



Contemporary Indian English

Variation and change

Andreas Sedlatschek

John Benjamins Publishing Company

Contemporary Indian English

Varieties of English Around the World (VEAW)

A companion monograph series devoted to sociolinguistic research, surveys and annotated text collections. The VEAW series is divided into two parts: a text series contains carefully selected specimens of Englishes documenting the coexistence of regional, social, stylistic and diachronic varieties in a particular region; and a general series which contains outstanding studies in the field, collections of papers devoted to one region or written by one scholar, bibliographies and other reference works.

General Editor

Edgar W. Schneider
Department of English & American Studies
University of Regensburg
Universitätsstraße 31
D-93053 REGENSBURG
Germany
edgar.schneider@sprachlit.uni-regensburg.de

Editorial Assistant

Alexander Kautzsch

Editorial Board

Laurie Bauer
Wellington
Manfred Görlach
Cologne
Rajend Mesthrie
Cape Town
Peter Trudgill
Fribourg
Walt Wolfram
Raleigh, NY

General Series, Volume G38

Contemporary Indian English. Variation and change.
by Andreas Sedlatschek

Contemporary Indian English

Variation and change

Andreas Sedlatschek

University of Freiburg

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Sedlatschek, Andreas.

Contemporary Indian English : variation and change / Andreas Sedlatschek.

p. cm. (Varieties of English Around the World, ISSN 0172-7362 ; v. G38)

Includes bibliographical references and index.

1. English language--Variation--India. 2. Linguistic change--India. I. Title.

PE3502.I61S43 2009

427'.954--dc22

2008054506

ISBN 978 90 272 4898 5 (HB; alk. paper)

ISBN 978 90 272 9012 0 (EB)

© 2009 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Map	IX
Abbreviations	XI
List of figures and tables	XIII
Acknowledgments	XIX
1. Introduction	1
2. Theoretical, methodological and descriptive framework	7
2.1 Background information: English in India – spread and current situation	8
2.1.1 The transplantation of English to India: 1579–1757	8
2.1.2 English becomes stabilized: 1757–1813	9
2.1.3 English takes root: 1813–1857	11
2.1.4 From ‘English in India’ to IndE: 1857–1947	14
2.1.5 The institutionalization of English after 1947	18
2.2 Previous descriptions of IndE	24
2.2.1 Describing IndE from within	28
2.2.2 Describing IndE in an international context	30
2.2.3 IndE – caught between the norms: A look at Indian usage handbooks	32
2.2.4 Previous corpus-based descriptions of IndE	34
2.2.5 Conclusion	40
2.3 Methodological and descriptive framework	40
2.3.1 Database: Primary Corpus	41
2.3.2 Database: Online corpora	44
2.3.3 Descriptive framework	46
3. Vocabulary	51
3.1 Loanwords	51
3.1.1 The distribution of loanwords in the Primary Corpus	52
3.1.2 The use of loanwords in IndE press reviews	54
3.1.3 The use of loanwords outside the Press Reviews section	59
3.1.4 More variation: A look at South Asia	64
3.1.5 Case study: The high numerals <i>lakh</i> and <i>crore</i>	71

3.2	More different words and meanings? In search of lexical Indianisms	76
3.2.1	In search of neologisms	77
3.2.2	From words to word formation	95
3.2.3	A note on meaning change	104
3.3	Lexical style variation	114
3.3.1	Function words: <i>among/amongst, while/whilst, amid/amidst, until/till</i>	116
3.3.2	Contractions	124
3.3.3	Even more variability? Near-but-not-identical words	126
3.4	Americanisms and Briticisms in the IndE lexicon	132
3.5	Chapter summary	145
4.	Lexicosyntax	149
4.1	Particle verbs: Verbs combining with <i>up, down, out, away, and off</i>	149
4.1.1	The distribution of unrecorded particle verbs in the Primary Corpus	150
4.1.2	Pinpointing the status of unrecorded verb-particle combinations	152
4.2	More variation in the use of particle verbs	161
4.2.1	Distribution in the Primary Corpus	161
4.2.2	Use of 'different' particles	163
4.2.3	Use of 'additional' particles	166
4.2.4	'No' particles	174
4.3	Even more variation in IndE lexicosyntax? Case studies	180
4.3.1	<i>Enjoy</i>	180
4.3.2	<i>Contest</i>	182
4.3.3	<i>Cater</i>	184
4.3.4	<i>Protest</i>	187
4.3.5	<i>Want</i>	190
4.3.6	The <i>to</i> -infinitive vs. <i>N-ing</i> divide	191
4.4	Chapter summary	195
5.	Morphosyntax and grammar at the sentence level	197
5.1	Noun phrase: Articles	197
5.1.1	Quantitative analysis	198
5.1.2	Qualitative investigations: Article usage in the student essays	202
5.1.2.1	Zero article for definite article	204
5.1.2.2	Zero article for indefinite article	211
5.1.2.3	Definite article for zero article	213
5.1.2.4	Indefinite article for zero article	224
5.1.2.5	Exchange of surface articles	226
5.1.3	Summary	227

5.2	Noun phrase: Countability and number	227
5.2.1	Quantitative analysis	229
5.2.2	Qualitative investigations	231
5.2.3	Summary	243
5.3	Nouns, pronouns and verbs: Concord with collective nouns	244
5.3.1	Quantitative analysis	245
5.3.1.1	Verbal concord	246
5.3.1.2	Pronominal concord	249
5.3.2	Case study: <i>Police</i>	252
5.3.3	Summary	256
5.4	Verb phrase: The past perfect	257
5.4.1	Quantitative analysis	260
5.4.2	Qualitative investigations	265
5.4.3	Summary	274
5.5	More variation in tense and aspect? Two observations	275
5.5.1	<i>Since</i>	275
5.5.2	Adverbials of the past	278
5.6	Verb phrase: The mandative subjunctive	281
5.7	Sentence level: Inversion in direct and embedded <i>wh</i> -questions	289
5.7.1	Direct <i>wh</i> -questions	292
5.7.2	Embedded <i>wh</i> -questions	296
5.8	More variation at the sentence level? <i>Only</i> and <i>itself</i> as focus elements	299
5.9	Chapter summary	308
6.	Conclusion	311
	References	317
	Appendices	331
	Appendix 2	331
	Appendix 3	332
	Appendix 4	345
	Appendix 5	355
	Index	361



Courtesy of the University of Texas Libraries, The University of Texas at Austin
(http://www.lib.utexas.edu/maps/cia08/india_sm_2008.gif).

Abbreviations

AmE	American English
AusE	Australian English
BaE	Bangladeshi English
BrE	British English
IndE	Indian English
NZE	New Zealand English
PakE	Pakistani English
SAfrE	South African English
SgE	Singaporean English
SrLE	Sri Lankan English
CAM	Cambridge International Dictionary of English
COB	Collins Cobuild English Dictionary for Advanced Learners
COD	Concise Oxford Dictionary
MW	Merriam-Webster's Collegiate Dictionary
MWDEU	Merriam-Webster's Dictionary of English Usage
OALD	Oxford Advanced Learner's Dictionary of Current English
OED	Oxford English Dictionary
BROWN	Brown University Corpus (American English, 1961)
FLOB	Freiburg-LOB Corpus (British English, 1991)
FROWN	Freiburg-Brown Corpus (American English, 1992)
LOB	Lancaster-Oslo-Bergen Corpus (British English, 1961)
KOL	Kolhapur Corpus (Indian English, 1978)
ACE	Macquarie Corpus (Australian English, 1980s)
WC	Wellington Corpus (New Zealand English, 1980s)
ICE	International Corpus of English
ICE-IND	Indian English sub-corpus of ICE
ICE-GB	British English sub-corpus of ICE
IND1978	Primary Corpus: Indian English, 1978
IND2000	Primary Corpus: Indian English, 2000
UK1961	Primary Corpus: British English, 1961
UK1991	Primary Corpus: British English, 1991

US1961	Primary Corpus: American English, 1961
US1992	Primary Corpus: American English, 1992
Broad	Primary Corpus: broadcast section
Disc	Primary Corpus: discussions section
Edit	Primary Corpus: editorials section
Ess	Primary Corpus: student essays section
I/D	Primary Corpus: interviews and discussions sections
Int	Primary Corpus: interviews section
News	Primary Corpus: news section
Press	Primary Corpus: press sections
Rep	Primary Corpus: reportage section
Rev	Primary Corpus: reviews section
DA	The Dawn (Pakistan), online edition
DH	The Deccan Herald (India), online edition
GU	The Guardian (UK), online edition
HI	The Hindu (India), online edition
HT	The Hindustan Times (India), online edition
IE	The Indian Express (India), online edition
IN	The Independent (Bangladesh), online edition
ST	The Sunday Times (Sri Lanka), online edition
TE	The Telegraph (India), online edition
TR	The Tribune (India), online edition
WP	The Washington Post (USA), online edition
.au	Internet top-level country domain: Australia
.in	Internet top-level country domain: India
.sg	Internet top-level country domain: Singapore
.uk	Internet top-level country domain: United Kingdom
.us	Internet top-level country domain: United States of America
.za	Internet top-level country domain: South Africa

List of figures and tables

Figures

- 2.1 The Primary Corpus
- 2.2 Supplementary data: online newspapers
- 2.3 Supplementary data: top-level country domains
- 2.4 Chapter overview

Tables

Chapter 3

- 3.1 Loanwords from Indian languages
- 3.2 Collocations of *Vidhan Sabha*
- 3.3 *Lathi(-)charge* vs. *baton(-)charge*
- 3.4 *Rasta roko* in IndE newspapers
- 3.5 *Rasta roko(s)*, *road rocko(s)* and *road blockade(s)*
- 3.6 *Lakh* and *crore* in currency statements [Internet: newspapers]
- 3.7 *Lakh* and *crore* in currency statements [KOL, Primary Corpus, ICE-IND]
- 3.8 Collocations of *lakhs of* and *crores of*
- 3.9 *Eve-teasing* and *harrasing*
- 3.10 *Time(-)pass* and *pastime*
- 3.11 *Petrol bunk*, *petrol pump*, *petrol station* and *gas station*
- 3.12 *Upgradation*, *upgrading* and *upgrade*
- 3.13 *Delink* vs. *disconnect*
- 3.14 *Delink*, *upgradation* and variants
- 3.15 *Uplift* vs. *upliftment* [Internet: newspapers]
- 3.16 *Uplift* vs. *upliftment* [Internet: country domains]
- 3.17 *Dematerialisation/dematerialization* vs. *demat* [Internet: newspapers]
- 3.18 *Dematerialisation/dematerialization* vs. *demat* [Internet: country domains]
- 3.19 *Hydel* vs. *hydro-electric*
- 3.20 *Charge(-)sheet*

- 3.21 *Undertrial(s)* and *incharge(s)*
- 3.22 Copulative compound nouns (N-*cum*-N)
- 3.23 Copulative compound adjectives (adj-*cum*-adj)
- 3.24 Adjectival compounds (N+V-*ed*) in attributive position
- 3.25 Adjectival compounds (N+V-*ing*) in attributive position
- 3.26 Prefixation: *pre-*, *post-*, *pro-* and *anti-*
- 3.27 Nominal compounds vs. prepositional constructions
- 3.28 *Sympathy wave* vs. *wave of sympathy* [Internet: newspapers]
- 3.29 *Sympathy wave* vs. *wave of sympathy* [Internet: country domains]
- 3.30 *Year-end*, *month-end*, *January-end*
- 3.31 Compounds: N+N vs. adj+N
- 3.32 *Telephone/telephonic; parliament/-ary* [Internet: newspapers]
- 3.33 *Telephone/telephonic; parliament/-ary* [Internet: country domains]
- 3.34 Standard and extended senses of *loot* and *shift*
- 3.35 Collocations of *loot* and *shift*
- 3.36 Mediopassive *release* [Internet: newspapers]
- 3.37 Mediopassive *release* [Internet: country domains]
- 3.38 *Meet* vs. *meeting* [Internet: IndE, BrE, AmE newspapers]
- 3.39 *Meet* vs. *meeting* [Internet: PakE, BaE, SrLE newspapers]
- 3.40 Spelling: <pandit>/<pundit>, <mogul>/<moghul>/<mughal>
- 3.41 *Among* vs. *amongst* [KOL, BROWN, FROWN, LOB, FLOB]
- 3.42 *Among* vs. *amongst* [Primary Corpus, ICE-IND]
- 3.43 *While* vs. *whilst* [KOL, BROWN, FROWN, LOB, FLOB]
- 3.44 *While* vs. *whilst* [Primary Corpus, ICE-IND]
- 3.45 *Amid* vs. *amidst* [KOL, BROWN, FROWN, LOB, FLOB]
- 3.46 *Amid* vs. *amidst* [Primary Corpus, ICE-IND]
- 3.47 *Amid* vs. *amidst* [Internet: IndE, BrE, AmE newspapers]
- 3.48 *Amid* vs. *amidst* [Internet: PakE, BaE, SrLE newspapers]
- 3.49 *Amid* vs. *amidst* [Internet: country domains]
- 3.50 *Amid* vs. *amidst* [OED quotations base]
- 3.51 *Until* vs. *till* [KOL, BROWN, FROWN, LOB, FLOB]
- 3.52 *Until* vs. *till* [Primary Corpus, ICE-IND]
- 3.53 *Until* vs. *till* [Internet: IndE, BrE, AmE newspapers]
- 3.54 *Until* vs. *till* [Internet: PakE, BaE, SrLE newspapers]
- 3.55 *Until* vs. *till* [Internet: country domains]
- 3.56 *Until* vs. *till* [OED quotations base]
- 3.57 Contractions in IndE, BrE, and AmE [Primary Corpus; press]
- 3.58 Contractions in IndE and BrE [Primary Corpus; broadcast]
- 3.59 *Tense* vs. *tensed* [Internet: newspapers]
- 3.60 *Tense* vs. *tensed* [Internet: country domains]
- 3.61 *Alternate* vs. *alternative*

- 3.62 *Lecturership* vs. *lectureship* [Internet: IndE, BrE, AmE newspapers]
- 3.63 *Lecturership* vs. *lectureship* [Internet: PakE, BaE, SrLE newspapers]
- 3.64 *Lecturership* vs. *lectureship* [Internet: country domains]
- 3.65 *Flammable* vs. *inflammable* [Internet: newspapers]
- 3.66 *Flammable* vs. *inflammable* [Internet: country domains]
- 3.67 *Film* vs. *movie* vs. *picture* 'moving pictures' [Primary Corpus; press]
- 3.68 *Film* vs. *movie* vs. *picture* 'moving pictures' [Primary Corpus; broadcast]
- 3.69 *Cinema hall* vs. *movie hall* vs. *picture hall*
- 3.70 *Talkie* vs. *bioscope*
- 3.71 *Toward(s)*; *afterward(s)* [KOL, BROWN, FROWN, LOB, FLOB]
- 3.72 *Toward(s)*; *afterward(s)* [Internet: newspapers]
- 3.73 *Taxi stand* vs. *taxi rank*
- 3.74 *Bus stand* vs. *bus stop*
- 3.75 *Rubbish* vs. *garbage*
- 3.76 *Price rise* vs. *price hike*
- 3.77 *Fully fledged* vs. *full-fledged*
- 3.78 *Okay*, *I guess*, *gonna*

Chapter 4

- 4.1 Unrecorded particle verbs with *up*, *down*, *away*, *off* and *out*
- 4.2 *Find (out) a solution* [Internet: IndE, BrE, AmE newspapers]
- 4.3 *Find (out) a solution* [Internet: PakE, BaE and SrLE newspapers]
- 4.4 *Find (out) a solution* [Internet: country domains]
- 4.5 *Lower (down)* [Internet: IndE, BrE, AmE newspapers]
- 4.6 *Lower (down)* [Internet: PakE, BaE and SrLE newspapers]
- 4.7 *Lower (down)* [Internet: country domains]
- 4.8 *Chart out* [Internet: IndE, BrE, AmE newspapers]
- 4.9 *Chart out* [Internet: PakE, BaE, SrLE newspapers]
- 4.10 *Chart out* [Internet: country domains]
- 4.11 *Chalk out* [Internet: IndE, BrE, AmE newspapers]
- 4.12 *Chalk out* [Internet: PakE, BaE, SrLE newspapers]
- 4.13 *Chalk out* [Internet: country domains]
- 4.14 *Take out* [Internet: IndE, BrE, AmE newspapers]
- 4.15 *Take out* [Internet: PakE, BaE, SrLE newspapers]
- 4.16 *Take out* [Internet: country domains]
- 4.17 'Different' particles [Primary Corpus]
- 4.18 'Different' particles [Internet: IndE, BrE, AmE newspapers]
- 4.19 'Different' particles [Internet: country domains]
- 4.20 *Admit*, *admission* and *translate*
- 4.21 *Discuss*, *emphasise*, *stress*, *urge*, *request*, *investigate* [ICE-IND]
- 4.22 *Discuss*, *emphasise*, *stress*, *urge*, *request*, *investigate* [Internet: country domains]

- 4.23 *Call/rename/term (as)* [KOL, Primary Corpus, ICE-IND]
- 4.24 *Call/rename/term (as)* [Internet: IndE, BrE, AmE newspapers]
- 4.25 *Call/rename/term (as)* [Internet: PakE, BaE and SrLE newspapers]
- 4.26 *Call/rename/term (as)* [Internet: country domains]
- 4.27 *Comprise (of)* [Internet: IndE, BrE, AmE newspapers]
- 4.28 *Comprise (of)* [Internet: PakE, BaE, SrLE newspapers]
- 4.29 *Comprise (of)* [Internet: country domains]
- 4.30 *Bunk (off)* [Internet: IndE, BrE, AmE newspapers]
- 4.31 *Bunk (off)* [Internet: PakE, BaE, SrLE newspapers]
- 4.32 *Bunk (off)* [Internet: country domains]
- 4.33 *Avail (of)* [KOL, BROWN, FROWN, LOB, FLOB]
- 4.34 *Avail (of)* [ICE-IND]
- 4.35 *Avail (of)* [Internet: IndE, BrE, AmE newspapers]
- 4.36 *Avail (of)* [Internet: PakE, BaE, SrLE newspapers]
- 4.37 *Avail (of)* [Internet: country domains]
- 4.38 *Enjoy* [Primary Corpus]
- 4.39 *Enjoy* [ICE-IND]
- 4.40 Complementation of *cater*
- 4.41 *Cater for* vs. *cater to* [Internet: country domains]
- 4.42 *Cater for* vs. *cater to* [Internet: newspapers]
- 4.43 *Cater for* vs. *cater to* [KOL, BROWN, FROWN, LOB, FLOB]
- 4.44 Complementation of *protest* [KOL, BROWN, FROWN, LOB, FLOB]
- 4.45 Complementation of *protest* [Internet: newspapers]
- 4.46 Complementation of *protest* [Internet: country domains]
- 4.47 Complementation of *want* [Primary Corpus]
- 4.48 Complementation of *want* [Internet: newspapers]
- 4.49 Complementation of *want* [Internet: country domains]
- 4.50 *Look forward, contemplate, with a view* [Internet: newspapers]
- 4.51 *Look forward, contemplate, with a view* [Internet: country domains]
- 4.52 *With a view to* + infinitive/V-ing [OED quotations base]

Chapter 5

- 5.1 Definite articles per 1,000 words [Primary Corpus; press]
- 5.2 Definite articles per 1,000 words [Primary Corpus; broadcast and essays]
- 5.3 Indefinite articles per 1,000 words [Primary Corpus; press]
- 5.4 Indefinite articles per 1,000 words [Primary Corpus; broadcast and essays]
- 5.5 'Different' uses of articles [Primary Corpus; essays]
- 5.6 *(A/the) majority of* [ICE-IND]
- 5.7 *(A/the) majority of* [Internet: country domains]
- 5.8 *(A/the) majority of* [Internet: newspapers]
- 5.9 *(A) lot of* [Primary Corpus, ICE-IND]

-
- 5.10 (A) *lot of* [Internet: country domains]
 - 5.11 Definite/zero article with *police* [Primary Corpus; press]
 - 5.12 Definite/zero article with *police* [Primary Corpus; broadcast]
 - 5.13 Definite/zero article with *police* [ICE-IND]
 - 5.14 Definite/zero article with *police* [Internet: newspapers]
 - 5.15 Definite/zero article with *police* [Internet: country domains]
 - 5.16 Definite/zero article with *society* [Primary Corpus; press]
 - 5.17 Definite/zero article with *society* [Internet: newspapers]
 - 5.18 Definite/zero article with *society* [Internet: country domains]
 - 5.19 Definite/zero article with names [Primary Corpus; press]
 - 5.20 Definite/zero article with names [Primary Corpus; broadcast]
 - 5.21 Definite/zero article with names [Internet: newspapers]
 - 5.22 Definite/zero article with *hospital* [Primary Corpus; press]
 - 5.23 Definite/zero article with *hospital* [Primary Corpus; broadcast]
 - 5.24 Definite/zero article with *hospital* [ICE-IND]
 - 5.25 Definite/zero article with *hospital* [Internet: newspapers]
 - 5.26 Definite/zero article with *hospital* [Internet: country domains]
 - 5.27 *Last but not (the) least* [Internet: newspapers]
 - 5.28 *Last but not (the) least* [Internet: country domains]
 - 5.29 Unrecorded count uses of nouns in IndE
 - 5.30 Count and noncount nouns: selected collocations
 - 5.31 *Legislation(s)*
 - 5.32 *Equipment(s)*
 - 5.33 *Agitation(s)*
 - 5.34 *Advice(s)* and *tuition(s)*
 - 5.35 *Content(s)*, *damage(s)*, *resistance(s)*
 - 5.36 *Machinery/-ies*
 - 5.37 *Aircraft(s)* [Internet: country domains]
 - 5.38 *Aircraft(s)* [Internet: newspapers]
 - 5.39 Count and noncount uses of nouns: lexicalized phrases
 - 5.40 *At all cost(s)*
 - 5.41 Singular and plural verbal concord with collective nouns
 - 5.42 Verbal concord with collective nouns [Primary Corpus; press]
 - 5.43 Verbal concord with collective nouns [Primary Corpus; broadcast]
 - 5.44 Singular and plural pronominal concord with collective nouns
 - 5.45 Pronominal concord with collective nouns [Primary Corpus; press]
 - 5.46 Pronominal concord with collective nouns [Primary Corpus; broadcast]
 - 5.47 Concord with *police*
 - 5.48 Verbal concord with *police* [OED quotations base]
 - 5.49 Verbal concord with *police* [Internet: newspapers]
 - 5.50 Verbal concord with *police* [Internet: country domains]

- 5.51 Past perfect [Primary Corpus; press]
- 5.52 Backshift in reported speech after past-tense *said*
- 5.53 Suspended backshift
- 5.54 Choice of tense with *after*
- 5.55 Past perfect [Primary Corpus; broadcast and essays]
- 5.56 'Different' uses of the past-perfect in IndE
- 5.57 Temporal *since* [Primary Corpus; press]
- 5.58 Temporal *since* [Primary Corpus; broadcast and essays]
- 5.59 Tense with *ago* and *yesterday* [Primary Corpus]
- 5.60 Tense with *ago* and *yesterday* [ICE-IND]
- 5.61 *Ago, back* and *before* [Primary Corpus; press]
- 5.62 *Ago, back* and *before* [Primary Corpus; broadcast and essays]
- 5.63 Mandative subjunctive [LOB-BROWN corpus family]
- 5.64 Subjunctive vs. *should* with selected verbs [LOB-BROWN corpus family]
- 5.65 Subjunctive vs. *should* with selected verbs [Internet: country domains]
- 5.66 Subjunctive vs. *should* with selected verbs [Internet: newspapers]
- 5.67 Inversion in direct *wh*-questions
- 5.68 Inversion in embedded *wh*-questions
- 5.69 Subjunct *only* [Primary Corpus; press]
- 5.70 Subjunct *only* [Primary Corpus; broadcast and essays]
- 5.71 Subjunct *only* in selected collocations [Internet: country domains]
- 5.72 Subjunct *only* in selected collocations [Internet: newspapers]
- 5.73 *Itself* as focus element

Acknowledgments

This book would not have seen the light of the day without the help of a great many people and institutions in India and at home. I would like to express my sincerest gratitude to the headmasters, staff and students of The Doon School and Carman School in Dehradun for allowing me to be a part of their lives on and off campus during my first field trip to India in the spring of 2000. I am also grateful to the headmistress, staff and students of The Vivekandanda School in Dehradun for their assistance over the years. K. N. Bahuguna's knowledge opened doors and helped me tremendously in getting started with my field work, which was greatly appreciated. I would also like to thank Birendra and Gouri Das for their hospitality during my stay in Lucknow and Ravi Goel and his family for their friendship.

Over here in Germany, I would like to thank my supervisor Christian Mair at the University of Freiburg for his never-ending patience, encouragement and expertise that were so essential for completing this work. I also wish to acknowledge my appreciation to Monika Fludernik, who triggered my academic interest in Indian-English literature in her classes at Freiburg University. I would further like to thank Andrea Sand for discussing several drafts of this book with me, giving me valuable feedback, and helping me stay focused. A big 'thank you' also goes to Dagmar Deuber for her encouragement and continued interest in my research, to Sandra Hestermann for reading and commenting on the manuscript, and to Birgit Waibel, who assisted me in proofreading the final version of this book and compiling the Index. I wish to extend my thanks to the Landesgraduiertenförderung Baden-Württemberg, whose financial support between 2000 and 2002 was indispensable, and to Shiva-Kumar and Ursula Sharma and the Indian School Society, Bonn, as well as Henning and Ursula Brandis and the Sir Dietrich Brandis Foundation, Bonn, for their guidance and support throughout this project. I am also grateful to Edgar W. Schneider for accepting this book into the *Varieties of English Around the World* series.

Last but certainly not least, I would like to thank my family and friends, whose support kept me in one piece during the sometimes difficult process of completing this research. I would particularly like to thank Trish and Stan Ballard for starting my interest in the English language, Arne, Michaela, Ingo, Ulla, Claudia and Michael for being there from beginning to end, Carmen, Evi, Friederike, Heike, Hubbi and Martin for their encouragement during the final stages of writing this book, and Ursula Brandis for all of this and more. I would finally like to thank my parents, Dorothee and Gerhard Sedlatschek, and my sister Gabi for all their love and support throughout the years. There are simply no words to express how grateful I am for having you. It is to you that I dedicate this book.

CHAPTER 1

Introduction

The spread of English over the past four hundred years has led to the emergence of transplanted varieties of English in variegated sociocultural and linguistic contexts. While English was used by an estimated 5 to 7 million people at the beginning of the seventeenth century, the majority of whom were living in the British Isles, there are an estimated 377 million users of English today who speak English as a first language and another 235 million people using it as a second language. Many more are exposed to English regularly (Crystal 1997: 60; Jenkins 2003: 3–5). Even though the spread of English set in well before the seventeenth century, it was in the wake of the British Empire that English started to move from England, its original home, across the world with those who settled abroad for commercial, religious and political reasons. When the British Empire drew to its close three centuries later, English turned into one of its legacies. Many colonies became independent, but English often remained an essential part of their environment. The rise of the United States, a former British colony, and its economic and political influence after World War II have secured the position of English in the world as an international *lingua franca* and global language. Today, many different regional varieties of English or ‘Englishes’ exist around the globe and are slowly but steadily gaining recognition. Indian English (IndE) is one of the oldest.

In India, English has been in use for more than four centuries, first as the language of the early merchants, missionaries and settlers, later as the language of the British colonial power, and finally – after India’s Independence in 1947 – as the so-called associate official language of the Indian Union. Although it has repeatedly been attempted to abolish the former colonial language, English has remained. Sixty years after Independence, it continues to play an important, albeit controversial, role in numerous domains of Indian life. Besides its status as the associate official language of the Indian Union next to Hindi, India’s national language, English is the official language in several Union states and territories. English-language newspapers and magazines are published in all states and are second in number and circulation only to Hindi publications. India has grown into the third largest producer of English books in the world, and private as well as state-run broadcasting companies offer English-language programs on a daily basis. In the educational sector, too, English is a major factor. While secondary education is offered country-wide in Hindi, English and a regional language according to the Three-Language Formula, English-medium education enjoys great prestige, with higher education often accessible only through

English, especially in the fields of science and technology. The teaching of English has, in fact, turned into a profitable business sector, not least because English is becoming increasingly important in a society whose fast-growing IT sector attracts companies from all over the world seeking to capitalize on the skills of English-speaking Indians. While the number of speakers of English as a mother tongue is comparatively low (at about 180,000 according to the 1991 Census), the number of Indians using English as an additional language is much higher, with estimates varying from 3% to 5% of India's population (cf. D'souza 1997) up to a staggering 30%, the equivalent of 330 to 350 million users (cf. Kachru 2005: 15; Kamdar 2007: 4). Numerical guesswork of this kind is interesting as it reflects the problems involved when trying to identify the group of English-using Indians, who differ in degrees of competence, regional and educational backgrounds, and the extent to which they employ English. If there is one appropriate generalization, it is that there is a strong association of the affluent and influential sections of Indian society with English, which also explains the high social prestige that Indian society at large attaches to this language these days. In the words of the linguists Kapil Kapoor and R. S. Gupta,

[n]o major Indian language today has the same 'paying potential' as English has in every sphere of life: in trade and in commerce, in administration, in education and in science and technology ... You cannot become a doctor, a scientist, a technocrat, a top-level business executive or a high-ranking bureaucrat if you do not know English. (Kapoor/Gupta 1991: 17)

The conceptualization of IndE as a linguistic entity has posed challenges, and its existence as a variety in its own right has repeatedly been questioned. Although linguists nowadays agree widely that IndE has established itself as an "independent language tradition" (Gramley/Pätzold 1992: 441) not to be mistaken for an impoverished version of the 'Queen's English', the question of just how unique or different IndE is as compared to other varieties of English is open. Should IndE be treated as an autonomous language system (Verma 1978, 1982)? Should it be treated as "normal English" with more or less learner-specific deviations" (Schmied 1994: 217)? Or should it be treated as a "modular" (Krishnaswamy/Burde 1998), "national" (Carls 1994) or "international" (Trudgill/Hannah 2002) variety? It is surprising to see that in spite of the plethora of publications from theoretical, historical and sociolinguistic perspectives (cf. Carls 1979; Leitner 1985; Ramaiah 1988), comparatively little empirical linguistic research has been conducted on the structure and use of IndE that would help us put the available hypotheses to test. A look at Krishnaswamy/Burde's (1998: 1–24) twenty-four-page-long survey of previous research on the topic is insightful in this regard: the information listed under the heading 'Data-based Attempts' comprises a meagre sixteen lines of running text. The few studies that have so far set out to describe IndE on the basis of authentic data are either dated (e.g. Labru 1984; Dubey 1989; Parasher 1991; Shastri 1988a, 1992, 1996) or skewed toward the written domain (e.g. Rogers 2002; Olavarria de Ersson/Shaw 2003). More common are characterizations of IndE

based on impressionistically gathered data: the well-known ‘feature lists’ (e.g. Bakshi 1991; Lukmani 1992; Bhaskararao 2002). Those usually include long lists of lexical items, syntactic structures and phonological features that are claimed to be markers of IndE. The tacit assumption underlying those listings is that the chosen ‘Indianisms’ characterize contemporary IndE in some way – but rarely do we find specific information on where exactly the characteristicness of the selected features can be expected to lie. Are we dealing with independent innovations found exclusively in IndE? Are we dealing with features shared by IndE with other varieties? Are there qualitative or quantitative differences observable between IndE and other varieties or across registers and text types within IndE? And if that should be so, how relevant are those differences and what do they signify? Hence, although it has been argued that IndE is on its way toward carving out a regional standard variety, it is still unknown to date what the characteristics of this regional standard really are.

This situation is unfortunate as more empirical knowledge on the stability of variety-specific forms would be a precondition for codifying local IndE usage norms in handbooks, dictionaries and grammars. That there is, in fact, an urgent need in post-Independence India for such books is seen in the great number of glossaries that have been published regularly over the decades (cf. Nihalani et al. 1979, 2004; Hawkins 1984; Lewis 1991; Muthiah 1991; Crowther 1996; Yadurajan 2001; Hankin 2003). But none of the available publications has been based on extensive empirical research. A solid data-based description of IndE, however, could be used to provide Indian learners of English with a comprehensive and realistic model of their own English within the context of world Englishes. As Tickoo 1991b points out correctly, “a scientific description of educated usage is the only road to giving standard IndE general viability” (145).

The need for more empirical research on IndE has long been recognized. Consider the following statements:

No study on IE [IndE] has attempted to show any systematic relationship between other (native) varieties of English and IE. (Daswani 1978: 115)

Insgesamt muss jedoch festgestellt werden, daß eine umfassende Darstellung des IE [Indischen Englisch] mit der sich herauskristallisierenden Norm eines “Standard Indian English” als zentralem Thema noch immer aussteht. Was vor allem fehlt, sind empirische Erhebungen ... (Spitzbardt 1987: 128)¹

IndE has not yet been defined or, for pedagogical purposes, devised as a model. Nor is there a book or manual that provides access to acceptable Indianisms.

(Tickoo 1991b: 138)

1. Translation: All in all, it has to be stated that a comprehensive account of IE [IndE] which would focus on the emerging norms of a “standard Indian English” as its central topic is yet to be written. In particular, there is a dearth of empirical investigations ...

The performance and frequency oriented early work on 'IE' [IndE] syntax is, in my view, badly in need of empirically rich studies that can tell us what the 'IE' norms might be. (Singh 1994: 378)

[W]e have no local reference materials available (at least not in any serious sense) and are forced to turn to dictionaries and grammars based on other varieties when in doubt about items of usage and grammar. This only heightens the sense of insecurity about the language. (D'souza 2001b: 245)

There is as yet no large-scale study of spoken or written South Asian English. Nor has any serious attempt been made to distinguish the features in terms of the proficiency scale, the register-specificity of the features and the distribution of grammatical features with reference to the regions. (Kachru 2005: 48)

Stabilization, homogenization, and codification are lagging far behind ... Works like Yule and Burnell 1886, for lexis, and Nihalani et al. 2004, for general usage issues, may be regarded as forerunners, but they are clearly a far cry from an adequate descriptive account of Indian English. (Schneider 2007: 172)

It is against this background that the present study aims to make a contribution toward describing IndE and its emerging regional standard in a corpus-linguistic framework. When taking a closer look at the available descriptions of IndE, it soon became clear that any new descriptive approach would have to be more specific about the status of those numerous claimed 'Indianisms'. Stable uses would have to be separated from unstable ones, and contemporary IndE would have to be viewed in a comparative framework to describe the variety dynamically in terms of its degree of nativization and in terms of its relation to other varieties of English. Early on, the idea had been born to compile a carefully designed micro-corpus of high-proficiency IndE speech and writing that would lend itself to diachronically sensitive comparative research. It was hypothesized that if used alongside larger text collections such as the Kolhapur Corpus, the Indian component of the International Corpus of English (ICE-India) and the World Wide Web, such a micro-corpus could be employed systematically to test the validity of the available descriptions and address the question to what extent IndE was on its way toward regional standardization. This set the agenda for the present study.

The main objective of this study is to test the validity and accuracy of the available feature-list descriptions of IndE lexis and syntax. In particular, the aims are (1) to describe IndE in terms of feature-specific variational profiles across registers and modes, using a micro-corpus of contemporary spoken and written IndE (the so-called Primary Corpus) and supplementary text collections; (2) to compare the variational profiles of IndE to those of other varieties and to the codified lexical and grammatical norms of standard English in order to determine quantitative and qualitative similarities and differences; and (3) to interpret those findings meaningfully in terms of their characteristicness for contemporary IndE and their relevance for pinpointing directions of linguistic change in that variety.

Chapter 2 provides the theoretical and methodological background information to the linguistic analysis of the corpus data and presents a survey of previous empirical research on IndE structure and use. Chapter 2 also describes the corpora used in this study and lays out the analytical framework for the case studies on IndE vocabulary and syntax. The results of the linguistic analysis are presented in Chapters 3 (vocabulary), 4 (lexicosyntax) and 5 (morphosyntax and grammar at the sentence level). Chapter 6 integrates the individual findings and presents the conclusion.

Theoretical, methodological and descriptive framework

One of the most influential conceptual frameworks for categorizing world Englishes is Braj B. Kachru's Three Circles model. Originally proposed by Kachru 1985, the model classifies Englishes broadly as to their historically grown position in a community, with three partially overlapping circles representing the spread of English from Britain to countries of the Inner Circle by migration (e.g. Australia), to the Outer Circle by colonization (e.g. India) and to the Expanding Circle by means of its function as an international language (e.g. Japan).¹ While Inner Circle Englishes are spoken by the majority of their users as a first language, English is used primarily as a foreign language in the territories of the Expanding Circle. In the Outer Circle, English is used in multilingual contexts as an additional language for intranational communication. IndE, according to Kachru's (1985) model, is a variety of the Outer Circle. Used mainly as a second language today, IndE shares with other varieties in this category several important features: a history of transplantation and "broken transmission" (Kandiah 1998: 36), the status of an institutionalized variety with a wide range of functions in society, signs of nativization in structure and use, and the emergence of local usage norms.² Before looking at the structure and use of present-day IndE in detail in Chapters 3, 4 and 5, this chapter provides important background information to the linguistic analysis. In the following, a profile of IndE as a variety of the Outer Circle is presented from a historical and contemporary perspective (2.1) before reviewing previous descriptions of the variety (2.2). This chapter also outlines the methodology and data applied in the present study, sketches out the descriptive framework underlying the case studies and discusses the nature of structural nativization in IndE (2.3).

1. For more recent representations of the model cf. Svartvik (1998:16–21) and Jenkins (2003:15–8). Jenkins 2003 also provides numbers of speakers in the Inner, Outer and Expanding Circle territories. Cf. Kachru (2005:13–6) for a more detailed description of the position of Asian Englishes within the Three Circles model.

2. For a discussion of the limitations of Kachru's tripartite classification of world Englishes and his choice of terminology cf. Kandiah (1998:6–8), Jenkins (2003:17–8) and Schneider (2003:237). It is understood that Kachru's (1985) model can only serve as a very broad categorization of varieties of English. When using the terms 'Inner Circle', 'Outer Circle' and 'Expanding Circle', it is not implied that the linguistic situation is uniform in all countries within any Circle.

2.1 Background information: English in India – spread and current situation

“Language change,” says S. V. Shastri, the founding father of the Kolhapur Corpus of IndE,

is a historical fact and the development of language varieties – regional, national, social and other – is inevitable. But the direction which these changes take is determined by various conscious and unconscious attitudes of the users of the language in question apart from the new setting in which it comes to be used.

(Shastri 1988c: vii)

This chapter traces the development of the English language in India and the directions of change from its early beginnings as a language of the first European settlers in the late sixteenth century to its implementation as associate official language of the Union of India after Independence in 1947. It is understood that any attempt to chronicle the history of English can only be an approximation to the complex processes of language contact and change involved in the emergence of a new variety.

2.1.1 The transplantation of English to India: 1579–1757

English was the third European language (after Portuguese and Dutch) to reach India in the sixteenth century (Parasher 1981: 330; A. Gupta 1991: 62; Das 1994: 1–12). The arrival of the Jesuit missionary Thomas Stephens in Goa in 1579, presumably one of the first Englishmen to settle in India, has been mentioned as the beginning of direct language contact between English and the languages of the Indian subcontinent (Lewis 1991: 14; Mehrotra 1998: 3). When Queen Elizabeth granted merchants in London a charter to trade with India two decades later, on December 31, 1600, the stage was set for the East India Company to develop into a major political force. With the initial activities of the East India Company, as yet “a small company of adventurous and enterprising merchants” (Kachru 1994: 502), English was introduced as a means of communication in trading posts, where merchants from Britain negotiated business with local rulers and trade agents. The first trading station of the East India Company was set up at Surat in 1612.³ Contact with English in the beginning years was probably limited to a small number of Indians.⁴ English traders were assisted by

3. By the end of the seventeenth century, trading stations had been established at Madras (now Chennai, 1639–40), Bombay (now Mumbai, 1674) and Calcutta (now Kolkata, 1690), cf. Kachru (1994: 502).

4. An early example of an English-speaking Indian is reported by Rahman (2002: 174). Stating that “[n]othing is known of the first Indian, or Indians, who learned English” (174), Rahman 2002 goes on to quote the seventeenth-century historian Khafi Khan who “tells us that he had an interview with Yakut Khan, the commander of an island off Bombay, who described how

peons, translators, who have been considered as “the earliest users of English in India” (Mehrotra 1998: 3). The first missionary schools imparted education in the languages spoken in the areas of settlement (e.g. Portuguese) and occasionally taught English as a foreign language. Later, English also began to be used in the military by Indian recruits (Mehrotra 1998: 3). But the first English-medium missionary schools were founded only in the early eighteenth century (Carls 1987: 116) after the Charter for the East India Company had been renewed in 1698 and a ‘missionary clause’ been included to regulate the passage of missionaries to India (Parasher 1991: 29).⁵ All in all, English must have had a rather limited impact on India during the first one hundred and fifty years of British involvement. Schneider 2007, applying his evolutionary model of new Englishes, speaks of the Foundation Phase, during which “English gradually gained more, but still limited, influence in India” (163).⁶ Mukherjee 2007, referring to Schneider’s (2003, 2007) model, argues that Indians must have viewed English as a foreign language in that period and “the colonizers still considered themselves genuinely British people” (164).

