of socio-political development are prone to. Hindi has a lexicon with multiple subsets of borrowings from different sources for administrative, cultural and collective purposes. The sources are, in the main, Sanskrit, Perso-Arabic and English, as well as the languages of its users in the bazaar. It is known as 'Hindi', with a moderate mixture of all these stocks, a common syntax, and the Devanāgarī script. With a tilt towards the vocabulary and morphology of Perso-Arabic, and the use of the Perso-Arabic script, it is known as Urdu, an official language of some states and the official language of the neighboring Pakistan. A tilt towards Sanskrit, however, is considered to yield a Sanskritized style of Hindi. The word 'Hindustani', used sometimes as a variant name for both Urdu and Hindi, and sometimes as the colloquial variety of either, is gradually disappearing, albeit not without the rapprochement of its supporters for the status of the official language of India. The chances of Hindi holding ground as the official language, however, are strong, what with the government continuing to pursue the policy of looking up to Sanskrit for the vocabulary of administration and scholarship, the name 'Hindi' having found currency in national/international political and cultural discourse, and the name simply being short and familiar, sounding like a preferred clipping, like 'hankie' and 'Paki'. It is noteworthy that all the names, Hindi, Urdu, and Hindustani, are of Perso-Arabic origin. In spite of there being scope for controversy regarding the forms and varieties of Hindi on account of its dynamic use

across regions and social strata, there is little doubt that forces of inertia have lent the language sufficient stability and form.

The present book by Yamuna Kachru (YK) offers an account of the stable form of modern standard Hindi, with painstakingly thorough documentation and illustrations.

Chapter 1 deals with some of the marked features of Hindi and its stylistic variants and offers a brief introduction to the history of the language and literary tradition.

Chapter 2 presents a sketch of the sound system-vowels, consonants, consonant clusters, syllable structures, the inherent vowel, morphophonemic alternations, stress and intonation.

Chapter 3 is a brief introduction to the orthographic devices of Devanāgarī in representing speech sounds.

The treatment of morphology and syntax starts with Chapter 4 ('Parts of Speech'). The chapter contains detailed classifications of word classes with copious data illustrating the inflectional morphology of the word classes posited on the basis of both formal and functional criteria. The word classes include Noun, Determiner, Pronoun, Adjective, Verb, Adverb, Postposition, Conjunction, Particle and Interjection. The chapter is rich in detail and by far the most thorough description of the morpho-syntax of Hindi.

Chapter 5 is devoted to word formation processes involved in the derivation of complex and compound nouns, adjectives and verbs. It also includes a section on reduplication.

Chapter 6 discusses the internal structures of simple and complex noun phrases. A simple noun phrase consists of a determiner and various quantifiers, while a complex noun phrase may consist of either a complement clause or a relative clause / participial modifier.

Chapter 7 elaborates the structure of the verb phrase in Hindi. Subject to the valency of the verb, the verb phrase may be a simple finite verb without an object and with (aspect) – (tense) – mood – (agreement) markers, or it may take 'a direct object, or both a direct object and an indirect object', or 'a direct object and a complement that refers back to the direct object.'

The account of the structure of the simple sentence in Hindi in Chapter 8 begins with a description of word order (essentially a verb-final language, with features of a verb-medial language) and principles of agreement (involving grammatical gender, number, case, aspect, tense and mood distinctions) between modifier and head, and noun phrase and verb. Word order in Hindi is shown to be rigid (with occasional violations to satisfy metrical re-

quirements, especially in poetic tests), but the constituents can occur freely with respect to each other. The obligatory (subject, predicate, aspect and tense agreement markers) and optional elements of the structure of a simple sentence are exhaustively covered. Explicit descriptions are provided for some of the construction types of Hindi discussed in the literature, such as the 'ergative' sentence constructions, and the dative subject type sentences. It is clearly shown that the verb in Hindi does not agree with a noun that is in the oblique case and followed by a postposition.

Chapter 9 presents complex and compound type sentences. Hindi is shown to have complex sentences formed on account of complement clauses, correlatie constructions (relative, appositive, adverbial and quality and quantity clauses), participial constructions (the present, the past, and the pan Indian conjunctive), and conditional clauses. Compound sentence types are five in number- co-ordinate, adverbial, disjunctive, concessive and antithetical.

The remaining two chapters that follow, Chapter 10 ('Information Structure') and Chapter 11 ('Discourse Structure'), have a more restricted scope, largely on account of limited work in the areas. The discussion of the information structure of a sentence in Hindi centers on the notion of Theme and Focus. The Theme is shown to be the function of any of the following constituents in a declarative sentence- the subject, the direct object, the complement, adverbial and the subject and the adverbial together. An interesting feature of the Topical Theme in Hindi is shown to be its zero realization.

The discussion of the discourse structure of a text in Hindi taken up for analysis is limited to those aspects 'that intersect with grammatical choices'. These include (i) the devices of cohesion, dependent on a number of processes, (ii) expression of politeness, dependent on grammatical and lexical choices, (iii) markers of discourse, and (iv) conventions governing speaking and writing. The chapters are followed by short Appendices, References, Select Bibliography and Index.

The book undoubtedly makes a significant contribution to the field of Hindi linguistics. One can't think of a better existing reference grammar of Hindi for the modern scholar interested not only in theoretical linguistics but also in applied linguistic areas such as computational linguistics (e.g. machine translation), language teaching and translation studies. The present description provides a rich source of material for soliciting students from these disciplines. The general utility of the grammar, in my opinion, is on account of the balanced approach of the author to linguistic description,

drawing from the best of the grammatical traditions on both sides of the Atlantic as well as from the Indian tradition (pp. 211–213). The author's original understanding of grammar and grammatical issues are evident from the elegant generalizations and presentation of relevant data. Of special mention are the chapters on morphology and syntax (Chapters 4–10), which compare with any reference descriptive grammar of a language, with most types of constructions discussed in the literature having found inclusion and careful exposition. The end chapters have made a start for the study of texts in Hindi, and are expected to show the way for further work on text linguistics in the language.

There are, however, certain aspects of the book that strike me to draw critical attention, and that I point out below.

Chapters 1 and 2 are a bit too sketchy, notwithstanding their limited scope. More specifically, Chapter 1 could have touched upon Hindi as lingua franca that has developed in different parts of the republic, such as the north-east and the Andaman-Nicobar islands. Chapter 2 could be more consistent in its treatment of the sound system of Hindi. The chapter contains 9 pages of discussion on types of consonant clusters, but 4 pages on syllables, stress, morphophonemic alternations and intonation, all put together.

The data are presented consistently using the IPA, with a couple of differences. YK uses /y/ for the IPA /j/ and /j/ for the IPA /ɟ/. Some of the citation forms are transcribed keeping the orthographic rather than the spoken form in mind. In particular, the retroflex /ʂ/ is not found to occur in isolation in spoken Hindi, although kept in the written form. The transcription of certain English borrowings could be more accurate, for example, /es.pi/ for /es.p/ (p. 144) or /ti.ʊi/ for /tv/ (p. 149).

The author uses the labels S ('Sanskritized'), PA ('Perso-Arabic') and E ('Englishized') for borrowed vocabulary. Forms that are not labeled are assumed to be indigenous to Hindi. The distinction between borrowed and indigenous is occasionally unclear for the S and PA items. Thus, most commonly used words like /vyapar/ 'trade' and /ʃyam/ 'a name' are treated as S on the one hand, and words such as /malik/, /sahəb/ and /lekin/ of P-A origin are without a label (p. 7), implying thereby that they are indigenous.