2.1.2 English becomes stabilized: 1757–1813

The position of English began to change in the course of the political developments throughout the eighteenth century, when Britain overcame competition from other European nations in its attempt to gain control over the Indian subcontinent. The victory of Lord Clive in the Battle of Plassey (1757) confirmed the position of Britain

the English had conquered a ship in which an acquaintance of his was travelling. This acquaintance told Yaqt that the Englishmen took revenge for an earlier defeat they had suffered at his (Yaqt’s) hands. Their words, in English, were translated to the acquaintance by a fellow traveller.” Rahman 2002 adds that “[i]t is not quite clear who the translator was and how much English he knew. However, if he had learned English, then this must be a very early instance of doing so because it occurred in Aurangzeb’s days (r. 1658 – 1707).” (174)

5. English-medium schools were set up in Madras (St. Mary’s Charity School, 1715), Bombay (Charity School, 1719) and Calcutta (Charity School, 1720–31), cf. Kachru (1994: 503).

6. In his ‘dynamic model of the evolution of new Englishes’, Schneider (2003, 2007) describes the emergence of new varieties of English as a uniform process in five stages: foundation, exo-normative stabilization, nativization, endonormative stabilization and differentiation. Schneider describes those stages in terms of the changes in the interaction of the two groups participating in that process – the colonizers (forming the so-called settler or STL strand) and the colonized (forming the so-called indigenous or IDG strand) – and the resulting reconstruction of group identities. As the two groups become more and more intertwined, a process of linguistic accommodation sets in (cf. Schneider 2003: 243–4, 2007: 33–6). In the Foundation Phase, contact between speakers of the STL strand and the IDG strand is as yet highly restricted, “impeded by the inability to understand each other and by different concerns and needs” (Schneider 2003: 244). There is, for example, little influence of the indigenous languages on the English spoken in the settler communities apart from the beginnings of lexical borrowing.

as the leading power on the subcontinent (cf. Spear 1965:81–92; Sinha 1978:150–1; Schneider 2007:163). The military power in Bengal was now in the hands of the East India Company, which used it to the benefit of British merchants, who, according to Spear 1965, “were given a free run of the country’s internal trade” (85). Clive’s activities also paved the way for the eventual transfer of authority over Bengal and Bihar from the Mughal rulers to the East India Company eight years later (Diwani Act 1765), which was followed by several years of intrigues and financial irregularities within the East India Company (cf. Spear 1965:86–7). The Regulating Act of 1773, which installed Warren Hastings as Governor-General, stabilized the situation when it came into force in the following year and has been seen as “the real beginning of British administration in India” (Spear 1965:88). With the passing of the India Act of 1784, the East India Company was finally under full political control from Britain, with a minister in London and the Governor-General in Calcutta supervising its activities (cf. Spear 1965:94). In the following years, English acquired importance especially for those Indians who were working for the British administration. This confirmed the presence of the English language in the domains of administration and bureaucracy (cf. Krishnaswamy/Burde 1998:81–7). In their analysis of written documents from the late eighteenth century, Krishnaswamy/Burde (1998:83–7) also point out that the English used by Indians around that time was largely “unstable English” (86), frequently marked by a “lack of clarity at the discourse level” (86):

Since there was no formal education in English available during the first phase, the inadequate knowledge of English, the use of Persian for royal, administrative, and legal purposes with its royal, ostentatious formality, and the deeply entrenched feudal mentality in the Indian subcontinent must have resulted in the ‘affective vocabulary’ which is found in the petitions written during this phase. The ornate style used in the earlier forms of English must have also contributed to this register. (Krishnaswamy/Burde 1998:82)

Those observations suggest that English, from an Indian perspective, was probably still widely perceived as a foreign language to be acquired directly from the British administrators, whose English set the norm. The administrators’ English itself had already started to incorporate loanwords from indigenous Indian languages to capture what Schneider 2003 calls “the additional flavour of the extraterritorial experience which those who stayed ‘home’ do not share” (246). But Lewis 1991 notes that the influx of Indian loanwords, which had begun in the early seventeenth century, had come to a temporary halt by the beginning of the eighteenth century, which suggests a phase of linguistic stabilization around that time:

With the passing of the seventeenth century, already richly abundant in its harvest of Anglo-Indian words of the common sort, and to a lesser degree of those belonging to more exalted spheres, the first half of the eighteenth century was marked by a sharp decline in the number of acquired Anglo-Indian expressions.

Anglo-Indian English had almost reached a temporary saturation point of the more immediately useful acquisitions. (Lewis 1991: 20)

It was only from the second half of the eighteenth century on that “a considerable addition of scholarly words” (Lewis 1991: 20) took place. However, as this probably did not affect the use of spoken English in the subcontinent significantly, it is likely that English was, by and large, still going through a phase of ‘exonormative stabilization’ around the year 1800, to use Schneider’s (2003, 2007) terminology, with the English of both the European settlers and the indigenous people continuing to be oriented largely toward British usage norms.⁷ This is also the view taken by Mukherjee (2007: 165), who evaluates the linguistic situation throughout that period in reference to Schneider 2003 by stating that “[d]espite the influx of Indianisms in the English language in India, the standards and norms of the English language in general ... remained British and, thus, exonormatively set” (165).

2.1.3 English takes root: 1813–1857

Historically, the geographical boundaries of India under British rule were set up in the first half of the nineteenth century (cf. Marshall 1996: 46–51). Linguistically, this period has been characterized as the time during which English became established further in numerous domains of life that, to date, continue as major domains of use for contemporary IndE: bureaucracy, commerce, the print media, academic and literary writing and, at the heart of the matter, education (cf. Krishnaswamy/Burde 1998: 87–98). When the Charter for the East India Company was renewed in 1813, education was recognized as the responsibility of the state. However, concrete political measures in this regard were taken only from 1823 onward (A. Gupta 1991: 66). The opening of several missionary schools and colleges in the subcontinent between 1780 and 1830 resulted in the arrival of more British settlers and was in part a reaction to the increasing demand for Indian clerks speaking and reading English (cf. Mehrotra 1998: 3).⁸ As

7. According to Schneider (2003: 245–7, 2007: 36–40), in the phase of exonormative stabilization, the settlers’ community stabilizes politically, which results in the stabilization of the position of English, for which an external usage norm still prevails. As Schneider 2003 puts it, “[t]he external norm, usually written and spoken British English as used by educated speakers, is accepted as a linguistic standard of reference, without much consideration given to that question” (245). In that phase, members of the STL strand are beginning to develop what Schneider calls an “English-cum-local” (246) identity. At the same time, more and more members of the IDG strand start to learn English. It is against the background of those developments that the structural nativization of English must be expected to commence: “in this phase the earliest structural features typical of local usage emerge, if only slowly” (Schneider 2003: 246).

8. English-medium schools were set up in Madras (Lady Campbell’s Female Orphan Asylum and the Male Asylum 1787) and in the south of India, e.g. Tanjore (1772), Ramnad (1785) and Sivaganga (1785), cf. Kachru (1994: 503).

a growing number of Indians became involved in handling the bureaucratic affairs of the Empire, English spread in India's urban areas and, having become a precondition for success, must also have gained in prestige. At the same time, education in local languages was losing ground (Parasher 1981: 330). The more widespread use of English in schools coincided with the introduction of the printing press (Masani 1987: 44; Sanjay 2006: 13), which set forth the publication and spread of textbooks in English, of creative writing in English by Indian pupils, and of other texts related to English education. Krishnaswamy/Burde 1998 see in those developments the very beginnings of English-language print-media communication in India:

The print-media was used for publishing essays of school boys, annual reports of colleges, and write-ups that deal with topics like criticism, defence of an English teacher and English education, bad influence of English education of Indian youth, proposals for publishing translations of some ancient Hindu texts, etc. ... English education, besides feeding into the bureaucratic domain, induced communication needs for which the space of print-media was made available to Indians. (Krishnaswamy/Burde 1998: 89)

By 1830, English had become a dominant language in public business and in government departments (A. Gupta 1991: 66). English, hence, was no longer a foreign language but beginning to develop into "Indians' English" (Krishnaswamy/Burde 1998: 89). To speak with Schneider 2003, this must have had "immediate linguistic consequences" (247) characteristic of the beginnings of the phase of 'nativization' in the life cycle of new varieties of English, for example the emergence of new discourse styles.⁹ Krishnaswamy/Burde 1998 observe:

Thus, English education and bureaucracy emerge almost simultaneously and become a decisive combination for the use of English in India. English education enabled Indians to use English, and the only way of putting it to use was to get a job in the government and write bureaucratic documents. The use of English in the bureaucratic domain was the beginning of Indians' English. It is not surprising, therefore, that the bureaucratic manner of writing dominated even the non-bureaucratic writing ... (Krishnaswamy/Burde 1998: 89)

Opinions on the desirability of English in India's education were not unanimous, however. A debate on the appropriate role of English between the 'Orientalists' (favoring

9. According to Schneider (2003, 2007), the third developmental stage of nativization sees a fundamental redefinition of the relationship between the STL and IDG strands, with the gap between the two narrowing, as "[b]oth parties regard themselves as permanent residents of the same territory (though differences in status, prosperity, and lifestyle persist) and thus realize the need to get along with each other." (Schneider 2003: 247) Linguistically, this phase is characterized by "the heaviest effects on the restructuring of the English language itself" (Schneider 2003: 248), especially in lexicosyntax.

Indian languages) and the ‘Anglicists’ or ‘Occidentalists’ (speaking in favor of English) had set in at the end of eighteenth century that divided British administrators and Indian intellectuals alike (cf. Edwardes 1967: 110–27; Sinha 1978: 151–2; Parasher 1981: 331; Carls 1987: 118–20; John 2007: 18; Schneider 2007: 164).¹⁰ The debate ended when Thomas Babington Macaulay, strongly convinced of the superiority of the Empire, passed his much-discussed *Minute on Indian Education* on February 2, 1835, in which he called for the use of English in order to train an English-speaking Indian class of interpreters:

I think it is clear that we are free to employ our funds as we choose; that we ought to employ them in teaching what is best worth knowing; that English is better worth knowing than Sanskrit or Arabic; that the natives are desirous to be taught English, and are not desirous to be taught Sanskrit or Arabic; that neither as the languages of law, nor as the languages of religion, have the Sanskrit and Arabic any peculiar claim to our encouragement; that it is possible to make natives of this country thoroughly good English scholars, and that to this end our efforts ought to be directed ... We must at present do our best to form a class who may be interpreters between us and the millions whom we govern; a class of persons, Indian in blood and colour, but English in taste, in opinions, in morals, and in intellect.

(qtd. in Edwardes 1967: 126)

Despite the disapproval from the Orientalist group, the Minute was accepted by the Governor-General. English, hence, became the official language of education and, two years later, of the government (Annamalai 1994: 263). The impact of those decisions was tremendous. Sociolinguistically, Macaulay’s Minute, according to Carls (1987: 120), finally shifted the status of English from being a foreign language to being an official second language. Socioculturally, it formed “the beginning of a systematic and widespread bilingual education in India throughout the nineteenth century” (Schneider 2007: 164) and encouraged the rise of an “Anglophone subculture” (Hickey 2004: 542), which, as Rahman 2002 reminds us, would eventually also turn English “into a marker of elitist schooling and the key to powerful government jobs

10. In this context, Raja Rammohan Roy’s letter to Lord Amherst, written in 1823, is frequently cited as one of the earliest examples of local Indian support for the spread of English (cf. Krishnaswamy/Burde 1998: 164–5). In his letter, Rammohan Roy “oppose[d] the proposal for a Sanskrit college in Calcutta on the grounds that such an establishment “can only be expected to land the minds of youth with grammatical niceties and metaphysical distinctions of little or no practical use to the possessors or to society”” (Aramavudan 2007: 41). In the frame of Schneider’s (2007) evolutionary model, Rammohan Roy’s letter to Lord Amherst represents “a direct expression of a characteristic identity construction of a member of the indigenous population (or at least its leadership), combining local roots with the putative advantages that the English language and through it access to western culture have to offer. In contrast, it is clear that the British in India saw themselves as representatives of their country and agents in her interest.” (Schneider 2007: 164)

and high social status” (169). Mukherjee 2007, applying Schneider’s (2003) model, sees Macaulay’s Minute as “the first step toward the beginning of the nativization of the English language in India” (Mukherjee 2007: 165). It must be noted, however, that some opposition and mistrust to the use of English would remain long after the Minute had been passed, especially among Muslims:¹¹

Muslims did not respond to the new education with any enthusiasm. In fact, a vigorous protest was signed by eight thousand Calcutta Muslims when English was made the official language of government. Although an English-language class had been established in the Calcutta Madrasa as early as 1826, only two students passed the junior scholarship examination during the next twenty-five years. It was only after the Mutiny of 1857 that the Muslims became aware of the advantages they had let slip. (Edwardes 1967: 118)

2.1.4 From ‘English in India’ to IndE: 1857–1947

The Great Rebellion of 1857 and 1858 (the so-called Indian Mutiny), a massive revolt against British rule that had started out with a rising in the Bengal army and spread all across northern India, resulted in the abolishment of the East India Company and the handover of control to the Crown (Spear 1965: 145–57; Marshall 1996: 50). According to Spear 1965, the rising had important psychological consequences for both rulers and ruled. The British had suffered “a severe shock to the self-confidence” (143) which had characterized their attitude in the first half of the century. In Spear’s (1965) words,

[b]y and large there was a change of attitude in two directions. It was realised that the government should be closer in touch with, and more sensitive to, Indian opinion, particularly the established classes who could control the general mass of the people, and there was a new caution in implementing the westernizing policy. (Spear 1965: 144)

The Indian side, on the other hand,

liked the impinging west no better but realized that it now had to be lived with and absorbed into Indian life, as many foreign intrusions before. As a corollary it lost confidence in the traditional leaders, rajahs, chiefs, *zamindars*, and began to turn, without openly admitting it, to the new westernized class.

(Spear 1965: 144)¹²

11. Cf. Rahman (2002: 168–202) for an analysis of the attitudes of Muslims in north India toward English from its introduction to the present day.

12. A *zamindar* is “a landlord responsible for collecting money and paying it to the government” (Crowther 1996: 1475).

This outcome, a gradual coming to terms of rulers and ruled, was significant for the future development of English in India in that it enabled further-reaching nativization. The nativization of English in a new territory, according to Schneider (2003, 2007), presupposes awareness by settlers and indigenous people alike of a fundamental change in their relationship toward one another. This is what had taken place in India after the Great Rebellion. As Schneider 2003 puts it, nativization is

the central phase of both cultural and linguistic transformation in which both parties involved realize that something fundamental has been changing for good: traditional realities, identities, and socio-political alignments are discerned as no longer conforming to a changed reality, and the potentially painful process of gradually replacing them with something different, a new identity reflecting a changed reality, combining the old and the new, is in full swing. This process has immediate linguistic consequences, for the drastically increased ranges of communication between the parties involved now make language use a major practical issue and an expression of new identity. (Schneider 2003: 247)

In the course of restructuring and expanding the bureaucracy of British India after 1858, more government departments opened up across the subcontinent, which led to an increase in the number of English-using Indians. This development was furthered when the codification of a uniform legal system, begun in the early 1860s and conducted in English, progressed (cf. Masani 1987: 18–22). The growing need for English-educated Indians around that time is also reflected in the rising number of English-medium schools and colleges. In the year of the rebellion, in 1857, three universities had been founded in Bombay (now Mumbai), Calcutta (now Kolkata) and Madras (now Chennai). Functioning as examining boards, those universities also supervised the educational institutions in their vicinity:

By 1859, Calcutta had eleven affiliated colleges, Bombay two, and Madras one. Though statistics were by no means accurate for all schools, they show that there were thirteen government colleges with 1,909 students and four aided colleges with 878 students; seventy-four superior government schools with 10,989 students, and 209 aided schools of the same or somewhat lower grade, with 16,956 students; twenty-five normal schools with 2,241 scholars; and sixteen colleges for special subjects, with 1,154 scholars. (Edwardes 1967: 236–7)

In 1882, Kachru (1994: 507) reports, more than 60 percent of the primary schools in the major cities were English-medium. The spread of English-medium schools, colleges and universities ultimately also encouraged the use of English in more social domains. Debating societies and clubs, in which English became the primary means of communication among Indians, sprang up (cf. Masani 1987: 51–70; Annamalai 2001: 103–4). In this environment, too, Indian literature in English took root (cf.

Mehrotra 1989; Tharu 1994; Paul 2003; Dayal 2006; Singh 2006).¹³ The spread of English in the second half of the nineteenth century also extended to the sphere of the media, as more and more English-language newspapers were founded in the major cities across the subcontinent.¹⁴ Normally run by British owners, those newspapers had a small but highly influential and politically active readership (cf. Masani 1987: 44).

Krishnaswamy/Burde 1998, on analyzing written documents from the late nineteenth and early twentieth centuries, come to the conclusion that by the end of the nineteenth century the use of 'Indians' English' had become firmly established in the domains of bureaucracy and administration. But it had also begun to grow into a vehicle for personal communication among English-educated Indians and into a medium deemed appropriate for creative writing. Indians, it appears, had developed an ease at using English in formal and increasingly also in informal contexts. This widening range of functions to which English was put to use supports the view that the end of the nineteenth century can be considered as marking "the beginnings of the evolution of "educated" Indian English" (Mukherjee 2007: 167). Interestingly, there is also evidence for growing awareness among the British in India that new varieties of English were emerging further down the lectal range. However, as the following entry in Yule/Burnell's (1994 [1886]) *Hobson-Jobson* exemplifies, the attitude of the rulers toward varieties such as 'Butler English' (cf. Chapter 2.2.1) was clearly one of condescension:¹⁵

BUTLER-ENGLISH: The broken English spoken by native servants in the Madras Presidency; which is not very much better than the Pigeon-English [sic] of China. It is a singular dialect; the present participle (e.g.) being used for the future indicative, and the preterite indicative being formed by 'done'; thus *I telling* = 'I will tell'; *I done tell* = 'I have told'; *done come* = 'actually arrived.' Peculiar meanings are also attached to words; thus *family* = 'wife.' The oddest characteristic

13. Cf. Dayal 2006, who points out that while "a tradition of creative writing [in English] has also gradually developed since the founding of Hindu College in Calcutta in 1817" (304), it was only after founding the University of Calcutta in 1857 that "the link between new learning and a new literature" (304) strengthened.

14. The following major newspapers were founded in the nineteenth century: *The Times of India* (1838), *The Statesman* (1875), *The Hindu* (1878), and *The Tribune* (1881); cf. Beliya (2001: 11). India's first English-language newspaper had been Hickey's *Bengal Gazette*, founded in 1779 and first published in 1780 (cf. Sanjay 2006: 13; Singh 2006: 42).

15. Cf. Aravamudan 2007: "An Indian form of English – and therefore its development as a new South Asian vernacular rather than just as an imperial echo – first acquired recognition, paradoxically, when representatives of high Victorian imperialism dismissed it as bureaucratic cant of the native functionaries and interpreters of the Raj, a "Baboo English" or "Cheechee English," to be literally ridiculed and disparaged. Even lesser variants began to be recognized, such as Butler English, Bearer English, Box-Wallah English, Kitchen English, and Hinglish (Hindi-English) ..." (4)

about this jargon is (or was) that masters used it in speaking to their servants as well as servants to their masters. (Yule/Burnell 1994 [1886]: 133–4)

Despite its increasing number of users, English remained a minority language throughout the nineteenth century. In Krishnaswamy/Burde's (1998) words, it was also "the restrictedness of the use of English in India" (99) that got stabilized during that period.

The use of English saw further consolidation in the domains of education and bureaucracy after 1900. The need for English-speaking Indians continued to surge as the Indian Civil Service, the postal and the railways services expanded (Carls 1987: 121). The number of English-language newspapers was rising further, too, as was the number of their readers. The increase in sales figures brought about an increase in the number of advertisements, which in effect resulted in more English-using "editors, journalists, reporters, correspondents, printers, publishers, writers and readers" (Krishnaswamy/Burde 1998: 111). But English was acquiring yet another function, as those who were opposed to British rule began to use English against the rulers themselves. The National Congress had been founded in 1885 and had spread across India by the turn of the century (cf. Spear 1965: 170). Its members were using English as their primary means of communication to share their ideas across the subcontinent (cf. Carls 1987: 121; Krishnaswamy/Burde 1998: 112; John 2007: 123–57). At the same time, the number of English-educated college students went up from an estimated 36,000 students in 1911 to approximately 145,000 students by 1939, which provided the nationalist movement with a steadily growing number of followers (Masani 1987: 89–106). As Masani 1987 explains,

[b]y the 1920s it was clear that the very education with which the Raj had hoped to train loyal Indian subordinates had turned against it in the demand for self-government. This development was not altogether unexpected. It had been anticipated in the fierce debate between the protagonists and opponents of Western education during the first half of the nineteenth century; and Britain's more enlightened colonial educationists had always welcomed the prospect of Indians some day graduating through British education to democratic self-government. What they had not anticipated was that the pace and manner of India's political graduation might be beyond British control. (Masani 1987: 89)

The position of English was only called into question again openly when Gandhi advocated the use of local languages in education (Sinha 1978: 153). As a result, numerous institutions opened up across the country, which aimed at promoting indigenous education in the years before Independence (Mehrotra 1998: 5). This, according to Mehrotra 1998, led to a temporary estrangement with English among parts of the educated elite that would come to a halt only after India's Independence in 1947 when Hindi was declared the official language of the Indian Union.

2.1.5 The institutionalization of English after 1947

In post-Independence India, English passed through an eventful history of rejection and acceptance after 1947 before it would gain permanent official recognition. The National Movement under Mahatma Gandhi had called for Hindustani as a national language to foster integration and promote a sense of unity (Sinha 1978: 153–4; Svartvik/Leech 2006: 118; Schneider 2007: 166).¹⁶ Eventually, Hindi (written in the Devanagari script), mother tongue to more Indians than any other language spoken on the subcontinent, was assigned the status of national language in the new constitution (Parasher 1991: 33).¹⁷ However, since English had become established firmly in the domains of administration and bureaucracy and as the language of higher education and the legal system, it could not be replaced instantly. As Agnihotri/Khanna 1997 sum up succinctly,

[b]y the time India became independent, English had already consolidated its position in the school and university education. Leaders of the freedom movement – Gandhi, Nehru, Tilak, Gokhale and Bose among others – used it extensively in their discussions and writings though they could reach the masses only in their own languages. Creative writing had also discovered Indian voices in Mulk Raj Anand, R. K. Narayan and Raja Rao among others.

(Agnihotri/Khanna 1997: 30–1)

When the new constitution was passed in 1950, it therefore made the provision that English was to be used as an additional language for another fifteen years before replacing it with Hindi, which itself was to undergo planning and modernization to meet the demands of its new role (Carls 1987: 122; Vaish 2008: 18). With the year 1965 nearing, however, resistance to Hindi as sole national language became fiercer – especially in southern India, where Dravidian languages were spoken and English was perceived as a more neutral linguistic choice than Hindi.¹⁸ As a result, an “unplanned

16. Cf. Kachru 2008: “Hindi and Urdu have a common form known as Hindustani which is essentially a colloquial language (Verma 1933). This was the variety that was adopted by Mahatma Gandhi and the Indian National Congress as a symbol of national identity during the struggle for freedom. It, however, never became a language of literature and high-culture” (81). Note that while Gandhi promoted Hindustani, “his message to the elite was generally expressed in English.” (Gargesh 2006: 93)

17. According to the 1991 Census, Hindi was spoken by approximately 337 million Indians as a first language, about one third of India’s population (cf. Vijayanunni 1997: 11; Annamalai 2008: 225).

18. In the language debates of the 1950s and 1960s, the development of Hindi was often criticized for its extensive borrowing from Sanskrit. See the following anecdote reported by L. K. Advani: “Pandit Nehru used to frequently take exception to the Hindi used by Akashvani (All India Radio). He would complain that the Hindi was too complicated and could not be

development” (Annamalai 1994:261) set in. In 1963, two years before the projected date for abolishing English, Parliament passed the Official Languages Act, which provided for the continuing use of English “as co-official language” (Vaish 2008: 19). But hostility toward Hindi in the South, which had resulted in a series of violent protests in previous years, went on and culminated in January 1965, when language riots took place in Tamil Nadu that cost dozens of Indians their lives (Annamalai 1994:266; Mehrotra 1998:6; Vaish 2008:23). Mukherjee (2007: 168) even interprets those outbursts of dissatisfaction in the framework of Schneider’s (2003) model as an ‘Event X’ – “some exceptional, quasi-catastrophic political event” (Schneider 2003:250) catapulting English into the fourth evolutionary phase of ‘endonormative stabilization.’¹⁹ In his view, it was “the language riots [that] made the political parties readjust their stance on language policy and ensure the continuing use of the English language in India” (Mukherjee 2007:168). The Official Languages Act, when passed in 1967 in its amended form, finally assigned English the permanent status of associate official language of the Indian Union that was to be used for the proceedings in the Supreme Court (the highest court in the Indian Union), in the High Courts (the highest courts in the states) and in Parliament. Moreover, all legislative and judicial documents in the Union were to be drafted in Hindi and English (Annamalai 1994:266).²⁰ Hence, English, as a language of administration, had come to stay. Today, next to its legal status as associate official national language, English is the official language of four Union states (Manipur, Meghalaya, Nagaland and Tripura) and of eight Union territories that are under direct control from New Delhi (the Andamar and Nicobar

understood by the common people. ‘We should use only simple Hindi,’ he would insist. By this he generally meant that there should be more Urdu words in it because, according to him, Urdu was more widely understood in North India.” (Advani 2008: 134) Those favoring the Sanskritization of Hindi, on the other hand, would point to the “immense richness and adaptability of Sanskrit” and the fact that “in at least three of the four South Indian languages – Telugu, Kannada and Malayalam – there are many Sanskrit words that are common to Hindi” (Advani 2008: 135). However, skepticism toward using Sanskritized Hindi prevailed among Hindi speakers as well, who would often continue to prefer English.

19. According to Schneider (2003, 2007), the phase of endonormative stabilization is characterized by “the gradual adoption and acceptance of an indigenous linguistic norm, supported by a new locally rooted linguistic self-confidence” (2003:249). Schneider 2003 argues that this phase typically follows political independence and may be triggered by a “quasi-catastrophic political event” (250). Local usage norms are accepted as appropriate also in formal contexts, the new variety now functioning as a vehicle for a “new regional identity” (251).

20. The Official Languages Act also recognized major regional languages as national languages to be used within the Union states, for example as languages of the administration or in the state educational systems (cf. D’souza 1987:63–4).

Islands, Arunachal Pradesh, Chandigarh, Dadra and Nagar Haveli, Delhi, Lakshadwip, Mizoram and Pondicherry; cf. McArthur 2002: 313).²¹

Another response to the growing resentment against Hindi in non-Hindi-using areas was the so-called Three Language Formula, which has become a crucial part of the Indian educational system, although its implementation has proved difficult. Originally proposed in 1956, amended in 1961 and finally accepted in 1968, the Three Language Formula prescribes the compulsory teaching and learning of one indigenous regional language (the mother tongue), of Hindi (for non-Hindi speakers) or another modern Indian language (for Hindi speakers), and of English in schools across India (Aggarwal 1988: 290; Kapoor 1993: 90–1; Biswas 2004: 107; Vaish 2008: 14–16).²² The position assigned to English in the Three Language Formula is, therefore, neither that of an Indian language nor that of a foreign language; English is to hold a special status in between. Today, although English is no longer the compulsory medium of education at the primary and secondary levels, English-medium education is highly popular and available throughout the subcontinent, especially in privately-run schools. In all states and Union territories, English is an alternative medium of education for primary and secondary education (cf. Annamalai 1994: 269–70). In universities around the country, the number of students learning English has also increased steadily over the years – a development which has been accelerated by the growing importance of English as an international language of science, technology, politics and commerce (cf. Agnihotri/Khanna 1997: 139–44).

As regards the development of English in the Indian media after Independence, a look at India's national press is insightful as it allows a glimpse of the current position of English alongside India's indigenous languages (cf. Mehrotra 1998: 11; Kachru

21. The development of Hindi as an official language has proceeded simultaneously, cf. Vaish 2008: "In 1950 the Ministry of Education sponsored the Board of Scientific Terminology 'which was assigned the job of preparing 350,000 new terms in Hindi, of which 290,000 were produced by 1963' (Dasgupta 1970: 165). Today this board exists as the Commission for Scientific and Technical Terminology. Similarly, the Kendriya Hindi Samiti (Central Hindi Committee) and the Hindi Salahakar Samiti (Hindi Advisory Committee) help to propagate and develop Hindi, reviewing the progress made in Hindi as an official language in various government ministries, and submit reports. The Department of Official Language publishes Rajbhasha Bharti (India's language), manuals, posters and calendars depicting the use of Hindi in various spheres. There is a Central Translation Bureau to translate manuals, codes and nonstatutory literature for government offices and banks, from English to Hindi and vice versa ..." (Vaish 2008: 19–20).

22. There are marked differences in the ways that individual states have implemented the Three Language Formula. Cf. Vaish 2008: "[T]he teaching of Hindi is obligatory in all Indian states except in Tamil Nadu, Tripura and the Karaikal region of Pondicherry. Similarly, though the teaching of English is obligatory, the state of Bihar refuses to comply. The state of West Bengal refuses to teach Hindi. The stage at which the second and third languages are introduced is dependent on the resources and ideology of the state. In general English is compulsorily taught in Grades 6–9." (15)

2005: 15–6; Gargesh 2006: 100–1; Sanjay 2006: 20–3). According to official statistics, newspapers in India in the year 2000 were published in 101 languages (cf. Beliya 2001: 2–9). Of those, the highest number of newspapers was published in Hindi (19,685). English-language newspapers came in second in number (7,175), two thirds of which (4,437) were published in the four major metropolitan cities Delhi, Kolkata, Mumbai and Chennai. In terms of circulation, Hindi newspapers were in the lead in the year 2000 (54,439,135 copies) followed by English newspapers (22,962,607 copies). Among India's dailies, the number of English publications (390) came in third after Hindi (2,393) and Urdu (525) dailies. In the category of 'single-edition daily newspapers', the *Hindustan Times* (published from Delhi) had the highest circulation figures (847,346 copies) followed by *The Hindu* (79,851 copies, published from Chennai). In the category of 'multi-edition dailies', *The Times of India* was the most widely sold newspaper (168,799 copies). Though the circulation figures for English newspapers are, hence, lower than those for Hindi newspapers, it has been pointed out that English-language newspapers tend to reach the more affluent sections of Indian society. In the early 1980s, 53% of the share in advertisement revenues went to English-language newspapers as opposed to 15% to Hindi newspapers (Parasher 1981: 336–8; Annamalai 1994: 273). In this light, there can be no doubt that English-language publications play an influential and important role within the mass media in post-Independence India.

Mehrotra (1998: 11–2) identifies nine domains, in which English is nowadays employed in India on a regular basis. The first three domains of use listed by Mehrotra 1998 are the traditionally established domains of 'trade and commerce' (e.g. in correspondence or when reading financial newspapers and magazines), 'administration' (e.g. when writing official notes, reports or announcements) and 'education' (e.g. when lecturing in the classroom). In addition, Mehrotra 1998 names a further six domains to document the wide functional range that contemporary IndE has acquired, namely 'family' (e.g. when talking to people from other states and countries or interacting with people at gatherings), 'recreation' (e.g. when watching TV programs), 'hotel and restaurant' (e.g. when interacting with guests and customers), 'sports' (e.g. when listening to sports commentary or reading sports magazines), 'politics' (e.g. when delivering lectures, issuing orders and instructions) and 'religion' (e.g. when reading religious works). What a survey like Mehrotra's (1998) does not capture, of course, is the very embeddedness of English in a multilingual context, which is bound to render any allocation of English to single domains or functions an incomplete representation of the complex linguistic realities in India. From a user's perspective, this reality may look like this:

A multilingual educated Indian (meaning thereby an 'English-knowing' Indian) uses his native mother-dialect with close relatives and in intimate family circles, he uses the major regional language while interacting with the 'ordinary folk', acquaintances, vendors and traders, but when he enters a deluxe supermarket or the lobby of a five-star hotel he uses English. If he travels by bus or by ordinary

rail-coach he uses one of the regional Indian languages, but when he travels first-class or on a plane he uses English and nothing but English.

(Kapoor/Gupta 1991: 16–7)

However, English in post-Independence India is also more than Kapoor/Gupta's (1991) poignant comment on English being the language of the 'deluxe supermarket' and the 'five-star hotel' suggests. Kapoor/Gupta 1991 are certainly right in implying that English, due to its being a dominant language of administration, education and the media, has emerged as the language of a highly influential group of Indians who are often referred to as India's 'educated elite'. This has frequently resulted in associating English rather superficially with power, success and prestige.²³ But it must also be acknowledged that English, the chosen medium of communication of millions of Indians, has made noticeable inroads into their lives in less formal domains of use, for example 'friendship'. This is indicated by the few available systematic studies on the use of English by bi- and multilingual Indians (cf. Parasher 1980; Sridhar 1982; Sahgal 1991; Agnihotri/Khanna 1997). Parasher 1980, for example, studied the use of English by 350 bilinguals of various language backgrounds living and working in Hyderabad and Secunderabad (aged 25 to 45). In communication situations falling in the domain 'family', the informants clearly preferred their mother tongues. In the domain 'friendship', on the other hand, English was the dominant language, which Parasher 1980 related to the possibility of English being the only shared language in more extended multilingual friendship circles. Similar results were produced by Sridhar 1982, who investigated the use of English, Hindi and Kannada (the informants' mother tongue) by 299 university students and 88 employees in Karnataka (South India), and by Sahgal 1991, who obtained information from 45 speakers from three ethnically different communities living in Delhi with Hindi, Bengali or Tamil as their first languages. Interestingly, in Sahgal's (1991) study, English also emerged as an important link language in the domain 'neighbourhood' for Tamilians and Bengalis, whereas Hindi speakers primarily used Hindi. Sahgal 1991 explained this by

23. This has been criticized by Krishnaswamy/Burde 1998, who claim that English has not become an Indian language, since, in their view, it is used only by an urban minority elite: "English and English education have been with a section of the urban population so long that it has uprooted quite a few, and the result is that they are neither here nor there; this section of the population has no mother tongue and no cultural roots in the conventional sense; they are comfortable only with English but that English is not 'native'" (127). Similar critical assessments of the position of English in India are found in Dasgupta 1993, Mathew 1997 and Tully 1997. The opposite view is represented, for example, by D'souza (1997, 2001b) and Das 2002, who argue in favor of looking at English as an Indian as well as global language. According to D'souza 1997, "English now belongs to the world, and multilinguals who speak English do not need to apologize for not being monolinguals ... English can only be understood if it is seen as indigenous" (102). Similarly, Das 2002 calls on Indians to "accept English as our own language" (22) and demands further: "Everyone must learn it along with his mother tongue. Let us shed our hypocrisy and eliminate the distinction between English and vernacular-medium schools." (22)

stating that Bengali and Tamilian migrants in Delhi, more than Hindus, tended to live in mixed neighbourhoods. Sahgal 1991 also noticed that English was more of a second choice in the 'family' domain for Tamil and Bengali speakers than for Hindi speakers, which was interpreted as evidence that English was causing a reallocation of languages in the repertoire of migrants living in Delhi (cf. Sahgal 1991: 302–3). As for language use in the public sphere, Parasher 1980 and Sridhar 1982 both found that English was a dominant language for a great majority of their respondents. In Parasher's (1980) study, English emerged as the language used most frequently in discourse on topics related to work, education, science, and technology, and it was the preferred language choice in several situations characterizing everyday life, for example when ordering food in restaurants, when talking to a doctor, or when making enquiries in public places. Supplementary interview data indicated that English was perceived by many respondents in this study as the most straightforward (and possibly also the most prestigious) way of drawing attention in such situations. In addition, according to Sridhar 1982, one of the most important motivations for students to learn English was their hope to improve their chances for later employment: 52% of the students expected their job prospects to be excellent if the medium of instruction was English vis-à-vis 7% for Hindi and 28% for Kannada (cf. Sridhar 1982: 144).

Throughout the 1990s, India saw an upsurge of interest in learning English, which has been interpreted as a consequence of the economic reforms of 1991 and the growing need for English-speaking Indians by multinational companies (cf. Varma 2004: 96–132). Agnihotri/Khanna 1997, who conducted a questionnaire survey in 1993/1994 with 1,128 informants of various age groups from across the country, came to the following conclusion:

As compared to old people, young students are more positively inclined towards English and have stronger instrumental motivation to learn English. They also use more English, claim higher levels of proficiency and recommend the use of English in different spheres of education more strongly. This clearly shows that the perception of the importance of English has increased in the post-independence period and the young people see it performing new roles.

(Agnihotri/Khanna 1997: 141–2)

Since the beginning of the new millennium, English has been developing further along those lines in a complex network of local and global communicative needs. Today, as Sheorey 2006 observes,

English seems to have become so entrenched a language and is such an integral part of India's multilingual, urban culture that it can hardly be ignored or dislodged. Even the most vociferous pro-English and anti-English voices have been quietened considerably, if not totally silenced, by a sense of indispensability of English in the national interest.

(Sheorey 2006: 17)

The present situation of English in India is therefore that of a constitutionally recognized additional language that is used as an intranational link language by middle- and

upper-class Indians in public domains such as education, administration, business and the media, but also in more private domains among family, friends and neighbors. English is viewed as a mark of culture and prestige and regarded as “the pathway to modernity” (Aravamudan 2007: 5), as a “window to the world” (Das 2002: 22) and as “a source of great competitive advantage in the global economy” (Das 2002: 22). In fact, the prominent role of English in global commerce, in science and technology and in the electronic media has been reinforcing the use of ‘international English’ in India at a staggering pace, and it has even been speculated that India might soon have the largest number of English speakers in the world (cf. Das 2002: 19; Varma 2004: 115).

2.2 Previous descriptions of IndE

Due to its contact with Indian languages and its use by Indians of varied linguistic backgrounds, English has undergone structural nativization, i.e. processes of change that have affected “the structure of language (e.g., sound system and rhythmic patterns, vocabulary and sentence structure) and use of language (e.g., conventions of speaking and writing)” (Kachru/Nelson 2006: 31). At the same time, English has remained embedded in an international context. Various metaphors and models have been suggested to describe the current position of IndE between local and global communicative needs. Leitner 1992a, for example, argues that IndE, like other new Englishes, is nowadays “caught in a web of conflicting tensions” (Leitner 1992a: 225). He identifies four such tensions that are said to determine variation and change: (1) ongoing orientation of users toward their parent variety, resulting in “linguistic continuity” (225); (2) ongoing nativization caused by “social, cultural, and environmental demands” (225), resulting, for example, in processes such as borrowing and code-mixing; (3) ongoing nativization caused by learner situations, resulting in “the independent creation of linguistic forms” which “may, but need not, lead to unsystematic differences between varieties of English” (225); and (4) global linguistic pressures arising “from similar needs in all societies, which are reinforced by the possibilities of technology and communications” (226–7). In this web of tensions, Leitner 1992a argues, users must decide whether they want to orient themselves toward internal or external usage models. Such decisions will not simply be idiosyncratic in nature, but be determined “by the social needs and benefits that are associated with the options” (Leitner 1992a: 227). Regional varieties like IndE are eventually shaped by the decisions taken in these situations. Similarly, Mukherjee 2007 describes the situation in IndE in terms of “conflicting forces” (174) acting on the variety today. He distinguishes between ‘progressive’ and ‘conservative’ forces at the linguistic, functional and attitudinal levels. Progressive forces, according to Mukherjee 2007, are innovative in that they encourage the use of English and, thus, contribute to the emergence and acceptance of local norms of usage. Such forces are, for example, seen at work in society in the use of English for creative writing and as a medium of instruction in Indian schools (cf. Mukherjee 2007: 175–6). Structurally, the emergence of local usage norms can

nowadays be felt strongly in pronunciation and the lexicon, for example in the use of “many loanwords” (Mukherjee 2007: 174) or “new lexical items and compounds” (174). In syntax, such effects are said to be less obvious. Conservative forces, on the other hand, work at the functional and attitudinal levels to discourage the use of English in India, which in effect also hinders users from accepting local norms. In modification of Schneider’s (2003, 2007) evolutionary model of new Englishes, Mukherjee 2007 also hypothesizes that IndE has reached the stage of endonormative stabilization but is not likely to diversify further regionally or socially, as the co-existence of progressive and conservative forces has resulted in what he interprets as a state of equilibrium:²⁴

[I]n spite of all the creative and innovative processes at all linguistic levels, Indian English will never become the dominant language in India with the vast majority of Indians turning into a predominantly English-speaking people, and will thus never enter phase 5, that is, differentiation, although dialectal variation may to some extent occur. (Mukherjee 2007: 182)

In this “web of conflicting tensions” (Leitner 1992a: 225), the label ‘Indian English’ itself has been applied in different ways with different underlying attitudes by language professionals and non-professionals alike. There have been ‘progressive’ voices using the name ‘Indian English’ in a unifying sense to refer to all the English sounds, words, meanings and structures spoken and written, heard and read in India today – often to emphasize that English, in the context of Indian multilingualism, has evolved as a viable regional brand of English capable of fitting the communicative demands of its Indian users without aping any other nation’s English. See, for example, David Crystal’s lively account of his impressions of IndE after his visit to India in 2005:

A couple of month ago, my wife and I returned from a two-week lecturing tour of India, sponsored by the British Council. We visited Chennai (or Madras, as it used to be called), Delhi, Kolkata (or Calcutta), Pune, and Mumbai (or Bombay), and we found ourselves surrounded by English everywhere. The roads into the city centers from the airports would pass through some very poor areas, but even the smallest shops and stalls would have an English sign or poster nearby. Nor were the slums exempt: on the corrugated walls of a straggling Mumbai complex was a series of ads for vitrified tiles, all in English. Outside the Red Fort, a Hindi-speaking teacher was marshalling a class of 30 Hindi-speaking teenagers, and giving them instructions about where to meet and when their bus would

24. According to Schneider (2003, 2007), the final evolutionary phase of ‘differentiation’ sees the emergence of sub-varieties within the newly born national variety, which function “as carriers of new group identities within the overall community: regional and social dialects, linguistic markers (accents, lexical expressions, and structural patterns) which carry a diagnostic function only within the new country” (2007: 54). Phase 5 is, thus, best seen as “a vigorous phase of new or increased, internal sociolinguistic diversification” (Schneider 2007: 54).

leave, but all in English. Outside St. Thomas' Cathedral in Chennai we met a group of primary-school Tamil children coming out of the local school. As soon as they saw us they waved excitedly – we were the only fair-skinned people to be seen – and we received a chorus of 'Hello', 'Hi', 'How are you?' 'Fine, thanks, how are you?' we replied. 'Oh, we're fine too,' they said. Seven-year-olds, we marvelled, on a confident career-track towards English. Towards Indian English, of course. India has had a longer exposure to English than any other country which uses it as a second language, and its distinctive words, idioms, grammar, rhetoric and rhythms are numerous and pervasive. (Crystal 2005)

Others have used the label 'Indian English' in a much more 'conservative' sense in reference to a type of English that is characterized predominantly by errors. The implicit claim here is that 'Indian English' is the result of a dangerous lowering of language standards in the post-Independence era that might threaten international intelligibility and is itself symbolic of a major identity crisis. The following comment by the former Press Secretary of the President of India, Pavan K. Varma, illustrates this stance:

Indian English is littered with instances of spelling errors, grammatical mistakes, pronunciation howlers and incorrect phraseology. The continued alienness of the language is brought out by constructions such as this one, prominently displayed in a photo studio in Srinagar: WELCOME TO BE BLOWN UP IN KASHMIRI DRESS. The author of that invitation could – like many others – read and write English, but knew nothing of the language. Those who think they know English take pride in their distance from their mother tongues; those who speak English inadequately claim to know it well. The result is a nation of linguistic 'half-castes', insecure in English and neglectful of their own mother tongue.

(Varma 2004: 114)

What, then, is IndE, and how can it be described?

Any linguistic description seeking to characterize IndE must take cognizance of its highly variable nature, as it comes "in a range of varieties, both regional and social" (Schneider 2007: 168). According to Kachru 1994, there are three main factors creating variation in what he terms 'South Asian English': users' proficiency (measured in terms of the parameters 'language acquisition' and 'years of instruction'), users' regional backgrounds and the influence of those regions' dominant languages on their English (which has given rise to varieties such as Tibeto-Burman Indian English; cf. Wiltshire 2005), and users' ethnic backgrounds (which allows one to speak of varieties such as Anglo-Indian English, the native variety of English used in India; cf. Coehlo 1997).²⁵ The influence of those factors has resulted in a cline of proficiency, ranging from 'educated South Asian English' at one end to 'Broken English' at the other. Along

25. Kachru 1994 uses the term South Asian English as a cover term for the varieties of English spoken in South Asia, recognizing that there are several varieties within this variety, for example IndE or PakE (cf. Kachru 1994: 508).

this cline, various sub-varieties of IndE can be located, for example Babu English, Butler English or Boxwallah English. Babu English is a highly formal variety of contemporary IndE with roots in the English spoken by clerks in Bengal in the nineteenth century (cf. Chapter 2.1.4). It is nowadays said to be used “in most of north India, in Nepal, and in some circles in South India” (Kachru 1994: 509). Butler English was spoken in the days of the British Empire in the Madras Presidency by servants in English households (cf. Chapter 2.1.4) and continues today as a restricted contact variety which can be encountered “in major metropolitan cities in South Asia where English-speaking foreigners live” (Kachru 1994: 511; also cf. Hosali 1997, 2004, 2005 and Mehrotra 2000). Boxwallah English is a subvariety of Broken English used by door-to-door sellers of goods in hotels and well-off neighbourhoods in metropolitan areas across South Asia. Kachru 1994 points out that the very presence of registers such as Babu English, Butler English and Boxwallah English must be seen as evidence for the high degree of institutionalization of English in South Asia today, i.e. its wide range of functions and its considerable social penetration. He also conceives of Babu English and Butler English as two poles of a lectal range, between which ‘educated South Asian English’ is to be positioned, ranking “high” (Kachru 1994: 513).

It is the educated variety of contemporary South Asian English used in India that the present study aims to describe. According to Kachru’s (1994) model, educated IndE is determined by the high proficiency of its users, its wide range of functions in Indian society, and supraregional intelligibility. Another appropriate name denoting the usage range under investigation in the present book would be ‘standard IndE’, with the notion ‘standard’ defined in the sense described by Kortmann/Schneider 2008 as that

variety, or set of closely related varieties, which enjoys the highest social prestige. It serves as a reference system and target norm in formal situations, in the language used by people taking on a public persona (including, for example, anchorpersons in the news media), and as a model in the teaching of English ...

(Kortmann/Schneider 2008: 2)

Whenever the label IndE is used in the present study, it is intended as a reference to this usage range.

This chapter looks at previous approaches to IndE. First, two descriptions of IndE structure and usage are discussed (2.2.1 and 2.2.2). This is followed by a closer look at the treatment of IndE in two handbooks intended for Indian users of English (2.2.3). The two descriptions and the two guidebooks have been valuable in the context of the present study in terms of generating numerous working hypotheses on the impact of nativization in the lexis and syntax of IndE. Afterwards, the available corpus-linguistic research on IndE is surveyed (2.2.4) before the conclusions regarding the tenets of the present corpus-based description of IndE are drawn (2.2.5).

2.2.1 Describing IndE from within

What is known about the distinguishing features of the local IndE usage norm? Almost thirty years ago, Pattanayak 1978 passed the following critical remark: “Unfortunately, the present dominant tendency seems to be to lump all the local peculiarities under the label ‘Indian English’ without examining their relevance to the totality of the speaker’s system and the domain of their use.” (Pattanayak 1978: 186) As has been pointed out in the Introduction, the quality of many of the available descriptions of IndE has not improved significantly since Pattanayak 1978 made this comment. All too often, descriptive approaches to IndE structure and usage have not progressed beyond the stage of compiling lists of features that are claimed to represent variety-specific usage patterns but lack information on their distributional properties within and across different contexts of use. Let me illustrate in reference to one concrete example from those lists, namely the use of invariant tag questions in IndE, that this is, in fact, a shortcoming.

According to the *Cambridge History of the English Language* (Kachru 1994), educated South Asian English usage is characterized by a restriction in the choice of tag questions, which is explained in terms of L1 interference:

The structure of tag questions in South Asian English is identical to that of many other non-native institutionalised varieties of English (e.g. West African, South-east Asian). In the native varieties of English, the tag question is attached to a statement. There is a contrasting polarity in such structures: a positive main clause is followed by a negative tag and vice versa. The parallel structure, such as in Hindi-Urdu, consists of a single clause with a postponed particle *na*. In British English the tag questions form a set, out of which an appropriate choice has to be made according to the context. In South Asian English generally, that choice is restricted to *isn’t it*. *You are going tomorrow, isn’t it?*, *He isn’t going there, isn’t it?*

(Kachru 1994: 519–20)

While invariant tag questions have frequently been attested for IndE and are certainly not unheard of in this variety (cf., e.g., Daswani 1978: 122; Pattanayak 1978: 190; Kachru 1986: 40; McArthur 2002: 321; Trudgill/Hannah 2002: 134; Melchers/Shaw 2003: 133; Gargesh 2006: 104; McArthur 2006: 389; Sheorey 2006: 188; Kirkpatrick 2007: 94; Schneider 2007: 85), it is Kachru’s (1994) final statement on the nature of educated South Asian English being “generally” restricted that contrasts with the few available empirical findings. Analyzing the one-million-word Kolhapur Corpus of written IndE from 1978 (cf. Shastri et al. 1986), Schmied (1994: 223), for example, found only one instance of *isn’t it* functioning as an invariant tag question alongside ten instances that were in accordance with the codified norms of standard English. The example that Schmied 1994 came across in the Kolhapur Corpus was: *You thought I would be late, isn’t it?* A closer look at the context in which this utterance occurred revealed that the given example appeared in the text category of fictional writing within

a quotation of direct speech. Hence, while *isn't it* might well qualify as a feature recurrent in spontaneous spoken IndE or even in IndE literary writing, it was apparently not common enough in other written text genres around 1978 to be attested in a text collection comprising one million words. Of course, this finding cannot be taken to mean that invariant tag questions are largely absent from IndE today. Even a one-million-word corpus is much too small to allow for such generalization, all the more so as the Kolhapur Corpus comprises only written texts and tag questions should be expected to be more common in speech. However, I would say that Schmied's (1994) finding, small as it is, does call the claim into question that the choice of tag questions is 'generally' restricted in educated South Asian English or, more specifically, IndE.

A noteworthy explanation for this near-absence of invariant tag questions in IndE writing is provided by Sahgal/Agnihotri 1985, who found that invariant uses of *isn't it* and *no*, "so frequently heard in Delhi's colloquial English" according to the authors (Sahgal/Agnihotri 1985: 123), were rarely considered 'good English' by their group of informants, 42 residents of "four residential colonies and three schools in South Delhi, a relatively westernized and smart area of Delhi" (118), who had been asked to judge the acceptability of a total of 27 linguistic patterns. In other words, there appeared to be a difference between the normative behavior of the informants in Sahgal/Agnihotri's (1985) study and their actual performance – an attitudinal stance typical of users of Outer Circle Englishes who are confronted with the co-existence of mostly uncoded local usage norms and coded external language models (which themselves are predominantly based on BrE and AmE conventions). As Kachru 2005 himself points out,

[i]n South Asia, as in other parts of the world, there is a difference between linguistic behaviour and the idealized norm. Traditionally, for historical reasons, southern British English has been the norm presented to the South Asians through the BBC, a small percentage of the English administrators and some teachers. In the written mode the exocentric norm came in the form of British literature and newspapers. In reality there is a wide gap between the perceived norm and the performance of users. (Kachru 2005: 55)

However, Sahgal/Agnihotri's (1985) study also brings to light that there were significant differences in the acceptability ratings for *no* (which was rejected by 63.64% as 'wrong', accepted by 34.09% as 'good enough for informal English' and by 2.27% as 'good English') and *isn't it* (rejected only by 36% as 'wrong' and accepted by 39% as 'good enough for informal English' and 25% as 'good English'). This could indicate that *isn't it* was on its way toward becoming more acceptable whereas *no* was clearly unacceptable around the mid-1980s.

What are the lessons to be learned from those findings regarding any new description of contemporary IndE? If we combine Schmied's (1994) corpus-based research result with Sahgal/Agnihotri's (1985) acceptability measurements, we arrive at the following conclusion. What is frequently cited as a characteristic feature of IndE

is a linguistic phenomenon that may be common in colloquial IndE speech (further evidence would be needed to back up this claim) but one that is extremely rare in educated writing (at least as far as the texts genres represented by the Kolhapur Corpus are concerned). While *no* was unacceptable to a majority of educated users of IndE in the mid-1980s, *isn't it* was more acceptable and possibly on its way toward further integration. In the light of this interpretation, Kachru's (1994) description of IndE can be enhanced. It is not so much the restriction of choices in the use of tag questions that characterizes contemporary IndE but its variational profile across speech and writing, seeing the use of invariant tag questions in one domain of use and their absence in others. Generally speaking, in order to describe IndE 'from within,' it is necessary to capture the dynamics between nativized and internationally established usage patterns, as it is this interplay, rather than the use of isolated features, that makes IndE 'uniquely Indian.' It is fortunate that resources such as the International Corpus of English (ICE) and the World Wide Web, which have become available only over the past decade, offer exciting new research possibilities to do so successfully.

2.2.2 Describing IndE in an international context

Trudgill/Hannah's (2002) approach to IndE in their textbook *International English* aims to place IndE in an international context as a variety of standard English and, as such, is intended to be comparative in outlook. In their chapter on standard IndE, Trudgill/Hannah 2002 classify the variety as a second-language variety of English and refer to the role of English in India "as one of the languages of education and wider communication" (129). The authors' comparison of IndE to the "main standard variety" English English (Trudgill/Hannah 2002: 1), however, results in another problematic attempt at assessing the characteristicness of the Indian standard on the structural level. Acknowledging the existence of forms in IndE "which are used by the majority of educated speakers and can also be found in newspapers" (129) but differ from English English, the authors state at the outset of their description that it is yet unclear for most of those forms whether they belong to "the Indian standard" (129). The lexicon and grammar of IndE are characterized by two lists of features, one each for morphology/grammar and lexis. The vocabulary list comprises ten loanwords out of context which are said to illustrate "that there is a substantial lexical borrowing from Indian languages into English" (135), and another twenty words out of context which are claimed to represent the "extension or alteration of meaning of EngEng [English English] words, retention of archaic forms or innovations" (135). No further attempt is made to capture any nuances of language use, for example in terms of describing stylistic, register or regional distinctions. Nor does the listing convey any sense of proportion as to how often the chosen lexical items are employed in their claimed variety-specific ways. Hence, Trudgill/Hannah's (2002) list of lexical features, like Kachru's (1994), does not allow readers to determine the exact impact of nativization on the

IndE lexicon and its influence on the projected “Indian standard” (129). A similar feature list is supplied for IndE morphosyntax (130–34). Here, IndE is characterized as displaying “[d]ifferences in count noun-mass distinctions” (130), “[a]n extended use of compound formation” (131), “[t]he use of nominal rather than participial forms of some words when used as adjectives” (131), “[a] difference in use of prepositions in verb-preposition collocations” (131), “the use of *itself* and *only* to emphasize time or place where EngEng speakers would usually use intonation to provide emphasis” (132), “[t]he use of adverbial *there* for ‘dummy’ *there*” (132), “[d]ifferent use of some auxiliaries” (132), “several differences from EngEng in the usage of tense and aspect” (133), “[t]he absence of subject-verb inversion in direct questions, and the use of such inversions in indirect questions” (134), “the use of a universal, undifferentiated tag question” (134), “[d]ifferences in complement structures with certain verbs” (134), and “[a] non-English use of *yes* and *no*, as in WAfrEng [West African English]” (134). Each category listed by the authors is illustrated by example sentences based on “research *in situ*” (ix). But since no further information on the actual contexts of use has been provided, the given evidence remains sketchy. By stating infelicitously that the patterns of the listed types “occur sometimes in the English of even some educated Indians and in English-language newspapers in India” (130), the authors imply that their long list of grammatical features, entitled “IndEng morphology and grammar” (130), might rather be more typical of the language of ‘un-educated’ users. Moreover, by labeling several features “non-English” (134), IndE is stripped of its ‘Englishness’, which invites interpretations of IndE as a deviant learner variety.

A description like Trudgill/Hannah’s (2002) is insufficient for various reasons. While a comparison of IndE with other varieties and the codified norms of standard English is in principle legitimate, such a comparison must not limit itself to interpreting observed usage differences in terms of deviations from ‘English English’. To place IndE properly within the larger context of world Englishes, it would be more useful to assess the local usage norms of IndE first. In a next step, IndE should be compared to other varieties by identifying similarities, differences and possibly also interdependencies. The two reference points used by Trudgill/Hannah 2002 to describe the inroads of structural nativization in IndE, the ‘educated speaker’ and ‘newspapers’, are too general to provide orientation for an understanding of a variety of English that, as the authors themselves state, “is an official language and is used as one of the languages of education and wider communication” (Trudgill/Hannah 2002: 129). Since there is no way of determining whether the features encountered in Trudgill/Hannah’s (2002) lists represent stable usage patterns, one is left to speculate what exactly there is to be learned from their account apart from finding out that IndE ‘has different words’ or ‘uses different structures’. As this could probably be said about any variety, the description ultimately fails its mission as a characteristic of IndE in the context of world Englishes.

2.2.3 IndE – caught between the norms: A look at Indian usage handbooks

If Kachru 1994 illustrates the challenges involved when describing the nativization of IndE qualitatively from within and Trudgill/Hannah 2002 those when viewing IndE in a comparative framework, a look at the two major usage handbooks of English published for the Indian market, Nihalani et al. (1979, 2004) and Yadurajan 2001, is informative as both guide books are indicative of the dilemmas to be encountered when the nuances of IndE usage, molded by the impact of both local and international influences, are described intuitively without any reliable database of authentic spoken and written IndE close at hand. While India can look back on a long list of glossaries (e.g. Yule/Burnell 1994 [1886], Subba Rao 1950, Hawkins 1984, Lewis 1991, Muthiah 1991), the most recent ones being the supplementary glossary to the Indian edition of the *Oxford Advanced Learner's Dictionary* (Crowther 1996) and the latest edition of Nigel Hankin's *Hanklyn-Janklin* (Hankin 2003), there is to date no comprehensive dictionary or style guide that would document the extent of lexical-grammatical nativization in the Indian subcontinent and be up to the standards of present-day lexicographical research. The handbooks by Nihalani et al. (1979, 2004) and Yadurajan 2001 are the only two available usage handbooks intended especially for Indian users.²⁶

Nihalani et al. (1979, 2004), *Indian and British English: A Handbook of Usage and Pronunciation*, is a compilation of about 1,000 words, phrases and usage patterns that are claimed to differ from BrE. In addition, the authors have provided variety-specific pronunciations for another 2,000 words.²⁷ As the authors explain in their introduction, it is their aim to

provide teachers and learners of English in India with information about the way in which certain forms and patterns of English used in India differ from the contemporary version of the native speaker model to which Indian English is closest, namely British Standard English. (Nihalani et al. 2004: 4)

The information presented by Nihalani et al. (1979, 2004) in their handbook is “by no means a comprehensive description of English as it is used in India” (1979: 3, 2004: 3) but is intended as a selection of items that the authors say “are used in a distinctive manner by large numbers of educated Indian speakers of English” (3). The data in the handbook were not drawn from language corpora but from individuals like “university lecturers, school teachers, journalists, radio commentators, and leaders of opinion in the society” (5) who were considered “likely to influence the English of Indian learners of the language” (5). The criteria for including or omitting items are outlined only vaguely, however, with frequency of use and acceptability among

26. There is, of course, a wealth of IndE teaching materials, e.g. Mason 1978 or Dhillon 1998.

27. The second edition of the usage guide was published in 2004. Nihalani et al. 2004 differs only mildly from Nihalani et al. 1979 in that a few entries have been added or changed.

educated users being two factors hinted at. As the authors point out, “[s]ome of the items are extremely frequent in the English used in India, some of them less so” (5). Moreover, “[i]tems which all speakers of English would agree to be sub-standard have been omitted unless they are so widespread as possibly to influence serious learners of the language by their very ubiquitousness” (5). Although Nihalani et al. (1979, 2004) do not state clearly what they mean by ‘sub-standard’ (a point already mentioned in Shastri’s (1983) review), their intention is to describe rather than prescribe. As they put it carefully, “[w]hen comprehensive descriptions of the English used in India come to be written, as they will, questions of a prescribed standard will also have to be solved. Until then we have hesitated to do more than describe.” (Nihalani et al. 1979: 7, 2004: 7) In spite of the authors’ carefulness, however, many of the comments found in their handbook on topics as variant as grammar, lexis, idiom, style, social/cultural context, collocations, meanings, register, loanwords and neologisms (those are the labels attached to the individual entries; cf. Nihalani et al. 1979: 11–3, 2004: 10–2) remain implicitly prescriptive in lieu of a well-defined database. Throughout the guidebook, IndE usage is, for example, repeatedly characterized in terms of its being ‘non-British’. The comment by the authors under the entry *upgradation* exemplifies this problem. *Upgradation* is first illustrated by the sentence ‘The upgradation of a number of posts has been postponed because of the financial crisis’ and subsequently commented on as follows: “This word does not exist in BS [British Standard English], where ‘upgrading’ would serve in the above context” (Nihalani et al. 1979: 186–7, 2004: 188). As is shown in the present study (cf. Chapter 3.2), *upgradation* has become established firmly in contemporary IndE and all of South Asia and is widely preferred to *upgrading*, with the latter word continuing life in the Indian subcontinent as a minor variant. By referring to *upgradation* as a ‘non-British’ item, Nihalani et al.’s (1979, 2004) handbook not only does not do justice to the usage preference of Indian users but also tacitly (though unintentionally) devalues IndE as being deviant from an external British usage norm. But this is not to say that the observations recorded by Nihalani et al. (1979, 2004) are generally meaningless. On the contrary, as Shastri 1983 once noted, “the large majority of items are, all of us would agree, common and distinctively Indian” (3). However, it is necessary to test the validity of Nihalani et al.’s (1979, 2004) observations against a larger corpus of spoken and written IndE to arrive at a clearer understanding of the IndE norms of usage before comparing those to BrE in a second step (if this is felt to be necessary).

In contrast to Nihalani et al. (1979, 2004), *Current English: A Guide for the User of English in India* by K. S. Yadurajan (Yadurajan 2001) has been designed as a prescriptive usage guide and represents a predominantly ‘conservative force’ in present-day IndE, to use Mukherjee’s (2007) terminology. Covering several hundred words, expressions, phrases as well as comments on grammar and style, Yadurajan’s (2001) handbook is meant to be “a supplement to a good dictionary, not a shorter version of one dealing with selected items” (x). The author posits the existence of a monolithic Standard English, from which items in IndE may deviate and are therefore in need of clarification. As Yadurajan 2001 states in reference to the cline of proficiency in IndE,

[t]he audience to which the book is addressed is fairly high on this 'cline' of English in India. They are not interested (not being specialists) in knowing all about Indian English. Their interest is largely limited to knowing whether a certain expression which they have been regularly using is, in fact, admissible as Standard English. So I have chosen to concentrate only on items which the IE [IndE] user genuinely believes are part of Standard English but are not. (Yadurajan 2001:x)

Further down in Yadurajan's (2001) introductory text, Standard English is equalled with BrE, to which the author ascribes a model function for IndE: "the standard accepted in India by tradition and convention is British English" (xi). Though Yadurajan 2001 points out that the description of IndE usage in his handbook "[is] based on authentic examples" (xi), with his source material comprising "chiefly current periodicals and newspapers" (xi) from India, he also says that some examples were collected in the United States, for instance "in advertisements by Indian stores and establishments, and news publications brought out by Indians" (xi). Yet despite the eclectic nature of his source material, the conclusions drawn by Yadurajan 2001 are usually general and delivered with prescriptive force, as when Yadurajan 2001 claims that "[t]he use of a preposition with *comprise* (*comprise of*) is incorrect" (41), although both IndE and BrE, the claimed prestige model, do make regular use of *comprise of* today, though in slightly different contexts of use (cf. Chapter 4.2.1). While for those reasons Yadurajan 2001 does not qualify as a reliable "guide for the user of English in India" (as the subtitle of the handbook has it), many of the author's observations on IndE usage are nonetheless noteworthy as they point to potential areas of intrusion of nativized (as well as international) forms into educated IndE. As is the case for Nihalani et al. (1979, 2004), this calls for supplementary evidence.

2.2.4 Previous corpus-based descriptions of IndE

"A corpus-linguistic description," Schmied 1994 writes, "is only as good as the corpus it is based on" (229). When beginning work on the present study in early 2000, the Indian component of the International Corpus of English (ICE) had not been released yet, and the Kolhapur Corpus of IndE was the only available text corpus of IndE. Therefore, a closer look was taken at the design of the Kolhapur Corpus (cf. Shastri et al. 1986) and the descriptions of IndE based on Kolhapur to see what directions any new research into IndE would have to take in order to improve on the status quo of previous descriptions. Although scant, the available research on IndE based on the Kolhapur Corpus suggested that it was possible to minimize the shortcomings of feature-list descriptions and impressionistic accounts of the variety when drawing on a well-defined and carefully-designed database of authentic IndE usage.

The general advantages of using a corpus for describing linguistic variation in English are twofold (cf. Mair 1996:70; Kennedy 1998:7–12; McEnery et al. 2006:3–12). For one, a corpus allows for the qualitative study of authentic language in different contexts of use. When using a corpus, words, meanings and patterns can be

interpreted in reference to the communication contexts in which they occur, and interdependencies between language use and context can be uncovered. As the simple example of the invariant tag question *isn't it* in IndE illustrates (cf. Chapter 2.2.1), a corpus-based investigation can arrive at a rounder description of IndE by showing that invariant tag questions do not appear as indiscriminately in contemporary IndE as previous accounts of the variety imply. In this instance, it is not the restriction of grammatical choices in the use of tag questions but the overall expansion of communicative possibilities across registers that emerges as being characteristic of contemporary IndE. The example also suggests that the structural nativization of IndE may be proceeding at different speeds across different registers and text categories. The dynamics of such a process can only be documented in reference to a database that ideally comprises 'representative' specimens of as many types of natural language use as possible – but it can barely be described in reference to a static list of features predominantly based on introspection and intuition.²⁸ A second major advantage of working with language corpora lies in the very possibility of quantifying observations. As the example of the lexical item *upgradation* (cf. Chapter 2.2.3) illustrates, IndE is capable of carving out variety-specific local usage preferences that should not be considered as deviations or errors but as manifestations of local usage norms. Such preferences can be pinpointed empirically when comparing the frequencies of competing variants (that should ideally be fully compatible in function) in clearly-defined contexts of use. In a next step, the comparison of proportions of variety-specific forms across registers and text-types can lead to more insights into the extent to which individual items have become established within a variety, which may even suggest further directions of change. Again, a static feature-list description cannot possibly capture such nuances.

S. V. Shastri of Shivaji University, Kolhapur, must be given credit for being among the first to apply the corpus-linguistic method to IndE. Shastri designed and compiled the one-million-word Kolhapur Corpus in 1978 to match the 1961 LOB and BROWN corpora of BrE and AmE with a view to intravarietal as well as crossvarietal analysis (cf. Shastri 1983, 1985; Shastri et al. 1986; Schneider 2000). Restricted to written text genres, the Kolhapur Corpus runs largely parallel to the LOB and BROWN corpora and consists of fifteen text categories of edited IndE prose, including newspaper English (reportage, editorials, and reviews), fictional writing (six categories), popular lore, belles lettres, learned and scientific writing, religious writing, miscellaneous reports (for example government documents) and a category entitled 'Skills, trades and hobbies'.²⁹ Work on the Kolhapur Corpus has added to a fuller understanding of IndE by encouraging context-sensitive investigation 'inside IndE' and its comparison to the prestige varieties BrE and AmE. In a series of papers, Shastri has reported the results of his own investigations into lexical and grammatical differences in IndE, BrE

28. On the notion of representativeness in corpus design cf. Kennedy (1998: 62–6) and McEnery et al. (2006: 125–30).

29. For a detailed description of the text categories cf. Shastri et al. 1986.

and AmE (Shastri 1988a, 1992, 1996). The Kolhapur Corpus has also inspired other scholars to work on selected aspects in IndE lexis, morphosyntax and grammar (cf. Hirschmüller 1989; Sayder 1989; Schmied 1994; Schneider 2000; Sharma 2001). The majority of those studies have come to the conclusion that no major qualitative, but mostly quantitative differences exist between written IndE, BrE and AmE:

Our databank revealed that the structure of the vocabulary of the Kolhapur Corpus resembles closely that of the Brown Corpus and the Lancaster-Oslo/Bergen Corpus. It has about the same number of different words as its British and American counterparts (Kolhapur Corpus 45,859; Brown Corpus 48,989). Ranks one to fifteen of the fifteen most frequent words are shared by sixteen words in the three corpora, although there are marked differences in absolute frequencies ...

(Leitner 1991b: 221)

Our own corpus-based work in the Department of English, Shivaji University, has revealed that there is little difference between the native varieties and IE [IndE] as far as syntax is concerned ...

(Shastri 1992: 266)

The results of my empirical corpus investigations have shown that no fundamental, categorical difference between Indian English and any of the other national varieties was detected, but on the other hand there is also no full identity of patterns and preferences to be observed. Indian English shares the variability that characterizes all varieties of English worldwide, with largely comparable figures and proportions.

(Schneider 2000: 133–4)

But even if the structural nativization of IndE has not resulted in the emergence of any major lexical or syntactic differences so far, detailed analyses of the corpus material have produced numerous new insights into more subtle usage differences between IndE, its parent variety BrE and AmE that do not necessarily conform to the impressionistic accounts of IndE found in the available feature lists. Shastri 1988a, for example, is a convincing case study of how a careful qualitative investigation based on the Kolhapur Corpus can lead to a better understanding of lexical variation and innovation in IndE, which involves more than any reference to “substantial lexical borrowing” (Trudgill/Hannah 2002: 135) could possibly capture. Studying the use of loanwords from indigenous Indian languages, Shastri 1988a documents the functional range of loan vocabulary and code-mixing by drawing on authentic examples from different text genres. Shastri 1988a also observes differences in the density of loanwords across text categories in Kolhapur and hypothesizes that the lexical nativization of IndE via borrowing may affect text genres to variant degrees. He notes, for example, that loanwords occur more prominently in the Kolhapur Corpus in the context of religious writing, whereas they are nearly absent from learned and scientific prose, with other text categories hypothesized to fall on a cline from “most Indian” to “least Indian” (Shastri 1988a: 38). Since Shastri 1988a did not manage to quantify the interrelationship of text genres and borrowing in his study satisfactorily, his paper served

as a starting point for the investigations into the use of loanwords in contemporary IndE in the present study (cf. Chapter 3).

Shastri's (1992, 1996) work on IndE lexicosyntax, on the other hand, is indicative of the limitations of working with relatively small corpora such as the Kolhapur Corpus. On analyzing complementation patterns of a set of 29 words (including verbs, adjectives and nouns), Shastri 1996 found that only 14 items displayed the behavior in the Kolhapur Corpus that had been ascribed to South Asian English.³⁰ Of the 14 items, the 'unusual' patterns never outnumbered the standard uses of the items "as described in grammars" (Shastri 1996:81). Moreover, four of the 14 'unusual' patterns were attested also in LOB and BROWN, which called their status as absolute variety differentiators for South Asian English into question. Although Shastri (1992, 1996) concluded correctly on the basis of those findings that "the burden of determining variety features shifts to quantitative studies" (Shastri 1996:81), he was not able to pinpoint the stability of the attested rare patterns in any more detail because of the limited size of his database. This is unfortunate as the structural nativization in emerging varieties of English often happens at the interface of lexicon and grammar, which makes low-frequency observations of the types detected by Shastri 1996 particularly interesting for describing the emergence of local usage norms (cf. Mair 2002: 106; Schneider 2003:249; Mukherjee 2007:175). Note that Schneider 2004, on analyzing particle verbs in IndE and four other varieties of English, was facing similar difficulties when interpreting research findings in the low-frequency range. In reference to the possibilities of the one-million-word sub-corpora of ICE, Schneider 2004 remarked that

token frequencies of individual forms were too low throughout to allow for significance testing or hard-and-fast acceptance of the results ... They may represent emerging innovations, but whether they are is impossible to tell. This will require either substantially larger corpora or, simply, the test of time.

(Schneider 2004: 247)

Shastri 1992 himself concluded realistically that "[a]ll this points towards the need for systematic quantitative studies based on larger databases before we can say anything decisive about variety features" (274). It is for this reason that linguists have recently started to appropriate significantly larger corpora to investigate issues in IndE lexicosyntax, thereby challenging the claim that syntactic differences between IndE and BrE are only quantitative in nature. Olavarria de Ersson/Shaw 2003, Mukherjee/Hoffmann 2006 and Hoffmann/Mukherjee 2007, for example, looked at complementation patterns of several sets of verbs in IndE online newspaper corpora, which has produced convincing

30. The items discussed by Shastri 1996 are: *interested* (adj.), *prevent*, *fail*, *failure*, *like* (v.), *forbid*, *hesitate*, *eligible*, *want* (v.), *suggest*, *deplore*, *announce*, *inform*, *tell*, *interest* (n.), *busy*, *capable*, *prepare*, *succeed*, *successful*, *look forward to*, *stress* (v.), *avoid*, *contemplate*, *let*, *beseech*, *ban* (v.), *discourage*, and *remind* (74).

evidence that systematic structural differences might exist between IndE and BrE in this area of the language system.³¹ Lexicosyntactic aspects of this type are discussed in detail in the present book as well, which turns to closed corpora like Kolhapur and significantly larger Internet databases to test numerous hypotheses on IndE lexicosyntax (cf. Chapter 4).

Next to the limited size of the Kolhapur Corpus, the difference in 'age' between the Kolhapur Corpus (compiled in 1978) and LOB/BROWN (compiled in 1961) has been a hindrance to employing the corpora successfully for comparative research. When Shastri designed the Kolhapur Corpus in the 1970s, he made several changes as compared to the original design of LOB and BROWN that were "dictated mainly by logistic and practical considerations" (Shastri et al. 1986: 2). One of the decisions taken was to use material from 1978, whereas the material in LOB and BROWN had been published in 1961. Shastri et al. 1986 argued that the difference of seventeen years was not a shortcoming but, in fact, an enhancement of the value of the corpus, since "the Indianness of Indian English is a post-Independence phenomenon and may have reached a discernible stage in the thirty years after Independence" (3). Shastri et al. 1986 also speculated that "in the same thirty years the American and British English may not have undergone such changes" (3). But Leitner 1991b raises the valid point that it is "not clear whether one may compare quantitative data from corpora from 1961 (the BROWN and Lancaster-Oslo/Bergen corpora) and 1978 (the Kolhapur Corpus)" (229) and states that Shastri's claim that "a significant level of institutionalization was hardly reached before 1978 (Shastri et al. 1986) would need exemplification" (229–30). Schmied 1994 comes to a similar conclusion:

Irrespective of the questions as to how representative the text types are for IndE as a whole and how representative the texts chosen are for the type they are supposed to stand for, there is a clear restriction to the written form and a clear limitation in time: the texts [in the Kolhapur, LOB and BROWN corpora] were published 15 and over 30 years ago and they are more than 15 years apart.

(Schmied 1994: 229)

31. Olavarria de Ersson/Shaw 2003 study the behavior of nine verbs: *provide, furnish, supply, entrust, present, pelt, shower, pepper and bombard*. They find, for example, that the most common complementation pattern with *pelt* in the online editions of *The Hindu* and the *Times of India* is *to pelt s.th. at s.o.*, whereas the pattern *to pelt s.o. with s.th.* is standard in BrE and AmE. The authors also find *pelt* to collocate frequently with *stone* (as in *they pelted stones at him*), which leads them to the conclusion that the 'new' complementation pattern is pragmatically an appropriate development, since *stone* does not carry much information and, accordingly, is not placed in focus position (also cf. Schneider 2007: 87). Mukherjee/Hoffmann 2006 and Hoffmann/Mukherjee 2007, on the other hand, study ditransitive verb complementation in BrE and IndE and find a tendency in IndE for ditransitive verbs to combine with a direct object. The construction types *give s.th.* and *give s.th. to s.o.*, for example, are found to be more common than the BrE-favored *give s.o. s.th.* (also cf. Schneider 2007: 87).

In fact, Sayder's (1989) Kolhapur-based study on the mandative subjunctive in IndE, BrE and AmE exemplifies the dangers of jumping to premature conclusions if diachronic aspects are ignored. Sayder 1989 found differences between IndE, BrE and AmE in the use of the mandative subjunctive to be "of a quantitative, not qualitative kind" (65) and argued that IndE was in an 'intermediate' position as compared with AmE and BrE. However, the comparison of IndE written texts from 1978 with BrE and AmE written texts from 1961 is problematic in this very instance, because the mandative subjunctive was in the process of becoming more popular outside AmE (where it had traditionally been strong) in the three decades between 1961 and 1991, possibly due to the postcolonial force of AmE (cf. Övergaard 1995: 51, 54; Hundt 1998b: 159). A static comparison of late 1970s IndE with early 1960s BrE and AmE cannot account for the dynamics of such a process. Similarly, Sharma 2001 drew conclusions on the alleged formality of IndE vis-à-vis BrE and AmE after comparing the Kolhapur, LOB and BROWN corpora. Her findings, too, turn out to be untenable when considering more recent developments in Englishes worldwide in the second half of the twentieth century. Both Sayder's (1989) claims on the position of the mandative subjunctive in IndE vis-à-vis BrE and AmE and Sharma's (2001) remarks on IndE style are rectified in the present study (cf. Chapters 3 and 5).³²

Corpus-based research on IndE has been given a fillip with the release of the Indian component of the International Corpus of English (ICE), which was published in late 2002 (cf. Sand 2004; Schneider 2004; Mukherjee/Hoffmann 2006; Schilk 2006; Hoffmann/Mukherjee 2007; Lange 2007). Compiled by S. V. Shastri (Kolhapur) and Gerhard Leitner (Berlin) to match its twin corpora in the ICE family (cf. Greenbaum 1996), ICE-India comprises one million words of IndE from the 1990s and includes a wide range of text genres representing written and spoken language use in public as well as private communication situations.³³ Sand 2004 and Schneider 2004, for example, compared ICE-India with other ICE components representing Englishes from the Inner and Outer Circles and uncovered noteworthy differences in the use of definite articles and particle verbs worldwide. Sand's (2004) and Schneider's (2004) results are revisited in this study and discussed in the light of new evidence from larger online databases (cf. Chapters 4 and 5). Lange 2007, on the other hand, employed the Kolhapur Corpus and ICE-India side-by-side to investigate real-time change within IndE, thereby producing empirical evidence for IndE to have evolved stable new functions of *only* and *itself* as presentational focusing devices. Her findings, too, are reconsidered in the present book (cf. Chapter 5). Gerhard Leitner's own corpus-based research on IndE has generated several more working hypotheses. Leitner 1991b noticed the use of AmE-based spelling variants and lexical items in the Kolhapur Corpus, which led

32. Similar problems were encountered when comparing the Macquarie Corpus of Australian English (ACE) and the Wellington Corpus of Written New Zealand English (WC) with LOB and BROWN, cf. Peters (1998: 89).

33. Cf. Nelson 1996 for a detailed description of the text categories included in ICE.

him to speculate that the influence from AmE on IndE might have been on the rise in IndE in the late 1970s. The interdependencies of IndE and AmE are investigated in greater detail in this book (cf. Chapter 3). In an earlier paper, Leitner 1989a challenged the claim that syntactic differences between IndE and AmE and BrE were merely of a quantitative kind. His qualitative analysis of the IndE past perfect, using a letter to the editor of an IndE newspaper, suggested that the nativization of IndE might already have affected the grammatical core of English in that particular area – another hypothesis to which this study returns (cf. Chapter 5).

2.2.5 Conclusion

The available descriptions of IndE are marred by several shortcomings that a new description of IndE must set out to minimize. The major weakness of feature-list descriptions of IndE lies in their presentation of data in a decontextualized fashion without documenting the stability of individual items, their domains of use and their overall relevance for IndE as a linguistic entity. The available IndE usage guides are mostly based on introspection rather than factual research and tend to disregard (and, in the case of Yadurajan 2001, also occasionally disrespect) local usage norms, which is likely to instill a sense of linguistic insecurity in users rather than provide ‘guidance’. Corpus-based descriptions of IndE, on the other hand, have been more successful in pinpointing the status of individual Indianisms, suggesting inroads of nativization, and illustrating external influences, for example from AmE. However, the areas investigated have so far been restricted mostly to the written domain, and cross-varietal comparison has often been limited to BrE and AmE. In view of more recent research on BrE and AmE based on LOB and BROWN and their younger twin corpora FLOB and FROWN (representing BrE and AmE usage at the beginning of the 1990s), the explanatory potential of the Kolhapur Corpus appears to be all the more limited (cf. Mair 1998: 139–41; Mair 2002: 105–7). Comparative research using those parallel corpora has produced solid evidence that written BrE and AmE were undergoing change in the second half of the twentieth century, which calls Shastri et al.’s (1986) assumption on BrE and AmE being more ‘static’ than IndE clearly into question. All those points signal the need for more empirical research looking ‘inside IndE’ and ‘beyond’ on the basis of new data to arrive at a fuller understanding of the dynamics of variation and change in present-day IndE.