There are instances of older and now infrequent grammatical and lexical forms being treated as contemporary. For instance, the vocative plural is stated to be the oral vowel /o/, as distinct from the oblique plural /õ/. In contemporary Hindi, one hears both the vocative and the oblique plural as the nasal /õ/. Forms such as /səhuain/ 'merchant F'(p. 47) and /koʈʰəriya/

'small room' (p. 117), /ṭoṭa/ 'loss, damage' are rare. The correlative pronoun *ʋəh jo* is frequently *ʋo jo* in spoken as well as written form. The most significant aspect of grammar in this case is the 2nd P. pronoun (pp. 63–64, 81–82). The 2nd P. SG pronoun is throughout specified as *tū* in the direct case. The 2nd P. PL in the direct case is given as *tum.* There is a third common honorific form *ap* The PRES and PAST tense auxiliaries for the 2nd P. SG pronoun are also accordingly differently specified. I reproduce below the 2nd P. pronoun forms from YK (p. 63):

(1)

	DIR	OBL
2nd P. SG	tū	tuj^h
2nd P. PL	tum	tum
HON	ap	ap

The tense auxiliary forms for the 2nd P. pronouns are given as follows (p. 82)

(2)

	PRES				PAST			
	M.SG	M.PL	F.SG	F.PL	M.SG	M.PL	F.SG	F.PL
2nd P	hɛ	ho	hɛ	ho	t^{h}a	t^{h}e	t^{h}ī	t^{h}ĩ

This description is at variance with contemporary Hindi, which has, a threefold rather than two-fold distinction, namely, Honorific, Non-honorific, and unmarked (see also Annie Montaut 2004, *A Grammar of Hindi*, Munchen, Lincom Europa). The pronoun forms are as follows:

(3)

	DIR	OBL
2nd P. SG NH	tū	tuj^h
2nd P. PL NH	tum	tum
2nd P. SG	tum	tum
2nd P. PL	tum	tum
HON	ap	ap

The tense auxiliary forms for the 2nd P. pronoun can be revised as in (4).

(4)

	PRES				PAST			
	M.SG	M.PL	F.SG	F.PL	M.SG	M.PL	F.SG	F.PL
2nd PNH	hɛ	ho	hɛ	ho	tʰa	tʰe	tʰī	tʰĩ:
2nd P	ho	ho	ho	ho	tʰe	tʰe	tʰĩ:	tʰĩ:
HON	hɛ̃	hɛ̃	hɛ̃	hɛ̃	tʰe	tʰe	tʰĩ:	tʰĩ:

If YK has any considerations for presenting the facts about pronouns in Hindi in the way she has, then I have failed to notice them.

The Preface is dated 'Ramanavami, June 6, 2006'. The Ramanavami in 2006, incidentally, was on April 6.

A final point that has struck the present reviewer is the exclusion of some crucial references from the select bibliography; for instance, Ashok R. Kelkar (1968, *Hindi Word Phonology*: Part I, Poona: Deccan College) for the discussion on syllable structure types (Ch. 3), and S. K. Verma (1974, 'The semantics of 'caahiye'. *Foundations of Language*) for the discussion on *cahiye* (Ch. 7). The contribution of Kelkar is of general interest and has been discussed in recent theoretical works, such as John J. McCarthy and Alan R. Prince (1993, *Prosodic Morphology: Constraint Interaction and Satisfaction*, Rutgers University Center for Cognitive Science. Technical Report 3). There are other references that can easily be mentioned along this line. The point, however, is that omissions of significant references, even if inadvertent, are naturally not expected in a work of this measure and elegance.

P. G. Patel, Pramod Pandey and Dilip Rajgor (eds.) – The Indic Scripts: Paleographic and Linguistic Perspective.

New Delhi: D. K. Printworld (P) Ltd.
2007. xiii, 266 pp. Hc. Indices. Rs. 750 / US $ 37.50. ISBN: 81-246-0406-1.

Reviewed by *Tej K. Bhatia*

One of the fascinating aspects of the Indic scripts is the conceptual framework which set the stage for their representation and organization. The history of the Indic scripts follows the path which can best be characterized as diametrically opposed to a path followed by Semitic and other writing systems of the world. While ancient India was making astonishing advances in grammatical theory, phonetic sciences, metric and other area of linguistic sciences, it is puzzling to find a lag of centuries before the emergence of the Brāhmī script. Why so slow?

In addition to answering this question by pinpointing the unique socio-cultural context responsible for the origin and the development of the Brāhmī script and its variants, this book attempts to answer a number of other complex cognitive and linguistic issues pertaining to the decipherment, representation, classification (typological and others), organization, and acquisition of the Indic scripts.

The book is one-of-its-own kind as it attempts to integrate two main approaches – palaeographic and linguistic. It is refreshing and a welcome departure from the research paradigm which has been overwhelmingly palaeographic and historical in nature. Broadly speaking, six Chapters (1–3, 8, 10–12) fall in the category of linguistic approach, while the remaining six chapters deal with the palaeographic and historical aspects of the Brāhmī script and its derivative varieties.

The concept of *Akşara* which forms the basis of an orthographic unit in Indic scripts is the central theme of four excellent chapters (1, 10, 11 and 12). The notion of *Akşara* is so powerful and compelling that it found its currency not only in several Brāhmī -derived scripts, used for diverse language

families such as Indo-Aryan, Dravidian, Austro-Asiatic and Tibeto-Burman languages of South and South-East Asia, but also continues to find its validity in the modern current phonetic and phonological theory. Kapoor (Chapter 1) traces the concept of *Akşara* from the Vedas to Sanskrit linguistic works on grammar, phonetics and Indian philosophy. Patel (Chapter 10) provides further evidence from *Pratishākhyas* on the evolution of the concept of *Akaşara* which was originally conceived for Sanskrit. In chapter 11 and 12, Pandey convincingly argues for positing *Akşara* as a minimal speech unit and for the generative and phonetic-phonological aspects of Brāhmī. It is this generative nature, according to him, which gives Brāhmī its 'universal' characteristic.

Chapters 2 and 3 by Salomon follow a historical and typological approach to the Indic family scripts. Chapter 2 attempts to answer the question raised earlier, namely why Indians were so slow to develop writing systems. Contrary to the wide-spread western conception of a writing system as a basis for literacy and intellectual endeavors, India favors orality or memory over writing. That Panini did not use writing in composing the *Aşṭādhyāyī*, it does not necessarily follow from this that 'he was illiterate' (p. 22). Salomon goes on to argue that writing played a distinct socio-cultural role in ancient India as compared to any other ancient civilization; writing was not seen critical to scientific and intellectual pursuits. The suggestion of William Bright is also instructive in this context. Despite the fact that ancient Indians showed a clear preference for orality over writing, it is inconceivable that in a wealthy empire such as that of the Nandas, royal accounts were maintained exclusively through memory aids, there must be a tradition of business scripts, even if a peripheral one (for more on the business scripts of Punjab, see Bhatia, T. 2003. *Gurmukhi script and other writing systems of Punjabi: History, structure and identity*. In Bhaskarao (2003), pp. 181–213.). In chapter 3, Salomon classifies Indic scripts as 'alphsyllabic' or 'abugida' type, which is distinct and separate from alphabetic and true syllabic scripts. He notes that the pattern of alphsyllabic scripts, while it is dominant in India, it is rare in other parts of the world. While scripts such as Meriotic and Ethopic share the characteristic of the inherent vowel *schwa*; however, they ignore the problem of *schwa* deletion. 'Only in the Indic scripts do we find special mechanisms to explicitly and distinctively mark the absence of a vowel, namely the formation of consonantal conjuncts and, in limited cases, the use of vowel cancellation marker.' (p.42). This refinement can be credited to the advances in phonetic sciences in ancient India

cient India (for details: W. S. Allen, 1955. *Phonetics in ancient India*. London: Oxford University Press).

The paleographic/historical studies address the issue of the deciphering of the Indus Valley script and methodology employed for the dating and the classification of Brāhmī and the variants of the Indic script family. In Chapter 4, based on the chronological evidence drawn from Kharosti and Brāhmī, Falk claims that the Indic writing system must have evolved during the first year of Ashoka's rule. This position is challenged by Kak who claims that Brāhmī evolved from its parent form, the Saravati writing system. His conclusion is based on the comparative graphemic analysis of Brāhmī and Sarasvatī, in addition to the evidence drawn from recent archaeological excavations and the discovery of the Rigvedic astronomical code. Based on a comparative-historical methods, Rajgor (Chapter 8) posits five stages of Brāhmī – Harapan, Proto- Brāhmī, Pre-Mauryan Mauryan and Post-Mauryan. Contributions from Singh (Chapter 6), Mukherjee (Chapter 7) and Glass (Chapter 10) best exemplifies the use of paleography methods to resolve the issues of decipherment, and chronology.

This book represents a new milestone in research in Indic languages by bringing together the paleographic, historical, computational, linguistic and psycholinguistic dimensions of Indic writing system. It forges a strong link to the new phase of research initiated by the project on the writing systems of South and South East Asia by the Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies (for details, P. Bhaskarao (ed.), 2003. *Working Papers of International Symposium on Indic Scripts: Past and Future*. Tokyo: ILCCA.). In addition to the integration of research presented in this work, I hope it will inspire a new generation of theoretical and empirical research on the Indic scripts.

Kashi Wali – Marathi. A study in comparative South Asian languages.

Delhi: Indian Institute of Language Studies.
2006. xii + 323 pp. Rs. 700 / US $ 40. ISBN 81-86323-24-4.

Reviewed by *Alice Davison*

This volume brings together a collection of papers by Kashi Wali, representing work which she has done on Marathi and related languages since 1972. The articles and conference papers were published in many different journals, edited volumes and working papers, and this volume does a great service to linguists interested in getting an insightful and nuanced overview of the syntax and morphology of Marathi. From the papers written at different times and from different theoretical perspectives, some very important and interesting generalizations emerge, which I will discuss in more detail below. A language like Marathi poses some difficult challenges to commonly held generalizations, particularly those embodied in various versions of relational grammar and the evolving statement of the Chomskyan view of grammar. It is not that Marathi is an unusual or esoteric language. Rather, its patterns, though superficially resembling the syntax of Japanese or English in certain respects, on closer examination prove to have somewhat different properties. Each of the chapters clearly examines some accepted view and offers a detailed critique of its effectiveness in capturing the generalizations supported by the data of Marathi. For the most part, Dr. Wali is able to offer a restatement introducing some kind of parametric variation. Here I give a summary of some of the important questions which Dr. Wali explores in this collection.

The issues which these articles discuss are fundamental questions of clause structure, grammatical functions, case and agreement, operator scope and coindexing requirements. One of the major topics which Dr. Wali has been exploring since her 1976 dissertation (and even earlier papers) is the nature of reflexive pronouns in Marathi.

One of the most central topics in Dr. Wali's work is the nature of anaphors in Marathi. The antecedent-anaphor relation is not just structurally

based, but also lexically constrained. Contrary to what early generative work predicted, there are actually two reflexives in Marathi, *swataah*, which is locally bound within a (finite) clause, and *aapaN*, which may be bound across (finite) clause boundaries. Both must have c-commanding subject antecedents, yield only sloppy identity readings in coordinated sentences, and may not have split antecedents – these are defining properties which are typical of anaphors in many languages. Both reflexives contrast with pronouns, which may have non-commanding, non-subject and split antecedents. The two lexical anaphors are most clearly distinct when arguments are coindexed, and finite clause boundaries are involved. If the internal clause is non-finite, or a small clause, the *swataah* form may have both a local and a long-distance antecedent, like Hindi/Urdu *apnee-*. Both forms are possible if non-arguments (possessives, non subcategorized adjuncts) are coindexed with a local subject. This fact challenges the notion that the finite clause as domain of binding must be qualified to take into account the grammatical function of the anaphor as well as the antecedent.

The antecedent must be a (c-commanding) subject, defined independently of case and agreement. The category of subject in finite simplex clauses includes DPs with nominative, ergative and dative case marking, which behave similarly. For example, various participles with controlled null subjects require a subject antecedent in the matrix clause, regardless of case. Subjects with this range of case marking are antecedents of both lexical anaphors, as well as participles with controlled PRO subjects requiring a subject controller; the different properties of aspectual and conjunctive participles are explored at length in two of the chapters; see the summary of subject properties on p. 123.

The theoretical puzzles increase when relation-changing processes or constructions are involved, internally to a simplex clause, such as passive or causative formation ('Twin passives, inversion and multistratalism in Marathi', 'Oblique causee and the passive explanation', 'Two Marathi reflexives and the structure of causatives', 'Cause, causer, and causee'). In passive sentences, both the promoted direct object and the demoted subject, marked with the oblique *kaDuN* by, are possible antecedents of the local reflexive *swataah*, as might be expected in a monoclausal sentence. But the oblique subject may bind the non-local anaphor *aapaN* as well, suggesting that it some sense it is an adjunct, like a possessive or locative which allows the non-local anaphor, or that passive sentences have some properties of a non-local sub-domain. In Exceptional Case Marking/Raising to Object sen-

tences, the matrix subject binds both the local *swataah* and the non-local *aapaN* as subjects of the embedded clause, again showing that this construction is syntactically ambiguous between a mono-clausal and a bi-clausal structure.

In causative sentences, there is a different set of binding restrictions. Causatives have long been assumed to consist of a combination of two argument structure, a transitive or intransitive clause embedded below a transitive causative verb. In all cases in Marathi the matrix (causer) subject binds an anaphor. In Marathi, the transitivity of the embedded predicate determines binding possibilities. Causatives formed from intransitive verbs allow the causee to be *only* the local reflexive *swataah*, and not *aapaN*, unlike the Raising construction which is assumed to be structurally very similar to it in other languages. Causatives formed from transitive complements reject both reflexive forms in the intermediate causee, which in turn cannot bind the object of the transitive causative.

Wali's discussion of these problems within the context of different theoretical proposals is most interesting. She shows with detailed argumentation that these facts fail to be captured by cyclic transformations, and by other later standard analyses. In the paper co-authored with Carol Rosen, she shows what account can be given of the two passive constructions using the Relational Grammar notions of stratum and successive shifts of grammatical function. In the most recent paper, she discusses the proposal by Ura to think of grammatical functions such as subject not as holistic entities meeting a single universal set of criteria, but rather as bundles of construal properties which may be distributed over different arguments, allowing distinctions of constructions such as passive, Raising and causative sentences.

Clause structure and possible scope relations are the topic of three other papers. In 'A note on wh-questions in Marathi and Kashmiri, Dr. Wali contrasts the scope marking devices in Marathi and Kashmiri for wh questions of the type 'Who do you think that they saw __?'. Both languages have wh-in situ in the finite clause. Wide scope is marked in Kashmiri, as in Hindi/Urdu, with a matrix clause scope marker *kah* what?. Marathi does not have this option; instead in wide scope questions, the embedded clause is made transparent to wh-scope by being marked with suffixes. It can be the quotative *mhanun* so, having said, or a demonstrative, the choice of which is determined by the matrix verb. It can be a definite determiner *te,* an indefinite *asa* such. Factive matrix predicates allow *te* but not *asa*; propositional predicates allow *asa* but not *te.* Factives disallow narrow scope questions.

The nature of factivity is further explored in ‘Event’, co-authored with Philip Peterson.

In Negation-dependent idioms Dr. Wali argues that the modal verbs in Marathi have full non-finite clausal complements. They resemble other matrix verbs like try and want, which select either an infinitive or finite complement. Adverbs and negation may associate in two ways, with the modal/ matrix verb or the complement, as shown by word order and semantic interpretation. Marathi has negation-dependent idioms, a useful list of which is included in the paper. Negation may associate with these idioms across (several) non-finite clause boundaries, but association is blocked by finite complements.

A brief and speculative early paper on the relative clause constructions in Marathi contrasts two possible basic structures, a NP adjoined structure from which the relative clause may be moved, or base adjunction. While both analyses confront serious problems, the adjunction analysis seems to account in a better way for the range of relative properties, including multiple relatives.

Another set of papers deals with case, grammatical function and agreement in Marathi (‘Long shadows of ergativity in Kashmiri and Marathi’ (with Omkar N. Koul); ‘Non-nominative subjects in Marathi’, ‘On distinguishing AGR from agr: evidence from South Asia’ (with James Gair). There are three central problems: (1) ergative and dative subjects, which do not control agreement, even if the ergative case is not overtly realized in Marathi), (2) null-cased objects which, like nominative subjects, do control agreement, and (3) multiple agreement in Marathi, which has an invariant additional agreement for transitive 2nd person subjects. In Marathi, like Kashmiri and Hindi/Urdu, finite perfective aspect is required for ergative subject marking, and in Marathi subjunctive clauses have ergative marking. Otherwise subjects of non-experiencer verbs are nominative; in ‘Ergativity’, there is useful and important discussion of the three categories in Kashmiri of intransitive verbs, one of which has ergative subjects. In comparing different agreement facts in two varieties of Sinhala, Hindi and Marathi Wali and Gair propose that what licenses unmarked subjects is not the same functional projection which licenses null-cased objects. A language may have one or the other or both. These issues of case licensing have persisted and are the focus of much current theoretical work.

The papers in this volume are distinguished both by the perceptive descriptive accounts of South Asian languages, and the carefully worked out

analytical perspectives. There is a wealth of concrete information about Marathi and other languages, which is shown by complex argumentation to challenge various general theoretical assumptions. The volume bringing together these papers on related topics allows the reader to get a much more comprehensive overall impression of the linguistic properties of Marathi and other languages than if the papers were read separately. Some of them appeared in sources which are now difficult to obtain. The bibliography contains references to work which is not well known outside of India, such as dissertations, and studies in Marathi. While some of the papers are rather complex in organization, the reader is well rewarded by the wealth of data and the overall picture of how these issues play out in syntactic theory from early generative grammar to current versions of the Minimalist Program.