2.3 Methodological and descriptive framework

In view of the limitations of working with the Kolhapur Corpus (and the fact that ICE-India had not yet been made available when work on the present study commenced), a new micro-corpus of contemporary IndE speech and writing was put together at

the outset of this research to be able to describe IndE across various domains of use and, hence, make up for (at least) some of the shortcomings of the available descriptions. Since previous research on IndE had relied primarily on edited writing, spoken and unpublished written material was included in the corpus to make the selection of text categories and registers on which to base claims on IndE more representative. Moreover, the database was to lend itself to diachronically sensitive research of the type conducted at Freiburg University under the supervision of Christian Mair, which had managed to describe ongoing change in BrE and AmE on the basis of the parallel corpora LOB/FLOB for BrE and BROWN/FROWN for AmE (cf. Mair 1998, 2002). It was hypothesized that if the new micro-corpus of IndE were designed so as to run parallel to selected sections of the Kolhapur Corpus, IndE usage from 1978 could be compared to IndE usage from 2000, which might help to illustrate the influence from internal and external usage norms and maybe also indicate directions of change. At the same time, diachronic developments in IndE between 1978 and 2000 could be compared to diachronic developments in BrE and AmE in the second half of the twentieth century to describe the interdependencies among the varieties and gauge degrees of endonormative and exonormative orientation in IndE. On the other hand, it was also clear from the very beginnings that it would be difficult to analyze interesting mid- to low-frequency phenomena on the basis of a new micro-corpus alone. As Shastri's (1996) analysis of lexicosyntactic features based on the Kolhapur Corpus had brought to light, even a one-million-word corpus was much too small to yield satisfactory insights into the stability of rare features like particle verbs, complementation patterns or the use of neologisms and loanwords. A new micro-corpus would even be smaller in size and, thus, be more limited in scope than Kolhapur. Early on, the Internet was consulted whenever closed corpora like the Primary Corpus (the new micro-corpus), the Kolhapur Corpus and, after its release, ICE-India, too, would allow no further insights. In fact, the Internet emerged as the second most important resource employed in the present study.

2.3.1 Database: Primary Corpus

The Primary Corpus, on which large parts of this study are based, is a micro-corpus in the truest sense of the word. Consisting of 180,000 words of spoken and written IndE recorded and collected in the year 2000, it was put together specifically for this research project to represent IndE usage in the high-proficiency range in various domains of use. The corpus includes published press texts (40 texts of approximately 2,000 words each), published broadcast material (40 texts) as well as unpublished written material in the form of student essays (10 texts). See Figure 2.1, which illustrates its design.

The press section (IND2000-Press) comprises published texts from the year 2000 that were collected from the online editions of seven national newspapers from

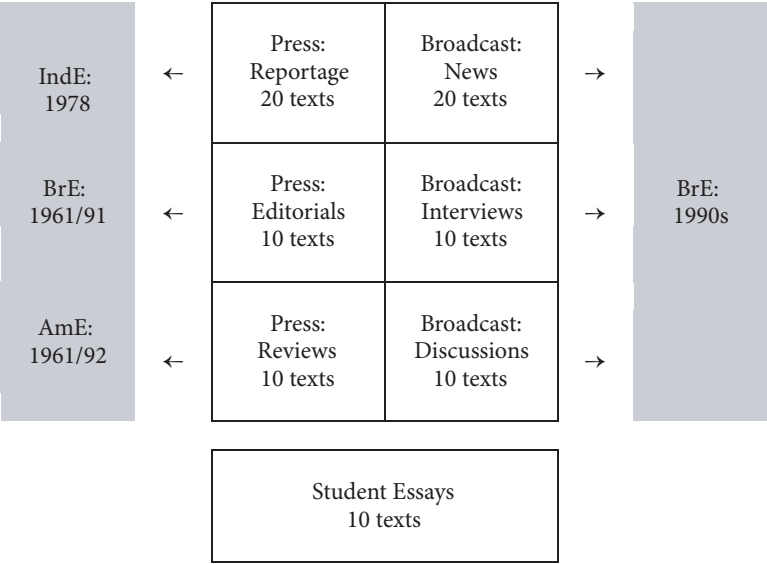


Figure 2.1 The Primary Corpus.

different parts of India.³⁴ In view of the aim to describe the standard usage range, no material from regional newspapers was included. The category of national newspapers was hypothesized to represent a form of edited IndE writing with supraregional model function, as India's national newspapers had a pan-Indian readership and were available nationwide. The texts included in the corpus are newspaper reports (20 texts of 2,000 words each), editorials (10 texts) and press reviews (10 texts). The press texts chosen for inclusion were not selected by random sampling. The primary goal was to match the new IndE press material as closely as possible in size and composition with newspaper texts in the 1978 Kolhapur Corpus, in the BROWN and FROWN corpora (representing written AmE from 1961 and 1992) and in the LOB and FLOB corpora (representing written BrE from 1961 and 1991). To achieve this, the principles outlined for the compilation of FLOB and FROWN (cf. Sand/Siemund 1992; Leech 2007) were closely followed. It was assumed that those six parallel micro-corpora would make it possible to test hypotheses on endonormative as well as exonormative tendencies in IndE in a diachronically sensitive fashion.³⁵

34. The texts were taken from the following newspapers: *The Indian Express*, *The Hindu*, *The Hindustan Times*, *The Times of India*, *The Statesman*, *The Sunday Statesman*, and *The Economic Times*. Cf. Appendix 2.1.

35. Cf. Appendix 2.1 for a complete list of parallel texts from the Kolhapur, LOB, FLOB, BROWN and FROWN corpora matching the press texts in the Primary Corpus.

The broadcast section of the Primary Corpus (IND2000-Broad) represents language use at the formal end of the stylistic spectrum in various public situations. During a three-month stay in India between April and June 2000, approximately 50 hours of television broadcast had been recorded, from which 34 of the 40 final corpus texts were sampled. An additional six texts were recorded from the Internet in October and November 2000 because the audio quality of the original tapes had partly been too low to allow for adequate transcription. The speech section contains 80,000 words, which is the equivalent of about 530 minutes of spoken IndE, and includes scripted as well as spontaneous speech in the form of television newscast, broadcast interviews and broadcast discussions. The news broadcast was recorded from three national television channels (Doordarshan, Star TV and Zee TV) and represents predominantly scripted monologues, although the news programs are occasionally also interspersed with passages of spontaneous conversation, for example when including excerpts from interviews. The broadcast interviews and discussions, on the other hand, represent predominantly spontaneous language use in the form of dialogues recorded from five popular English-language programs produced in India and televised nationwide, thus reaching out to large audiences in the country (*Question Time India*, *The Big Fight*, *News Hour*, *DD News*, *Face to Face*, *India Talks*). The English compiled here is that of Indian journalists, politicians, businesspeople, artists, athletes, musicians and celebrities – a spectrum of speakers from diverse regional and linguistic backgrounds that must be placed very high on the cline of proficiency. It was hypothesized that this range of speakers would be ideal to represent an “empirical standard” (Bartsch 1985:32), from which the features of the IndE usage norm could be extracted. To be able to compare IndE speech with BrE, a matching sample was drawn from the broadcast sections of the BrE sub-corpus of the International Corpus of English (ICE-GB), which had already been available in the year 2000 (cf. Greenbaum 1996).

The essay section (IND2000-Ess) was included in the Primary Corpus to strike a balance between the heavily edited press genres and the broadcast material. In contrast to the student essays included in ICE-India, this text category comprises unpublished writing by secondary-school students who were attending two prestigious English-medium public schools in Dehradun, Uttarakhand (then Uttar Pradesh), in their final year (class 12) and had, by the year 2000, received at least seven years of English-medium education.³⁶ English was the second language of all corpus contributors, Hindi being their first language. The essays were not written specifically for the present study, but were photocopied from the students’ exercise books after written permission had been granted by the students for the inclusion of their writing in the Primary Corpus.

36. The schools are The Doon School, Dehradun, and Carman School, Dehradun. Both schools are affiliated to the Council for Indian School Certificate Examination (CISCE). English is the primary means of instruction in all classes except for language classes in Hindi, Sanskrit, other Indian languages or foreign languages (cf. Dutt 1999). Also cf. Srivastava 1998 for a detailed analysis of the history and contemporary role of The Doon School in present-day India.

After the source material for the Primary Corpus had been collected, the corpus was built. The speech data were transferred from video tapes to audio tapes and subsequently transcribed orthographically. The press texts were downloaded from the Internet and saved in Word files, whereas the student essays were transcribed by hand. Throughout the investigations, it was the Primary Corpus which was consulted first to test previous hypotheses on the nature of IndE. Whenever the Primary Corpus would be too small to yield satisfactory answers, supplementary data were considered to arrive at a rounder description of IndE. Although slightly dated, the Kolhapur Corpus was useful to address several issues for which the Primary Corpus was too limited in size and had the additional advantage of running parallel to its BrE, AmE, NZE and AusE twin corpora. ICE-India allowed further insights into the patternings of IndE-typical forms in more informal usage domains, especially spoken language. Most importantly, the World Wide Web grew into an invaluable resource for placing low-frequency observations in a larger context.

2.3.2 Database: Online corpora

While closed corpora like the Primary Corpus, the Kolhapur Corpus or ICE-India have all the advantages of carefully designed and well researched databases that allow for systematic investigation across registers, modes and text types and accurate quantification of research findings, larger corpora are needed to study the use of rare linguistic features. The sheer endless material found on the Internet holds exciting new possibilities in store for investigating such phenomena across Englishes worldwide if the medium 'Internet' is tamed for systematic utilization and meaningful interpretation (cf. Thelwall 2005; Mair 2006b, 2007; Lüdeling et al. 2007). In the context of the present study, the Google Advanced Search Option turned out to be a very simple but highly effective tool to document the availability and behavior of numerous low-frequency phenomena in more detail than previous research had been able to accomplish.³⁷ Since the Google Advanced Search Option allows users to restrict Internet search queries to specific websites or sub-domains on the Internet, 'snapshots' from the Internet were collected, i.e. frequency measurements of competing lexical and grammatical variants returned by Google-based search queries that were conducted across selected domains and websites. In particular, search queries were run regularly across six online editions of Indian quality newspapers based in different parts of the country (*The Indian Express*, *The Hindu*, *The Hindustan Times*, *The Tribune*, *The Telegraph* and the *Deccan Herald*), across three quality newspapers from South Asia (Pakistan's *Dawn*, Sri Lanka's *Sunday Times* and Bangladesh's *Independent*) and across two quality newspapers from Great Britain and the United States

37. Cf. Thelwall (2005: 523–6) for a discussion of the potential disadvantages of using commercial search engines for linguistic research purposes. The restrictions outlined by Thelwall 2005 for search engines in general apply to the Google search engine as well.

publication	http://	abbr.
INDIA		
<i>The Indian Express</i> (Mumbai)	www.indianexpress.com	IE
<i>The Hindu</i> (Chennai)	www.hinduonnet.com	HI
<i>The Deccan Herald</i> (Bangalore)	www.deccanherald.com	DH
<i>The Hindustan Times</i> (Delhi)	www.hindustantimes.com	HT
<i>The Tribune</i> (Chandigarh)	www.tribuneindia.com	TR
<i>The Telegraph</i> (Kolkata)	www.telegraphindia.com	TE
SOUTH ASIA		
<i>The Dawn</i> (Pakistan)	www.dawn.com	DA
<i>The Sunday Times</i> (Sri Lanka)	www.sundaytimes.lk	ST
<i>The Independent</i> (Bangladesh)	www.independent-bangladesh.com	IN
UK		
<i>The Guardian</i> (London)	www.guardian.co.uk	GU
USA		
<i>The Washington Post</i> (D.C.)	www.washingtonpost.com	WP

Figure 2.2 Supplementary data: online newspapers.

country	top-level country domain
India	.in
Singapore	.sg
United Kingdom	.uk
South Africa	.za
Australia	.au
United States of America	.us

Figure 2.3 Supplementary data: top-level country domains.

(Britain's online *Guardian* and the US-American online edition of *The Washington Post*). This allowed for systematic comparative research within the domain of journalistic writing (cf. Figure 2.2).

To cross-check whether findings in the online newspaper corpora would have any relevance beyond the register of journalistic writing, both in India and elsewhere in the English-speaking world, snapshots were also collected from the wider World Wide Web. For that purpose, further Google-based search queries were run across six different top-level country domains, namely the Indian ('.in') and Singaporean ('.sg') domains, the Australian domain ('.au'), the South African domain ('.za'), the US domain ('.us') and, finally, the UK domain ('.uk'), cf. Figure 2.3.

In some cases, those snapshots from the Internet revealed that a feature which had been claimed to be an Indianism would simply not be found at all in the Indian online data, which was a strong argument against treating it as an IndE-typical feature, although care had to be taken not to jump to premature conclusions as the Web-based search queries were naturally limited to published written material only. In other cases, comparing snapshots from IndE newspapers with those from the 'in' domain pointed to the existence of domain-specific usage restrictions that called for context-sensitive interpretations of variation within IndE. In yet other cases, the comparison of Internet snapshots from around the world was indicative of broader Indian, South Asian or other regional usage preferences, thereby revealing (oftentimes unexpected) similarities of IndE with one or more of the other varieties represented by the snapshots. Applied flexibly but carefully as a supplementary resource, the Internet turned into a useful tool for describing variation within IndE, viewing IndE in a comparative framework, and describing the stability and status of individual low-frequency features more precisely than an analysis based only on closed corpora could have hoped for.

2.3.3 Descriptive framework

After compiling the source data, building the Primary Corpus and conducting initial experiments using the World Wide Web, the variables to be analyzed for this corpus-based description of contemporary IndE were chosen. The present study covers selected areas in vocabulary and syntax, but not phonology. Looking at the IndE lexicon was an obvious decision because the status of the majority of lexical items that had been claimed to be lexical Indianisms was yet unclear. This triggered my interest in exploring the ways in which the 'double-layered' approach of using the Primary Corpus alongside the Internet could be employed to describe the status and the behavior of the claimed Indianisms in more sufficient detail. Central to world Englishes, however, is the claim that varieties like IndE have made modifications to English grammar. The core of this study therefore revolves around the discussion of lexicogrammatical and morphosyntactic variation in IndE – two areas for which the feature-list descriptions and usage handbooks had postulated numerous hypotheses still awaiting the test of authentic corpus data. Also, the available corpus-based research had produced inconclusive results regarding the degrees of grammatical nativization in IndE in areas like the article system (Schmied 1994; Sharma 2005) and the past perfect (Leitner 1989b; Sharma 2001), which called for a fresh look at those issues. Although a phonological analysis of IndE would have been desirable, phonology was not included in the present study for a rather practical reason: the broadcast data collected in India made it nearly impossible to control the variable 'background language', which would have been crucial to interpret patterns of variation among speakers of IndE meaningfully.

The results of the corpus-based investigations are presented in three separate chapters on vocabulary (Chapter 3), lexicosyntax (Chapter 4), and morphosyntax and grammar at the sentence level (Chapter 5). Each chapter looks at variables considered relevant for a description of IndE “in a web of conflicting tensions” (Leitner 1992a: 225). The analysis of feature-list Indianisms settles the issue of their characteristicness for IndE and allows new insights into the degrees to which users of IndE were orienting themselves toward local usage norms around the turn of the millennium. The study of typical variety differentiators for BrE and AmE illuminates the interdependencies between IndE and the two prestige varieties. Did IndE adhere largely to BrE usage conventions around the year 2000? Or was it under the influence from AmE? Cf. Figure 2.4.

Chapter 3 first looks at loanwords, neologisms and meaning change (Chapters 3.1 and 3.2) and interprets the presence or absence of those features across text types and registers as a gauge for the impact of lexical innovation “caused by the need to express certain Indian concepts which have no equivalent in English tradition and hence no appropriate form of expression” (Lukmani 1992: 155–6). But as Leitner 1989b, for example, has pointed out, innovative tendencies may also have affected IndE syntax to a considerable extent, which is why the present study includes in-depth analyses of claimed grammatical Indianisms in lexicosyntax (Chapter 4), in the noun phrase (Chapters 5.1 to 5.3), in the verb phrase (Chapters 5.4 to 5.6) and at the sentence level (Chapters 5.7 and 5.8).

Next to independent innovation, interference from indigenous Indian languages has frequently been cited as a trigger for structural nativization in second-language varieties. L1 interference involves speakers carrying over features of their first language into their second language. In her study on language contact, Thomason 2001, for instance, refers to IndE as an illustrative example of a variety of English that is characterized heavily by L1 interference. In Thomason’s (2001) words, IndE “has numerous interference features of this type from indigenous languages of India” (74). Several variables discussed in the present study were chosen to measure the impact of L1 interference on IndE. Carls 1999, for example, says that interference from Hindi has resulted in IndE making frequent use of the word-formation pattern *N-cum-N*, because Hindi has a similar pattern revolving around “the element *aur* (‘and’) between the coordinated word-stems” (150). He also states that IndE makes frequent use of adjectival compounds of the pattern *N+V-ed* in attributive position, which he relates to the observation that “[t]he adjectival compounds of the model come very close to suffix-like word-formation elements [in Hindi]” (150). Both hypotheses are tested in Chapter 3.2. According to Sridhar 1992, speakers of Dravidian languages frequently use the particle *as* after the verbs *call*, *rename* and *term* in IndE, because in “Dravidian languages, speech material such as technical terms, names, and quotations have to be marked by the quotative particle” (142–3). This hypothesis is tested in Chapter 4.2. Spitzbardt 1976, too, follows this line of argumentation by characterizing IndE in terms of a “widespread Indian tendency to disregard the articles, both

endonormative tendencies expected	← →	exonormative tendencies expected
3.1 loanwords		3.3 function words
3.2 neologisms word formation semantic change		contractions near-but-not- identical words
		3.4 Briticisms/Americanism
4.1 verb particles, e.g. <i>up, down</i>		
4.2 particle verbs, e.g. <i>discuss about</i>		
4.3 complementation, e.g. <i>want + that</i> -clause		
5.2 count/number divide, e.g. <i>legislations</i> <i>equipments</i>	5.1 article use	5.3 concord with collective nouns
5.5 tense and aspect with <i>since, ago, yesterday</i>	5.4 past perfect	5.6 mandative subjunctive
5.7 direct and embedded <i>wh</i> -questions		
5.8 focus syntax, e.g. <i>only</i>		

Figure 2.4 Chapter overview.

definite and indefinite” (40). This, according to Spitzbardt 1976, is the case because major Indian languages (for example Hindi or Tamil) do not have an article system comparable to the English article system. Comrie 1985, on the other hand, claims that the past perfect is more frequent in IndE than in other varieties because it is used with preterite meaning on the analogy of a tense form in Hindi and Urdu that “most closely corresponds formally to the English pluperfect” (69). Both Spitzbardt’s (1976) and Comrie’s (1985) claims on the distribution of surface articles and the past perfect in IndE are put to test in Chapters 5.1 and 5.4.

Another trigger for second-language varieties of English to nativize has been said to lie in their very nature as ‘second-language’ Englishes that are learned mostly in a school setting, a factor also responsible for the emergence of shared features among new Englishes worldwide, the so-called “Angloversals” (Mair 2003a: 84). As Kachru 1986 once stated for South Asia, the ‘South Asianness’ of South Asian English is in part reflected in “the conditions under which it is acquired in various parts of the sub-continent” (38). It is in reference to its acquisitional context that IndE has, for example, been characterized as being ‘bookish’ and ‘formal’ (cf. Pattanayak 1978: 185). This hypothesis is taken up in Chapter 3.3, which also zooms in on other aspects of lexical style variation for which Englishes worldwide are variable and which have so far been neglected in studies on IndE: the use of function words (*among/amongst, while/whilst, amid/amidst, until/till*) and the use of contractions. In addition, the very ‘learnability’ of forms and structures has been pointed out as a determining factor for variation in Outer Circle Englishes. According to this argument, areas that are relatively difficult to acquire for learners of English may be particularly likely to undergo change (cf. Thomason 2001: 76). This argument is, for example, found in Agnihotri et al.’s (1988) analysis of the IndE past perfect, which arrives at the conclusion that the past perfect is rare in IndE student writing because “students find the use of the past perfect more difficult than the use of the present perfect” (94). The past perfect, itself used variably in BrE and AmE (which is why the variable has been placed near the center in Figure 2.4), is studied in detail in Chapter 5.4. Learner difficulties have been said to affect variation in IndE also in the use of near-but-not-identical words (Chapter 3.3), particle verbs (Chapters 4.1 and 4.2), verb complementation (Chapter 4.3), the use of definite, indefinite and zero articles (Chapter 5.1), the use of count, noncount and collective nouns (Chapters 5.2 and 5.3), the use of tense forms (Chapters 5.4 and 5.5) and question formation (Chapter 5.7).

The variables positioned on the right in Figure 2.4 were chosen to measure the extent to which IndE could be said to be exonormatively oriented. The use of lexical Briticisms and Americanisms in the IndE lexicon is studied in greater detail in Chapter 3.4. Concord with collective nouns (Chapter 5.3) and the mandative subjunctive (Chapter 5.6) were chosen as syntactic variables to describe the interdependencies between IndE, BrE and AmE in a diachronically sensitive fashion. In the case of concord with collective nouns, contemporary BrE differs from AmE in using plural verbal concord with many collective nouns that are normally used with singular verbal concord in AmE. This synchronic difference between the two prestige varieties has been related to a BrE-internal development in the late nineteenth and early twentieth centuries. Could IndE today be shown to orient itself toward BrE usage practices in adopting plural verbal concord? In the case of the mandative subjunctive, on the other hand, it is the postcolonial force of AmE that has been said to be responsible for reintroducing the mandative subjunctive on a worldwide scale. Would IndE turn out to be affected by this development?

What looks rather tidy in Figure 2.4 posed complex analytical and interpretative challenges throughout the investigation of the data. Especially in the field of morpho-syntax, neat categorizations of findings were usually impossible. Variables like the definite and indefinite article or the past perfect were the most challenging from an analytical point of view because endonormative as well as exonormative factors interacted in a highly complex manner to create the variability patterns in the corpus data. Likewise, features that were found to be shared for example by IndE and AmE did not necessarily represent the postcolonial force of AmE, but also had to be given a diachronic comparative reading to see whether they were possibly the result of divergent diachronic developments in BrE. In some cases findings in contemporary IndE could be given a historical interpretation by turning to the *Oxford English Dictionary* (OED) and its quotations base. Other distributions await a final interpretation and remain a linguistic riddle to be solved by future research.

CHAPTER 3

Vocabulary

It is well-known that the potential open-endedness of the English lexicon invites innovation and regional variation around the English-speaking world. But while it is comparatively easy to list individual words and hold them up as lexical peculiarities of IndE, it is more difficult to address issues such as the following: What is the currency of a lexical item on a pan-Indian scale? What is its relative position in the IndE lexicon alongside its synonymous or near-synonymous alternatives? And what are its grammatical properties and its communicative load? Since much of the available lexicographical research on IndE has not been based on corpus data, those questions are yet to be answered for many of the items that have been suggested as Indianisms (cf. Nihalani et al. 1979, 2004; Hawkins 1984; Muthiah 1991; Crowther 1996; Yadurajan 2001; Hankin 2003). This chapter takes a closer look at those issues with the help of the Primary Corpus, the Kolhapur Corpus, ICE-India and the online corpora. The focus is on loanwords from Indian languages (3.1), neologisms and IndE-specific meanings (3.2), lexical style variation (3.3), and the use of lexical Britishisms and Americanisms in IndE (3.4).

3.1 Loanwords

The study of the IndE lexicon has in large parts been a study of its loanwords (e.g. Yule/Burnell 1994 [1886]; Subba Rao 1950; Kachru 1965, 1982b, 1983, 2005; Shastri 1983, 1988a, 1992; Bakshi 1991; Lewis 1991; Muthiah 1991; Carls 1994; Hankin 2003), which is not surprising considering the long history of language contact in the Indian subcontinent. Words from Sanskrit, Persian and Arabic started out their journey into the English lexicon here, as did numerous items from foreign languages such as Portuguese with which English has come in contact in South Asia throughout the centuries (cf. Kachru 1986: 33–6, 1994: 500–8; Leitner 1991a; Hansen et al. 1996: 210–6; Kachru/Nelson 2006: 153–5). Today, it is the indigenous languages of India that provide additional lexical resources for Indian users of English to draw on, and IndE has been described as a *khichri* of words (Kohli 1989: 1332), i.e. a mixture of words like *brahmin* (< Sanskrit), *loot* (< Hindi), and *thug* (< Hindi), which have long entered the international common core of English lexis, as well as *bandh* ‘general strike’ (< Hindi), *hammal* ‘porter’ (< Arabic), and *jawan* ‘ordinary soldier, private’ (< Urdu), which are more restricted in currency to the Indian subcontinent. While it is uncontroversial that borrowing from India’s background languages is a productive process

contributing to the innovation of the IndE lexicon, contradictory statements have been put forward regarding its exact impact. Some have viewed borrowing as a major driving force of lexical innovation, as “an important contributory factor in the process of Indianization” (Shastri 1988a:38) which is not “arbitrary” in nature but “register-specific” (Kachru 1978:85). Others have warned that the overall influence of lexical borrowing should not be overstated. Dubey 1989, for example, finds that loanwords occur in “minuscule proportions” (93) in his micro-corpus of English-language newspapers from India, and Görlach 1995, too, claims with respect to IndE broadsheets that “the number of words from Indian languages contained in them is quite small – smaller in fact than, say, the number of English loanwords in German newspapers” (49).¹ But are loanwords from Indian languages really as rare in contemporary IndE as Dubey 1989 and Görlach 1995 claim with respect to journalistic prose?

3.1.1 The distribution of loanwords in the Primary Corpus

Tackling this question on an empirical basis is not a straightforward task as any given answer crucially depends on how the line is drawn between “loanwords,” unintegrated “foreign” words and fully assimilated “English” words.² To be able to make a quantitative statement, the Primary Corpus was consulted first and the vocabulary from indigenous Indian languages in the database categorized with the help of the following set of criteria. Since my interest was in the proportion of variety-specific loanwords that would qualify as characteristic lexical features of contemporary IndE, it was decided a priori to exclude all those words from consideration that could be said to be fully assimilated in international English. To identify those, four major dictionaries of English (*Concise Oxford Dictionary* [COD], *Merriam-Webster's Collegiate Dictionary* [MW], *Collins-Cobuild English Dictionary for Advanced Learners* [COB] and *Cambridge International Dictionary of English* [CAM]) were chosen as benchmarks.³ Among the words of South Asian origin excluded from consideration are, for example, *brahmin* and *guru*, *karma* and *mantra*, and *sari* and *pardah* – items which are no longer restricted in use to South Asia only but are firmly established inter-

1. Cf. Leitner (1991b:222), who observed that “nativized, i.e. Indian words, such as *bandh* ‘strike’, occur very infrequently” in the Kolhapur Corpus.

2. Dubey (1989, 1991) and Shastri 1988a fail to clearly outline the criteria on which their counts were based.

3. An item was treated as belonging in the international common core if it was found in both COD and MW (both dictionaries representing the two major varieties of English today) and at least in one of the other two dictionaries (COB and CAM). Multi-word tokens were counted as single items when forming semantic units. For example, *Panchayati Raj* was treated as one item because of its function as a generic term for Indian administrative units at the village level. In contrast, the combination *panchayat samiti* was counted as two words (also cf. Shastri 1992:265).

Table 3.1 Loanwords from Indian languages.

	Types / text	Tokens / text	Tokens / 1,000 words	Type/token ratio
IND1978 Press				
Reportage	2.2	8.4	4.2	0.26
Editorials	2.6	8.3	4.2	0.31
Reviews	19.5	34.0	17.0	0.57
IND2000 Press				
Reportage	2.8	9.1	4.6	0.30
Editorials	4.7	7.6	3.8	0.61
Reviews	16.2	28.1	14.1	0.58
IND2000 Broad				
News	2.4	7.2	3.6	0.34
Interviews	2.1	5.0	2.5	0.42
Discussions	1.4	2.8	1.4	0.50
IND2000 Ess				
Student Essays	1.8	2.3	1.2	0.78

nationally, this way ceasing to function as characteristically IndE lexical features in absolute terms. In addition, unique names as well as indigenous vocabulary in quoted passages (as for example in book titles) were also excluded.⁴ Table 3.1 gives the distribution that was arrived at on those grounds.

Considering the numerous assumptions underlying the count, Table 3.1 must be read with caution, and the limited size of the database does not allow far-reaching generalizations to be made either. But even against that background, it seems safe to conclude that loanwords from Indian languages are rare in the material under investigation – a finding which supports Dubey's (1989) and Görlach's (1995) claims that lexical innovation in IndE does not happen extensively through borrowing. In fact, when comparing the newspaper material from 1978 and 2000, it appears that the proportion of loan vocabulary had slightly decreased in that particular domain during the given timespan. To leave it at that, however, would mean to miss the point that the distribution of loanwords is not arbitrary. A comparison of the key figures 'Types per Text' and 'Tokens per Text' across the text categories reveals that loanwords do not stratify evenly in the Primary Corpus but tend to cluster in one particular text category, namely the press reviews. In the datasets from 1978 and from 2000, the number of types and tokens per text is notably higher in the reviews section than in the other text categories,

4. The line between proper names and common nouns is obviously not clear-cut, for instance in the case of *Rashtrapathi Bhawan*, which functions both as proper name (denoting the residence of the President of India in Delhi) and common noun (when used metonymically to refer to the President of India and his political activities) on analogy of *White House* or *Downing Street*.

which suggests a diachronically stable affinity of those types of texts to the use of loan vocabulary. In the following, the IndE press reviews are therefore analyzed first.

3.1.2 The use of loanwords in IndE press reviews

One explanation for the comparatively high number of loanwords in the press reviews section lies in the fact that this text category subsumes numerous texts that communicate aspects of Indian life which are specific to Indian culture and tradition, i.e. texts that “emphasize local Indian traditions, such as food, music, religion, etc.” (Leitner 1991b: 222). This in turn has necessitated the frequent use of loanwords to refer to concepts and properties for which the English lexicon provides no straightforward equivalents, for example in domains such as religion, philosophy, art, and music. Shastri 1992 labels those texts “heavily Indian” (265), but the following examples from the Primary Corpus illustrate that the functions of loanwords in such textual environments are more variegated than this all-encompassing national label suggests.

The first example from the reviews section is from the review of a classical CD and features several lexical borrowings that refer to the universe of Indian classical music. Consider the following passage:

- (1) EXCERPTS FROM the live concert of P. S. Narayanaswamy, a senior member of the Semmangudi school, held at the Narada Gana Sabha, have been recorded in two albums by Gramophone Company. V. Thiagarajan is the violinist, K. V. Prasad the mridangist and on the kanjira is V. Nagarajan. The Kothavasal Venkatarama Iyer varnam ‘Sarasuda’ in Saveri sung in two speeds, facilitates a spirited beginning. The inlay card gives the name of the composer as Kothavasal Natesa Iyer. The Huseni kriti of Tyagaraja ‘Rama Ninne Namminanu’ with a spate of sarvalaghu swara prastharas rendered in the mode of the artiste’s preceptor accelerates the momentum. The raga alapana of Bilahari with all the requisite ingredients present in equal measure is of commendable quality. Thiagarajan plays an essay with pertinent prayogas that help to clearly reveal the raga’s swaroota.

[IND2000 Press Rev 02]

Several items (e.g. *mridangist* ‘musician playing a *mridangam*, a long drum made of wood or clay’; *kanjira* ‘open drum with small cymbals woven into the frame’) and concepts that are important in Carnatic music (e.g. *raga* ‘a pattern of notes used as a basis for improvisation’; *varnam* ‘a piece of music played at the beginning of a raga concert’; *kriti* ‘musical format related to raga’; *swaroota* ‘form, spirit of a raga’; *prayoga* ‘particular musical phrase that is characteristic of a raga’; *swara* ‘a type of musical sound measured in terms of its relative (higher or lower) position to a note’; *sarvalaghu swara prasthara* ‘special combination of swaras’; *Bilahari*, *Huseni*, *Saveri* ‘names

for special types of ragas') are referred to by loanwords. Note that the long-windedness of the annotations contrasts sharply with the effectiveness of the loanwords, which function as precise referents to the world of raga music, for which English has no exact words. But there is more to the choice of the loanwords in the passage above than the filling of lexical gaps. Their choice adds to the elevated style of the review, which is seen elsewhere in the passage in the careful use of formal expressions such as *the artiste's preceptor*, *requisite ingredients*, and *pertinent prayogas*. In that context, the high density of loanwords can be interpreted as signaling the very learnedness of the reviewer about one of the highest forms of musical art in India. Note too that no annotations are provided anywhere in the passage, which suggests that the text was intended for a readership familiar with Indian classical music. For those unfamiliar with Carnatic music, the review is likely to be unintelligible.

What is of general significance here is the fact that next to illustrating the effectiveness of loanwords for filling gaps in the English lexicon, example (1) also shows loanwords to be capable of taking over contextually relevant expressive functions, as that of marking a highly learned style. This is usually overlooked in classifications of lexical loans that are based on semantic criteria alone (see, for example, Kachru (1978, 1983) or, for a more recent semantic classification of loanwords in IndE, Carls 1994). The example above is particularly insightful in that it puts a question mark on Verma's claim that "[f]ormal style and language mixture are normally mutually exclusive." (qtd. in Mehrotra 1982: 162)⁵ Corpus evidence shows that extensive borrowing and formal diction may go together well, as is the case in the passage above. Note that reviews of classical concerts and classical recordings are attested in the 1978 reviews section, too, and, generally speaking, are easy to come by when browsing India's quality broadsheets today, which suggests that this text genre has become firmly established in IndE journalism.⁶

That loanwords may carry a significant expressive potential in more casual contexts as well is exemplified in the second passage, which is from the opening paragraph of a book review. Consider:

5. A similar generalization is found in Pandharipande 1987, who states that "loan vocabulary ... is readily used in the informal variety but not in the formal variety of Indian English." (154)

6. Cf. Kachru (1982b: 334–5) for further examples of IndE concert reviews from Indian quality newspapers. Mehrotra (1982: 162), too, observes loanwords and "language-mixing" to be particularly common in the reporting of musical concerts. Other registers and text types not represented in the Primary Corpus but described by Mehrotra 1982 as being open to loanwords from Indian languages include "college slang, film reporting, matrimonial advertisements, astrological forecasts, religious discourse, private letters, market reviews, casual conversation, cookery [books] ... Indian English fiction, and question-box in Indian periodicals." (162) Cf. Mehrotra 1998 for examples of those text types.

- (2) The twenty articles which appear in this collection were originally written for India Magazine, under Frieze's editorship, and which was, as Frieze puts it, "the most readable Indian magazine you've never read"[sic]. A sign of our times. Whom gods love, get left in the cupboard. Or worse still, are reincarnated as mournful paper bags, *thongas*, doomed to sheltering *moong dal*, *maida*, *bhelpuris*, *pakor*as, onions, potatoes and sometimes even cauliflowers.

[IND2000 Press Rev 06]

With the exception of *thonga* 'paper bag', the loanwords in the passage above fall in the semantic domain of food and cooking (*moong dal* 'dish made of lentils'; *maida* 'refined wheat flour'; *bhelpuri* 'dish of puffed rice, spices and hot, sweet and sour chutney, often served with small puris'; *pakora* 'fritter, usually of vegetables, meat or cheese, cooked with spices'), a word-field that has been observed to be particularly open to adopting loanwords from indigenous languages the English-speaking world over (cf. Platt et al. 1984: 89). But what accounts for the choice of loanwords in the example above is not so much this insight as the fact that their enumeration is part of a deliberately employed literary strategy, which, in a jocular fashion, aims to situate the review article, in its opening paragraph, firmly in the Indian scene. The line that the reviewer draws between the concept of reincarnation on the one hand and the making of paper bags out of old magazines is primarily meant to entertain and capture readers' interest in what is to be presented in the main body of the review, namely the criticism of a collection of short stories by Indian writers. In this context, the reference to Indian dishes can quite literally be said to add a local taste to the introductory paragraph. So does the deliberate addition of *thonga* to its English equivalent *paper bag*, an informal Hindi item (cf. Wagenaar et al. 1993: 667) that contributes to the mildly casual tone permeating the entire passage – if readers are familiar with the word, that is. The fact that *thonga* is rarely ever found in IndE quality broadsheets suggests a restricted currency of the word in IndE today and appears to have made the presence of the English synonym in the passage necessary to guarantee the success of this light-hearted introduction (cf. Appendix 3.1).

If the first two passages illustrate the functional potential of loanwords in formal and mildly informal contexts, the third set of examples from the press reviews section marks the most informal end of the stylistic cline on which the loanwords in that text category fall. The following loans made their way into the database through movie reviews.

- (3) Salman Khan: *Filmi* stage show Hrithik ishtyle [IND2000 Press Rev 04]
 (4) The film does nothing to highlight the Kashmir issue. Just another noble attempt gone wrong. Watch it for action *masala*. [IND2000 Press Rev 04]
 (5) ... this is a total downer, considering the *baap* of action flicks is associated with it. John Woo *saab* goes haywire with weird motorcycle chases and leaps in the air that defy gravity. Slow motion *zindabad*! [IND2000 Press Rev 04]

- (6) In an industry where the dance director is God and chalu music adds up to instant popularity of a film, 'Astitva' comes as a breath of fresh air.

[IND2000 Press Rev 09]

Unlike in the previous two examples, the loan items in examples (3) through (6) cannot be said to lack straightforward English equivalents, which would in fact be easily available. What has motivated the choice of the nouns *filmi* 'film star' (example (3)) and *masala* 'spices' (example (4)), the interjection *zindabad* 'long live!' (example (5)), and the adjective *chalu* 'cunning' (example (6)) in the first place are the strong associations of casualness and colloquial speech that they are likely to evoke. All items have been codified in Crowther 1996 and, from that point of view, qualify as established loan items. However, there are major stylistic restrictions to their currency. Crowther 1996 labels *filmi*, a hybrid Hindi-English coinage made up of the English noun *film* and the Hindi derivational suffix *-i*, a slang item (1447).⁷ *Chalu*, a borrowing from Hindi, is assigned the label "informal" (1441), and the noncount use of *masala* (also borrowed from Hindi) in example (4) is given a separate entry that points to the strong sexual connotation of the word in contexts like the one exemplified above ('the power to cause sexual desire in people', 1458). Interestingly, the impression of great informality triggered by those lexical choices is furthered at the syntactic level by means of using elliptical constructions (*Just another attempt gone wrong*; example (4)) and imperative forms (*Watch it for action masala*; *Slow motion zindabad*; examples (4) and (5)), which are strongly reminiscent of colloquial speech. In this context, the utterance *slow motion zindabad* in example (5) makes for a particularly oral ring by recreating the word order of Hindi.⁸ In fact, it is their proximity to colloquial IndE that sets examples (3) through (6) most clearly apart from the examples discussed previously. Their 'pseudo-oral' written texture functions as a highly stylized representation of a code-mixed amalgam of Hindi and English that is meant to involve and affect readers by recreating the sounds and rhythms of colloquial Hinglish.⁹ Note that the movie reviews from which examples (3) through (6) are taken are, on the whole, closer to the English than the Hindi pole of the continuum. There are, for example, no loanwords in their main paragraphs. Only at carefully selected points within the texts, namely in headlines (example (3)) and in closing statements (examples (4), (5), and

7. Both *film* and *filmi* are used in Hindi today, cf. Khan (1994: 159): "[the] English word 'film' has undergone morphological change through suffixation as per Hindi grammatical rules. It is a case of lexical borrowing as well as morpho-syntactical adaptation; 'film' is fully Hindiized." For an attestation of *film* in spoken Hindi cf. Singh (1995: 133).

8. *Zindabad* has also been discussed in reference to PakE by Baumgardner (1992: 132–3).

9. The term *Hinglish*, D'souza (1996: 251) explains, has become an established term in India to denote a wide range of mixed forms of Hindi and English found in the Hindi-speaking territories in India, comprising predominantly Hindi-based manifestations at one end of the cline all the way to English-dominant forms at the other, with a continuum of intermediate degrees of language mixture in between.