The only problematic feature of this volume is that the composition of the volume was not done with enough attention to formatting, fonts and proofreading. There are typographical errors and sudden font shifts, as well as misaligned glosses. Since the papers are so well organized, for the most part, there is no loss of intelligibility, but occasionally the reader is forced to reconstruct the intended text or glosses.

Dialogue

Whole Word Morphology and its Indian Critics*

Rajendra Singh

The purpose of this intervention is to open a debate, particularly in this annual, regarding the nature and structure of morphology. I shall attempt to do so by answering questions about and objections against Whole Word Morphology (WWM), the theory of morphology outlined in Ford et al. (1997) and further substantiated in Singh & Ford (2000), Singh & Dasgupta (1999), and several other related publications, that have been raised by Indian linguists. I hope the remarks that follow will revive discussion of morphological matters in South Asia, particularly in India, which has a very long tradition of morphological inquiry. I shall focus on the objections that stem from the construal of morphology as a search for recurrent partials (cf. Amritavalli 1998), the misconceived 'computational explosion' argument (cf. Sengupta 1998), the commonly voiced neo-Paninian objection that WWM is too weak, and the standard formalist claim that morphonological, *dyotaka* modifications should be treated in phonology. The answers to these objections will, I hope, also show that in addition to providing an adequate account of the sorts of morphological systems that exist in India, as already demonstrated in Singh & Agnihotri (1997), Singh & Ford (2000) and Singh (2001 & 2003), WWM helps us come to terms with several difficulties faced by one or two tier (cf. Mohanan & Mohanan 1999) morphous or weakly amorphous theories of morphology such as Anderson's. These difficulties include the well-known cases of syncretism in Sanskrit and other Indian languages as well as the ones pointed out by Tirumalesh (1997 – *contra* Aronoff & Sridhar's 1988 attempt to justify a distinction between derivation and compounding in Kannada) and Dasgupta (1989 – against the traditional analysis of the constant elements in so-called complex verbs in Bangla and other Indic languages). They will also reveal why the use of variables and constants in WWM does NOT, as some have suggested, tantamount to a synchronic recognition of what WWM considers to be quintessentially diachronic, etymological constructs such as roots (cf. Dasgupta, Ford & Singh 2000) and why a multipartite analysis of morphological com-

plexity, which gives rise to contradictions like ‘lightweight roots’ and ‘heavy-weight, word-like affixes’.

I shall conclude with some reflections on why the (immanent and ancient) Indian critique of Paninian morphology, which culminates in Bhartrihari and provides several substantial arguments in favour of WWM, has been ignored both in India and the West, and on why, given Chomsky’s (1995) repeated endorsement of Jespersen’s (1924) assessment of the possibility of Universal Morphology, morphologists still committed to sub-lexical units may wish to revisit these arguments.

Let me begin with the theory of WWM It is given in (A), (B) and (1) below, taken, with minor modifications, from Singh & Ford 2000. (A) is the preamble and (B) the post-script to (1), which has the brevity it does because we believe that nothing more of substance can really be said about morphology (or, alternatively, what more is said can be shown to be either pedagogical, as (A) and (B) in fact are, or indefensible (cf. Singh 2000).

(A) 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 (1) below. We refer to these instantiations as W(ord) F(ormation) S(trategies) because as generalizations drawn from known particular facts, they can be activated in the production and understanding of new words (cf. Ford & Singh 1991 and Ford, Singh & Martohardjono 1997) WFS’s must be formulated as generally as possible, but, and this is crucial, only as generally as the facts of the matter permit.

(1) /X/a ⟷ /X’/b

where:

a. /X/a and /X’/b are words and X and X’ are abbreviations of the forms of classes of words belonging to categories a and b (with which specific words belonging to the right category can be unified or on to which they can be mapped).

b. ‘represents (all the) form-related differences North of automatic phonology between /X/ and /X’/

c. a and b are major lexical categories[1]

d. the ⟷ represents a bidirectional implication (if X, then X’ and if X’, then X)

e. /X'/b is interpretable as a semantic function of the interpretation of /X/a and vice versa.

f. ' can be null iff a =/= b

(B) It should be obvious that we consider 'morphophonology' (or its more fashionable *avatara* 'lexical phonology') to be an integral part of morphology (cf. b above) and we do not think that either intra-linguistic (inflections vs. derivation, affixation vs. compounding etc.) or inter-linguistic (flectional, isolating etc.) morphological diversity can affect (1) in any fundamental way. (1) offers a unified account of what have sometimes been seen as different types of morphologies. The diversity that exists can be read off the system of strategies that instantiate (1) above, but it does not need to be expressed as a difference in type. As any word can be exhaustively parsed into non-overlapping constant and variable subcomponents with respect to some particular WFS, (1) encapsulates our rejection of multipartite analysis of words and denies any theoretical status to units such as 'morpheme' 'root' 'stem', and 'suffix'[2]. The constant sub-component is always specified in the WFS and has no autonomous status outside the WFS in which it figures. Sometimes some of the variable sub-component needs to be specified on both sides of the arrow, but being specified on both sides is precisely what distinguishes it from the constant sub-component, which, for obvious reasons, is specified on only one side. As all morphological relationships can be expressed by strategies instantiating (1), morphology has little or no architecture and, to change the metaphor, no traffic rules (such as *krt* (level 1 affixes) before *taddhita* (level 2 affixes)). Representations of the speaker's knowledge of the patterns of morphological relatedness in her language, Morphological Strategies (= instantiations of (1)) are invoked only in moments of crisis, i.e., when the speaker needs to analyze or fashion a word she needs for the purpose at hand, often to meet a syntactically enforced requirement.Their exploitation, of course, helps her to bridge the gap between the actual words she happens to know and the possible words she can be said to know – actually their existence makes the known merely a subset of the knowable. When they ARE invoked to produce what will become words, their "outputs" are seamless wholes, with no brackets, boundaries, or a-cyclic graph fragments in them. They are not there to be deleted; they are just not there. WFS's certainly do not supply them. And neither the strategies nor their "outputs" have any syntactic constituency relationships marked in them

in any fashion whatsoever. In both the active and the passive mode, they license the words a speaker has or may come up with (in the "on line" mode).

Having presented the theory above, I shall first take up Amritavalli's review of Ford *et al* 1997, a book-length attempt to motivate the departures enshrined in (1). It is an elegant and interesting review. Although she sees some of the advantages of "relocating morphology 'a step upward' from the 'morpheme" and, from her point of view, of the potential of WWM for what she calls "shaping," she is understandably reluctant to give up most of what generative phonology and morphology have popularized during the last half century. That it is her attachment to these popular generative beliefs – anyone not so impressed by the generative tradition in phonology and morphology as she seems to be could not really say, as she does, that we have a "crucial argument for 'morphonology' over conventional morphophonology" for we actually argue against anything in between phonology and morphology – that is responsible for her very first question. She asks why we do not include Aronoff (1976) in our list of advocates of word-based morphology before us. As the answer to her question is quite straightforward, I shall provide it without much fan-fare: Aronoff's theory, despite the label the paradigm Amritavalli seems most familiar with confers on it, is NOT really a word-based but a stem-based theory of morphology. The fact that the distinction stem/word is not clear in English is no reason to let Orwell take over Plato.

As for her criticisms, I shall first take up the one that stems from her belief, shared by many, that whereas we are concerned with description, the more visible, neo-Paninian paradigm(s) actually explains certain things in morphology. I would like to submit that if what needs to be described can't even be described, there is really nothing to explain, and that a representational *tour de force* is NOT an explanation. Consider the case of morphonology. Our descrptivist obsession, she claims, makes us insensitive to economy and explanation. Here what needs to be described is a set of non-automatic alternations inextricably implied in a set of morphological operations and what needs to be explained is the fact that 'morphonological rules' do not generalize to become across the board phonological rules. There is nothing to explain about the fact that there are subtle representational differences, perhaps of no empirical consequence, amongst various generative accounts of umlaut in German or of aspiration throw-back in Sanskrit (cf. Joseph & Janda 2002). In accusing us of an excessive concern with descriptive exhaustiveness, she is really only revealing that she too

takes representational ingenuity to be explanation. We are, of course, interested in explanation, but we insist on explaining what needs to be explained and do not take 'explanatory adequacy' to be a theory-internal game. Our explanation for the diachronic non-generalizability of socalled morphophonology is quite straight-forward: it does not generalize because it is a part of morphology, and no part of morphology, affixal or non-affixal, ever does so (cf. Comrie 2001). I would like to submit that what Amritavalli's preferred paradigm has offered over the years are impressive ways of hiding what in the final analysis cannot be hidden – the fact that 'morphophonology' is NOT a part of phonology. It walks like morphology, acts like morphology, and IS morphology.