(6)), do the journalists employ explicitly literary strategies to symbolically suggest the flow of colloquial code-mixing (cf. Mair 1992a), for example by playing with the graphic surface of their texts. The phonetic spellings <saab> (for *sahib* 'title added to the name of the position somebody holds') and <baap> (for *baba* 'father') contribute to creating an *effet de réel* by triggering associations of rapid speech, and the creative spelling <ishtyle> (for *style*) carries a particularly strong connotation to a South Asian English accent by alluding to a phonetic feature not uncommon in the English of many speakers of Indo-Aryan languages such as Hindi, Punjabi, or Kashmiri, namely the epenthetic vowel before consonant clusters like *st* or *sk*.¹⁰ Interestingly, the status of <saab> and <baap> as nonstandard spellings is made explicit in examples (4) and (5) by their typographical highlighting through italics, as if the reviewers wanted to signal their awareness of the fact that their spelling was not in accordance with standard orthography. Note that the strategy of graphic highlighting has been extended in those examples to the informal loan items *masala* and *zindabad* as well. By contrast, examples (3) and (6) do without any typographical marking, and *filmi* and *chalu* function more unobtrusively as representatives of highly informal Hinglish.

Taken together, the six examples discussed in detail above illustrate the enormous functional potential of loanwords in the press reviews section of the Primary Corpus – a potential that has contributed to the relatively high number of loanwords in that particular text category. What is especially interesting in this regard is the following diachronic observation. There seems to be a mild difference between the data from 1978 and from 2000 as far as the occurrence of informal loan vocabulary in the corpus is concerned. While the reviews from the year 2000 yield more than a handful of examples of informal items such as *thonga* or *chalu*, most in the context of what Khan 1994 terms “cine-journalism” (169), no single such informal loanword can be found in the sample from 1978. In other words, the comparison of the 1978 and 2000 reviews suggests the slow spread of informal loanwords into India's quality broadsheets after 1978. Although this finding cannot be given a statistical backing, it appears to reflect a slightly more relaxed, more self-confident attitude on the part of IndE quality broadsheets toward using such items in the year 2000 than twenty-two years before. This, in fact, seems to be a linguistic repercussion of a larger trend in IndE that has been observed among young urban professionals, college students, artists and celebrities in India for quite some time, namely “[t]he explosion of Hinglish” (Schneider 2007: 171). The growing popularity of Hinglish and other mixed forms as an identity-marking expression of a new, modern Indian lifestyle has been linked to the arrival of satellite and cable television in India in the 1990s and the ensuing impact of youth channels such as MTV India and Channel V (cf. Mehrotra 1998: 14; Schneider 2007: 171; Bhaya Nair 2008: 490). According to Trivedi 2006, Hinglish has also been on the rise in Hindi cinema “over the past decade” (80), where it is nowadays particularly common in Hindi films with foreign settings reflecting “an upper-class cosmopolitan Indian mobility” (80). Similar developments have been observed

10. Cf. Kachru (1986: 39); McArthur (2002: 321).

in Southern Indian and Sri Lankan popular cinema by Dissanayake 2008, who points out that code-mixing between English and languages like Tamil and Sinhalese has become “widely popular” (403) recently. In advertising and in commercials, too, code-mixing is employed extensively these days – even in rural areas, where, according to Bhatia 2000, Hinglish oftentimes functions to create a “[m]odern, western, scientific, innovative” (211) product appeal. It will be informative to follow the developments in IndE in this regard more closely in the future and address the question “whether ‘Hinglish’ or any other mixed code can fill the vacuum of a non-existing pan-ethnic informal identity carrier and can be viewed as the penetration of an English-derived variety into the lower echelons of society” (Schneider 2007: 171). In light of the evidence discussed above, it can be established that India’s quality broadsheets had started to open up to Hinglish by the beginning of the new millennium.

3.1.3 The use of loanwords outside the Press Reviews section

When we move outside the press reviews section, loanwords occur less frequently in the Primary Corpus, as the quantitative analysis above has brought to light (cf. Table 3.1). Moreover, a qualitative analysis of the loanwords in those text categories reveals that the expressive functions to which the loans are put to use are more restricted. In particular, markedly creative uses of the types illustrated in examples (3) through (6) are absent. Let us look at the text categories in more detail.

In the categories of spontaneous speech, casual uses of loanwords are rare. Loanwords and loan expressions which previous research has observed to recur frequently in spontaneous IndE speech (see e.g. Valentine 1991:331), for instance the Hindi agreement marker *accha* ‘okay’ or the commonly heard expression *teekh hai* ‘all right’, are noticeably absent in the Primary Corpus, which suggests that their use in contemporary IndE speech is yet restricted to more private and more casual communication situations.¹¹ If we do come across such items, it is, for example, in the context of rapid language use, as in the following exchange between speaker C (a former Union minister; female) and speaker D (a university professor; male), which is taken from a discussion on India’s population policy. In the passage below, speaker D employs the item *chota* ‘small’ twice for a marked shift in tone that can be said to function both as an audible sign of disagreement and, at the discourse level, as an attempt to reclaim the floor from speaker C.¹²

11. For *accha*, this is corroborated by evidence from the larger ICE-India. Of the 242 tokens of *accha* in ICE-India (all found in spoken language), 201 tokens (83.4%) appear in texts representing private dialogues (category S1A), 40 tokens (16.6%) in texts representing public dialogues (category S1B) and only one token in spontaneous commentary (category S2A).

12. *Chota* is listed as Indianism by Carls (1994:210) and Kachru (1994:516) and has also been attested for PakE by Baumgardner (1993c:93). For a survey of the discourse functions of Hindi-English code switching also cf. Gupta 1991.

- (7) <\$C><#>you know I I think he has we have a fundamental difference in our interpretation of what is a politician<,> <#>a politician is not just a minister<,> <#>the person in the zilla parishads and the mandals<,> at the gram panchayati levels<,> are politicians¹³
 <\$D><#>yes
 <\$C><#>they're opinion-makers
 <\$D><#>>no I mean the politicians in Delhi
 <\$C><#>they have the they have the vehicle the the the way of going into that and<,> orienting those people to accept this
 <\$D><#>right
 <\$C><#>so <./> you can't do without the politician<,>
 <\$D><#>no no these are the chota politicians
 <\$C><#>I mean you give me an alternate agency and I'm willing to step back
 <\$D><#>no I was not talking of chota politicians I was talking of the people in Delhi
 [IND2000 Broad Disc 07]

While such uses are exceptional in the Primary Corpus speech data, loan items functioning in more formal contexts as deferential forms of address are common. A case in point is the affix *-ji*, which is employed recurrently by users of IndE whenever they wish to indicate their respect for their addressees (cf. Crowther 1996: 1453; McArthur 2002: 322; Hankin 2003: 238). The affix *-ji* combines with first names and last names, as is illustrated in the following two examples:

- (8) <\$D><#>Kesriji<,> has<,> had a very long<,> career of service<,> to the country ...
 [IND2000 Broad News 11]
 (9) <\$F><#>Margaretji there is here a clarification from the PMO<,> saying that he has never made this statement<,> <#>I don't know whom to believe
 [IND2000 Broad Disc 08]

In example (8), which is taken from an interview given by speaker D on the occasion of the death of a well-known politician, the choice of *-ji* after the last name signals speaker D's respect for the deceased. In example (9), which is taken from a broadcast discussion in which the participants have so far referred to each other by their first names only, speaker F adds *-ji* to the name of his addressee for a similar communicative purpose: the choice of the honorific makes his critical statement on his addressee's credibility sound less offensive or 'face-threatening'. In similar ways, *sri* and *sahib* may be drawn on in IndE when polite forms of reference are sought for. While *sri* is used before the names of persons, *sahib* is added to names and is commonly found with (but is not exclusive to) Muslim names (cf. Crowther 1996: 1467–8). Consider:

13. *Zilla* 'small district'; *parishad* 'village'; *mandal* 'division'; *gram* 'village'; *panchayat* 'council'.

- (10) <\$B><#>I think that would be uh very unfair to the memory of uh Sri Rajesh Pilot<,> <#>uh<,> we must uh carry out from where he left
[IND2000 Broad Int 01]
- (11) <\$B><#>my father was a star Hafiz Ali Khan sahib <#>it was my privilege that I was born<,> <#>and he was my guru he was my father
[IND2000 Broad Int 05]

In more informal registers, *sahib* may nowadays also be used in a joking fashion to connote that “someone has become westernized” (Crowther 1996: 1468), which is a play on the common use of the word in reference to “male Europeans”. Such a use has been illustrated in example (5), where the reference is to a Hollywood-based movie director from Hong Kong (also cf. Verma 1978: 210, 1982: 180; Trudgill/Hannah 2002: 135; Hankin 2003: 455–6; Gargesh 2006: 105; Schneider 2007: 165).

The vast majority of the loans in the press and broadcast data are nouns denoting aspects of political and public life. If we take Kachru’s list of semantic fields (1978: 93–6) as a yardstick, the domains of politics, administration, and religion are represented more conspicuously here, whereas loanwords referring to more domestic areas of life such as clothing and food are almost absent. See the following representative examples:

- (12) Local militants head the councils set up in every tehsil.
[IND2000 Press Rep 02]
- (13) The Kutch police late on Wednesday seized 22 kg of plastic explosives (PEK) ... from Kamaguna village bordering Bhuj and Nakhtrana talukas.
[IND2000 Press Rep 14]
- (14) <\$A><#>employees of the water-supply department along with technicians from the panchayat samitis<,> ... are working to repair the remaining unused hand pumps ...
[IND2000 Broad News 13]
- (15) <\$A><#>with the opposition laying down arms the Vajpayee government today managed to push through the Information Technology Bill<,> with near consensus in the Lok Sabha<,> <#>the bill dubbed the mother bill for the new millennium will be tabled in the Rajya Sabha on Wednesday
[IND2000 Broad News 20]
- (16) Though the main task of a Vidhan Sabha is to legislate it hardly gets enough time for it.
[IND1978 Press Edit 09]

Tehsil (example (12)) and *taluka* (example (13)) are long-standing administrative terms in IndE, each denoting ‘a smaller division of a district’. According to the *Oxford English Dictionary*, *tehsil* is a loanword from Urdu (cf. Baumgardner 1993: 117),

whereas *taluka* is of Arabic origin.¹⁴ Today, the status of Hindi as India's national language (cf. Chapter 2.1.5) tends to favor the use and spread of Hindi administrative vocabulary like *panchayat samiti* 'village council' (example (14)), *Lok Sabha* 'lower house of the Indian parliament', *Rajya Sabha* 'upper house of the Indian parliament' (example (15)), and *Vidhan Sabha* '(lower) house of parliament in a union state or territory' (example (16)) on a pan-Indian scale (also cf. Labru 1984: 17; Carls 1994: 206; Hankin 2003: 290–1, 508). But it is important to note that the process of borrowing words from Hindi into IndE does not take place in a political vacuum. It is precisely for the reason that Hindi is not widely spoken in India's south that Hindi terms cannot be expected to gain pan-Indian currency automatically. Shastri (1988a: 36) and Dubey (1989: 89–90) both remark that the item *Vidhan Sabha* had not been established India-wide by the late 1980s. The Primary Corpus does not point to any further spread of the word by the turn of the millennium. On the contrary, while in the 1978 press section *Vidhan Sabha* occurs six times (four instances of which are found in one editorial, though), it is not attested in the data from the year 2000. ICE-India, too, contains only one example of *Vidhan Sabha*.¹⁵ When references to local parliaments are made in the Primary Corpus and ICE-India, they are made by means of the English item *assembly*:

- (17) <\$K><#>the issue was initiated yesterday in the Uttar Pradesh assembly<,>
by Mr Mishra who sought the permission of the chair to clarify his position
in respect to the charges levelled against him [IND2000 Broad News 05]

A look at the Internet puts the conclusion drawn from the closed corpora into perspective, though: *Vidhan Sabha* can be found recurrently in the online editions of India's major national dailies today as are *panchayat samiti*, *Lok Sabha* and *Rajya Sabha* (cf. Appendix 3.3). Moreover, an additional unrestricted Google search on the entire Internet (February 14, 2001) for collocations of the types 'X Vidhan Sabha' and 'X assembly' (with X standing for the names of India's states and union territories) yielded numerous further attestations of the Hindi word in reference to state parliaments across India in other IndE text categories as variant as official government

14. Both items are found in the *Oxford English Dictionary*, the first entry of *tehsil* dating back to 1846 and that of *taluka* to 1793. Yule/Burnell (1994[1886]: 894) only included *talook* in the *Hobson-Jobson*. According to the OED, <tehsil> has an alternative spelling in <tahsil>. Corpus evidence shows that this variant is nowadays exceptionally rare in IndE journalism; cf. Appendix 3.2. By contrast, the two spelling variants listed in the OED, <taluka> and <taluk>, are common in Indian press texts, with individual newspapers displaying clear preferences for one or the other version; cf. Appendix 3.2. The spellings <taluq>, also listed in the OED, and <talook> are rare.

15. The only example of *Vidhan Sabha* in ICE-India is found in news broadcast (text S2B-010) in reference to the state of Himachal Pradesh.

Table 3.2 Collocations of *Vidhan Sabha*.

	Assembly	Vidhan Sabha	Assembly : Vidhan Sabha
Chhattisgarh ...	8	13	1 : 2
Himachal Pradesh ...	115	123	1 : 1
Haryana ...	535	433	1 : 1
Punjab ...	2070	1200	2 : 1
Uttaranchal ...	503	210	2 : 1
Madhya Pradesh ...	344	77	4 : 1
Rajasthan ...	274	39	7 : 1
Uttar Pradesh ...	1570	79	20 : 1
Delhi ...	1690	80	21 : 1
Tripura ...	86	4	22 : 1
Jharkhand ...	336	10	34 : 1
Mizoram ...	173	5	35 : 1
Gujarat ...	1390	29	48 : 1
Bihar ...	1260	26	48 : 1
Maharashtra ...	802	15	53 : 1
Nagaland ...	73	1	73 : 1
West Bengal ...	772	2	386 : 1
Jammu and Kashmir ...	2010	4	503 : 1
Goa ...	2530	3	843 : 1
Manipur ...	1020	1	1020 : 1
Kerala ...	3190	3	1063 : 1

publications, academic writing, and private websites. Consider the distribution of the respective variants in Table 3.2.

What is particularly interesting in reference to Table 3.2 is that the individual collocations are not distributed evenly on the World Wide Web. In fact, a noteworthy regional distribution of the variants can be observed. The snapshot from the Internet shows that *Vidhan Sabha* is used on the Internet at least once in every ten references to local parliaments only in combination with the names *Chhattisgarh*, *Himachal Pradesh*, *Haryana*, *Punjab*, *Uttaranchal*, *Madhya Pradesh*, and *Rajasthan*. The fact that the search query did not return a single attestation of *Vidhan Sabha* collocating with place names such as *Andaman and Nicobar Islands*, *Andhra Pradesh*, *Arunachal Pradesh*, *Assam*, *Dadra and Nagar Haveli*, *Daman and Diu*, *Karnataka*, *Lakshadweep*, *Meghalaya*, *Orissa*, *Pondicherry*, *Sikkim* and *Tamil Nadu* suggests that the item has next to no foothold in states and territories in which Hindi is not spoken as a first language (cf. Appendix 3.4). But even with names of states belonging to the Hindi heartland, for instance Uttar Pradesh, the term *Vidhan Sabha* appears not to be preferred. The item *assembly*, on the other hand, functions as a pan-Indian and international

lexical link, which is indicated by the fact that for each state and Union territory we get numerous examples of “X Assembly”.¹⁶

What does this snapshot from the Internet amount to? The case of *assembly* vs. *Vidhan Sabha* illustrates that IndE lexical variability may well be exploited differently across the subcontinent to meet region-specific communicative needs, with Hindi loanwords not automatically gaining ground in non-Hindi-speaking areas. In this light, it is interesting to observe that official publications by India’s Ministry of Statistics normally employ *Vidhan Sabha* in reference to all state assemblies across India (cf., for example, Singh 2000). This qualifies as a marked usage which may be said to function symbolically to assert a national identity through Hindi, India’s official national language. According to Carls 1994, loanwords from Hindi used “for events on a national scale or for institutions at the level of the Union” (207) are particularly likely to take over such symbolic function. The fact that IndE journalism usually prefers making references to state parliaments through the item *assembly*, on the other hand, renders the lexical style of IndE journalism more neutral than that of government publications in that particular instance. Note in this context that IndE news reporting is characterized by the lowest type/token ratio for loanwords in the Primary Corpus (cf. Table 3.1). In other words, the news sections are the least diversified text categories in the database as far as the density of different loanwords per text is concerned. This might reflect a general concern with supraregional (and international) intelligibility on the part of Indian news reporters, who can be expected to aspire to an unhampered information flow in order to appeal to audiences nationwide.

3.1.4 More variation: A look at South Asia

The case of *Vidhan Sabha* illustrates the need for more register-sensitive areal lexicographical research into the use of loanwords across India and, for that purpose, all of South Asia. Generally speaking, lexical similarities can be expected to exist among South Asian Englishes like IndE, PakeE, SrLE, or BaE today for various reasons. For one, loanwords borrowed into English before 1947 may still be in use across South Asia because of the common historical roots tying the subcontinent together. Also, the supraregional influence of shared background languages and the function of English as a South Asian lingua franca (cf. Ferguson 1996: 36) are likely to result in common lexical features among South Asian varieties of English.¹⁷ Since little is known at present about the nature of regional variation within and across the vocabularies of South Asian Englishes, it was decided to conduct a case study to see whether or not

16. Note that local parliaments may have two houses – a *Vidhan Sabha* and a *Vidhan Parishad*. Strictly speaking, “assembly” is synonymous with “House(s) of Parliament.”

17. Examples include Hindi/Urdu for India and Pakistan, Tamil for India and Sri Lanka, and Bengali for India and Bangladesh. Cf. Kachru (1994: 497–500) for a map of languages spoken across South Asia.

similar meaningful patterns of regional variation as those for *Vidhan Sabha* could also be uncovered for South Asia by consulting the Internet corpora. The words chosen for the analysis have in common that they denote various aspects of public protest and police intervention, a word field that has been discussed in reference to borrowing into IndE as well as PakE (e.g. Baumgardner 1993c: 185). Consider first the following set of examples from the Primary Corpus:

- (18) The farmers regrouped and tried to stage a dharna for the second time.
[IND2000 Press Rep 11]
- (19) <\$A><#>meanwhile shops businesses and commercial establishments in Kollam district<,> downed their shutters in response to a call for a hartal ...
[IND2000 Broad News 04]
- (20) The one-day Bombay bandh ... was marked by violent incidents and arson.
[IND1978 Press Rep 11]
- (21) The Minister for Youth Affairs and Small Industries, Mr. T Sitaram, was gheraoed by Congress and CPI(M) workers at Tekkali today ... The gherao, led by Youth Congress leaders, Mr. Ch. Ganapati and Mr. S. Satyam, Mr. N. Shanmukha Rao of CPI(M) and the former MLA, B. Narayana Swamy, continued for about half an hour.
[IND2000 Press Rep 11]
- (22) Farmers go on rampage, lathicharged
By Our Staff Reporter
ARMOOR, JUNE 24. Irate farmers protesting against the steep hike in power tariff went on the rampage on Saturday evening damaging over 50 RTC buses and private vehicles, after the police resorted to a lathicharge to disperse agitating protestors staging 'rasta roko' at Mamidipalli crossroads on national highway – 7. Several protesters sustained bleeding injuries in the lathicharge to clear them from the area after the 'rasta roko' continued for over eight hours. One person, identified as Mallepalli Gangaram of Armoor town, sustained grievous injury and was shifted to the District Headquarters Hospital, Nizamabad, as his condition turned serious.
[IND2000 Press Rep 11]

A look at the *Oxford English Dictionary* reveals that *dharna* 'sit-in protest', *lathi* 'long bamboo stick used by the police' and *hartal* 'strike involving all or most workers' are long-standing loanwords from Hindi that came into the English lexicon well before the partition of the subcontinent.¹⁸ *Gherao* 'encircling of a person to extract a concession',

18. The first attestation of *dharna* in the OED dates back to 1793, that of *lathi* to 1850 and that of the compound *lathicharge* to 1930. *Dhurna* and *lattee* are also entries in the *Hobson-Jobson* (cf. Yule/Burnell 1994 [1886]: 315–6; 510), but not *hartal*. The arrival of *hartal* in English might closely be related to an important political event in the history of India, namely Gandhi's opposition to the Rowlatt Act of 1919, to which Gandhi responded by proclaiming a

bandh 'general strike' and *rasta roko* 'road blockade', on the other hand, seem to have passed into common IndE usage only after 1947.¹⁹ The loan items in the examples above are codified in Crowther 1996, which suggests that they have become firmly established in the IndE lexicon.²⁰ But would it also be possible to trace the use of those items beyond the Indian borders across South Asia, which would justify speaking of a 'South Asian English' lexicon at least as far as English-language journalism is concerned? To address this issue, the six IndE online newspapers were consulted to obtain the frequencies of the six words alongside those of their nearest English equivalents. Those distributions were then compared to a set of data collected from three South Asian online publications from Bangladesh, Pakistan, and Sri Lanka, namely *The Independent* (Dhaka), *The Dawn* (Lahore), and the *Sunday Times* (Colombo). The comparison brought to light that while there is indeed a shared communicative need in the Bangladeshi, Indian, Lankan and Pakistani newspapers to report on various forms of public protest and police intervention, the nine publications consulted make variant use of the loanwords under investigation. Those distributions are discussed in more detail now.

The most straightforward differences among the online corpora from India, Pakistan, Bangladesh and Sri Lanka are observed for the items *bandh*, *dharna* and the hybrid formation *lathicharge*. Those items are attested hundreds of times in the IndE online newspapers and clearly outnumber their English near-equivalents, whereas the reverse picture emerges for the newspapers from Pakistan, Bangladesh, and

day of *hartal*, i.e. the '[o]rganized shutting of shops and cessation of business' (OED); cf. Spear (1965: 190–1). Its first attestation in the OED dates from 1920.

19. Only *gherao* has been recognized in the second edition of the OED, its first attestation dating from 1967. Tharoor 2007 humorously calls *gheraos* "India's contribution to the art of industrial disputes" (459) and goes on to add: "The notion of getting your own way by blockading your opponent in his office may have little in common with that of the self-sacrificial fast, but as a tactic of coercion it is used at least as often in India. Regrettably, there is no equivalent Indian invention on the conciliation side of the process." (459)

20. *Dharna* is also discussed by Baumgardner (1993c: 90) in reference to PakE and by Hankin (2003: 118) in reference to IndE. *Lathicharge* and/or *lathi* are listed as lexical Indianisms in Mehrotra (1982: 161), Labru (1984: 131), Kachru (1986: 42–3), Leitner (1989a: 170–1), Gramley/Pätzold (1992: 443), Carls (1994: 211), Kachru (1994: 516), Görlach (1995: 82, 89), George C (1998: 17), McArthur (2002: 322), Viereck et al. (2002: 210), Hankin (2003: 288) and Melchers/Shaw (2003: 142). It has also been discussed in reference to PakE by Baumgardner (1993c: 116–9). *Hartal* is mentioned as lexical Indianism by Shastri (1992: 264), Hansen et al. (1996: 22), Trudgill/Hannah (2002: 135) and Hankin (2003: 190); it has also been attested for PakE by Baumgardner (1993c: 90). For references to *gherao* in IndE cf. Mehrotra (1982: 161), Shastri (1988a: 40), Carls (1994: 206), Kachru (1994: 516), George C (1998: 17), Trudgill/Hannah (2002: 135), Hankin (2003: 162–3) and Melchers/Shaw (2003: 142). For references to *gherao* in PakE cf. Baumgardner (1992: 90, 1993c: 130–6). *Bandh* is listed as lexical Indianism in Carls (1994: 214), Trudgill/Hannah (2002: 135) and Hankin (2003: 32–3).

Table 3.3 *Lathi(-)charge* vs. *baton(-)charge*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
<i>lathicharge</i>	77	203	34	157	83	70	1	–	3
<i>lathi(-)charge</i>	286	60	55	151	615	7	30	1	7
<i>batoncharge</i>	–	1	–	–	–	–	1	–	2
<i>baton(-)charge</i>	5	4	3	12	22	–	259	12	15
<i>lathicharged</i>	27	88	9	177	89	19	1	–	1
<i>lathi(-)charged</i>	22	14	7	43	132	–	12	–	2
<i>batoncharged</i>	–	–	–	–	–	–	2	–	–
<i>baton(-)charged</i>	1	15	1	14	15	–	190	10	8

Sri Lanka. Consider as a representative example the distributions for *lathicharge* vis-à-vis *batoncharge* (March 15, 2001).

Table 3.3 shows that the hybrid compound *lathicharge* outnumbers its English counterpart *batoncharge* clearly in the IndE corpora, while at the same time it is nearly absent from the papers from Bangladesh, Pakistan, and Sri Lanka. The identical picture emerges for *bandh* and *dharna*.²¹ The near absence of *bandh*, *dharna*, and *lathicharge* in the Bangladeshi and Lankan papers most likely has to do with the fact that Hindi is not a major language in those countries, which is of course a major constraint for the spread and use of Hindi vocabulary across South Asia in general (cf. Kachru 2005:30). The figures for PakE cannot be given such a straightforward interpretation, however, as *bandh*, *dharna* and *lathi* are also nouns in Urdu, one of Pakistan's national languages, and have been attested for PakE before (Baumgardner 1993c:118, 185). In the case of *lathicharge*, the data from the Internet are in line with Kachru's (1983:154) observation that this hybrid compound outnumbers its English equivalent *batoncharge* in IndE but is virtually absent from PakE. On the other hand, the figures contrast sharply with Baumgardner's (1993c:118) statement that *lathicharge* and *batoncharge* are distributed equally in contemporary PakE. What may account for the distribution in the *Dawn* could be publication-specific conventions sanctioning the use of Urdu loanwords in high numbers, for example for reasons of international intelligibility. But given the preponderance of the three items in the IndE online corpora, another plausible explanation comes to mind. It is not unlikely that those words are widely perceived as Indianisms throughout South Asia today, in which case they might be avoided intentionally in communication contexts that would make connotations to India undesirable. This would nowadays most probably be the case in Pakistan's national broadsheets. Note in this context

21. To be able to compare the relative frequencies of *dharna* and its near-synonym *sit-in* in functionally compatible contexts, the online search was restricted to strings of the type *a/the dharna* and *a/the sit-in*. Verb uses of *sit* are thus not included in the count. Cf. Appendix 3.5.

Leitner's (1989a: 170–1) observation that *batoncharge*, and not *lathicharge*, was used in Benazir Bhutto's autobiography *Daughter of the East* from 1988. Whether or not that particular lexical choice reflects a deliberate decision on the part of the author to avoid a salient lexical Indianism or, more practically, the editor's orientation toward an international buying public cannot be reconstructed in hindsight, of course, and both factors may have played a role in that very instance. What can be said for certain on the basis of the evidence here is that, as of today, *lathicharge* as well as *bandh* and *dharna* have a stronger foothold in IndE journalistic writing than in the three comparable datasets from elsewhere in South Asia.

The interpretation of the search results for *hartal* and *gherao* is less straightforward, as those two words are generally less common in all nine online corpora than the high-frequency items discussed above (cf. Appendix 3.6). It is interesting to see, however, that both *hartal* and *gherao* have a comparatively strong foothold in Bangladesh's *Independent*. Likewise, the relatively high frequency of *hartal* in Sri Lanka's *Sunday Times*, which did not yield a single occurrence of either *bandh*, *dharna*, or *lathicharge* and only one example of *gherao*, is noteworthy. But the figures returned by Google are, all in all, too small to allow any final judgment to be passed on the nature of the differences between IndE journalistic writing on the one hand and Bangladeshi and Lankan English on the other. To be able to delve more thoroughly into the issue of regional variation, a more comprehensive areal survey covering more registers beside the one of journalistic prose would be necessary at this point. What is striking once again, however, is the fact that *gherao*, which has been discussed in reference to lexical innovation in PakE newspapers (Baumgardner 1992: 129), is outnumbered in the *Dawn* by the English near-synonymous word *picketing*. It turns out that among the nine newspapers looked at, Pakistan's *Dawn* is most reluctant to employ any of the loanwords studied here.

Of the six IndE loanwords tested for their regional currency across South Asia, *rasta roko* is the only item attested exclusively in the Indian newspapers around the year 2000. Whenever references to forms of protest involving the blocking of road traffic are made in the three non-Indian broadsheets at all, they are usually realized by the all-English nominal compound *road blockade* (cf. Appendix 3.7). Going back to the Hindi noun *raastaa* 'path, way' and the imperative form of the Hindi verb *roknaa* 'to block', *rasta roko* translates literally as 'block the way!' Today, *rasta roko* is a set phrase in contemporary Hindi and, as such, has been borrowed into IndE, where it is in the process of establishing itself as a noun alongside the English near-equivalent *road blockade*. What makes *rasta roko* stand out among the examples from the Primary Corpus is the fact that it is still set in inverted commas – as if to signal the uneasiness of the journalist regarding the appropriateness of this lexical choice. Further qualitative research on the basis of the six IndE online corpora (March 20, 2001) shows that this textual treatment of *rasta roko* in the Primary Corpus is not a coincidence but recurrent practice in IndE journalistic prose; cf. Table 3.4.

Table 3.4 *Rasta roko* in IndE newspapers.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE
rasta roko (total) of those:	43	63	41	10	62	2
singular	38	55	41	10	58	1
plural	5	8	–	–	4	1
annotations	–	–	–	–	–	–
inverted commas	2 (4.7%)	31 (49.2%)	1 (2.4%)	2 (20.0%)	37 (59.7%)	–

* Excluded from the count were uses of *rasta roko* within quotations.

Interestingly, according to Table 3.4, *The Hindu* and *The Tribune* are least likely to drop inverted commas around *rasta roko*, a finding which may point to house-internal stylistic conventions working against the full integration of the item in those two publications. But the fact that *rasta roko* is used without any further annotations or glosses in all newspapers also suggests that it is expected to be understood throughout the country by the turn of the millennium. The grammatical properties of *rasta roko* are still in flux, however. While Crowther 1996 recognizes noncount and singular uses of *rasta roko*, corpus evidence suggests that *rasta roko* is beginning to be used in the plural, too, albeit only rarely so far:

- (23) As people in the area received news of the murder, there were angry protests and rasta rokoks for a brief period of time.

[*Indian Express*; September 24, 2000²²]

The analogous hybrid formation *road roko*, which has not been codified anywhere yet, is also attested repeatedly in the online corpora in both singular and plural uses:

- (24) A savage case of arson, incidents of picketing and stoning of buses, besides road roko by AIADMK partymen and volunteers marked the widespread protests in the State ...

[*The Hindu*; February 4, 2000²³]

- (25) Faulting the state government for failing to get tough with the Dalit Panthers for its alleged acts of violence in the district, he said the police should book those resorting to road rokoks there under the National Security Act.

[*The Hindu*; June 29, 2000²⁴]

22. www.indian-express.com/ie/daily/2000/09/24/ina24015.htm, online March 20, 2001.

23. www.hinduonnet.com/thehindu/2000/02/03/stories/04032231.htm, online March 20, 2001.

24. www.hinduonnet.com/thehindu/2000/06/29/stories/04292234.htm, online March 20, 2001.

Table 3.5 *Rasta roko(s), road rocko(s) and road blockade(s).*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE
SINGULAR:						
rasta roko	38	55	41	10	58	1
road roko	1	92	1	11	5	–
road blockade	27	46	18	16	78	19
PLURAL:			–			
rasta rokos	5	8	–	–	4	1
road rokos	2	6	–	1	1	–
road blockades	17	13	8	12	45	9

When we compare the absolute figures for the singular and plural forms of *rasta roko(s)*, *road roko(s)*, and *road blockade(s)* across the six online corpora (March 20, 2001), another interesting distribution emerges. Table 3.5 reveals that the six Indian newspapers today are stylistically quite varied in their selection among the singular forms. *Rasta roko* turns out to be the most common lexical choice in the *Indian Express* and the *Deccan Herald*, *road blockade* in the *Hindustan Times*, the *Tribune* and the *Telegraph*, and *road roko* exclusively in the *Hindu*. In other words, unlike in the case of *lathicharge* or *dharna*, there is no shared preference among India’s quality broadsheets for any of the three lexical alternatives in the singular. In the plural, the situation is different. All six online corpora clearly prefer the all-English compound *road blockades*, whereas *rasta rokos* and *road rokos* so far function only as minority variants. On the basis of those search results, it can be concluded that *rasta roko* and *road roko* are yet on their way toward full integration in the IndE lexicon. To employ Platt et al.’s (1984: 88–9) terminology, *rasta roko* and *road roko* were “stabilizing” rather than “stabilized” lexical Indianisms around the year 2000. It will be interesting to watch the future development in IndE to see whether the plurals *rasta rokos* and *road rokos* will become more frequent in India and whether this will be accompanied by the full textual integration of the items in publications such as *The Hindu* or *The Tribune*. As of today, those two publications appear particularly reluctant to employ *rasta roko* and *road roko* in an unobtrusive fashion.

To conclude, corpus evidence indicates that regional differences do exist in the use of loanwords within and across South Asian varieties of English – but no easy generalization suggests itself. Among the six loanwords looked at within the register of journalistic prose, *rasta roko* comes closest to functioning as an ‘absolute’ lexical variety differentiator for IndE today. *Rasta roko* is a fairly recent IndE innovation that appears not to have spread across the Indian borders, being yet on its way toward becoming fully established in IndE itself. The other five items looked at are attested in the Pakistani, Bangladeshi, and Lankan newspaper corpora as well and must be considered shared lexical features of the South Asian Englishes studied here. However,

since the newspaper corpora show the items to be much more frequent in IndE than elsewhere, they are variety differentiators for IndE in quantitative terms – at least as regards the text category of newspaper journalism.

3.1.5 Case study: The high numerals *lakh* and *crore*

The discussion of loanwords in IndE concludes with a look at another two loanwords suggesting themselves for further analysis as they are the two most frequently recurring single lexical loans in the Primary Corpus: the cardinal numbers *lakh* ‘one hundred thousand’ and *crore* ‘ten million’ (cf. also Leitner 1991b: 222; McArthur 2002: 322; Hankin 2003: 101, 283–4; Melchers/Shaw 2003: 142; Gargesh 2006: 103). Next to describing the functions of the two words in context, this chapter is meant to illustrate the fruitfulness of utilizing structured corpora and online newspaper corpora side-by-side to produce a rounder lexicographical description of the two loanwords in IndE than previous studies have achieved.

When it comes to the formal features of *lakh* and *crore*, a look at the Primary Corpus suffices to illustrate the high degree to which both loanwords are integrated in IndE today. For example, both words can be used as numerals in determiner position (cf. examples (26), (27)) and in noun-phrase head position (cf. examples (28), (29)).

- (26) <\$B><#>currently there is more than a hundred lakh tonnes of excess foodgrain in government godowns all over the country

[IND2000 Broad News 14]

- (27) <\$A><#>an estimated seventy-five to eighty per cent of the one point two crore voters<,> cast their ballots in the eleventh parliamentary polls

[IND2000 Broad News 03]

- (28) <\$C><#>mostly one lakh and ninety thousand have gone back<,> <#>now almost sixty to seventy thousand refugees are here

[IND2000 Broad Disc 06]

- (29) <\$A><#>so far the garment industry enjoyed only small-scale industry status<,> which meant<,> that there was a limit of one crore on plant and machinery investment

[IND2000 Broad News 14]

Lakh and *crore* are also used as quantity nouns, in which case they take the s-plural (examples (30) through (33)), can be premodified (example (30)), and be followed by *of*-phrases (examples (32) and (33)):

- (30) A generous “resident,” the Associated Cement Companies (ACC), provided the kind of space that would have cost several lakhs at the end of a few months.

[IND2000 Press Rep 05]

- (31) <\$C><#>apart from creating new designs<,> to attract customers<,> card and gift companies spend crores on advertising [IND2000 Broad News 20]
- (32) <\$C><#>they were in the jungles suffering<,> lakhs and lakhs of Tamils<,> <#>they are looking forward to go to their own homeland to their homes ... [IND2000 Broad Disc 06]
- (33) <\$A><#>the fire at Bharatpur depot which supplies arms to the seventh command<,> has cost the nation crores of rupees and a few human lives ... [IND2000 Broad News 07]

As is possible for *hundred*, *thousand*, and *million*, *lakh* and *crore*, too, can be preceded by the indefinite article *a*, which functions as a slightly less emphatic variant of *one* (cf. also examples (28) and (29) above).

- (34) About a lakh tonne of rice was also available in both the states, the release added. [Indian Express; April 25, 2000²⁵]
- (35) The morning after the whole country watched Harshavardhan Nawathe win a crore on television, the 27-year-old 'IT boy' arrived at photographer Atul Kasbekar's Worli studio to have some pictures taken. [Indian Express; October 21, 2000²⁶]

As a look at the online corpora reveals, *lakh* and *crore* may also be used in combination, one *lakh crore* equaling 10^{12} :

- (36) This kind of ideological and politically-motivated squandering of the assets of the nation, running to over one lakh crore rupees, must be stopped. [Indian Express; January 11, 2000²⁷]

On the analogy of the adjective *multimillion*, IndE also has *multicrore*, which is attested once in the Primary Corpus:

- (37) <\$A><#>Mr Yadav's announced yesterday that the case against him resulted from a wrong evaluation of his property<,> and the case was not related to the multicrore fodder scam [IND2000 Broad News 13]

When considering the evidence presented so far, it can be established that *lakh* and *crore* participate in the grammatical behavior of English high numerals in the ways described by Quirk et al. (1985:393–8). But the corpora hold more information in store. The Primary Corpus, for example, reveals that *lakh* and *crore* are found most

25. www.indian-express.com/ie/daily/2000/04/25/ifr25016.htm, online March 16, 2001.

26. www.indian-express.com/ie/daily/2000/10/21/ina21044.htm, online March 16, 2001.

27. www.indian-express.com/ie/daily/2000/01/11/ied11028.htm, online March 16, 2001.

frequently in the context of Indian currency statements.²⁸ Here, the name of India's currency, the *rupee*, is usually written in its abbreviated (plural) form *Rs* before the numeral. In speech, it is spoken after the numeral, cf.

- (38) After the transfer was stayed, Mr Sharma is alleged to have fraudulently and dishonestly withdrawn Rs 39,92,09,095 (about Rs 40 crore) between July 1993 and March 1994 and Rs 37,70,39,743 (about Rs 38 crore) between April 1994 and March 1995. [IND2000 Press Rep 12]
- (39) <\$C><#>and despite reservations even the finance ministry may give its nod to the proposal<,> as even maintaining buffer stocks in government godowns<,> costs over four thousand crore rupees per year [IND2000 Broad News 14]

Note that currency names are occasionally said before numerals in IndE speech, too, as in the following two examples:

- (40) <\$A><#>in Jhansi six masked men attacked a convent of the Daughters of the Sacred Heart on Tuesday night<,> <#>according to reports they assaulted the nuns there<,> and stole rupees twelve thousand in cash ... [IND2000 Broad News 16]
- (41) <\$E><#> ... that was a very big tournament for me <#>that was a dollar ten thousand [IND2000 Broad Int 01]

This usage has met with prescriptive skepticism (Smith-Pearse 1968 [1934]:7) and might, for that reason, be exceptionally rare in the Primary Corpus. What makes this observation relevant for a description of the formal properties of *lakh* and *crore* is the fact that an interesting case of morphosyntactic variation is observed in the corpus data which can be related to the interpretation of the grammatical role of currency names in contexts such as the following:

- (42) Robbers looted Rs 26 lakh from the Hill Cart Road branch of UCO Bank this morning. [IND2000 Press Rep 12]
- (43) The board may have spent over Rs 20 lakhs on the Calcutta meeting but the end result was not what the senior board members have been professing for some time. [IND2000 Press Rep 17]
- (44) Far from making money, the airline has projected a loss of Rs 40 crore on a fleet of 12 (now 11) planes. [IND2000 Press Edit 03]
- (45) Mr. Paswan maintained that the cost of this package for the employees would be Rs. 68 crores a year only ... [IND2000 Press Rep 19]

28. Note the IndE-typical way of using the comma in writing for bundling digits in numbers greater than or equal to one lakh.

In examples (42) and (44), the numerals *lakh* and *crore* function as determiners ‘in disguise’, i.e. in currency statements of the abbreviated written form *Rs. X lakh/crore*, which are conventionalized representations of noun phrases of the form *X lakh/crore rupees*, in which *rupees* functions as the noun-phrase head. When we look at examples (43) and (45), it becomes clear that IndE also uses abbreviated currency statements of the form *Rs. X lakhs/crores*, in which the pluralized numerals function as noun-phrase heads and the abbreviated currency names as premodifiers in much the same way as *rupees* and *dollar* do in examples (40) and (41). Note that the vacillation between zero and s-plural is not restricted to Indian currency statements alone but occasionally affects international currency statements featuring the high numerals *million(s)* and *billion(s)*. See the following example from the Primary Corpus:

- (46) The new American interest in natural gas has given a big surge to U.S. investments in Bangladesh – from about \$20 millions of cumulative investment until 1996 to about \$700 millions now. The U.S. estimates that this could easily rise to about \$ three billions in the next few years, if Bangladesh adopts appropriate strategies. [IND2000 Press Rep 01]

The variation described above has not gone unobserved among style critics in India. Yadurajan (2001: 99–101) looks at two quality broadsheets, the *Economic Times* and *The Hindu*, and mentions the possibility that the s-versions could be a peculiarity of *The Hindu*. The results of a Google-based search conducted in the six IndE online corpora (March 16, 2001) lend support to Yadurajan’s hypothesis; see Table 3.6.

With the exception of *The Hindu*, all Indian quality newspapers consulted clearly prefer the s-less variants in Indian currency statements today, which suggests house-style conventions as the main factor determining the variation at stake. But with the help of the Primary Corpus, the Kolhapur Corpus and ICE-India, it is further possible to show that there is an interesting diachronic dimension to the variation observed. Consider Table 3.7.

The figures from the Kolhapur Corpus (which comprises specimens of journalistic prose plus fourteen other written text genres) reveal that the s-variant used to be widely preferred in written usage around 1978 (column 1). But when zooming in on the 1978 press material sampled from Kolhapur (column 2), it turns out that the s-variant had already begun to lose ground within the Kolhapur press section by that year. Twenty-two years later, in the year 2000, the s-variant was only a minority variant in IndE journalism (column 3) – a development which could be interpreted as a move toward the morphosyntactically more regular s-less version, i.e. that version which is in line with the international English convention treating currency names as noun-phrase heads and numerals as determiners. In the press material from ICE-India, too, the preference is clearly with the zero plural (column 4). Outside the domain of journalistic writing, on the other hand, the s-variant appears to be going stronger (column 5). However, 13 of the 14 attestations of *Rs. X crores* in the written component of ICE-India (excluding the press texts) are found in a single student essay, which narrows down the

Table 3.6 *Lakh* and *crore* in currency statements [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE
Rs 5 lakh	489	33	150	740	1540	93
Rs 5 lakhs	28	170	5	51	17	4
Rs 25 lakh	281	27	83	355	652	60
Rs 25 lakhs	24	127	8	36	10	–
Rs 75 lakh	46	7	20	72	145	7
Rs 75 lakhs	4	25	2	2	3	–
Rs 5 crore	358	50	92	210	733	74
Rs 5 crores	6	174	4	9	5	2
Rs 25 crore	459	42	64	133	290	48
Rs 25 crores	8	104	5	4	3	–
Rs 75 crore	109	20	32	84	132	23
Rs 75 crores	2	50	–	4	1	–

Table 3.7 *Lakh* and *crore* in currency statements [KOL, Primary Corpus, ICE-IND].

	IndE KOL 1978	IndE IND1978 Press	IndE IND2000 Press	IndE ICE-IND Press	IndE ICE-IND other writing
Rs. X crore	20	8	28	78	–
Rs. X crores	97	30	3	32	14
Rs. X lakh	10	7	10	11	–
Rs. X lakhs	31	4	4	7	–

difference between journalistic and non-journalistic writing in this particular instance considerably. When we tie in the findings from the online corpora with those from the structured corpora, we arrive at the following conclusion. What looks like a stylistic peculiarity of *The Hindu* today is not a house-style-induced independent innovation, but encases a more conservative developmental stage within IndE (journalistic) writing. It is possible that *The Hindu* is nowadays still closer to general IndE written usage outside the domain of journalism than the other publications are. Within the domain of journalistic writing, *The Hindu* is less advanced than the other online publications. Only time can tell whether (and if so, at what pace) *The Hindu* will eventually follow the observed trend toward morphosyntactic regularization.

Interestingly, in the one instance where variety-specific usage patterns can be discerned at the morphosyntactic plane, namely in the case of pluralized high numerals within the context of currency statements, IndE seems to be in the process of converging toward, rather than diverging from, the international English norm, at least as far as the register of newspaper prose is concerned. What may have started this

Table 3.8 Collocations of *lakhs of* and *crores of*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE
thousands of rupees	35	24	10	2	357	1
lakhs of rupees	114	124	56	10	1210	4
millions of rupees	20	10	6	4	109	3
crores of rupees	345	355	110	13	1740	17
thousands of dollars	28	11	2	3	45	1
lakhs of dollars	1	–	–	–	1	–
millions of dollars	122	101	35	18	227	5
crores of dollars	–	–	–	–	3	–

trend could simply be the willingness of Indian newspaper editors to conform to internationally established grammatical standards of English as closely as possible. But one question remains: what are the factors motivating the use of *lakh* and *crore* in the first place? A final look at the corpus material provides the following answer: next to the subtle semantic differences of the English and the borrowed numerals (*thousand* < *lakh* < *million* < *crore*), *lakh* and *crore* differ from their English near-synonymous counterparts in that their use is restricted to what may be termed ‘Indian contexts’, i.e. contexts that can be situated in the geographical, social, and cultural realities of India. In explicitly non-Indian contexts, for example in news reports covering foreign affairs, *lakh* and *crore* are rarely ever found, and *thousand* and *million* function as more natural choices. This difference can, for example, be observed in the ways IndE exploits the quantifying phrases *thousands of*, *lakhs of*, *millions of*, and *crores of* to denote large sums of Indian and foreign money respectively (Google search dated March 17, 2001). As is shown in Table 3.8, the Hindi numerals are hardly ever used in reference to sums of money in dollars, which makes collocations such as *lakhs of dollars* and *crores of dollars* exceptionally rare (but not impossible). When it comes to sums of money in rupees, on the other hand, *lakhs of* and *crores of* are clearly the preferred quantifiers. It thus appears that users of IndE are well aware of the etymological origins of *lakh* and *crore* and, accordingly, opt for the use of those numerals only in contexts that make strong connotations to the Indian context seem desirable and natural.

3.2 More different words and meanings? In search of lexical Indianisms

Next to borrowing, the coining of new words and semantic change have been claimed to contribute to the lexical innovation of varieties of English in general, and IndE is no exception (cf. Kachru 1983; Dubey 1989, 1991; Carls 1994, 1999; Görlach 1995). But even though glossaries and usage handbooks have been published to document the impact of those processes on IndE (cf. Nihalani et al. 1979, 2004; Hawkins 1984;

Muthiah 1991; Crowther 1996; Yadurajan 2001; Hankin 2003), most of the items that have been said to be lexical Indianisms in those works to date remain, in the words of Görlach 1995, “ambiguous as to status: many conspicuous formations are likely to represent individual adhoc solutions (or adhocisms ...), effects of first-language interference, or literary coinages made for specific stylistic effects ...” (79). In view of this gap in the available lexicographical research, it was a natural decision to trace those claimed lexical Indianisms in the available corpus material and describe their status more context-sensitively than had been accomplished before, using the structured corpora and the online corpora side by side. In the following, individual lexical coinages (3.2.1) and selected word-formation patterns that have been claimed to be particularly productive in IndE (3.2.2) are dealt with. Words with IndE-specific meanings are also discussed (3.2.3).

3.2.1 In search of neologisms

According to Carls 1999, word formation is “by far the most frequently used method of creating new lexical items” (141) in IndE. Since there is a constant need for users of English in India to adapt their vocabularies to their very own socio-cultural, physical and psychological realities, it is to be expected that they should continually draw on the possibilities of English word formation to expand their communicative range and create new words. But when scanning the Primary Corpus, the Kolhapur Corpus and ICE-India for examples of the numerous claimed lexical Indianisms in context, only a small number of them was found in the data. Partly attributable to the comparatively small size of the micro-corpus, this observation also meant that the proportion of lexical Indianisms was not to be overestimated for the usage range under investigation. A closer inspection of the few attested items, on the other hand, revealed interesting new insights into their status in contemporary IndE and into the general dynamics of lexical innovation and variation in that variety.

‘*Speed money*’. In light of the available data, the nominal compound *speed money* ‘a bribe given to sb. to make sure that one gets quick service’, which has been codified in Crowther (1996: 1470) and Hankin (2003: 468), qualifies as an independent IndE lexical innovation proper. Consider its only attestation in the Primary Corpus, which was found in a student essay on the topic of corruption:²⁹

- (47) Once political corruption in India goes or is kept under check, the present shameless exercise in “speedmoney”, “bakshish”, and bribe will disappear by and by. [IND2000 Press Ess 03]

The attestation above is informative in various respects. There are several hints in the passage indicating that *speed money*, although codified, is still in the process of establishing itself in the IndE lexicon. Note that the student did not rely on the compound

29. The compound *speed money* is found neither in the Kolhapur Corpus nor in ICE-India.

alone to communicate his ideas; he also employed a near-synonym, the loan item *bakshish* ‘tip’ (cf. Hankin 2003:31) to make sure that the meaning of his sentence would become clear to the reader. Also, the careful use of inverted commas around the compound seems to reflect some reluctance in employing the coinage, a textual treatment typical of items that are not yet fully integrated (cf. Chapter 3.1). A look at the online newspaper corpora confirms that the example from the student essay is representative for written IndE at large around the year 2000. Inverted commas are found repeatedly around *speed money* in the online data, too, where the item appears in all of the six Indian online publications tapped (cf. Appendix 3.8). In contrast to the Primary Corpus, the online newspapers usually spell the compound as two words, which renders the formation an establishing rather than established compound.³⁰ Consider three typical examples from the online editions of *The Hindu*, *The Tribune* and *The Indian Express*:

- (48) It took some ‘*speed money*’ to extricate themselves from a ticklish situation.
[*The Hindu*; June 21, 2001³¹]
- (49) Getting a job done with the help of “*speed money*” has become cheaper than taking the cumbersome routine course. [*The Tribune*; April 11, 2001³²]
- (50) A sample of 312 weak borrowers in Tamil Nadu by the Bank showed that “incidental expenses” and “*speed money*” took up around Rs 21 for every Rs 100 of subsidy. [*Indian Express*; February 9, 1999³³]

There were no hints to be found in the data regarding the origin of this compound. It is likely that *speed money* was formed simply for reasons of creativity and expressiveness and in due course has successfully caught on without losing its informal ring. Note that in contrast to a loanword such as *bakshish*, an all-English compound such as *speed money* has the obvious advantage of being understood straightforwardly across linguistic borders, but only time can tell whether or not the compound will spread further into other registers and text categories. As of today, it is safe to say that *speed money* is recurrent in IndE journalistic prose. The fact that the compound has not yet spread noticeably across the Indian borders into other South Asian varieties of English qualifies the item as a lexical Indianism proper (cf. Appendix 3.9).

30. As is the case for several of the loanwords discussed in Chapter 3.1, it is again the Chandigarh-based *The Tribune* that is most reluctant in using a new word without inverted commas, which supports the view that a publication-specific house-style is followed in the editorial offices of that particular newspaper (cf. Appendix 3.8).

31. www.hinduonnet.com/thehindu/2001/06/21/stories/13211065.htm, online August 4, 2001.

32. www.tribuneindia.com/2001/2001/04/11/mailbag.htm, online August 4, 2001.

33. www.indianexpress.com/ie/daily/1999/02/09/ige09064.htm, online August 4, 2001.

'*Eve-teasing*'. Like *speed money*, the item *eve-teasing*, too, qualifies as a creative coinage that appears to have originated in the Indian subcontinent. It has long been recognized as a South Asian regionalism and subsequently been codified. Today, on-line newspapers across South Asia are using it recurrently.³⁴ The word has been described by Nihalani et al. (1979:75; 2004:75) as referring to "[t]he almost universal pastime of 'teasing the girls'" (also cf. Carls 1994:212; McArthur 2002:323; Trudgill/Hannah 2002:135; Hankin 2003:143). While *eve-teasing* may indeed have started out as a jocular term, South Asian journalism nowadays employs the item in a more neutral fashion in reference to all acts of publicly annoying women "by offensive language and behaviour" (Crowther 1996:1447). Consider its only attestation in the Primary Corpus:

- (51) Sir, – It required a Tuticorin theatre fire to replace thatched roof with fire-proof asbestos roof, a Sarika Shaw's death to enact anti-*eve teasing* laws and a Jyotsna's death to instal emergency switches in escalators.

[IND2000 Press Edit 07]

In example (51), a reader is complaining in a letter to the editors of one of India's national English-language newspapers about missing emergency exits in the public buses of Pondicherry, claiming that city officials are not doing enough to improve the disastrous situation. Angrily he is referring to several gruesome incidents which could have been prevented if only public officials had acted sooner. In this context, the reference is to *anti-eve teasing* laws which were put into action in Pondicherry only after a woman had been murdered. The use of the item *eve-teasing* in this example is neutral in style, and no light-hearted tone seems intended. This is typical of the use to which the item is put in South Asian journalistic prose today. Consider the distribution of the coinage vis-à-vis its common-core equivalent *harassing* across the South Asian online corpora (November 27, 2003); cf. Table 3.9.

The fact that the word, which is alternatively spelled as two or with a hyphen, is found in all nine online corpora suggests that it has spread quite evenly across the subcontinent, where it is used alongside near-synonymous forms, for example *harassing*.³⁵ This finding puts Nihalani et al.'s (1979:75, 2004:75) claim that "the word is

34. Cf. S. Upendran on the origin of *eve-teasing* in *The Hindu* (December 8, 2003): "What do you normally find outside a girl's college? Boys, of course! A whole bunch of boys making fun of the girls and trying to get their attention. That's what eve teasing is; members of the male species making fun of the opposite sex. The fun that these roadside Romeos (or Adams?) indulge in isn't always innocent either. The expression "eve teasing" is Indian in origin; you won't find it in most dictionaries. Native speakers of English don't use it. Since it is Indian, should we change the name "Eve" to something else?" (www.hinduonnet.com/thehindu/thscrip/print.pl?file=2003120800161504.htm, online February 2, 2004)

35. Two attestations of *eve teasing* are found in ICE-India in a private conversation (text S1A-27).

Table 3.9 *Eve-teasing and harrasing.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
eve(-)teasing	20	351	15	22	594	47	2	4	2
harrasing	17	634	124	77	7330	190	711	4	18

largely unknown to speakers of other varieties of English” into perspective: *eve-teasing* should be labeled a South Asianism rather than Indianism.³⁶ What is equally interesting is that the Internet newspaper corpora also returned several metalinguistic statements questioning the appropriateness of using this item. In an editorial on the topic of sexual harassment at the workplace, for example, a journalist complains indirectly about the frequent euphemistic use of the word in *The Hindu* by noting that “form[s] of sexual harassment (that the readers of this paper may be more used to seeing described as ‘eve-teasing’), defined by the archaisms ‘insulting the modesty of a woman’ or ‘outraging the modesty of a woman’, are offences under the Indian Penal Code” (*The Hindu*; August 18, 2002).³⁷ In another essay published in *The Hindu* on the problem of sexual harassment at Indian colleges and universities, its author remarks that “[o]ften these incidents are dismissed under that inappropriate term “eve-teasing”” (*The Hindu*; June 16, 2003).³⁸ More recently, Shashi Tharoor has, likewise, warned against the use of *eve-teasing*:

Anyone who has seen eve-teasing in operation in Delhi knows that the term masks sordid and often vicious behavior by depraved youths against victims often in no condition to resist. Calling it “assault” or “molestation” would be more honest and might do more to raise public consciousness against it.

(Tharoor 2007: 455)

In light of this criticism, it will be informative to follow the future developments in South Asia to see whether *eve-teasing* will keep recurring in high numbers in IndE journalistic prose or whether its use will eventually be sanctioned on account of the pressure exerted by language critics these days.

‘Timepass’. In contrast to *speed money* and *eve-tasing*, the compound *timepass* ‘pastime’ had not been codified in any of the available dictionaries or handbooks by the turn of the millennium although it was recurrent in contemporary IndE. The only

36. See the following illustrative example from Pakistan’s *The Dawn* (November 19, 2000): “A rescue-15 ASI nabbed a youth for eve-teasing from outside the girls college on Saturday.” (www.dawn.com/2000/11/19/local11.html, online November 27, 2003)

37. www.hinduonnet.com/thehindu/2002/08/18/stories/2002081800281600.htm, online November 27, 2003.

38. www.thehindu.com/thehindu/mag/2003/06/15/stories/2003061500240300.htm, online November 27, 2003.

attestation for the item in the Primary Corpus was found in the reviews section, a text category which has been shown to be particularly open to lexical borrowing as well (cf. Chapter 3.1).³⁹ Consider:

- (52) All in all, catch this one only if you are desperately in need of timepass.
[IND2000 Press Rev 09]

Timepass is used in example (52) in a context where standard English would arguably have *pastime* or *entertainment*. Online research shows that this is not an exceptional but a common word in IndE journalistic writing today, where it is usually spelled solidly without inverted commas around it – two small findings that may suggest a slightly higher degree of integration of *timepass* than that of *speed money* despite the fact that the word has not been codified (cf. Appendices 3.10 and 3.11). This reading of the data is also supported by the more variegated functional potential of *timepass* that is at display when studying the item in context. While it is used most often as a noncount noun (as in example (52)), *timepass* is also beginning to be used as a count noun, as in the excerpt below which is taken from an article published in the electronic *Hindustan Times*:

- (53) Unable to make the moments in the relationship a part of her growing up experience, it all became a game ... a timepass.
[*Hindustan Times*; December 18, 2002⁴⁰]

Timepass may also occur flexibly as determinant in other compound nouns:

- (54) No sir. Guys are just guys. Uncomplicated timepass companions. People you can watch endless amounts of TV with, play pool and drink vast quantities of beer.
[*Indian Express*; January 7, 2000⁴¹]

And the item is available in colloquial Hinglish:

- (55) As I walked out of a shop with a bruised ego, I overheard a familiar voice commenting: “Time pass karne ke liye aaya hai, sala.”⁴²
[*Deccan Herald*; November 9, 2001⁴³]

39. ICE-India includes one token of *timepass* found in a private conversation: “<\$A><#>Of course evenings are meant <,> evenings are meant for that na <#> just for timepass” [ICE-India S1A-38]. No example is found in the Kolhapur Corpus.

40. www.hindustantimes.com/news/specials/proj_tabloid/ladylovers.htm, online March 19, 2003.

41. www.indian-express.com/ie/daily/2000/01/07/ien07072.htm, online March 19, 2003.

42. Translation: “This stupid guy has come for timepass.”

43. www.deccanherald.com/deccanherald/nov09/middle.htm, online March 19, 2003.

Table 3.10 *Time(-)pass* and *pastime*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
time(-)pass	10	20	5	23	15	1	1	–	–
pastime	6	278	39	–	10	97	–	–	3

But given that standard English has *pastime* and *entertainment*, the question remains why there should be a communicative need for *timepass* to begin with. While linguistic playfulness seems to have triggered the birth of *speed money* and *eve-teasing*, interference from Hindi might be the main driving force behind the spread of *timepass*. Although *timepass* is attested for all six IndE online newspapers and once for Pakistan's *The Dawn* as well (which, strictly speaking, renders the item a South Asianism rather than Indianism), the online corpora (March 19, 2003) reveal that *timepass* is going somewhat stronger in the Hindi-speaking areas of India's north and north-west (represented by the *Tribune* (TR) and the *Hindustan Times* (HT)) than in the south (*Deccan Herald*; DH) and east (*Telegraph*; TE); cf. Table 3.10. The reason for the distribution could be that *timepass* is a loan translation from Hindi which probably found its way into the IndE lexicon by passing through the code-mixed amalgam of Hinglish (cf. example (55)). It could be its very affinity to Hindi that may explain why *timepass* has come to be a preferred choice alongside *pastime* only in India's north and north-west whereas it appears to be rarer elsewhere in the subcontinent.⁴⁴ Interestingly, *timepass* has also already begun to appear in small numbers abroad under the 'uk', 'us' and 'au' top-level country domains, where it is found exclusively in contexts that can be related to the Indian subcontinent in content (examples (56) and (57)) and/or authorship (examples (57) and (58)):

(56) Music wise the film has a couple of *timepass* numbers.

[Origin: Britain; review of a Bollywood movie⁴⁵]

44. That *timepass* had not yet become fully integrated in the South by the turn of the millennium is also suggested by a reference to the word in S. Upendra's language column in *The Hindu* (Chennai), dated September 4, 2001, to which a reader from Tamil Nadu (M. Aruna, Thiruvananthapuram) sent in the question "Does the word *timepass* exist?" Upendra's answer was: "It certainly does; at least it is alive and kicking in Indian English. We often use this word in our country to mean a good way of spending one's time. For example, we hear people say *The movie was a good timepass. It was just timepass, I always take a book with me when I travel. It's good timepass.* The word does not exist in the native varieties of English. The native speaker would normally say "pass time". *The movie was an enjoyable way of passing one's time. I always take a book with me when I travel. It's a nice way of passing one's time. I passed time watching the kids at play.*" (www.hinduonnet.com/thehindu/2001/09/04/stories/13040376.htm, online March 19, 2003)

45. www.asiafreenet.co.uk/AspE-Zone/FilmEntertainment.asp?ArticleID=262&Status=LIA, online March 19, 2003.

- (57) HEIGHT OF TIMEPASS: A person sending email to himself; HEIGHT OF EXPECTATION: Sending Indian cricket team an e-mail, wishing them to win a match ... [Origin: United States; private website⁴⁶]
- (58) Board games I can play anytime anywhere. Sports. The ultimate timepass is sports. [Origin: Australia; private website⁴⁷]

It will therefore be informative to keep track of this item in the future, as this is likely to reveal more insights into the openness of English to lexical innovations originating in India's north and north-west. As of today, *timepass* appears to be common especially among Hindi-using Indians.

'*Petrol bunk*'. If *timepass* exemplifies a local coinage that is particularly popular in India's north and north-west, the compound noun *petrol bunk* 'petrol station' represents the case of a formation apparently limited in use largely to India's south today. Not attested in any of the closed corpora, the compound aroused my interest rather accidentally while searching the Internet for collocations of the verb *to bunk (off)* (cf. Chapter 4.1). See the item in context:

- (59) He said Reliance is interested in setting up over 10,000 petrol bunk outlets in Maharashtra and Gujarat for dispensing gas/petrol. [Indian Express; November 8, 2000⁴⁸]
- (60) A five-member dacoit gang attempted to strike at a petrol bunk owner's house in Sanjaynagar police station limits in the wee hours of today. [Deccan Herald; December 21, 2001⁴⁹]

Online research (January 6, 2002) shows that *petrol bunk* is nowadays used alongside the internationally more commonly available compounds *gas station*, *petrol station* and *petrol pump* only in *The Hindu* and the *Deccan Herald*, where it is the preferred variant. It turned out to be exceptionally rare in the other four national quality broad-sheets, cf. Table 3.11.

A closer reading of the 93 attestations of *petrol bunk* in the *Indian Express*, *The Hindu*, and the *Deccan Herald* showed further that around the year 2000 the item was used in IndE press texts only in reference to filling stations in Tamil Nadu,⁵⁰

46. www.hills.ccsf.cc.ca.us/~gsaxen01/frame2.htm, online March 19, 2003.

47. web.dis.unimelb.edu.au/undergrad/2001/yogesh/homepage/Interests.htm, online March 19, 2003.

48. www.expressindia.com/fe/daily/2000/12/08/fco08037.htm, online January 6, 2002.

49. www.deccanherald.com/deccanherald/dec21/ct5.htm, online January 6, 2002.

50. The references in the data downloaded from the Internet are to the cities of Nagapattinam, Madurai, and Coimbatore.

Table 3.11 *Petrol bunk, petrol pump, petrol station and gas station.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
petrol bunk	2	46	42	–	3	–	–	–	–
petrol pump	147	39	15	506	400	21	237	1	12
petrol station	16	6	3	39	24	2	28	5	1
gas station	16	7	6	46	29	–	21	2	4

Karnataka⁵¹, Andhra Pradesh⁵² and Pondicherry. But what exact motivation lay behind the coinage of the item or whether *petrol bunk* is on its way into or out of general IndE usage cannot be clarified on the basis of the evidence produced above, and more diachronically sensitive research is called for to solve this linguistic puzzle. Note in this context that not only IndE but new Englishes in general are quite variable in this semantic field. As Schneider 2007 points out, “a British *petrol station* is a *petrol kiosk* in Singapore, a *gas(oline) station* in the Philippines, and a *petrol pump* in India” (81). Hankin 2003 speculates that the IndE use of *bunk* in reference to filling stations might have been derived “from *bunker*, meaning a fuel store, e.g. a ship’s bunker” (64). In general, Indian users of English in the South do not seem to be aware of the compound being a South Indian regionalism, which is suggested by the fact that not a single metalinguistic statement on the use of the item was found anywhere in the online databases.

‘Upgradation’ and ‘delink’. While linguistic creativity and playfulness likely triggered the creation of *speed money* and *eve-teasing* and India’s multi-language background possibly that of *timepass*, it is more difficult to arrive at a straightforward explanation for the frequent occurrence of the next two formations to be discussed, *upgradation* and *delink*, which have been listed in Nihalani et al. (1979: 64, 186; 2004: 63, 188) and codified in Crowther 1996. Not attested in the *Oxford English Dictionary* quotations base nor codified in any of the international dictionaries of standard English today, the noun *upgradation* and the verb *delink* are common in contemporary IndE, where they are employed unselfconsciously alongside their respective standard English near-equivalents *upgrading/upgrade* and *disconnect*. It is, in fact, not unlikely that both items may have evolved independently in the Indian subcontinent. See the following two attestations first:

51. The references in the data downloaded from the Internet are to the cities of Bangalore, Pandavapura, Arsikere, Hungund, Hosur, Hoskote, Haveri, Bijapur, and Belgaum.

52. The references in the data downloaded from the Internet are to the cities of Hyderabad, Warangal, Visakhapatnam, and Kurnool.

Table 3.12 *Upgradation, upgrading and upgrade.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
upgradation	35	992	236	103	7100	178	773	–	12
the upgradation of	6	150	19	13	545	25	161	–	–
the upgrading of	–	31	4	2	47	4	28	–	1
the upgrade of	6	2	1	4	8	8	6	–	–

Table 3.13 *Delink vs. disconnect.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
delink	46	65	7	119	110	11	23	–	2
delinked from	35	31	4	25	69	9	18	2	2
disconnected from	1	18	–	7	24	2	19	9	3

- (61) The learning process involves emphasis on literacy, but not that only: it also stresses the importance of functional upgradation and of raising the level of awareness regarding their predicament among the poor and the illiterate.

[IndE KOL A 40]

- (62) In the budget for 1999–2000 the Finance Minister, concerned with the cost of their servicing, delinked these from the budget. [IND2000 Press Edit 01]

As regards the item *upgradation*, the Kolhapur Corpus returned two and ICE-India five tokens of the word in context (four of which were found in spontaneous speech).⁵³ The analysis of the online press corpora revealed further that *upgradation* is not only recurrent in Indian but also in Pakistani and Bangladeshi journalism today, where it is preferred to its synonyms *upgrading* and *upgrade* for example in noun phrases of the form *the N of*. In this particular pattern, the item is conspicuously absent from the Sri Lankan online database (April 5, 2002); cf. Table 3.12.

Similarly, *delink* qualifies as a South Asianism (also cf. Görlach 1985: 10; Carls 1994: 214, 2000: 35).⁵⁴ Recurrent in the press material from India, Pakistan and Bangladesh, the verb is even more common than its common-core companion *disconnect* in environments of the type *V-ed from* (April 5, 2002); cf. Table 3.13.

53. In ICE-India, the five attestations of *upgradation* are found alongside their standard English companion *upgrading*, which is attested three times (once in a press report and twice in public dialogue).

54. There is one example of *delink* in ICE-India, which is found in a text representing public dialogue (text S2B-46). The word is not attested in the Kolhapur Corpus.

Table 3.14 *Delink, upgradation* and variants.

	.in	.sg	.uk	.za	.au	.us
delinked from	133	29	54	66	33	6
disconnected from	117	157	5860	452	4030	2950
the upgradation	345	2	2	–	4	–
the upgrading	90	905	7960	3110	7770	2990
the upgrade	133	392	15500	1690	14800	9090

What may have contributed to the rise of *upgradation* and *delink* in South Asian Englishes and IndE in particular could be the forces of analogy. Both *upgradation* and *delink* fully comply with the rules of standard English word formation, but happen not to be used extensively the English-speaking world over – for no straightforward reason. Note that both items do occasionally appear in other varieties of English as well. But their distribution across top-level country domains on the Internet (April 5, 2002) suggests that it is only in South Asia that they have established themselves firmly as parts of the regional core vocabulary (cf. Table 3.14). On that basis, *upgradation* and *delink* qualify as quantitative lexical South Asianisms.

Upliftment. The formation *upliftment* ‘uplift’ has been singled out as a lexical Indianism before (cf. Nihalani et al. 1979: 187, 2004: 188–9; Carls 1994: 211, 2000: 35; Hickey 2004: 546). Upon closer investigation, however, *upliftment* qualifies as a lexical item shared by IndE with other varieties of English – but not with the major varieties BrE and AmE. Attested first in the *Oxford English Dictionary* quotations base for 1926, *upliftment* has apparently not found its way into the international prestige varieties, which is suggested by its absence from British and American dictionaries today. In spite of that, according to Nihalani et al. (1979: 187; 2004: 188), *upliftment* continues to be “very frequent in IVE [Indian variants of English] newspapers and official writing.” A closer look at ICE-India, for example, returned five tokens of *upliftment* and seven tokens of the internationally established synonym *uplift*. A glance at the online corpora revealed that numerous examples of *upliftment* can be found not only in IndE but all over South Asia, where *upliftment* is regularly employed alongside its companion *uplift* (cf. Table 3.15). While there is no clear-cut preference pattern discernible for IndE, it is noteworthy that *upliftment* is particularly strong in the Sri Lankan database (March 25, 2001).

Upliftment is also available in IndE speech, for example in broadcast news reports. See example (60) from the Primary Corpus:

- (60) <\$A><#>addressing the national convention of the Dalit Students Federation in New Delhi today<,> Mr Paswan said if necessary<,> the sena would be given a new name to make it a vibrant organisation<,> for the upliftment of dalits<,> on the lines of regional political parties

[IND2000 Broad News 01]

Table 3.15 *Uplift* and *upliftment* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
the uplift	20	56	20	31	392	19	223	55	81
the upliftment	37	29	29	25	149	–	55	112	1

Table 3.16 *Uplift* vs. *upliftment* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
the uplift	174	21	1510	47	634	338
the upliftment	562	26	87	1660	61	10

But labeling *upliftment* a South Asianism would still be inaccurate as the word is remarkably common also in parts of South East Asia and in South Africa today. Consider the distribution across top-level country domains (March 25, 2001) in Table 3.16. Hence, it appears that not only South Asian Englishes but also South East Asian and African varieties of English (represented by the data from Singapore and South Africa) have retained a degree of variability that was probably more common across the English-speaking world in the early twentieth century. While BrE and AmE have settled on the noun *uplift* as standard variant, IndE continues to be more variable to date – as are SAfrE and SgE. It will be particularly informative to follow the future developments in Englishes worldwide in this instance to measure the influence exerted by the international prestige norms of BrE and AmE. Will *upliftment* eventually be shed in IndE? Only time can tell.⁵⁵

'*Deptt.*', '*asstt.*', '*demat*', '*senti*', '*enthu*' and '*funda*'. It has repeatedly been claimed that IndE shows a characteristic liking for acronyms, abbreviations and other shortened forms (cf. Spitzbardt 1976: 39; Kachru 1983: 39, 1986: 45; Shekar/Hedge 1996: 61; Görlach 1995: 90). While statements of such general nature should always be read with caution, there can be no doubt that IndE has long been drawing on the multiple possibilities in English to abbreviate or shorten existing forms upon demand. Let me point out some illustrative examples from the corpora. IndE users, for instance, recurrently employ *deptt.* (for *department*) and *asstt.* (for *assistant*) in place of *dept.* and

55. Not attested anywhere in the online newspaper corpora are the formations *unproper*, *debtful* and *dissentment*, which have been discussed by Nihalani et al. (1979, 2004) in reference to the lexical innovativeness of IndE. *Debtful* has also been listed as a lexical Indianism by Mehrotra (1982: 160), *debtful* and *dissentment* by Carls (1994: 207). There is no evidence in the available data that those items should be any more common in IndE than in other varieties of English today. Cf. Appendices 3.12 and 3.13.

asst., which are the more common and codified choices elsewhere in the English-speaking world.⁵⁶ See the following examples from IndE:

- (61) Row over transfer duty security in MCD meeting – Excise Deptt. yet to withdraw controversial order
[headline found in *The Hindu*; September 26, 2002⁵⁷]
- (62) Asstt. Coach: Mahamaya Samaddar (Railways)
[caption found in the *Telegraph*; August 28, 2002⁵⁸]

Since both forms are frequently used also in PakE and BaE journalistic writing today, they qualify as quantitative lexical South Asianisms (cf. Appendices 3.14 and 3.15). While the double <t> in *deptt.* and *asstt.* bears close resemblance to the spellings of the original words, <department> and <assistant>, the double <t> in *govtt.* (for *government*), as illustrated in example (63) below, seems to have evolved on the analogy of *deptt.* and *asstt.* This abbreviation, which was stumbled upon coincidentally when searching for examples of *deptt.* and *asstt.*, turned out to be rather rare in South Asian English writing.⁵⁹

- (63) Govtt. provides a subsidy amount of Rs. 7500/- (Maximum) or 50% of the loan amount whichever is less.
[IndE; example found in a scientific paper⁶⁰]

In contrast to *deptt.* and *asstt.*, which function in writing as short representations of the given nouns but are intended to be read out in full, *demat* (for ‘dematerialization’) is a clipping proper, cf.

- (64) THE DEMATERIALIZATION in the debt market has got a shot in the arm, with two leading bond issuers in the country, ICICI and the Industrial Development Bank of India, deciding to provide the demat option to investors for all future wholesale and retail offerings. [IND2000 Press Rep 15]

New clippings frequently get their lease of life as part of “the in-house or in-group *jargon* of a particular institution or social group” (Peters 1995: 139). The fact that *demat* is particularly frequent in India’s quality broadsheets in financial contexts discussing the use of electronic (rather than ‘material’) shares (while being absent from the

56. Kolhapur Corpus: *dept./deptt.*: 0/1, *asst./asstt.*: 0/1; ICE-India: *dept. /deptt.*: 45/28, *asst./asstt.*: 2/4.

57. www.hinduongnet.com/2002/09/26/14hdline.htm, online November 25, 2002.

58. www.telegraphindia.com/1020828/asp/sports/story_1144297.asp, online November 25, 2002.

59. Kolhapur Corpus: *govt./govtt.*: 11/0; ICE-India: *govt./govtt.*: 14/0.

60. www.dantewada.gov.in/dev.htm, online November 25, 2002.

Table 3.17 *Dematerialisation/dematerialization* vs. *demat* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
dematerialisation	79	10	–	4	20	7	–	–	–
dematerialization	–	–	–	–	–	–	7	–	–
demat	227	52	1	57	81	11	–	–	–

Table 3.18 *Dematerialisation/dematerialization* vs. *demat* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
dematerialisation	449	8	663	704	305	1
dematerialization	44	6	261	40	64	15
demat	797	1	52	16	12	7

Kolhapur Corpus and ICE-India) suggests the jargon of stock brokers and investment bankers as the origin of this fairly young coinage. Used heavily by Indian journalists around the year 2000, *demat* was absent from PakE, BaE and SrLE newspapers (November 26, 2002), cf. Table 3.17. On a worldwide scale, *demat* was most common in the Indian subcontinent, where it was on its way toward replacing the original longer word. Consider the snapshot from the Internet across top-level country domains (November 26, 2002) in Table 3.18. Elsewhere in the English-speaking world, *demat* was also in use, but apparently only as a minor variant, which suggests that IndE was most advanced in terms of integrating the clipping into its vocabulary around the turn of the millennium. Whether or not the coinage originated in the Indian subcontinent can, of course, not be reconstructed on the basis of the available evidence, although the frequency distributions in Table 3.18 point in this direction. It remains to be seen whether *demat* will stand the test of time.

On scanning India's quality broadsheets on the Internet in search of more IndE-specific lexical items, several clippings turned up that are likely Indian in origin. Not yet available outside South Asia, *sentī*, *enthu* and *funda* can be heard frequently in more private communication situations among younger Indians today, but they are rare in more public and formal forms of speech and writing, which is indicated by their absence from the Primary Corpus and ICE-India (also cf. D'souza 2001a: 150–2; Bhaya Nair 2008). In India's online newspapers, *sentī*, *enthu* and *funda* are found only in more informal textual environments such as film reviews (cf. Chapter 3.1 for the use of loan vocabulary in that particular text category). *Sentī*, short for the adjective *sentimental*, is occasionally accompanied by glosses or put in inverted commas to mark the item explicitly as a non-established variant. Consider examples (65), (66) and (67) from the reviews of *The Tribune* and *The Hindu*:

- (65) Some of the officers got so senti (emotional) that they clutched on to their wives and started howling in the hall itself,” a bemused Cdt Harsh says.
[*The Tribune*; July 7, 1999⁶¹]
- (66) Other paced numbers are, “Loaded” and “Jezabel”. Emotion-filled “senti”-tracks too are rendered equally well in “Come To Me” and “Nobody Wants To Be Lonely”.
[*The Hindu*; December 1, 2000⁶²]
- (67) For, Bandhan relies on the ‘bandhan’ of emotional, deeply senti family bonds.
[*The Tribune*; October 3, 1998⁶³]

Enthu, short for *enthusiasm*, is found in similar contexts, as the next set of examples from *The Hindu* illustrates, cf.

- (68) So it was quite an action filled afternoon at the campus, when students went for the record with doses of ‘enthu’, or rather dosas of masala.
[*The Hindu*; September 1, 2001⁶⁴]
- (69) The auditorium packed with youthful students echoed as they yelled out of enthu for each and every argument and counter argument by the students.
[*The Hindu*; March 14, 2002⁶⁵]
- (70) There was cheering, bonding and most definitely infectious enthu in the air.
[*The Hindu*; April 1, 2001⁶⁶]

So is *funda*, short for *fundamental*, which can be used as an adjective in attributive and predicative positions, as in examples (71), (72) and (73) from *The Hindu* and the *Deccan Herald*, cf.

- (71) No high funda statements have been issued. [*The Hindu*: May 12, 2003⁶⁷]

61. www.tribuneindia.com/1999/99jul07/nation.htm, online September 15, 2003.

62. www.hinduonnet.com/thehindu/2000/12/01/stories/1301110q.htm, online September 15, 2003.

63. www.tribuneindia.com/1998/98oct04/sunday/bollywood.htm, online September 15, 2003.

64. www.hinduonnet.com/thehindu/2001/09/01/stories/0401401b.htm, online September 15, 2003.

65. www.hinduonnet.com/thehindu/lf/2002/03/14/stories/2002031401010200.htm, online September 15, 2003.

66. www.hinduonnet.com/thehindu/2001/04/01/stories/09010655.htm, online September 15, 2003.

67. www.hinduonnet.com/thehindu/mp/2003/05/12/stories/2003051200390100.htm, online September 15, 2003.

(72) “Pranab Da’s ‘hi-funda’ ideas scared us ... but then we gradually felt what he was saying,” she adds. [Deccan Herald; March 10, 2002⁶⁸]

(73) WHAT’S ‘FUNDA? Karting. What’s trendy? Spaghetti straps. What’s hot? Cappuccino. [The Hindu; August 2, 2001⁶⁹]

Functionally the most versatile of the three clippings and found in PakE newspapers as well (cf. Appendix 3.16), *funda* may also be used as a noun denoting any ‘fundamental, important or noteworthy idea’:

(74) The funda is that people celebrated the beginning of the new year on April 1 when the pre-Gregorian calendar was being followed ... [The Hindu; April 1, 2001⁷⁰]

(75) As one pleased-looking person said “we might not have made it, but we knew the answers. That’s the whole ‘funda’ about quizzing after all.” [The Hindu; August 17, 2003⁷¹]

Funda has also given rise to a plethora of creative coinages denoting ‘activities or events that are deemed important or entertaining’. In the examples below, the creative coinage *E-funda* is used in reference to a computer course (example (76)) and *balloon funda* in reference to an activity for children at a fun fair (example (77)), cf.

(76) To spread Internet awareness among the kids of Chandigarh the exercise called E-funda has been started. [The Tribune; July 14, 2000⁷²]

(77) ... younger students participated in activities such as ‘Rainbow chain’, ‘Balloon funda’, Paper plate puppet making and spray-painting. [The Hindu; March 8, 2003⁷³]

Also see the following headline found in the *Hindustan Times*, which revolves around playing with the two similar-sounding words *funda* and *fun*:

68. www.deccanherald.com/deccanherald/mar11/yh5.asp, online September 15, 2003.

69. www.hinduonnet.com/thehindu/2001/08/02/stories/13020463.htm, online September 15, 2003.

70. www.hinduonnet.com/thehindu/2001/04/01/stories/09010656.htm, online September 15, 2003.

71. www.hinduonnet.com/lf/stories/2003081707280200.htm, online September 15, 2003.

72. www.tribuneindia.com/2000/2000/07/15/chd.htm, online September 15, 2003.