Although I understand her desire to have a unified treatment of what she sees as the complementary distribution of voicing in Enlish plurals like *knives* and in what she calls 'denominal verbs' – a derivationalist term which may in fact be psychologically misleading and diachronically incorrect in some cases – like *knifes,* it is not clear if it indeed is what she sees it as, and it is equally unclear why anything other than avoidance of homonymy is needed to explain what she wants explained. Although in the examples she discusses the same class of phonemes is required by two morphological rules to undergo voicing, the 'outputs' are kept lexically distinct. Such structural complemetarity as exists is between the general rule that specifies nothing in particular and the rule that specifies the class of phonemes in question in each case. Although I would be perfectly willing to listen to any formulation that may 'advance the cause', and Amritavalli offers none, I do not feel compelled to meet the demand for something more structural than avoidance of homonymy as a unified 'explanation' for the bits of phonic substance involved in disparate morphological processes at the cost of explaining diachrony.

As for her 'nagging doubt' that we make too much of 'a rather restricted set of data', I wonder if in allowing herself to be guided by the assumptions that she rightly points out have guided much current work in morphology she too has missed the point that there is a systematic ambiguity in the literature she seems familiar with between various interpretations of what she calls 'the lexicon'[3]. Is it really hard to find dialects of English or other languages in which the forms NOT listed in the dictionaries linguists look up when they make their lists actually show up? We have never denied that some rules of morphology are more productive than other rules of morphology, only that this distinction should not, even if it can be, captured by

postulating two different kinds of rules because at the level of potentiality, both in fact have exactly the same status. Unless we make THAT assumption, we cannot explain why some Americans are as happy with *mongooses* as other Americans are with *mongeese* or why *musikism* seems as elegant to some of them as *musicism* does to others. The point about morphology that seems to have escaped contemporary structuralism, in which I include generativism despite its rhetoric to the contrary, has to do with the fact that it is perfectly possible to come up with identical rules of morphology from different I-lexica (or should I say lexicons?) and it is this similarity in morphological rules that allows one to see that the speaker who says *geese* and *mongeese* is, orphologically speaking, not doing anything very different from the speaker who says *geese* and *mongooses*. The existence of doublets, particularly in individual idiolects, makes the same point even more dramatically. Yes, there are productivity differences, but it is not clear if they should be built into grammatical rules. It is obvious that the rule that says that a singular noun terminating in *f* is related to a plural noun in which a *v* appears in its place is *ipso facto* less productive than a rule that says that a singular noun, any singular noun, is related to a plural noun which can be parsed as containing that singular noun and 'the plural marker' *z*. The reason that the latter is more productive is actually quite simple: the singular nouns that have or 'take' a *z*-plural have nothing in common (in any imaginable I-lexicon), and it is this lack of sharedness that allows speakers to generalize this rule almost mindlessly, as it were.

As Amritavalli seems to share the standard generative view that WWM is too unconstrained, I shall take that objection up a little later. For now, I want to comment on her desire to have sound-shape regularities of the sort one finds in English *when-then,what-that,* and *where-there* accounted for in morphology. I'm afraid she is being too charitable in giving us credit for allowing "this kind of morphology to happen". Given our insistence that the ' or the difference between X and X' be systematically exploited in a language, we can't allow it, and I am not sure if anybody should, but perhaps these things, like Hindi *puut-kapuut* 'son-illbegotten son', can potentially lead to WFS's. One must, however, be aware of the fact that the search for recurrent partials has been the bane of morphology, and it is time to realize that morphology, like anything else in language, is a matter of contrasts and their systematic exploitation. In denying any theoretical status to productivity, we are simply confirming that anything available for lexical contrast is in principle available for morphological exploitation. And this

provision, perhaps to the horror of generativists, includes not only classes of phonemes distinguished by features such as voicing (cf. English (*house*)N vs. (*house*)V) but also stress (cf. English (*protest*)N vs (*protest*)V, squarely a matter of English morphology (*pace SPE* and ALL its descendants). I think we do need to study what she refers to as shaping and quite possibly something that could be called lexology, but morphology may well have to be left to deal only with systematic exploitation of lexical contrasts (cf. Neuvel & Singh 2001).

Like Amritavalli, many people seem unhappy with our proposal because they say that they suspect that it may be the case that any language data could be accounted for by WFS's not because the theory is valid but because it is weak – too weak to be falsifiable. For reasons that are not clear to me, they also seem to think that WFS's, presumably unlike other mechanisms that have been proposed in morphology, create morphology: The objection seems to run as follows: "WFS's require morphology as a separate level or component of grammar. There does not seem to be any other evidence for it besides the fact that it is a requirement for WFS's". It seems to me that although well-formedness judgements are admitedly more uncertain in morphology than in phonology or syntax, the existence of morphology as a separate component of grammatical competence cannot really be doubted any more.

As for the other objection many share with Amritavalli, it is, for methodological reasons, important to point out that in formulating it one routinely adds the hommage to theory construction we have learnt to add to anything. I agree that any theory should allow us to imagine something that is not possible. We have always insisted that the quintessential difference between morphology and syntax is that the former does not allow movement. Given this constraint, it should not be too hard to find unarguable examples of movement in morphology, if they exist. It is important to point out that I am not talking about examples of the sort that have sometimes been analysed as involving movement but about examples that anybody would have to analyse that way (cf. the remarks on incorporation' below). Why aren't there, for example, any rules anywhere that take nouns ending in some particular substring and make verbs or adjectives out of them by merely shuffling that particular ending to the front of the nouns? A rule that could give us *itybanal* from *banality*, for example. Why, in other words, must a constant specified at the end of a word (as suffix, in neo-Paninian terms) never show up at the beginning of a word (as a prefix, as they say)?

There are actually very straightforward constraints built into our theory – they are actually constitutive of it. Consider, for example, the implications of the fact that we mention only two variables in (1), one on each side of the arrow. Having only one variable on each side of the arrow clearly implies that neither pole of a strategy may contain more than one variable, and that unambiguously says that no word-formation is possible without lexical subordination. Singh & Dasgupta (1999) provides evidence to show that even so-called compounding is subject to this overarching implication: one of the elements in socalled compounds is necessarily subordinated, even when it carries no mark of its *samasagata* status (cf. Hindi *ghoDa – ghuD*, as in *ghuDsavAr* 'rider' but not in *ghoDAgADi* 'horse-carriage'). Our analysis in fact confirms Paul's (1886) insight of more than a century ago that there is no sharp dividing line between derivation and compounding, an insight recently reaffirmed for German by Becker (2000). It allows us to understand why Tirumalesh rightly feels very uncomfortable with Aronoff & Sridhar's attempt to draw such a line for Kannada. He shows quite convincingly that the Kannada prefixes "discovered" by Aronoff & Sridhar are actually forms that obey the overarching two-mora constraint on Kannada words and are related, by truncation, to full words That is why old Kannada grammarians such as Keshiraja (11th century) treated Aronoff & Sridhar's prefixation as part of compounding and not because Kannada 'prefixes' had escaped their attention. Notice that our view allows us to say all this straightforwardly (avoiding even the terminological problems sometimes created by Tirumalesh's insightful resurrection of the older analysis). It also saves us from not even attempting to find a pragmatic explanation for the fact that whereas *doctor-appointment* is impossible in English, its word-level counterpart is impeccable in German. The standard characterization of compounds since Panini puts too heavy a burden on pragmatics, a burden it cannot handle. So-called 'compound-verbs' in South Asian languages show exactly the same thing, as has been known at least since Dasgupta (1988).

It is perhaps necessary to say something about the other case that allegedly involves two variables: 'incorporation'. As the languages analysed by scholars like Baker and Sadock fall outside the philological scope of this paper and as Ford and I have already spelled out the grounds for which their accounts of what they call 'incorporation' must be rejected (cf. Ford & Singh 1997), here I shall simply say that they seem not particularly interested in the fact that the 'incorporated form' is almost never the same as the 'non-incorporated form' and respond only to T. Mohanan's (1995) account

of what she refers to as 'incorporation' in Hindi. Her paper is a very unusual one. In my view, she is able to talk about 'incorporation in Hindi' partly because her lexicon doesn't contain some very ordinary words of Hindi. A substantial part of her evidence for 'incorporation' in Hindi relies on the absence of very ordinary Hindi words like *zild-sAz* 'book-binder', *lakaDhAra* 'lumberjack', *khAnsAma* 'cook', and *ghosi* 'supplier of milk', words that sometimes don't even contain what Paninians would call a suffix, let alone an NV structure. I wonder how many speakers of English say things like *He food-cooked in Montreal for a decade* instead of *He was a cook in Montreal for a decade* or even *He cut wood/wood-cut in British Columbia when he was young* instead of *He was a lumberjack in British Columbia when he was young.* The absence of such words in the lexicon of someone who claims to be a speaker of Hindi is somewhat intriguing, but I suppose I can't really quarrel with the claim that a language belongs to every one who claims to speak it. In thanking me for making her 'aware of the extent to which speaker judgements can differ' (p. 75), something I did when she was writing the paper, she very gracefully acknowledges the difference between her Hindi and mine She doesn't, however, tell the reader what these differences are based on. The only inference one can draw from her analysis is that I speak a variety of Hindi that does not obey the OCP! Given Bhatt's demonstration (2001) that the modality effects of the sort she is talking about can and should be derived without invoking incorporation for Hindi, it is perhaps best to leave the matter here.

The point here, of course, is that it is incorrect to say that our theory is too unconstrained. It is quite easy to imagine morphologies that we cannot describe. The fact that we can non-tortuously describe the morphological systems our critics seem to be familiar with should come as a relief and as a challenge to their somewhat unusual notion of 'explanation'.

Let me now turn to the 'computational explosion' argument. Although it is perhaps better to leave – to give Amritavalli her due – post-1976 publications that define the morpheme as 'a minimal meaningful sequence of letters or speech sounds' alone, I must, I'm afraid, take up Sengupta (1997) because it offers the Indian version of the computational explosion argument. It not only shows a clear preference for morpheme-based concatenative morphology but also takes it to be all of morphology, obviously ignoring what its author's tradition refers to as non-concatenative morphology. In comparing three models of morphological processing, Sengupta finds fault with some of the basic claims of WWM, in particular with its insistence on

listing of whole words and its rejection of constructs such as 'morpheme', 'root,' or 'stem' and of hierarchical structure within the word. These assumptions are justified and used in the source cited by him, Singh & Agnihotri (1997), a descriptive monograph devoted to applying WWM to Modern Hindi. As I am neither interested in nor have anything to say about computational linguistics, I shall concentrate on what Sengupta calls 'psychological plausibility'. It is interesting to note that he rejects what he calls the list model not on the grounds that it is in principle impossible to make a list of all the words of a language, the grounds on which WE reject any model that would require such a list, but on the grounds that if one were to make such a list – I would certainly like to know how one can – one would end up with huge lists, like the one that would have to contain 36,00,000 forms for 300 Finnish verbs. He notes: "Perhaps it is posssible to list all the words of English with their full inflected forms in a file. Yet the existence of languages like Finnish discredits listing as a model of morphological processing." (3). Although it is not clear why an indefinitely large number of words is easier to list than a fixed multiple of them, the argument about hugeness is a familiar one, and seems to have become a stock-in-trade for neo-Paninians since Hankamer (1989), who argued in a similar fashion on the basis of Turkish. Again, I repeat I don't know how one can list all the words of English – I don't even know what that expression means – and I certainly don't know anything about computer processing of morphology, but it is clear that no inferences can be drawn from this sort of observation for human processing of morphology. Even if I take Koskennieme's (1985) estimate at face value, under the assumption that it takes two pairs of words to motivate a WFS (cf. Neuvel & Singh 2001), the number of words I need to list for any I-lexicon of Finnish cannot be higher than 12000x the number of verb classes exhibited by Sengupta's 300 Finnish verbs x 4. Let us assume that these Finnish verbs exhibit 5 morphotactically different types. The maximum number of forms I have to list is 240,000 (and not Sengupta's 36,00,000). I'll, of course, have to add a certain number of WFS's to be able to project the remaining 3360000 forms Finnish speakers may come up with. The point is that there is nothing wrong with the computational explosion argument as far as computational linguistics, whatever that is, is concerned, but it has nothing to do with morphology. What we have proposed is that a speaker lists the words she knows (and on the basis of which she forms her WFS's) in their full glory, and to the best of my knowledge no one has ever proposed that speakers list all the words their language can

be said to have, as Sengupta acknowledges when he tells us that his list model is a straw-man. Speakers need to list only twice as many words as the number of strategies they give evidence for using, and in order to count the latter, one must take clustering and ganging of the sort discussed in Singh & Agnihotri (1997) into account. Again, I repeat once again that I am arguing against the ease with which what may be relevant for IBM, HP, or Microsoft is unjustifiably promoted to the status of an argument against full listing, which in linguistics can only mean the non-underspecified listing of the words an individual speaker can be said to have in her I-lexicon. Actually, if there is a psychological plausibility argument anywhere it resides in the fact that nobody has ever offered any evidence for the DE-listing of high frequency words used by speakers to initially come up with the WFS's that eventually allow one to see any actual lexicon as containing only a subset of the words made possible by these WFS's. How many trials and or exposures are needed before a word in entered in an I-lexicon is, of course, an empirical question, but largely unnatural language processing called NLP can't possibly throw any light on it. It is true that very frequent words are very often entered into individual lexcica, but it is not clear that there is a well-defined class of listable/unlistable words. The morphologically simple ones, which are either listed or not known at all, do not, of course, vitiate that claim. The point is that to say that the lexicon contains only words in their full glory is not to even remotely imply that it contains "all the words of a given language", whatever that expression means.

It is interesting to note that despite his appeal to psychological plausibility, Sengupta, in order to, according to him, validate Bloomfield's oft-cited tribute to Panini, would also like to see morphophonolgy analysed as phonology, something we don't allow precisely because it never generalizes historically and because speakers treat it as morphology in every bit of linguistic evidence we have ever seen. We don't question Panini's or Bloomfield's intelligence but only the validity of their analyses, at least for the objectives we all agree grammars must meet. Although from a linguistic point of view both Sengupta's definition of the morpheme and his text-book assumptions about the hierarchical structure of words, for which he provides no evidence, sound somewhat naïve, perhaps they are useful for the enterprise he is in, though morphologizers such as Neuvel's (2002) cast a shadow of doubt even on that possibility. What is perhaps really interesting about Sengupta's exploration is its refusal to come to terms with what I would call the limited full listing model – limited because it does not require every single form of

every word to be listed – just of enough of them to give rise to what his tradition calls word-formation rules – and full listing because it insists that that which is listed is listed in its full glory. At any rate, my point here is that none of his arguments have any validity against WWM, perhaps not even against its computational implementations (cf. Neuvel 2002).

Although he takes pains to demonstrate the inadequacy of some definitions of the root, he doesn't offer us one and doesn't quite tell us how his solution to the root-affix computational problem he discusses will or can handle English words like *receive*, *deceive*, *perceive*, *conceive,* words that in his tradition have received continued attention from Chomsky & Halle (1968) to Anderson (1992). The WWM formulation of the relationship between words like *receive*, *conceive*, *perceive* and *reception*, *conception*, and *perception*, a formulation that does not invoke the notions 'root' or 'suffix' is quite striaghtforward and shows the utter dispensibility of such constructs despite the importance Sengupta gives them in his morphemology. The WFS needed here MUST treat the substrings *ceive* and *ception* as constants (and NOT as variables) and the substrings *re, con*, and *per* as variables (and NOT as constants), making it clear that there is no substance to the objection, raised by several neo-Paninians, that our constants and variables are simply notational variants of Paninian *dhatus*, *angas*, and *pratyayas*. The point is NOT that our variables don't ever correspond to what neo-Paninians call roots and stems – they do in a very large number of cases – but that we see that as synchronically irrelevant. That WWM can do justice to facts that cannot be easily handled by the hierchical battery of sublexical arms that have een traditionally available for centuries – notice the addition of units such as augments, stem-forming suffixes, and affixoids to the standard Paninian battery, not to mention ploys such as Williams' simple assumption, as he calls it, that some TENSE markers in Swahili "are stems, not prefixes" (2000: 235; for some further discussion, see Singh 2001) – ought to, in my view, count as evidence for WWM. The point here is that words blend with other words over time and at any given point there are going to be cases where the Paninian knife can't slice things as neatly as Paninians would like us to believe, and this has been known at least since Kaunda Bhatt (cf. Rathore: 1998). Needless to add that parsing in terms of constants and variables presents no such problems and that the solution to the indeterminacy prblem created by Paninian slicing resides in our view not even in prototypicality theory but in abandoning the whole battery.