73. www.hinduonnet.com/thehindu/yw/2003/03/08/stories/2003030800s630300.htm, online September 15, 2003.

(78) Fun funda: Dublin rocks on ladies' night out

[*Hindustan Times*; July 18, 2003⁷⁴]

When new coinages like *senti*, *enthu* or *funda* are put to use in communication, problems of intelligibility may occasionally arise across generations. In the following excerpt taken from a science column in *The Tribune* (May 7, 2003), a university professor comments in a light-hearted fashion on the use of *funda* by his students and the potential misunderstandings resulting from that:

Is there any funda of white holes? I have put in the above question as received. Till three years ago I did not know the meaning of the word funda. There was no such word in any dictionary. Finally I had to ask some young students. They were amazed at my ignorance. When they started talking of fundamentals of physics and chemistry, I realised that funda is a short for fundamental, implying basic laws, even scientific reasons. So here is the "funda" behind the concept of white holes. Before we talk of white holes which probably do not exist, it might be useful to recount the science that led to the prediction of black holes.

(*The Tribune*; May 7, 2003⁷⁵)

Similarly, *enthu* has been brought to the attention of readers of *The Tribune* as an in-group marker among teenagers. The following excerpt is taken from an editorial on the use of English in India:

The world is changing and so is the English language. If you want to be 'with-it', try listening to teenagers since their expressions show the elements of change first. And only then will you be able to figure out that 'enthu' means enthusiasm and 'helu' is not a variant of hullo but refers to guys with dopey eyes. The poor father never gets his due: he is the 'old man' or 'geezer'. Helps though if he is 'loaded' with cash!

(*The Tribune*; January 9, 2001⁷⁶)

As is the case for the impact of mixed forms such as Hinglish (cf. Chapter 3.1.1), it remains to be seen to what extent younger generations of IndE will exploit the possibilities of English word formation to rejuvenate their vocabulary in the future. Given the evidence above, it can be established that mild repercussions of such innovative processes were felt in India's quality newspapers at the beginning of the twenty-first century. Other registers and text genres are likely to be more advanced in this regard, as Bhaya Nair's (2008: 472–89) more recent examples from IndE student writing and chat-room conversations illustrate. According to D'souza (2001a: 152), *funda* has already started to spread into the language of advertising.

74. www.hindustantimes.com/2003/Jul/19/674_313465,003100010002.htm, online September 15, 2003.

75. www.tribuneindia.com/2003/2003/05/08/science.htm, online September 15, 2003.

76. www.tribuneindia.com/2001/2001/09.27/edit.htm, online September 15, 2003.

Table 3.19 *Hydel* vs. *hydro-electric*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN
hydel	311	320	62	117	962	35	508	–	1
hydro-electric	109	81	28	21	362	27	269	–	3

‘Hydel’. While *demat*, *senti*, *enthu* and *funda* illustrate the productivity of the word-formational process of back-clipping in IndE, the item *hydel*, a shortened version of the compound *hydro-electric*, exemplifies a blend that probably originated in the Indian subcontinent. *Hydel* must be considered fully integrated in contemporary IndE while being virtually absent outside South Asia. Recorded by Nihalani et al. (1979: 99; 2004: 99) and Crowther (1996: 1451), the item appears once in the Primary Corpus in a broadcast news report:

- (79) <\$A><#>another important decision of the cabinet committee on economic affairs was to extend four per cent interest subsidies to thermal power projects<,> to be completed before March next year<,> and to hydel projects to be completed March two thousand and four [IND2000 Broad News 02]

In the example above, *hydel* is used in an unselfconscious way, and there are no indications that the blend was expected to cause any problems of intelligibility. Four similar examples of *hydel* in context were traced in ICE-India. In addition, online research (February 19, 2001) shows that *hydel* is also found in South Asian journalistic writing in high numbers today, where it is particularly frequent in Indian and Pakistani newspapers, cf. Table 3.19.

Both IndE and PakE journalists exploit the possibility of using the shortened form in place of its lengthy counterpart to the fullest by preferring *hydel* to the adjectival compound. As was the case for *upgradation* and *delink*, there appears to be little communicative need for BaE and SrLaE journalists to employ the two words.

‘Chargesheet’, ‘incharge’ and ‘undertrial’. The item *chargesheet* exemplifies the workings of functional shift in IndE (cf. Nihalani et al. 1979: 46, 2004: 45; Gargesh 2006: 103). A noun in standard English, *chargesheet* carries a function that must be considered specific to the Indian subcontinent, namely that of a transitive verb meaning “to accuse s.o. formally of committing an offence and to demand an official reply or defence” (Crowther 1996: 1441). Consider the following example from the broadcast section of the Primary Corpus:

- (80) <\$D><#>Mr Patak who was then a B-J-P cooperator and Mr Bhatt<,> an M-L-A<,> were chargesheeted by the police in nineteen eighty-five<,> but it has taken fifteen years for the case to now go on trial<,> in the sessions court [IND2000 Broad News 12]

Table 3.20 *Charge(-)sheet*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
chargesheeted	270	252	21	686	416	71	27	6	16
charge-sheeted	57	19	6	115	189	2	104	18	22
chargesheeting	12	11	1	27	33	1	–	–	1
charge-sheeting	7	1	–	1	21	–	3	3	–

Further research using the available corpora confirms that the item is widely available across South Asia today within the domain of journalistic writing. See the distribution of the inflected verb forms *charge(-)sheeted* and *charge(-)sheeting* in Table 3.20 (Google search dated September 22, 2001). Note the slight usage difference observable between the Indian data on the one hand and the Pakistani, Sri Lankan, and Bangladeshi data on the other. While the IndE online newspaper corpora have carved out a shared preference for the solid spelling variant, it is the hyphenated spelling variant that is preferred outside of India. This finding may be read as a graphic repercussion of slightly variant degrees of integration to which the functional shift has become established across South Asia, with IndE clearly leading the trend toward full integration.⁷⁷

Two more variety-specific conversions attested in the Primary Corpus are *incharge* and *undertrial*. Originally prepositional phrases, *in charge* and *under trial* have developed new variety-specific grammatical functions in IndE, *incharge* that of an adjective ('responsible') and count noun ('person in charge of s.th.') and *undertrial* ('the accused') that of a count noun (cf. Nihalani et al. 1979: 101, 186; 2004: 101, 187; Bansal 1983: 4; Hankin 2003: 498). See the following examples:

- (81) Briefing the Press about the Cabinet decision, the minister said the three ministers, who had been asked to visit the area immediately were the Health Minister Dr. Maalakaraddy, the Urban Development Minister, Mr. B. B. Chimmanakatti, and the incharge Minister of Dharwad District, Mr. A. H. Hindasgeri. [IND2000 Press Rep 11]

- (82) The central store had an executive engineer as its incharge. [The Indian Express; October 8, 1998⁷⁸]

77. Cf. Talaat (1993:61) for attestations of hyphenated *charge-sheet* from PakE newspapers dating from 1985 and 1986. Hickey (2004:552) lists *to charge sheet* (without hyphen) as a PakE lexical item.

78. www.indian-express.com/ie/daily/1998/10/08/28151574.htm, online September 26, 2001.

Table 3.21 *Undertrial(s)* and *incharge(s)*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN
undertrial	79	80	15	334	411	36	24	–	–
undertrials	106	109	20	490	346	31	5	–	–
the incharge	7	10	6	2	47	–	58	–	–
incharges	7	4	1	7	18	–	6	–	2

- (83) One third of our prisoners are undertrials awaiting disposal of cases for more than six months. [IND1978 Press Rep 18]

Unlike *charge(-)sheet*, *incharge* and *undertrial* are widely available only in IndE and PakE journalistic writing nowadays. In the SrLE and BaE corpora, the two items are exceptionally rare (September 26, 2001); cf. Table 3.21. Whether or not we have come across a genuine regional difference among South Asian Englishes cannot be answered on the basis of the available press data alone, as a comprehensive lexical field survey would be necessary to assess the availability of the items on a pan-South Asian scale more reliably.⁷⁹

3.2.2 From words to word formation

Previous research on IndE has attempted to characterize the variety not just in terms of individual variety-specific lexical items but also in terms of underlying preferences for several word-formation patterns (cf. Görlach 1995; Carls 1994, 1999, 2000). Not only has IndE been claimed to display a preference for nominal and adjectival compounding; extensive derivation using Latinate prefixes has also been singled out as a marked feature of the variety. But when taking a closer look at the corpus data to verify or falsify the existing claims, it proved difficult to substantiate the available hypotheses empirically without a database that had been tagged for parts of speech. In the following, the propensity of IndE to use nominal and adjectival compounding as well as prefixation using *pre*, *post*, *pro* and *anti* is discussed on the basis of the available evidence.

Copulative nominal and adjectival compounds: ‘N+cum+N’, ‘adj+cum+adj’. Written IndE comes closest to a quantifiable usage preference in the case of copulative compound nouns and adjectives of the patterns N+cum+N and adj+cum+adj. IndE

79. In contrast to *chargesheet*, *incharge* and *undertrial*, the conversions to *age-bar* (Nihalani et al. 1979: 19; 2004: 17) and to *by-heart* (Nihalani et al. 1979: 41; 2004: 39), which have been claimed to be lexical Indianisms, could not be traced on the World Wide Web, which raises doubt as to their contemporary status as lexical variety differentiators, cf. Appendices 3.17 and 3.18. Also cf. Görlach (1995: 89).

has been claimed to be fond of this pattern most recently by Carls 1999, who goes so far as to say that of the two models of copulative compound nouns available in standard English today, “the model “N+cum+N” (*assistance-cum-information*) seems to be preferred to the model “N-N” (149) in IndE. Ascribing this to “the general trend of IndE to use formal expressions” (150), Carls 1999 adds that “[t]he relatively frequent use of the model “N-cum-N” is supported by the existence of a nearly identical model in Hindi with the element *aur* (‘and’) between the coordinated word-stems” (150). Consider the following example from the Primary Corpus:

- (84) IF you weren't there at The Conclave between 20–22 October for Sapphire's exhibition-cum-sale of exclusive Tanjore paintings, contemporary art prints and sculptures, read this and weep. [IND2000 Press Rev 03]

In example (84), which is from a press review on the opening of an exhibition of contemporary art, *cum* combines the nouns *exhibition* and *sale* to form a compound referring to a function that is partly an exhibition of contemporary art and partly a sale. The dvandva compound *exhibition-cum-sale* is perfectly in line with standard English word formation, which according to the OED has been using *cum* (originally a Latin preposition that was first used in English to combine proper names only) to combine two common nouns since the nineteenth century. A look at the Kolhapur Corpus and its BrE and AmE twin corpora shows that formations based on the pattern N+cum+N are indeed more common in written IndE than BrE or AmE within the usage range under investigation – but their overall frequency should not be exaggerated. See Table 3.22, which compares varieties in terms of instances of tokens, types and hapaxes.

The AmE data did not return a single N+cum+N compound noun, whereas the BrE data returned a meagre four tokens for each of the 1961 and 1991 corpora.⁸⁰ In contrast, the Kolhapur Corpus of IndE returned 14 tokens. While it can be concluded that IndE does employ the N+cum+N model more commonly than BrE or AmE, which gives the hypothesis by Carls 1999 an empirical backing (also cf. Görlach 1995: 81), Carls' (1999) conclusion that “the model “N-cum-N” (*assistance-cum-information*) is preferred to the model “N-N”” (149) is most likely not true. Although this claim cannot be proved empirically (as this would necessitate identifying all copulative N+N formations in the untagged corpus material), the relatively low frequency of N-cum-N formations in the data makes it seem rather impossible that the productivity of the N-cum-N pattern should exceed that of the N-N pattern. The long list of N-cum-N formations provided by Carls (1999: 150), indeed, runs danger of misrepresenting IndE usage today, at least as far as learned writing is concerned. What is more interesting from a structural viewpoint is the fact that IndE users take greater freedom in employing *cum* also to form copulative adjectives. See the following example:

80. This distribution in the one-million-word corpora is in line with Algeo's (2006) observation that “[c]um is very popular in British, but much less so in American.” (165)

Table 3.22 Copulative compound nouns (N-cum-N).

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
N-cum-N					
tokens	14	4	4	–	–
types	14	3	4	–	–
hapaxes	14	3	4	–	–

Table 3.23 Copulative compound adjectives (adj-cum-adj).

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
adj-cum-adj					
tokens	8	–	–	–	–
types	5	–	–	–	–
hapaxes	5	–	–	–	–

- (85) The geographical compulsions flowing from close “Asian geo-political proximity” and the ethnic-cum-cultural ethos, emanating from a strong European heritage pull the country in different directions. [IND1978 Press Rev 08]

Even if such combinations should occur in BrE and AmE from time to time, they are too rare to be represented in the one-million word corpora BROWN, FROWN, LOB and FLOB. In contrast, the Kolhapur Corpus contains eight tokens of five types; cf. Table 3.23. See one more example from the Kolhapur Corpus:

- (86) The French government subsidises the Agence France Presse, France’s national-cum-transnational news agency. [IndE KOL F 05]

Adjectival compounds: ‘N+V-ed’ in attributive position. IndE has also been claimed to show a special liking for using adjectival compounds of the pattern N+V-ed in attributive position. Carls 1999 states that their “frequency of occurrence is remarkable” (150) in IndE, which he relates to a general stylistic preference for formal language and the fact that “[t]he adjectival compounds of the model come very close to suffix-like word-formation elements [in Hindi]” (150). Dubey 1989, too, claims that formations of this type qualify for the status of “Indianisms, though essentially in terms of their frequency than in those of their typical structure and meaning” (133). See example (87) from IndE, which displays the word-formation pattern in question (also cf. Hansen et al. 1996: 224):

Table 3.24 Adjectival compounds (N+V-*ed*) in attributive position.*

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
Tokens	14	35	17	23	9	25
Types	12	27	15	22	9	24
Hapaxes	10	22	13	21	9	23

* IND1978 vs. IND2000: $p < 0.005$; US1961 vs. US1992: $p < 0.01$.

- (87) While the refugees from Pakistan were absorbed without any discrimination in the middle class-led Indian democracy – one of them, I K Gujral, went on to become Prime Minister and another, L K Advani, is second-in-command in the ruling coalition – the landlord-dominated Pakistani polity did not care to absorb the Mohajirs. [IND2000 Press Edit 04]

Whether or not the choice of the two compounds *middle class-led* and *landlord-dominated* in the newspaper editorial above was influenced by the author’s mother tongue can of course not be reconstructed in retrospect. What is more likely is that the adjectival compounds were chosen for reasons of communicative and structural efficiency. The author of the editorial may have felt that alternative structures such as relative or participle clauses would have resulted in a less focused way of communicating his information, and typographical considerations may also have played a role. In fact, as Dahl 1993 puts it, a noun-plus-verb structure “manages to convey a lot of information in an economical way” (45) and must be expected to occur frequently in information-oriented text types in all varieties of English. But why should IndE, possibly because of the influence from Indian languages, still not use the pattern more often than other varieties? To see whether this question could be given an empirical answer, the press sections of the Primary Corpus were consulted to identify the relevant forms manually in the IndE, BrE and AmE sub-corpora and compute the frequencies of tokens, types and hapaxes. See Table 3.24 for the distribution of N+V-*ed* combinations in attributive position.

Interestingly, rather than reflecting an IndE-specific liking for the word-formation pattern, the distribution shows IndE to be participating in a process affecting international newspaper Englishes in general. The data show that newspapers from the US and India have seen a significant rise in the token frequencies of adjectival determinative compounds patterned on the model N+*ed* in the second half of the twentieth century. But there is no statistically significant difference for BrE along the time axis nor between IndE, BrE and AmE for either the earlier or the later developmental stages captured by the data. It must therefore be concluded that there is no numerical evidence for adjectival compounds of the N+*ed* pattern to be more common in IndE than BrE or AmE, at least not as regards the category of journalistic writing. Although it is possible that IndE employs the structure more frequently in other text categories,

Table 3.25 Adjectival compounds (N+V-*ing*) in attributive position.*

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
Tokens	15	11	14	25	8	22
Types	14	11	13	24	8	19
Hapaxes	13	11	12	23	8	16

* US1961 vs. US1992: $p < 0.025$.

this cannot be said for newspaper prose, which also falsifies Dubey's (1989) impressionistic claim of a "preference for *-ed*-suffixed adjectival clusters" (134) in India's English-language newspapers. If Indian background languages should exert influence on IndE in this area, this appears merely to reinforce stylistic trends observable in international press Englishes. Note, too, that when it comes to determinative adjectival compounds of the related pattern N+V-*ing*, the IndE press texts were found to have even fewer tokens, types, and hapaxes than their BrE and AmE counterparts. While AmE shows a statistically significant increase in token numbers over time, IndE and BrE do not (cf. Table 3.25).⁸¹ Unlike in AmE, there appears to be no greater communicative need in IndE or BrE journalism to employ this structure more extensively around the year 2000 than twenty-two years before. It might well be that US-American press English is currently in the process of diverging from IndE and BrE in that particular instance.

Latinate prefixes. Another claim by Carls (1994, 2000) concerns the use of Latinate prefixes, which are said to be particularly frequent in IndE. But as was the case for adjectival compounds of the type N+V-*ed*, no statistical evidence backing up this hypothesis could be produced when consulting the press corpora. Consider Table 3.26.⁸²

According to Table 3.26, IndE shares with BrE and AmE an increase in the number of forms prefixed with *pre-*, *post-*, *pro-* and *anti-* in the second half of the twentieth century. Note that this increase in token frequencies is statistically significant both for IndE and for AmE but not for BrE. Carls' (1994) statement that we are dealing with "[v]ery productive prefixes" (212) therefore does little in characterizing IndE in its own right, at least not in the category of newspaper prose. What the corpora do show is

81. See an IndE example of an N+V-*ing* compound in context: "Muslim rulers are referred to as temple-razing barbarians, the religious tolerance epitomised by Akbar, or the sophistication of culture is entirely forgotten." [IND2000 Press Edit 10]

82. After retrieving the hyphenated forms automatically by using the WordSmith tool, etymological formations (e.g. *pre-empt*; *anti-biotics*) were filtered out by hand in a second step, using the COB and OALD dictionaries as benchmarks. In a last step, the frequencies of tokens, types, and hapaxes per type were computed for each sub-corpus.

Table 3.26 Prefixation: *pre-*, *post-*, *pro-* and *anti-*.*

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
Tokens	24	41	25	28	27	55
Types	20	34	19	23	18	31
Hapaxes	17	31	17	20	12	21

* IND1978 vs. IND2000: $p < 0.05$; US1961 vs. US1992: $p < 0.005$.

that within the domain of journalistic prose IndE usage is somewhat closer to AmE usage than BrE usage, but the small figures do not leave much leeway for interpretation. Moreover, when looking at the individual formations, no noteworthy qualitative differences were uncovered among the varieties either. Note that, in contrast to Carls (1994: 212–3), the use of the structure *anti*+N/ProperN in complement position is not a lexical peculiarity of IndE alone. See the following three examples from IndE, BrE and AmE:

- (88) <\$B><#>on the crucial issue of budgetary recommendations for a cut in food-and-fertiliser subsidy<,> Prabhunath Singh said that the party should request Prime Minister Vajpayee to withdraw the cut as it was anti-poor
[IND2000 Broad News 06]
- (89) MORE than 100 viewers complained to the B B C last night that an American film, Britain Blood, Sweat, and Tears... Plus Twenty Years, was anti-British.
[UK1961 Press Rev 04]
- (90) This is an issue which boils down to a matter of opinion, depending on whether you're an American or National fan and anti or pro-Yankee.
[US1961 Press Rep 19]

In sum, while Latinate prefixes seem to team up with words somewhat more easily today than a language generation ago, there are no clear IndE-specific usage patterns observable in the data considered here.

Nominal determinative compounds: 'N+N'. IndE has also been claimed to show a preference for nominal determinative compounds (N+N) over prepositional constructions (N+pr+N) whenever English word formation allows for vacillation between the two possibilities, and the influence of Indian languages as well as a stylistic preference for nominal style have been quoted as possible influential factors (cf. Whitworth 1982 [1907]: 240–4; Lukmani 1992: 160; George C 1998: 16–7; Carls 1999: 146–7; Trudgill/Hannah 2002: 131; Melchers/Shaw 2003: 142; Gargesh 2006: 104). According to Carls 1999, the frequent use of compounds such as *freedom struggle* (vis-à-vis *struggle for freedom*) and *power centre* (vis-à-vis *centre of power*) can be taken to represent that trend in IndE. This hypothesis, too, must be taken with a pinch of salt in the light of

Table 3.27 Nominal compounds vs. prepositional constructions.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
freedom struggle	222	398	101	622	789	54	48	34
struggle for freedom	16	24	9	31	98	8	23	16
sympathy wave	40	45	9	188	71	23	–	–
wave of sympathy	2	1	4	5	7	2	9	2
power balance	14	7	1	9	29	7	28	8
balance of power	67	86	12	159	108	21	432	178
power centre/center	44	30	13	76	43	15	23	9
centre/center of power	21	12	5	13	5	9	35	8
aid-giver	–	1	–	1	3	–	1	–
giver of aid	–	–	–	–	–	–	1	–
by month-end	46	59	15	138	81	6	–	–
by (the) end of the month	–	1	–	–	–	–	–	–
by January-end	9	15	1	38	14	1	–	–
by (the) end of January	16	19	4	15	31	2	61	12
power shift	20	4	5	–	8	10	22	35
shift of power	9	–	–	1	10	2	50	5
staff strike	34	16	10	2	131	1	16	–
strike of the staff	–	–	–	–	1	–	–	–
delegation member	6	–	–	1	6	2	1	8
member of the delegation	21	6	1	2	71	8	8	16

the available corpus evidence. A look at the online databases from India, Britain and the United States (June 10, 2001) indicates that variation in IndE, BrE and AmE is more complex than Carls' 'trend' reading suggests. First, the Internet was consulted to determine the token frequencies of ten determinative compounds which had been listed by Carls 1999 as manifestations of the claimed trend and were attested for IndE in the Primary Corpus. When comparing those to the token frequencies of their related prepositional constructions, it was only in three out of ten cases that the IndE corpora showed a clear preference for the nominal compound where BrE and AmE did not. In the other seven cases, IndE was just as variable as BrE and AmE in choosing between the compound forms and the prepositional constructions. Table 3.27 gives the search results in detail.

Today, IndE, BrE and AmE newspapers alike use the compound *freedom struggle* more often than the expression *struggle for freedom*, which may be taken as a sign that *freedom struggle* is in the process of becoming fully lexicalized around the English-speaking world, not only in IndE. On the other hand, all newspapers prefer *balance of power* and *member of the delegation* to their corresponding compounds, which is strong evidence against any simplistic trend reading for IndE. It is certainly not the

Table 3.28 *Sympathy wave* vs. *wave of sympathy* [Internet: newspapers].

	PakE DA	BaE IN	SrLE ST
sympathy wave	6	–	11
wave of sympathy	4	1	5

case that Indian journalists automatically go for nominal compounds when there is the possibility to do so. IndE and BrE journalists share a preference for *staff strike* – an item which is not attested in the *Washington Post*, presumably because strikes are too rare in North America to be reported to begin with. *Power centre/center*, on the other hand, seems more frequent a compound in IndE and AmE journalistic prose than in BrE newspapers. As for *power shift*, IndE must be positioned in between AmE and BrE, with the *Washington Post* preferring the compound, the *Guardian* clearly favoring the N+pr+N structure and the six Indian newspapers showing mixed preferences. The item *aid-giver*, finally, is too rare in the newspaper corpora to allow for any meaningful conclusion to be drawn – apart from the fact that it is not a compound restricted in use to IndE alone. It is only for *sympathy wave*, *month end* and *January end* that IndE appears to display a preference for compound forms which is not shared by either BrE or AmE.

What are we to make of those figures? Rather than speaking of a trend toward nominal compounding, the snapshot from the Internet suggests that Indian users of English simply exploit a dimension of variation that is an inherent given in English word formation in general. In the cases of *sympathy wave*, *month end* and *January end*, the compound forms seem to meet a special communicative need in IndE, but they are not in any way unusual or exciting from a structural perspective. *Sympathy wave* should, in fact, be considered a South Asianism, as Table 3.28 (Google search dated June 11, 2001) reveals. Elsewhere in the English-speaking world, the compound appears to be exceptionally rare, which is also suggested by its near absence on the Internet across top-level country domains outside South Asia (June 11, 2001); cf. Table 3.29. Similarly, the compound *month-end* and combinations of the type NAME OF MONTH + *end* such as *January-end* qualify as South Asianisms today. It is possible that the workings of analogy have had a hand in the rise of those two forms, as the compound noun *year end* is available internationally (Google search dated June 13, 2001), cf. Table 3.30. But rather than reflecting a clear-cut trend toward determinative nominal compounding, the results produced above should be read as a warning-sign not to base claims on the productivity of individual word-formation patterns on a selection of individual lexical items. While the situation may look different in other

Table 3.29 *Sympathy wave* vs. *wave of sympathy* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
sympathy wave	4	–	1	–	1	–
wave of sympathy	1	3	124	7	82	5

Table 3.30 The compounds *year-end*, *month-end* and *January-end*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
by year end	32	119	22	534	97	7	20	24
by end of the year	2	3	1	1	4	1	–	–
by the year end	54	64	21	90	162	37	29	1
by the end of the year	70	94	25	215	181	24	846	232
by month end	46	59	15	138	81	6	–	–
by end of the month	–	1	–	–	–	–	–	–
by the month end	42	46	8	44	70	14	–	–
by the end of the month	25	22	7	39	79	11	131	64
by January end	9	15	1	38	14	1	–	–
by the end of January	16	19	4	15	31	2	61	12

written text types or in spontaneous speech, Carls' (1999) trend hypothesis must be refuted for newspaper prose.⁸³

A similar picture emerged when testing the claim that IndE users tend to employ nominal compounds of the pattern N+N in place of combinations of the pattern adj+N and vice versa (cf., e.g., Trudgill/Hannah 2002: 131). According to Carls (1999: 147, 2000: 41), IndE shows a preference for N+N formations, although he is careful enough to admit that some coinages appear to reverse the trend. The following examples from the Primary Corpus suggested themselves for further analysis in this regard as they must be considered unusual by standard English norms:

- (91) <\$A><#>as many as seventeen Pakistani fishermen were reportedly arrested at Jakhau Port in Kachchh district of Gujarat<,> officials of the Indian coastal guard said<,> that the Pakistanis were captured by coastal guards after they strayed into the Indian seas
[IND2000 Broad News 16]

83. Note in this context that the semantic opposition between the compound *beer bottle* (denoting the container) and *bottle of beer* (denoting the content), which has repeatedly been discussed in this context in reference to IndE (cf. Nihalani et al. 1979: 32–3, 2004: 31; Görlach 1995: 81; Carls 1999: 143, 2000: 35) is largely adhered to in India's quality broadsheets; cf. Appendix 3.19.

- (92) <\$C><#>after the attack by the suicidal squad of Lashkar-I-Toiba on Fifteen Corps headquarters<,> the security forces have doubled their vigilance
[IND2000 Broad News 20]
- (93) Badal is learnt to have had a telephonic conversation with Jagir Kaur on Saturday to discuss the brewing crisis. [IND2000 Press Rep 05]
- (94) <\$B><#>even if the Prime Minister has refused to commit himself to the opposition's demands<,> <#>ironically<,> the Congress activism comes just a day before the parliament session ends [IND2000 Broad News 20]

A Google-based search for more instances of *coastal guard* and *suicidal squad* in the online corpora (June 16, 2001) brought to light that both forms were exceptionally rare in the Indian data, with written IndE usage broadly in line with international usage conventions that treat the nominal compounds *coast guard* and *suicide squad* as standard variants. In contrast to those findings, the combinations *parliament session* and *telephonic conversation* turned out to be the preferred variants in the IndE quality newspapers, which makes IndE usage different from standard English in quantitative terms. See the distributions in Table 3.31 for details.

Note that *parliament session* and *telephonic conversation* are also available in PakE, BaE, and SrLE (June 16, 2001), which qualifies the items as South Asianisms, cf. Table 3.32. As more research on the Web (June 16, 2001) helped uncover, *parliament session* is even more widely available across Asia. The nominal compound is also used in Singapore, where it seems to be the preferred choice (cf. Table 3.33). Again, the lesson to be learned is that IndE may well carve out variety-specific preferences for certain compounds not used widely in the English-speaking world – but it is not necessarily the nominal compound of the N+N pattern that becomes preferred automatically. One should be careful not to overinterpret individual findings out of context.⁸⁴

3.2.3 A note on meaning change

Tracing differences in meaning across varieties of English is a complicated task, as Platt et al. 1984 observe in their pioneering comparative study on Englishes worldwide. “Words,” they say, “which may even appear at first to have ‘the same meaning’ in two varieties of English, e.g. British English and Nigerian English, ...

84. When consulting the Internet to check on the availability of several other combinations listed by Carls 1999, the combinations *schedule flight*, *charter bus* and *milched cow* were no more common in IndE than elsewhere. The combination *ice water* appears to be most common in AmE today (BrE and IndE using *iced water* alongside *ice water*), whereas *archival material* is a BrE favorite (AmE and IndE preferring *archive material*). Another lexical (South) Asianism appears to be *departmental store*, which is rarely found in BrE or AmE but which is common in India and, to a similar degree, also in Singapore (cf. George C 1998: 16). Cf. Appendices 3.20–3.22.

Table 3.31 Compounds: N+N vs. adj+N.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
coast guard	111	249	47	350	175	13	124	276
coastal guard	2	–	–	–	–	–	–	–
suicide squad	53	115	31	339	117	10	46	8
suicidal squad	–	–	1	6	5	–	–	–
telephone conversation	59	43	7	28	64	26	197	126
telephonic conversation	69	53	10	36	149	7	–	–
parliament session	117	138	14	333	141	22	3	5
parliamentary session	26	15	4	51	19	4	145	8

Table 3.32 *Telephone/telephonic; parliament/-ary* [Internet: newspapers].

	PakE DA	BaE IN	SrLE ST
telephone conversation	57	9	44
telephonic conversation	62	10	1
parliament session	26	78	14
parliamentary session	24	4	12

Table 3.33 *Telephone/telephonic; parliament/-ary* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
telephone conversation	60	102	5770	424	3160	5960
telephonic conversation	87	–	33	80	18	36
parliament session	109	335	286	20	64	1
parliamentary session	9	37	5190	719	1010	7

may have entirely different shades of meaning for an Englishman than they have for a Nigerian.” (105) In addition to that, “every word may have different shades of meaning for one particular person because of past experiences, pleasant or unpleasant” (105). This is why it remains difficult to pinpoint and quantify changes in meaning despite the availability of electronically readable text corpora. Although small, the Primary Corpus was found to contain several examples of English words that have been adapted innovatively by IndE users at the semantic level to meet their special communicative needs. In the following, those are illustrated.

In the semantic field of politics, which is heavily represented in the corpus Primary Corpus, the nouns *centre*, *union* and *congress* can be applied in variety-specific senses denoting the federal government (*Centre*), the Union of Indian states (*Union*) or the Congress Party (*Congress*), cf.

- (95) The Centre is sitting on silos overflowing with grain while the drought-hit are starving. [IND2000 Press Edit 02]
- (96) <\$A><#>hello and welcome to India Talks <#>I'm Paranjy Guha Thakurta in New Delhi<,> <#>the guest on our programme today<,> is Union Minister for Information Technology Pramod Mahajan<,> <#>he also happens to be holding the portfolio of parliamentary affairs and he's also the official spokesperson of the Union cabinet [IND2000 Broad Int 04]
- (97) <\$D><#>when the Congress was in power<,> uh certainly up to nineteen eighty-nine<,> we used our influence<,> and we used our very very considerable clout including our military clout<,> to bring about the only solution alternative to partition which is possible<,> which is devolution [IND2000 Broad Disc 03]

The adjective *backward* and the participle form *scheduled* (cf. Labru 1984: 132) carry variety-specific additional senses in IndE as well – they may be used to relate to groups, communities or castes that have officially been recognized by Indian authorities for special help and support, cf.

- (98) Under the influence of the Sasikala family (who are Thevars) the party came to be seen as pro-Thevar, a powerful backward community in the southern districts. [IND2000 Rep 03]
- (99) NEW DELHI: Back from a self-imposed political exile, BJP MP Uma Bharati on Sunday threatened to raise yet another banner of revolt, demanding that the BJP agree to a special quota for women belonging to other backward classes (OBCs) in the Women's Reservation Bill. [IND2000 Rep 05]
- (100) <\$L><#>the Andhra Pradesh government's decision to categorise scheduled castes into four groups<,> based on their population strength<,> was today upheld by the Andhra Pradesh High Court [IND2000 Broad News 12]

The verb *wish* may be used in IndE as a monotransitive verb in the sense of 'to greet' (cf. Nihalani et al. 1979: 195; 2004: 197). See example (101), the only one of its kind in the Primary Corpus found in a student essay:

- (101) Our way of wishing others, talking to others, eating habits and dressing ourselves all have a deep impact of Western culture. [IND2000 Ess 05]

While no such use is found in ICE-India, further examples can be traced quite easily in the online editions of India's quality newspapers, which suggests a fairly high

degree of integration in that particular domain. Example (102) is from a caption found below a photograph of Prime Minister Vajpayee shaking hands with an Indian cricket player:

- (102) Prime Minister Atal Behari Vajpayee wishing Sachin Tendulkar while his teammates Saurav Ganguly and Anil Kumble look on at the Prime Minister's residence in New Delhi on Thursday. [The Tribune; February 2, 2002⁸⁵]

Similar examples of semantic extension in IndE are provided by the verbs *loot* and *shift*. While in standard English *loot* 'plunder' usually collocates with inanimate nouns denoting a place, for example *shop* or *house*, IndE has extended the referential potential of the verb to refer to goods, money or to animate nouns (cf. Nihalani et al. 1979: 117; 2004: 117; Hosali/Tongue 1989: 54; Schneider 2007: 165). Consider:

- (103) Robbers looted Rs 26 lakh from the Hill Cart Road branch of UCO Bank this morning. [IND2000 Press Rep 12]
- (104) <\$C><#>see Barkha the pity is<,> do you mean to tell me only the Hindus should be looted<,> only the Hindus<,> <#>my point is this<,> because once a Christian is being looted<,> there's such a big hullabaloo<,> <#>and so many Indian women are raped<,> nobody bothers [IND2000 Broad Disc 05]

Likewise, while standard English employs *shift* most commonly to say that 'something moves or changes slightly,' as in

- (105) <\$A><#>however the bluechip and economy stocks showed some giant gains as investors shifted their money from the technology sector<,> to traditional sectors [IND2000 Broad News 17]

Indian users of English may employ the verb *shift* also to say that 'a person or a company goes to live or work in a different place' (cf. Nihalani et al. 1979: 159–60; 2004: 160), cf.

- (106) <\$B><#>Mr Vajpayee is likely to stay here for another week<,> for post-operating care<,> and later he's likely to shift to Raj Bhawan [IND2000 Broad News 03]
- (107) <\$B><#>oh yeah actually that was when when I joined college<,> ... <#>that's when we when we just shifted<,> <#>and by mistake instead of catching the fast train I caught the slow train and I landed up somewhere completely different [IND2000 Broad Int 03]

85. www.tribuneindia.com/2002/2002/02/02/sports.htm, online March 4, 2003.

Table 3.34 Standard and extended senses of *loot* and *shift*.

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess	IndE ICE-IND writing	IndE ICE-IND speech
loot (standard)	–	2	2	–	1	2
loot (extended)	–	–	2	–	–	1
shift (standard)	8	3	7	3	10	20
shift (extended)	1	2	3	–	16	14

However, although IndE has carved out variety-specific uses of those two verbs, it must be emphasized that the standard English uses have not been replaced. This is shown in Table 3.34, which compares the frequencies of occurrence of standard and extended functions of *loot* and *shift* in the Indian section of the Primary Corpus and in ICE-India. Interestingly, uses of *shift* in its extended function are particularly frequent in ICE-India in more private communication situations, namely in private conversations (nine tokens) and in non-printed letters (seven tokens). In more public communication situations, nativized uses of *shift* only cluster in the category ‘press reportage’ (six tokens). A look at the Internet (March 16, 2003) confirms that the latter finding is not a coincidence (cf. Table 3.35). IndE quality newspapers use *loot* recurrently today where standard English would have *rob* or *steal*, while *shift* may take over functions that are fulfilled by *move* in standard English. Will IndE see any further spread of such uses across text types and registers even into more formal and public domains of use? Exonormatively-oriented prescriptive pressure conforming IndE users to international usage conventions will certainly be a stumbling block to any such development in the future.⁸⁶

Next to *wish*, *loot*, and *shift*, which had been discussed in previous research on IndE, several other lexical items with extended functions that had not been reported yet were also uncovered. One such function concerns the verb *release*. When compiling the Primary Corpus, the following two examples aroused my interest:

- (108) This is clearly a winner amongst the four big releases this week. ‘Astitva’ is no doubt one of the better films to have released this year and we pray that Mahesh Manjrekar continues to make such films. [IND2000 Press Rev 09]
- (109) Dollar Dreams, scheduled to release in the city next week, is Kammula’s first effort. A low budget film, it draws from the experience of its director, telling the story of US-bound Indian youth. [IND2000 Press Rev 09]

86. Cf. Tongue/Hosali (1981:26) and Hosali/Tongue (1989:54), who follow BrE usage conventions by stating that *loot* “does not collocate with *passenger* or nouns denoting human beings.”

Table 3.35 Collocations of *loot* and *shift*.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
looted people	2	–	2	–	11	–	–	–
robbed people	1	2	–	–	7	–	3	1
looting money	–	2	3	–	7	2	–	–
stealing money	7	9	2	5	40	5	34	41
they shifted to + PLACE	9	10	3	1	25	4	–	–
they moved to + PLACE	12	24	5	4	31	18	96	107
shifting to CAPITAL	3	–	1	1	6	–	–	–
moving to CAPITAL	2	3	1	1	3	–	34	15

*CAPITAL stands for the names of the capitals of India, the UK and the US.

For one, *release* is used in both examples with a variety-specific nuance, namely to say that a film ‘is brought to the screen in a cinema’ at a particular point in time (also cf. Nihalani et al. 1979: 149–50, 2004: 150). BrE and AmE would arguably use the verb *open* to express this meaning. From a lexico-grammatical point of view, the examples are even more noteworthy in that IndE shows itself to be participating in a syntactic change affecting Englishes worldwide, the use of verbs ‘active’ in form but ‘passive’ in meaning, the so-called mediopassive verbs (cf. Hundt 1998a: 115–6; Hundt 2007). That the mediopassive use of *release* in reference to the distribution of films is a peculiarity of IndE becomes apparent when comparing the distribution of the collocations *the film/movie will release* vis-à-vis *the film/movie will open* across six of India’s major online newspapers and in the online editions of Britain’s *Guardian* and the US *Washington Post* (November 29, 2003), cf. Table 3.36.

Further research across the top-level country domains on the World Wide Web (November 29, 2003) supports the view that IndE is a stronghold for the use of *release* in that particular context, although the construction is occasionally found elsewhere in the English-speaking world, too, cf. Table 3.37. Interestingly, however, the examples of *the film will release* which were found outside the ‘in’ domain referred exclusively to the Indian movie world, which backs the hypothesis that we are dealing with an Indianism, cf.

- (110) The film will release in the UK on 28th November 2003 by Yash Raj Films. It’s a film about love and friendship.
[BrE; announcement published on the BBC homepage on the release of a Bollywood film⁸⁷]

87. www.bbc.co.uk/shropshire/films/bollywood/2003/10/kal_ho_na_ho_preview.shtml, on-line November 29, 2003.

Table 3.36 Mediopassive *release* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
the film will release	4	1	–	5	1	2	–	–
the film will open	–	4	–	1	1	3	2	6
the movie will release	–	–	–	–	1	–	–	–
the movie will open	–	–	–	–	–	–	–	4

Table 3.37 Mediopassive *release* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
the film will release	15	–	3	3	3	1
the film will open	–	3	197	26	365	21
the movie will release	19	–	3	–	–	–
the movie will open	–	2	9	4	18	14

Curiously enough, Hundt (1998a:115–6) found New Zealand English to use mediopassive *screen* in similar environments, which suggests that Englishes worldwide might generally be quite variable in this (admittedly limited) usage context (also cf. Bauer 2002: 54; Hundt 2007: 167–8).⁸⁸

Another example of IndE innovativeness on the semantic level concerns the noun *meet*. In standard English, *meet* denotes a ‘sports meeting’, as in example (111):

- (111) Bhupathi confident of good show at world doubles meet
[IND2000 Press Rep 08]

In IndE, *meet* can nowadays be used more generally in reference to political or business gatherings. In other words, the referential potential of the item has been extended, cf.