Sengupta is, of course, not the only Indian linguist who takes these objects of wonder seriously. Despite the demonstrations by Christdas and Kiparsky, Mohanan (1996), for example, wants us to account for Morpheme Structure Conditions, conditions which have no binding, constitutive import when they cannot be reduced to syllable or word-structure constraints, something Ford and I, following the insightful suggestions of Tranka of more than half a century ago, show in our response to his objections. He also takes constructs like 'root' and 'stem' quite seriously, though has, to the best of my knowledge, never offered any definitions of them. Needless to add that if one takes him to be operating with traditional definitions, he must accept the obligation of showing how the problems associated with these definitions are to be resolved.

Somewhat similar objections appear in Mohanan & Mohanan's (1998) notational translation of Hockett's (1947) distinction morph/morpheme that they offer as their theory of morphology, but before I take it up I must point out that their claim that "non-compositional morphology is lexeme-driven may be true of the other models of non-compositional morphology, it is certainly NOT true of WWM, as we explicitly pointed out as early as the early 80's. We reject the notions of 'base' and 'lexeme' just as much as we reject the notions 'root' and 'stem', the latter for the simple reason that it is a semantic notion that assumes a distinction between derivational and inflectional morphology, a distinction that cannot be sustained.

Given that they claim to have the "strong intuition that *went* and *picked* are parallel in their compositionality", it is understandably difficult for Mohanan & Mohanan to objectively comparae compositional and non-compositional morphology, something they claim to be doing in their paper. The fact that they assume that compositionality entails hierarchical organization – the assumed application of X-bar throughout begs the question, obviously – is of the same cloth. Be that as it may, let me turn to what they see as arguments. As they seem to be aware themselves, no appeal to Paninian constructs of the sort they use is required to get the right interpretations for nonce forms like re*blist* and *slarken* for they can be obtained from the interpretive rules associated with the WFS's given below:

/X/ ⟷ /riX/

/X/ ⟷ /Xen/

Actually, it is not difficult to see that both their observation that the Malyalam sequence /in/ "is part of neither case ending nor the stem, but is required as the glue that joins them," and their conditions such as singular nominatives ending in /n/ and plural nominatives ending in *kaL* and *maar* as well as feminines ending in *aL*, the non-nominative forms simply add the appropriate case-ending can and must be incorporated directly into WFS's for Malyalam case morphology. You see these conditions capture systematic correspondences between and among words, including syncretism, and were used by us in the early eighties to argue that they in fact constituted a particularly strong argument in favour of a word-based morphology because otherwsie one is forced to put them in what some have called 'meta-grammar'. It is, of course, a pity that Mohanan & Mohanan do not, preoccupied as they are with representational matters, even make an effort to see if their conditions can be integrated into what late Wurzel called Paradigm Structure Conditions (PSC), a worthy challenge to our morphology *sans* PSC's, though I obviously cannot address that issue here. It is also a pity that they do not show precisely how Semitic facts of the matter as analysed within their tradition are to be represented in their representational translation of Hockett. Needless to add that what they see as "additional complications" is also a result of the fact that they insist on looking at morphology as isolating 'recurrent partials' and arranging them in a preferably multidimensional tree rather than as linearly represented word-formation.

Unsurprisingly, the objections that have been raised against WWM by Indian linguists in print are only thinly disguised paraphrases of what is available in English, perhaps a predictable consequence of the fact that those that have raised them were actually trained in Anglo-America. I looked very hard to see if some objections had been raised in writings in other Indian languages, but did not find any objections or endorsements there, perhaps predictably a consequence of the fact that I have not written anything on these matters in any of those languages, clearly MY fault. I cannot, however, help wonder if the Indian familiarity with the Indian grammatical traditions is restricted to *obiter dicta* such as Bloomfield's comments about Panini. Perhaps now that Chomsky (1995: 20–21) is willing, despite his ambivalence towards a strong lexicalist position, to entertain the idea that the lexicon may contain redundancy rules and to insist on the correctness of Jespersen's (1924) dismissal of Universal Morphology, both Indians and non-Indians would find reason to not only seriously look at proposals such as WWM but also at the arguments against sub-lexical units and morpho-

logical architecture furnished by the Greco-Roman tradition and by the immanent critique of Panini, from Katyayana to Bhartrihari, in India (for an interesting and useful recent exploration of that sort, see Dasgupta 2007). Given that Panini and the Paninians of Ancient India do not have a general term to cover the range of parts they get their words from – the term 'morpheme' is a 19th century European invention – the revisit is bound to be rewarding. It is bound to take them beyond what, following Stampe (1999), might be called the "more what syndrome" they seem to exhibit now. When that happens, morphological matters would be taken up not by syntacticians, phonologists, and pragmaticists, who apparently think they can handle it with, to use an Indianism, their left hand, but by those who have seriously explored morphology. I shall look forward to that day.

Notes

* An earlier version of this paper appeared in *The J. D. Singh Festschrift*, published by the Linguistic Society of India. I am grateful to Alan Ford, Probal Dasgupta, Wolfgang Dressler, and Sylvain Neuvel for several useful discussions regarding the nature and structure of morphology. The usual disclaimers apply in full force.

1. In fairness to the Indian critics of WWM, I must point out that the formulations to which they were reacting did not contain the qualifying expression 'major lexical'. As no one made an issue of the quintessentially 'non-morphological' (non-wordformational?) nature of minor lexical categories, nothing in the argumentation is affected.
2. It is interesting to point out that Embick and Halle (2005) find that admitting stems in morphology is "un-necessary and problematic".
3. It is worth pointing out that it is this ambiguity that allows the generative tradition in morphology to postulate and hold on to what it calls Blocking. The fact that most middle-class speakers are familiar with most high frequency words cannot really be used to 'block' the output of a rule of word formation for someone who does not know the word most others can be said to know.

References

Anderson, Stephen A.
1992 *A-morphous morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark
1976 *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark & S. N. Sridhar
1987 Morphological levels in English and Kannada. Reprinted in *Rules and the Lexicon,* Edmond Gussman (ed.), 9–22. Lubin, Poland: Catholic University.
Amritavalli, Raghavachari
1999 Review of Ford et al., "Pace Panini: Towards a word-based theory of Morphology." *Indian Linguistics* 59: 155–161.
Becker, Thomas
2000 On the non-hierarchical structure of compounds. In *The Yearbook of South Asian Languages and Linguistics*, Rajendra Singh (ed.), 283–292. Sage: Thousand Oaks.
Bhatt, Rajesh
2001 Incorporation, word order, and long distance agreement. Paper presented at SALA, XXI, Konstanz, Germany, October 2001.
Chomsky, Noam
1995 *The Minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam & Morris Halle
1968 *The Sound Pattern of English*. Englewood: Prentice-Hall.
Christdas, Prathima
1988 The Phonology and Morphology of Tamil. Unpublished Cornell University Dissertation, 1988.
Comrie, Bernard
2002 Morphophonological alternations: typology and diachrony. In *Current Issues in Linguistic theory*, S. Bendajaballah, W. U. Dressler, O. E. Pfeiffer & M. Voeikova (eds.), Amsterdam/Philadephia: Benjamins.
Dasgupta, Probal
2007 Knowledge and Language in Classical Indian Linguistics: Some Observations. Unpublished Ms. Indian Statistical Institute, Kolkata.
Dasgupta, Probal, Alan Ford & Rajendra Singh
2000 *After Etymology:Towards a Substantivist Linguistics*. Berlin: Lincom Europa.
Embick, David & Morris Halle
2005 On the Status of Stems in Morphological Theory. In *Proceedings of Going Romance 2003*, T. Geerts & H. Jacobs (eds.). Amsterdam/Philadephia: Benjamins.
Ford, Alan & Rajendra Singh
1997 Reply to Mohanan. In *Trubetzkoy's Orphan,* Rajendra Singh (ed.), 166–170. Amsterdam/Philadephia: Benjamins.
Ford, Alan, Rajendra Singh & Gita Martohardjono
1997 *Pace Panini: Towards a Word-based Theory of Morphology*. New York: Peter Lang.

Hankamer, Jorge
1989 Morphological parsing and the lexicon.In *Lexical representation and process*, William Marslen-Wilson (ed.), 392–408. Cambridge, MA: MIT Press.

Hockett, Charles F.
1947 Problems of morphemic analysis. *Language* 23: 321–343.

Janda, Richard
1994 Form and content in a Functionalist Semiotic Model of Morphology: Comments on Dressle. In *Trubetzkoy's* Orphan, Rajendra Singh (ed.), 84–96. Amsterdam/Philadephia: Benjamins.

Jespersen, Otto
1924 *Philosophy of Grammar*. London: Allen & Unwin.

Joseph, Brian & Richard Janda
2002 Sanskrit as she has been misanalyzed prosodically. *The Yearbook of South Asian Languages and Linguistics*, Rajendra Singh (ed.), 59–90. Thousand Oaks: Sage.

Kiparsky, Paul
1982 Lexical Morphology and Phonology. In *Linguistics in the Morning Calm: Selected Papers from SICOL 1981*, Linguistic Society of Korea (ed.), 3–91. Seoul: Hanshin.

Koskenniemi, Kimmo
1983 *Two-level Morphology: A General Computational Model for Word-form Recognition and Production. Publication no 11*. University of Helsinki, Department of General Linguistics.

Mohanan, Karavannur P.
1994 Where does Morphonology belong? Comments on Ford and Singh. In *Trubetzkoy's Orphan,* Rajendra Singh (ed.), 140–154. Amsterdam: Benjamins.

Mohanan, Karavannur P. & Tara Mohanan
1998 Two tier morphology. Ms., NUS, Singapore.

Mohanan, Tara
1995 Wordhood and lexicality: Noun incorporation in Hindi. *Natural Language and Linguistic Theory* 13: 75–134.

Neuvel, Sylvain
2002 Whole word morphologizer: A non-stochastic computational approach. *Brain and Language* 81: 454–463.

Neuvel, Sylvain & Rajendra Singh
2002 Vive la difference: what morphology is really about. *Folia Linguistica* 35: 313–320.

Paul, Herman
1888 *Prinzipien der Sprachgeschichte*. Halle.

Rathore, Sandhya
1998 *Kaunda Bhatta's Vaiyakaranabhusanasara: An Analytical Study*. New Delhi: Indian Council of Philosophical Research (distributed by Munshiram Manoharlal, Delhi).

Sengupta, Gautam
1997 Three models of morphological processing. *South Asian Language Review* 7: 1–26.

Singh, Rajendra
2000 Morphology and your Miranda rights. Lecture, The Linguistic Society of Hawaii, Honolulu, August 2000.
2001a Constraints, preferences and context sensitivity in morphology. In *Constraints and preferences,* K. Dziubalska (ed.), 339–358. Berlin/ New York: Mouton de Gruyter.
2001b Review of "Wolfgang Dressler, Oskar Pfeiffer, Markus Pochtrager & John R. Rennison (eds.): Morphological Analysis in Comparison." *Poznan Studies in Linguistics* 39.
2003 Reduplication in Modern Hindi and the theory of reduplication. *The Yearbook of South Asian Languages and Linguistics*, Rajendra Singh (ed.), 155–172. Berlin/New York: Mouton de Gruyter.

Singh, Rajendra & Rama Kant Agnihotri
1997 *Modern Hindi Morphology: a word-based description*. Delhi: Motilal Banarasidass.

Singh, Rajendra & Alan Ford
2000 In praise of Sakatyayan: some remarks on whole word morphology. In *The Yearbook of South Asian Languages and Linguistics*, Rajendra Singh (ed.), 303–310. Thousand Oaks: Sage.

Singh, Rajendra & Probal Dasgupta
1999 On socalled compounds. In *The Yearbook of South Asian Languages and Linguistics 1999*, Rajendra Singh (ed.), 265–278. Thousand Oaks: Sage.

Stampe, David
1999 More what? Paper presented at the *32nd Poznan Linguistics Meeting*, Poznan, Poland, April, 1999.

Tirumalesh, K.V.
1997 Kannada Prefixation. In *Phases and Interfaces of Morphology*, Hariprasad, H. Nagarajan, P. Madhavan & K. G. Vijaykrishnan (eds.), 74–84. Hyderabad: The Central Institute of English and Foreign Languages.

Williams, Edwin
2000 Three models of the morphology-syntax interface. In *Morphological Analysis in Comparison*, Wolfgang Dressler, Oskar E. Pfeiffer, Markus Pochtrager & John R. Rennison, (eds.), 223–248. Amsterdam/ Philadephia: Benjamins.

Appendices

Announcements

The Gyandeep 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, 2008, for example, will be considered only for the 2009 issue of *ARSALL*.

Potential contributors are encouraged to send their initial submissions as word document files to the Editor (R. Singh). Those whose papers are accepted must, however, submit both hard and soft copies before the second deadline (March 1). The soft copy must be prepared using the pre-formatted template furnished by Mouton at:

http://www.degruyter.com/cont/imp/mouton/moutonAuthors.cfm

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

Ad Backus, Ph.D. (Tilburg) teaches in the Department of Languages and Minorities at the Katholik University of Tilburg, the Netherlands. He has published extensively on code-mixing and bilingualism.

a.m.backus@uvt.nl

Tej K. Bhatia, Ph.D. (Illinois) is Professor of Linguistics at Syracuse University, Syracuse, USA. He has published extensively in the domains of descriptive and applied linguistics (including sociolinguistics).

tkbhatia@mailbox.syr.edu

Rakesh M. Bhatt, Ph.D. (Illinois) is Associate Professor of Linguistics and Director of the Program in Second Language Acquisition and Teacher Education at the University of Illinois at Urbana-Champaign. His work focuses primarily on language contact, acquisition and use of English in South Asia, code switching, and language shift/loss of Kashmiri.

rbhatt@uiuc.edu

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

Alice Davison, Ph.D. (Chicago) is Professor of Linguistics at the University of Iowa. She is interested in the interrelation of Chomskyan linguistic theory and the particular structures and features of Hindi/Urdu and other South Asian languages. She received her PhD in Linguistics from University of Chicago, where she also studied Hindi and Sanskrit. She has taught at

SUNY Stony Brook, University of Illinois, Champaign-Urbana and University of Wisconsin.

alice-davison@uiowa.edu

Rajat Ghosh, Ph.D. (Tezpur University) was trained at the Central Institute of English and Foreign Languages, Hyderabad (M.Phil.) and has worked in Assam and West Bengal. He now teaches applied linguistics at Majan University College, Muscat, Oman.

ghosh.rajat@rediffmail.com

Ayesha Kidwai, Ph.D. (J.N.U.) teaches syntax and morphology at Jawaharlal Nehru University, New Delhi, and is the author of *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu* (2000, Oxford University Press, New York).

ayesh.kidwai@gmail.com

Claudia Lange is Lecturer in Linguistics at the Technical University, Dresden, Germany. She has published on reflexivity and intensification in the history of English as well as on focus marking in Indian English. She is currently working on discourse markers in spoken Indian English.

claudia.lange@mailbox.tu-dresden.de

Pramod Pandey, Ph.D. (Pune) is Professor of Linguistics at the Jawahar Lal Nehru University, New Delhi.He has taught linguistics for over a quarter century in different academic milieu. His areas of research interest include phonetics, phonology, morphology, historical linguistics, writing systems, and language teaching. He is currently working on a manuscript, *Sounds and their Patterns in Indic Languages*.

pkspandey@yahoo.com

John Peterson, Ph.D. (Kiel) completed his doctoral degree in linguistics while working as a research associate in Indology. He is currently working as an assistant professor at the department of Linguistics at the University of

Osnabrück (Germany). Together with Tej Bhatia (Syracuse) he co-manages the email-discussion group *Vyâkaran*, which he founded, and also manages the "Bibliography for Seldom Studied and Endangered South Asian Languages" under http://www.southasianbibliography.de.

jpeterso@uni-osnabrueck.de

Syed A. Sayeed, Ph.D. (Aligarh) is an independent scholar. He has lectured at Ailgarh Muslim University, Aligarh and the University of Hyderabad.

syedsayeed@rediffmail.com

Rajendra Singh, Ph.D. (Brown) is Professor of Linguistics at Université de Montréal, Montreal. He has written extensively on phonology, morphology, and sociolinguistics. He edits this annual for Mouton.

rajendra.singh@umontreal.ca