- (112) The meeting was held as part of Mr Pawar’s effort to forge a joint platform of sugar cooperatives. The International Finance Corporation had sent its representatives for this crucial meet.
[IND2000 Press Rep 09]

While this extended sense is not found in ICE-India (which has 14 standard uses of the noun *meet*), an online search query research testing the frequencies of occurrence of the collocations *political meet* and *all-party meet* (May 19, 2001) produced the following distribution indicating that the new extended sense may not be preferred but is still recurrent in Indian newspapers, cf. Table 3.38.

⁸⁸ Cf. Appendices 3.23 and 3.24 for frequencies of occurrence of mediopassive *release* on the Internet. ICE-India did not return any examples.

Table 3.38 *Meet* vs. *meeting* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
political meet	2	6	–	1	1	1	–	–
political meeting	7	22	2	2	31	7	33	11
all-party meet	74	280	35	887	37	18	–	–
all-party meeting	224	335	42	476	536	106	3	–

Table 3.39 *Meet* vs. *meeting* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
political meet	–	–	–
political meeting	18	5	–
all-party meet	2	2	16
all-party meeting	23	4	23

In fact, we might be dealing with a semantic South Asianism, as Pakistani, Bangladeshi, and Sri Lankan journalists also use the noun *meet* in its extended sense from time to time (Google search dated May 19, 2001), cf. Table 3.39.

A similar example of a semantic extension in IndE is provided by the noun *ticket*. In political contexts, *ticket* can be used in IndE in the sense of ‘candidature’ or ‘a position on a list of candidates representing a particular party or group in an election’. See the following examples found in the Primary Corpus:

- (113) Relatives of prominent politicians have even otherwise cornered many tickets from different parties ... Veteran Congress leader Chiranji Lal Sharma has managed to get tickets for his son Ashok Sharma from Sonapat.

[IND2000 Press Rep 13]

- (114) <\$A><#>both the major parties the B-J-P and the Congress<,> rejected the Chief Election Commissioner’s suggestion<,> to make political parties distribute a percentage of their tickets to women rather than reserving seats

[IND2000 Broad News 15]

- (115) <\$D><#>I support your point in fact<,> that if you increase the number of credible women candidates by reserving<,> tickets in the party<,> you are increasing the likelihood of many more women coming

[IND2000 Broad Disc 08]

BrE and AmE both use the word in slightly different ways. While in BrE *ticket* refers to ‘a party that a politician represents or the policies they support’, AmE employs the word most usually to denote ‘a list of candidates representing a particular political

party or group in an election' (cf. Benson et al. 1986: 148).⁸⁹ Considering that BrE and AmE have appropriated the functional potential of the item to their own ends, it is not surprising to find that IndE, too, should have evolved its own variety-specific sense.

A more informal example of semantic extension at work detected in the data comes from another student essay. In IndE, the blend *hi-fi* can be used as an adjective to denote that something is of 'high quality':

- (116) Our mother tongue is Hindi but now for doing every simple business or for joining a good post in any company the person should know high-figh English [sic]. [IND2000 Ess 06]

Since the example above was the only attestation found in the closed corpora, the Internet was checked to see whether we were dealing with a stable linguistic feature, which produced more examples of this variety-specific use, especially in Chandigarh's *Tribune*. In the citations below, *hi-fi* is employed in informal contexts to say that something is 'unusual', 'posh', or 'fancy':

- (117) It has become a status symbol to dress the kids in hi-fi clothes with matching boots, hats and purses. [The Tribune; August 31, 2000⁹⁰]
- (118) The 'karigars' working for designers, copy the designs and offer these to the hi-fi boutiques, who charge a lot of money from their clients ... [The Tribune; April 15, 2001⁹¹]
- (119) IT was just another of the hi-fi ideas of science fiction writers and first popularised by RCW Ettinger in 1964. [The Tribune; February 10, 2000⁹²]

While in the examples above *hi-fi* collocates with inanimate nouns, other examples show that *hi-fi* can also be used also in combination with animate nouns:

- (120) We are very Hi-fi guys and only buy foreign clothes. [The Tribune; July 10, 2000⁹³]

89. See the following example from the Primary Corpus, which shows that this sense is also available in IndE: "Anti-defection is one vital area needing specific changes. The relevant provisions can be modified to ensure that once a candidate gets elected on a particular party ticket, he/she stays with that party for the duration of the legislature." [IND2000 Press Edit 07]

90. www.tribuneindia.com/2000/2000/08/31/ldh1.htm, online May 5, 2001.

91. www.tribuneindia.com/2001/2001/04/15/ldh1.htm, online May 5, 2001.

92. www.tribuneindia.com/2000/2000/02/10/science.htm, online May 5, 2001.

93. www.tribuneindia.com/2000/2000/07/10/ldh.htm, online May 5, 2001.

(121) Why work for less? State may replace hi-fi bureaucrat Bongirwar
[*Indian Express*; June 1, 2000⁹⁴]

(122) Controversies do not seem to stop chasing the hi-fi Minister of State for
Tourism in Rajasthan [The *Tribune*; April 29, 2001⁹⁵]

This extension of meaning has resulted in an extension of the lexico-grammatical functions of the word. In IndE, *hi-fi* may nowadays be used in subject and object complement positions, as is displayed in examples (123) and (124):

(123) Children's Day will be hi-fi this year ... [The *Tribune*; November 14, 2000⁹⁶]

(124) Tamarind Court who, however, adds that he wasn't present at his particular party "lately I had stopped going there as I began to find the atmosphere too hi-fi ... [The *Tribune*; May 3, 1999⁹⁷]

Again, time must tell whether *hi-fi* will become established in its new functions more widely in IndE or whether it will remain restricted in use to more casual and informal contexts.

A final interesting set of examples is provided by the words *pandit* and *mughal*. Those items have in common that they were once borrowed from Indian languages into English and subsequently changed their meanings to fit their new linguistic and cultural contexts of use. Today, IndE shows a characteristic mix of etymologically older and younger senses used side-by-side. The original meaning of *pandit* (also spelled *pundit*) is that of a 'learned person in Hinduism', its first recorded attestation in English dating back to 1672 according to the OED (also cf. Yule/Burnell 1994 [1886]: 740–1; Hankin 2003: 364). Not surprisingly, this is still the most common use in IndE today. See example (125) from the Primary Corpus:

(125) Since the release of her first album ... Padmashree Shubha Mudgal, trained in the Gwalior gharana under the tutelage of Pandit Rameshwar Jha, has been doing consistently well ... [IND2000 Press Rev 09]

Deriving originally from a Persian-Arabic mispronunciation of the name *Mongol*, *mughal* (also spelled *moghul* and *mogul*) has been used as a name for the emperor of Delhi since the late sixteenth century and as an adjective 'of, pertaining or relating to, the Moguls, or the Mongol empire in India' (OED) shortly afterward (also cf. Yule/Burnell 1994 [1886]: 570–3; Hankin 2003: 321–2), as in the following quotation from an editorial:

94. www.indian-express.com/ie/daily/2000/06/01/ifr01038.htm, online May 5, 2001.

95. www.tribuneindia.com/2001/2001/04/29/nation.htm, online May 5, 2001.

96. www.tribuneindia.com/2000/2000/11/14/ldh2.htm, online May 5, 2001.

97. www.tribuneindia.com/1999/99may03/edit.htm, online May 5, 2001.

- (126) Another change is to deny the role of Mughal rule in the evolution of the country. [IND2000 Press Edit 10]

The OED informs us that both *pandit* and *mughal* have undergone similar broadenings of meaning in the process of spreading across the English-speaking world. *Pandit* has come to be applied not only in reference to learned Hindu scholars, but to ‘teachers’ or ‘experts’ in general. Likewise, *mughal* has seen a transfer in function to non-Indian contexts. Today, it can be used in reference to a ‘powerful person’, a ‘businessman’ or a ‘magnate’ without necessarily connoting an Indian context (also cf. McArthur 2002: 322). Online research brings to light that the broadened meanings have been borrowed back into IndE, where they are used in quality broadsheets throughout the country today – albeit in relatively small numbers, which might point to a certain reluctance on the part of IndE users to pick up the ‘imported’ meanings for words originally Asian, cf.

- (127) Conventional wisdom among political pundits in the Capital was that Congress had made a dramatic comeback in the last couple of weeks.

[*The Tribune*; September 25, 1999⁹⁸]

- (128) The media mogul, Mr. Silvio Berlusconi, is likely to come back to power in the polls to be held in April next. [*The Hindu*; April 19, 2000⁹⁹]

See the distribution of selected collocations in India’s online newspapers (October 3, 2000) in Table 3.40.

As can also be learned from Table 3.40, IndE newspapers today basically keep to the internationally established distinctions in spelling to differentiate between the original and the broadened meanings, with the variants <pandit> and <mughal> being more common choices in the first, <pundit> and <mogul> in the second case.

3.3 Lexical style variation

In previous lexicographical descriptions, the lexical style of IndE has been characterized predominantly in terms of a preponderance of formal, literary, and archaic vocabulary or, in Spitzbardt’s (1976) words, in terms of “poetic, archaic, bookish if not bombastic words”, a “penchant for the florid”, a “general Indian English tendency to archaism”, a “common liking for genteelisms, poetical-sounding archaisms, bookish and officialese expressions” and a “prevalence of Graecisms and Latinisms” (38). In such contexts, the use of stylistically marked Latin- and Greek-based vocabulary has been

98. www.tribuneindia.com/1999/99sep25/edit.htm, online October 3, 2000.

99. www.hinduonnet.com/thehindu/2000/04/20/stories/03200006.htm, online October 3, 2000.

Table 3.40 Spelling: <pandit>/<pundit>, <mogul>/<moghul>/<mughal>.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
Pundit Nehru	–	1	–	–	–	–	–	–
Pandit Nehru	30	59	37	16	136	4	8	–
Kashmiri pundits	6	6	2	5	11	1	–	–
Kashmiri pandits	127	124	20	111	371	18	–	1
political pundit(s)	18	15	2	7	54	4	21	13
political pandit(s)	1	3	–	–	6	–	–	–
mughal empire	9	23	–	3	65	1	1	–
moghul empire	2	3	–	1	6	–	11	–
mogul empire	–	–	–	–	1	–	–	–
media mughal	–	–	–	–	–	–	–	–
media moghul	4	9	–	2	3	4	–	–
media mogul	17	3	–	10	25	2	98	28

compared to the use of Sanskrit words in Indian languages, which suggests L1 interference as a trigger for stylistic variation in IndE (cf. Goffin 1934; Spitzbardt 1976: 38–9; Subramanian 1978; Kachru 1982a: 360; Mehrotra 1982: 164; Görlach 1985: 11; Parasher 1991: 182–93; Shekar/Hedge 1996: 61). References have also been made to the historical roots of IndE in the language of British colonial administrators, which possibly encouraged the spread of administrative and technical vocabulary into common usage (cf. Spitzbardt 1976: 37–8, 1987: 130; Mehrotra 1982: 163; Lukmani 1992: 162). Moreover, the acquisitional context of IndE, which often sees English being acquired in formal school settings through the study of textbooks and literary classics, has been mentioned as an influential factor (cf. Spitzbardt 1976: 36–7, 1987: 130; Verma 1978: 210–1; Das 1982: 144; Labru 1984: 14–5; Kachru 1986: 43; Gramley/Pätzold 1992: 444; Carls 1994: 212; Sharma 2001: 350; Hickey 2004: 547). But while there can be little doubt that contemporary IndE has its fair share of ‘formal’, ‘archaic’ and ‘literary’ vocabulary today for the reasons mentioned above, one should be careful not to overgeneralize the situation.¹⁰⁰ It must not be forgotten that Indian users of English are well capable of moving freely along the stylistic cline from formal English to informal English as their communicative needs demand it in concrete communication situations, and it is fortunate that with the advent of electronic corpora and the Internet it has become possible

100. Several formal lexical items, on testing their frequencies of occurrence on the World Wide Web, turned out to be more common in the IndE press material than in BrE or AmE, for example *detenu* ‘detainee’, *to demit* ‘to console’, *untoward* ‘unexpected’ and *thrice* ‘three times’ (cf. Labru 1984: 48; Viereck et al. 2002: 210), cf. Appendix 3.25. However, none of those items is found exclusively in IndE today, nor is any formal item generally preferred. Differences among varieties are merely of a quantitative kind.

to dismiss as wrong such long-standing impressionistic claims as that by Das 1982, who rather bluntly stated once that IndE “tends to use *ancient* rather than *old*, *demise* rather than *death*, *bosom* rather than *chest*, *comely* rather than *pretty*, *resplendent* rather than *dazzling*” (144).¹⁰¹ This chapter concentrates on aspects of lexical style variation that have so far been neglected in studies on IndE – the use of stylistically marked function words (3.3.1) and contractions (3.3.2) as well as the use of near-but-not-identical words (3.3.3). It is the aim of this chapter to contribute to a rounder description of the IndE lexicon based on corpus data, not on stereotypical impressions.

3.3.1 Function words: *among/amongst*, *while/whilst*, *amid/amidst*, *until/till*

The discussion starts by looking at four word pairs that style critics from around the world have pointed out as linguistic markers of regional stylistic preferences: *among/amongst*, *while/whilst*, *amid/amidst* and *until/till*.¹⁰² The working hypothesis was that if IndE usage was indeed characterized by an underlying tendency to use formal vocabulary more often than, say, BrE or AmE, this might also show up in a preference for formally marked function words like *amongst*, *whilst*, *amidst*, and *until* over their less formal equivalents *among*, *while*, *amid*, and *till*. When taking a closer look at the corpus material, however, no such general preference for the more formal items emerged. Instead, IndE was found to display a distinct variational profile across the stylistic range from ‘formal’ to ‘informal’ English that was more complex than any reference to the alleged formality of IndE could capture.

‘*Among*’ vs. ‘*amongst*’. In light of the available corpus data, IndE and BrE display a similar behavior in the way that the prepositions *among* and *amongst* are put to use. Generally speaking, both forms have the same meaning in standard English today, the difference being merely one of style, with *amongst* generally considered the more literary choice of the two. As the American style critic Wilson 1993 puts it, *amongst* is oftentimes felt to have “a rather dusty-genteel quality” (28) to it. A look at the one-million-word corpora reveals that American users of English nowadays share Wilson’s (1993) view more rigorously than users in India or Britain. While *amongst* is almost absent in written AmE today, it functions as a minority variant in IndE and BrE, cf. Table 3.41.

In view of Table 3.41, it is safe to say that *amongst* was on its way toward becoming an archaism in AmE writing at the end of the twentieth century, whereas BrE

101. None of the formal items referred to by Das 1982 is preferred in the IndE material on the Web, nor does the proportion of formal vocabulary appear to be notably higher in IndE than in BrE or AmE. While it is possible that the situation looks different in other text categories, the statement by Das 1982 does not capture the nature of IndE lexical variability in published writing. Cf. Appendix 3.26.

102. Cf. Burchfield 1996 on BrE, Peters 1995 on AusE and Wilson 1993 on AmE.

Table 3.41 *Among* vs. *amongst* [KOL, BROWN, FROWN, LOB, FLOB].*

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
<i>among</i>	443	313	249	370	450
<i>amongst</i>	43	45	51	4	1

* KOL vs. BROWN: $p < 0.001$; KOL vs. FROWN: $p < 0.001$.

Table 3.42 *Among* vs. *amongst* [Primary Corpus, ICE-IND].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess	IndE ICE-IND
<i>among</i>	46	22	24	5	234
<i>amongst</i>	–	2	4	1	51

was holding on to this variant somewhat more tightly.¹⁰³ IndE featured significantly more instances of *amongst* than AmE, while differences between IndE and BrE were marginal. A look at ICE-India and the Primary Corpus suggests that *amongst* continued to have a foothold in IndE around the year 2000, but only as minority variant, cf. Table 3.42.

Note that there is no evidence in the data that *amongst* is any more common in contemporary IndE than in BrE, which also goes to say that IndE does not automatically ‘overuse’ markedly formal items.

‘While’ vs. ‘whilst’. Slightly more distinct regional differences between IndE, BrE and AmE were uncovered when studying the word pair *while* and *whilst*. Corpus evidence suggests that contemporary IndE stands between BrE and AmE today by using *whilst* significantly less frequently than BrE but more often than AmE, for which no attestation was returned from the one-million-word corpora, cf. Table 3.43. *Whilst* has the strongest foothold in written BrE among the varieties investigated. In IndE, *whilst* was noticeably rare around 1978. Twenty-two years later, *whilst* was nearly absent from the Primary Corpus and ICE-India, cf. Table 3.44.

103. Algeo 2006 reports that “[t]he BNC [British National Corpus] has 4447 instances (17 percent) of *amongst* versus 22441 instances (83 percent) of *among*. The Michigan Corpus of Academic Spoken English (MICASE) has 19 instances (12 percent) of *amongst* versus 146 instances (88 percent) of *among*. ... Similarly, CIC [Cambridge International Corpus] has 13 percent of *amongst* versus 87 percent of *among* in British texts, but 1 percent of *amongst* versus 99 percent of *among* in American texts, the same percentages as in LOB and Brown.” (160–1)

Table 3.43 *While* vs. *whilst* [KOL, BROWN, FROWN, LOB, FLOB].*

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
while	693	590	674	680	673
whilst	9	66	54	–	–

* KOL vs. LOB: $p < 0.001$; KOL vs. FLOB: $p < 0.001$.

Table 3.44 *While* vs. *whilst* [Primary Corpus, ICE–IND].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess	IndE ICE–IND
while	68	84	58	10	496
whilst	–	–	1	–	1

Peters (1995: 808) made a similar observation in contemporary Australian English, for which she found *whilst* to be restricted in use to more formal text genres and to literary writing. If the general tide in the English-speaking world is with dropping *whilst*, IndE appears to be slightly more advanced than BrE in that particular instance, with AmE spearheading the trend (cf. Burchfield 1996: 845).¹⁰⁴

‘Amid’ vs. ‘amidst’. In contrast to the two variables discussed above, IndE turned out to be in a class by itself in its choice between the two prepositions *amid* and *amidst*. Corpus evidence suggests that it is only IndE users today who are particularly keen on using *amidst*, the more formal variant of the two by present-day standard English conventions, cf. Table 3.45.

Table 3.45 is noteworthy in that *amidst* is the preferred variant in the IndE material, whereas it is clearly outnumbered by *amid* in BrE and in AmE, with American users of English showing a slightly more pronounced preference for *amid*. Data from the Primary Corpus and ICE–India indicate that the situation had not changed around the turn of the millennium. Although attestations in the closed corpora are rather few, *amidst* is represented in fairly high numbers, cf. Table 3.46.

104. Algeo 2006 provides the following additional information on the distribution of the two variants in BrE and AmE: “[w]hilst is a popular form in British English although secondary to *while*, with 1388 versus 11,800 tokens of the two forms in the OED, 5775 versus 54,778 in the BNC [British National Corpus], and 379.8 versus 5890.1 iptmw [instances per ten million words] in CIC [Cambridge International Corpus] British texts, compared with 8.8 versus 6674.2 in American texts. The Michigan Corpus (MICASE) has no tokens of *whilst* versus 458 of *while* ... In American English *whilst* is rare.” (205)

Table 3.45 *Amid* vs. *amidst* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
<i>amid</i>	6	9	18	14	17
<i>amidst</i>	22	5	8	3	2

Table 3.46 *Amid* vs. *amidst* [Primary Corpus, ICE-IND].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess	IndE ICE-IND
<i>amid</i>	–	2	2	–	2
<i>amidst</i>	–	3	1	–	8

Table 3.47 *Amid* vs. *amidst* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
<i>amid</i>	1630	1640	210	1340	2190	341	6920	2369
<i>amidst</i>	1380	1160	192	920	2380	69	203	92
<i>amid</i> : <i>amidst</i>	1 : 1	1 : 1	1 : 1	1 : 1	1 : 1	5 : 1	34 : 1	26 : 1

In the online newspapers (January 26, 2001), too, *amidst* has a firmer stance in IndE than in BrE or AmE, although it is not necessarily the preferred variant, cf. Table 3.47. A look at South Asian online newspapers (January 26, 2001) shows a similar liking for *amidst* in the Sri Lankan database, whereas both Pakistan's *Dawn* as well as Bangladesh's *Independent* adhere more closely to BrE and AmE usage conventions, cf. Table 3.48. Across the top-level country domains (January 26, 2001), on the other hand, only IndE is found to give preference to *amidst*, cf. Table 3.49.

Given this evidence, the question arises why *amidst* should be going so strong in IndE today and, for that matter, across South Asia. A look at the OED proved insightful in this regard. Consider Table 3.50, which lists the absolute frequencies of *amid* and *amidst* in the OED quotations base from the mid-eighteenth century onward. It appears that the stylistic preference for *amid* in BrE and AmE is a comparatively young phenomenon that evolved around the turn of the eighteenth and nineteenth centuries. Hence, the variational profile of contemporary IndE can be said to display a usage preference that was more common in the English-speaking world around the time when English became rooted in the Indian subcontinent (cf. Chapter 2.1).

Table 3.48 *Amid* vs. *amidst* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
amid	3160	256	1030
amidst	570	86	1830
amid : amidst	6 : 1	3 : 1	1 : 2

Table 3.49 *Amid* vs. *amidst* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
amid	443	3200	111000	9080	31000	7890
amidst	1110	2800	35800	2840	14600	3910
amid : amidst	1 : 3	1 : 1	3 : 1	3 : 1	2 : 1	2 : 1

Table 3.50 *Amid* vs. *amidst* [OED quotations base].

	amid	amidst
<1750	102	118
1750–1799	34	53
1800–1849	147	80
1850–1899	169	105
1900–	78	22

Put differently, rather than representing a variety-specific independent development, IndE usage today represents an earlier developmental stage of English which has been superseded in the prestige varieties. While the exact reason for IndE to be holding on so tightly to *amidst* (as opposed to *whilst* or *amongst*) is yet to be uncovered, the very fact that there has been no communicative need to replace *amidst* in IndE suggests a fairly high degree of independence of IndE vis-à-vis the prestige varieties in that particular instance and serves as a reminder that IndE must not be treated as a carbon copy of BrE or AmE but is to be appreciated in its own right.¹⁰⁵

105. In contrast to Table 3.49 Algeo 2006 reports that “[i]n CIC [Cambridge International Corpus], British *amidst* is less than one-third as frequent as *amid*; but American *amid* is approximately 23 times more frequent than *amidst*” (160). Since the Cambridge International Corpus comprises spoken and written language from Britain and America (cf. Algeo 2006: 4) whereas Table 3.49 represents only written usage, the quantitative differences most likely reflect differences in corpus design, with written AmE usage under the ‘us’ top-level country domain showing a greater proportion of the more formal variant *amidst* than spoken and written AmE usage as represented by CIC.

Table 3.51 *Until* vs. *till* [KOL, BROWN, FROWN, LOB, FLOB].*

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
<i>until</i>	159	470	465	461	380
<i>till</i>	204	76	34	51	33

* Noun uses of *till* 'counter' were excluded from the counts.

'*Until*' vs. '*till*'. Another word pair that displays a characteristically IndE preference pattern from a synchronic point of view is *until* and *till*. In contemporary standard English, the difference between *until* and *till* is one between formal (written) and informal (spoken) English, *until* being the formal and *till* the informal choice (cf. Quirk et al. 1985:534). Nihalani et al. (1979:180; 2004:181–2) were the first to observe that IndE usage did not match this description but instead showed a curious liking for the more informal variant, *till*. A look at the structured corpora confirms that this is indeed the case. In the Kolhapur Corpus, representing IndE usage around the year 1978, *till* outnumbered *until* clearly, whereas in the comparable BrE and AmE databases *until* was the preferred choice, cf. Table 3.51.

The same picture emerges for the situation in IndE twenty-two years later, which suggests that we are dealing with a diachronically stable phenomenon both in writing and in speech, cf. Table 3.52. When consulting the online newspaper corpora (January 30, 2001), even more evidence was produced showing that contemporary IndE is characterized by a clear-cut preference for *till*, cf. Table 3.53. The situation is somewhat more blurry in newspapers across South Asia (cf. Table 3.54). Here, the PakE and BaE newspapers show a preference for *till*, whereas the Sri Lankan *Sunday Times* prefers *until* (Google search dated January 30, 2001). Across the top-level country domains (January 30, 2001), on the other hand, IndE is the only variety displaying a quantitative preference for *till*. In AmE, by contrast, *till* is remarkably rare. This has also been observed by Algeo 2006, who reports that "CIC [the Cambridge International Corpus of British and American English] has nearly 5 times more tokens of *till* in British texts than in American" (204); cf. Table 3.55.

In view of this evidence it can be concluded that the stylistic constraints operating on *till* in BrE and AmE do not hold in contemporary IndE, for which *till* functions as the preferred choice in all stylistic environments, text types and modes. Have we come across an independent South Asian innovation in this particular case, though? A look at the OED quotations base suggests a different reading of the data. What stands as a contemporary usage difference between IndE on the one hand and BrE and AmE on the other is again most likely the outcome of divergent diachronic developments in IndE, BrE and AmE over the past two hundred years. Consider first the distribution of the function words *until* and *till* in the OED quotations base in Table 3.56.

Table 3.52 *Until* vs. *till* [Primary Corpus, ICE-IND].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess	IndE ICE-IND
until	6	7	12	–	127
till	17	25	24	3	227

Table 3.53 *Until* vs. *till* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
until	5150	4100	956	1930	6660	614	34400	15400
till	10700	9120	1410	4040	17300	1790	2200	784
until : till	1 : 2	1 : 2	1 : 1	1 : 2	1 : 3	1 : 3	16 : 1	20 : 1

Table 3.54 *Until* vs. *till* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
until	7550	591	7450
till	8940	816	3760
until : till	1 : 1	1 : 1	2 : 1

Table 3.55 *Until* vs. *till* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
until	16700	44400	2280000	159000	1090000	1310000
till	20800	17800	341000	17500	142000	42400
until : till	1 : 1	2 : 1	7 : 1	9 : 1	8 : 1	31 : 1

Table 3.56 suggests that *until* started to outnumber *till* in contemporary BrE (and AmE) writing only in the late nineteenth century, a development which might have been triggered by the reinterpretation of *till* as a shortened, less formal variant of *until*. By the beginning of the twentieth century, *until* had become established as the more common choice in writing. It appears that IndE had already become independent enough in the Indian subcontinent by then and did not follow the prestige varieties in this development. Hence, what looks like a peculiar preference for the ‘informal’ item *till* in IndE vis-à-vis BrE and AmE today is, again, best treated as a repercussion of a linguistic state that once used to be common the English-speaking world over. At

Table 3.56 *Until* vs. *till* [OED quotations base].

	until	till
1850–59	168	259
1860–69	187	268
1870–79	227	265
1880–89	277	250
1890–99	225	241
1900–09	199	111
1910–19	140	83
1920–29	159	94
1930–39	238	94
1940–49	178	51
1950–59	282	54
1960–69	391	89
1970–79	468	88
1980–89	60	7

present, according to Table 3.55, IndE is most likely, AmE least likely to use *till*, with SgE, BrE, AusE and SAfrE usage falling between those two extreme poles.

Note, too, in this context that there is also some evidence (though scarce) supporting Nihalani et al.'s (1979:180; 2004:181–2) claim that *till* may currently be undergoing an expansion of its functional load. According to Nihalani et al. (1979, 2004), *till* may take over functions usually held by *while* (as in *Offer available till stocks last*, meaning that offers are available as long as stocks last) or *by* (as in *It has not been found till today*, meaning the item has not been found yet). Quantitative evidence from the corpora helps put those hypotheses into perspective. Of the 49 instances of *till* in the press and broadcast data in the Primary Corpus, two were found to display the functions described by Nihalani et al. (1979, 2004). In the following example from a book review, *till* is used in a temporal sense. The example is in line with the first example quoted in Nihalani et al. (1979, 2004):

- (129) “Till the time the British were traders and merchants, they presented a rosy picture of the Indian woman. They saw her as an epitome of beauty, grace, elegance and intelligence,” explains Nevile, adding, “Only after the British took on the mantle of rulers, they started making harsh statements about Indian women.” [IND2000 Press Rev 06]

In the next example, again from a press review, *till* has been stripped of its function of implying that something stops at a specific point in time. It corresponds to Nihalani et al.'s (1979, 2004) second example:

- (130) After graduating from Delhi College of Art in graphic design in '94 Parekh moved to Mumbai to assist fashion-photographer Ashok Salian. Which is where her affair with glossy fashion photography began and continues till today. [IND2000 Press Rev 10]

There is another nonstandard example in the Primary Corpus. In example (131), *till* is used in the spatial sense 'up to':

- (131) This water flowing over the Narmada dam will be supplied through the Narmada main canal till Timba, near Wanakbori reservoir. From here, through the Shodhi canal, it will be transferred to the Raska weir and Pariyej-Kanewal pipeline. [IND2000 Press Rep 02]

Once again, it remains to be seen whether those uses will become more popular in IndE in the future. Around the year 2000, they were still minority variants.

3.3.2 Contractions

If the comparatively high frequencies of the function words *amidst* and *till* in IndE remind us of the fact that Indian users of English today do not necessarily follow internationally established usage conventions automatically, the comparatively frequent use of contracted forms in IndE newspapers and broadcast around the turn of the millennium stands as a luminous example that IndE is not sealed off from contemporary international developments either. It is an established fact that standard English typically employs contracted forms in less planned and less formal writing (cf. Quirk et al. 1985: 122).¹⁰⁶ But corpus research has also shown convincingly that both BrE and AmE saw a significant increase in the use of contractions in the second half of the twentieth century and their subsequent spread also into traditionally more formal domains. This has been interpreted as a manifestation of a more general trend in world Englishes led by AmE, namely that toward colloquializing the written norms (cf. Hundt 1996; Svartvik/Leech 2006: 207–9; Mair 2006a: 183–93). Nowadays, as Peters 1995 points out in her Australian English style guide, “the informality that contractions lend to a style is these days often sought, in business and elsewhere, as something which helps to ease communication” (165). A look at the Primary Corpus reveals that IndE journalistic prose has been participating in this global trend, although it has been doing so at its own pace, which is slower than that of AmE (which is leading the trend) and BrE. This is another small but important finding showing that IndE is stylistically more variable than any stereotypical reference to the variety's supposed formal, archaic or literary flavor could possibly capture.

¹⁰⁶ For a survey of contractions used in BrE and AmE cf. Biber et al. (1999: 1128–32) and Algeo (2006: 19–24).

Table 3.57 Contractions in IndE, BrE, and AmE per 1,000 words
[Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
News reports	0.3	1.3	1.2	2.9	3.1	7.7
Editorials	0.3	0.6	1.0	1.5	2.0	8.4
Reviews	0.5	3.1	3.2	3.0	2.8	6.5

Compare, first of all, the figures for IndE, BrE and AmE journalistic writing in Table 3.57, which gives the accumulated token frequencies of the contracted forms of *not*, *be*, *have* (present and past), *will* and *would* per 1,000 words across the Primary Corpus.¹⁰⁷ It appears that contractions have found their way into the press registers in all three varieties over the past few decades, with AmE leading the trend. While contractions were extremely rare in Indian quality newspapers as late as 1978, which is also noted by Sharma (2001: 350), the trend of adopting contracted forms into writing had caught on in IndE by the year 2000. Although contractions were still less frequent in IndE around 2000 than in BrE around 1991, the gap between IndE and BrE had narrowed down considerably. Moreover, both in IndE and in BrE contractions were less of an option in newspaper editorials (for which contractions might still be felt to be too colloquial), but noticeably common in the press reviews sections, with the category of news reports falling in between. In AmE, by contrast, contractions were distributed almost equally across the text categories represented in the corpus. In fact, it is in the category of newspaper editorials that contracted forms showed the highest recurrency rate in the AmE material from 1992.

In the speech section, IndE turned out to be slightly less advanced than BrE in employing contracted forms, although contractions were generally more common here than in writing, cf. Table 3.58. The most striking difference between IndE and BrE is observed in the broadcast news section, where contractions feature approximately four times more often in the British than in the Indian news programs. Since the broadcast news section is more heterogeneous a text category than interviews and discussions in that it contains both spontaneous and non-spontaneous speech, this distribution must be interpreted with caution, though. It might well be the case that BrE broadcast news programs feature more excerpts from spontaneous speech (i.e. interviews and commentary) than their IndE counterparts.

From a diachronic viewpoint, we note an increase not only in the token frequencies of contractions in present-day IndE but also in the number of types of contractions, although not all text categories have been affected by this development to the

107. Cf. Appendices 3.27 and 3.28 for the frequencies of occurrence of the individual contractions in IndE, BrE and AmE.

Table 3.58 Contractions in IndE and BrE per 1,000 words [Primary Corpus; broadcast].

	IndE IND2000 Broad	BrE UK1991 Broad
News	4.5	16.9
Interviews	22.8	34.1
News	28.9	31.9

same extent. The Primary Corpus reveals that BrE and IndE speech nowadays make use of the same repertoire of contractions. In journalistic writing, however, IndE had allowed fewer types into common usage by the year 2000 than BrE and AmE had by the early 1990s. In the IndE press texts, not a single instance of contracted *has* can be found, although this form is already encountered in the 1961 BrE and AmE press data. The contraction is found in the spoken data, though, albeit in much smaller numbers than in the BrE data.¹⁰⁸ This may have to do with the potential ambiguity of *'s* as a marker of the possessive case as well as the marker of contracted *is* (in its uses as progressive and passive auxiliary). Similarly, IndE newspapers to date remain reluctant to contract the operator *had* (much in the way AmE and BrE were in the early 1960s), while *'d* is slowly spreading in BrE and AmE newspaper prose and becoming acceptable there. Again, the reason may be that *'d* is potentially ambiguous and therefore somewhat more difficult to master, especially for users of English as a second language.

3.3.3 Even more variability? Near-but-not-identical words

When scanning the available usage handbooks on IndE for further examples of lexical differences between IndE and other varieties of English, several claimed Indianisms fell into the category of near-but-not-identical words, for example the word pair *tense/tensed*, for which Nihalani et al. (1979: 176; 2004: 178) observe a confusion in IndE. Since the “[c]onfusion of near-sounding words or ‘minimal pairs’ is an ever-present possibility” (Burchfield 1996: 63) in all varieties of English and a particularly tricky lexical area for learners of English, observations like Nihalani et al.’s (1979, 2004) did not come as a great surprise. However, in order to test whether there was more to the lexical variability in IndE than the occasional confusion of similar items, the corpus material was scanned for authentic examples of such near-but-not-identical words in context. What was detected in due course was more evidence suggesting that the na-

¹⁰⁸ The exact distribution of contracted *has* in the broadcast sections of the Primary Corpus is: UK1991-News: 64, UK1991-Int/Disc: 50; IND2000-News: 11, IND2000-Int/Disc: 20; cf. Appendix 3.28.

ture of lexical style variation in contemporary IndE can be described adequately only in reference to the variety's independent history (cf. Chapter 2.1).

'*Tense*' vs. '*tensed*'. Let us first look at the word pair *tense/tensed*. Nihalani et al. (1979: 176; 2004: 178) notice the confusion between those two forms in IndE that sees the participle used in contexts in which standard English would have the adjective. The example provided by the authors is *the atmosphere became tensed*, on which they comment that "[t]he participial form of the verb 'to tense' is found in BS [British Standard English] but it is rare and would never replace the adjective 'tense' in contexts such as the one exemplified" (1979: 176; 2004: 178). When looking at the Primary Corpus, the IndE section returned one example of *tensed* replacing *tense* alongside two standard instances of *tense*, cf.

- (132) <\$A><#>Tundla in Ferozabad district in Uttar Pradesh remained tensed for a fourth day today<,> after the killing of four dalits here<,> and the subsequent killing of one more person by the police [IND2000 Broad News 04]

However, example (132) looks more like an exception when taking into account the much larger Kolhapur Corpus and ICE-India. While the Kolhapur Corpus does not include a single example of the participle form replacing the adjective but thirteen standard uses of the adjective *tense*, ICE-India returned one use of *tensed* replacing *tense* alongside five standard uses of the adjective *tense*. The following example must therefore be considered to be more representative of IndE usage at large, at least as far as the written domain is concerned:

- (133) When the situation became tense police was called in and six students were taken into custody. [IndE KOL A 35]

Further research using the online newspaper corpora from India, Britain and the US confirmed the hypothesis that contemporary IndE basically keeps to the distinction between *tense* and *tensed*. See the distribution of the two most common collocations of *tense* and *tensed* attested in the online newspaper corpora (July 1, 2002), cf. Table 3.59. A similar picture emerged across the top-level country domains (July 1, 2002), cf. Table 3.60. It is possible, of course, that *tense* and *tensed* are more readily confused in less heavily edited written texts or in speech, but this would most probably be true for all varieties, not only for IndE. In published writing on the Internet, the confusion is rare, and no IndE-specific behavior can be pinpointed.

'*Alternative*' vs. '*alternate*'. The two adjectives *alternative* and *alternate* make up another pair of words that has received wide attention from style critics of IndE and other varieties (cf. Nihalani et al. 1979: 21; 2004: 19–20). In standard English, the adjective *alternate* carries the ordinary sense 'every other', as in example (134) from IndE:

- (134) <\$B><#>and I used to get like pocket-money every alternate day to have an ice-cream<,> so like any normal child [IND2000 Broad Int 03]

Table 3.59 *Tense* vs. *tensed* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
tense situation	31	112	17	28	83	10	41	9
tensed situation	–	–	–	–	–	–	–	–
remains tense	1	15	–	4	16	2	11	6
remains tensed	–	–	–	–	–	–	–	–

Table 3.60 *Tense* vs. *tensed* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
tense situation	155	60	11900	357	845	1260
tensed situation	2	2	116	–	2	–
remains tense	13	14	1400	111	366	92
remains tensed	–	–	3	–	1	–

On the other hand, the adjective *alternative* is normally used in the sense ‘(of one or more things) available in place of another.’ See example (135):

- (135) In addition, non-government organisations should come forward to save forests by providing alternative ways of earning a living to tribals alongside an arrangement for their education. [IND2000 Press Edit 08]

Alternate can also be used in the place of *alternative*, a use common especially in AmE if we follow the labeling practice of the major English-language dictionaries today. Consider example (136) from AmE:

- (136) .. she uses it to monitor herself (thus freeing others for alternate forms of surveillance) ... [AmE FROWN G 65]

Interestingly, when scanning the Primary Corpus and ICE-India, IndE came closer to what has been described to be typical of AmE usage. In the Primary Corpus, there were three examples of *alternate* used in place of *alternative* (exclusively found in the broadcast section) alongside one use of *alternative*. Similarly, in ICE-India, there were four attestations of *alternate* ‘alternative’ vis-à-vis 16 attestations of *alternative*. See the following example from the Primary Corpus:

- (137) <\$A><#>Mr Sayeed also requested the Prime Minister to provide some financial help to these passengers<,> for their sustenance<,> till some alternate arrangement is made for their transportation [IND2000 Broad News 01]

Table 3.61 *Alternate* vs. *alternative*.

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
alternate(ly) 'every other'	10	9	3	10	6
alternate(ly) 'alternative(ly)'	7	1	–	6	5

A look at the Kolhapur Corpus and its accompanying BrE and AmE equivalents supports the view that IndE is slightly more variable than BrE in that instance, cf. Table 3.61. Burchfield (1996: 45) informs us that the use of *alternate* 'in place of another' has its origins in early twentieth-century AmE. Although the figures from the structured corpora are quite small, they do suggest that this use has become established internationally by now, with IndE being slightly more open to employing this form than BrE.

'*Lectureship*' vs. '*lecturership*'. While IndE displays greater affinity to AmE than BrE in the previous case, IndE is in a class by itself when it comes to the choice between the two lexical items *lecturership* and *lectureship*, which have also been referred to in previous descriptions of IndE (cf. Nihalani et al. 1979: 112, 2004: 112; Carls 1994: 214). Consider the following example of *lecturership* from an IndE press editorial first:

- (138) Recently the commission advertised four vacancies for lecturerships in journalism in four colleges; there were 10 candidates for these posts.

[IND2000 Press Edit 08]

What is unusual about example (138) by contemporary international standards is the very use of the word *lecturership*, for which the near-identical form *lectureship* would be a more natural choice to denote 'the post of a lecturer'. ICE-India contains a similar example of *lecturership* within a passage of spontaneous speech but no example of *lectureship*.¹⁰⁹ A look at the online newspaper corpora shows that IndE is, indeed, more variable than the codified international standard in that it employs both *lecturership* and *lectureship* side by side (July 9, 2002). However, BrE, too, displays some variability here, at least in the online edition of the *Guardian*, cf. Table 3.62.

Though not codified in the major dictionaries of contemporary English, *lecturership* is also found across South Asia today, for example in Pakistan's *Dawn* and Bangladesh's *Independent* (July 9, 2002), cf. Table 3.63. In fact, outside the domain of quality newspapers, *lecturership* seems to be more common a choice in South Asia than elsewhere in the English-speaking world (July 9, 2002), cf. Table 3.64.

109. ICE-India: "<\$C><#>And we got uh <,,> certificate for the eligibility of lecturership" [ICE-India S1A-21].

Table 3.62 *Lecturership* vs. *lectureship* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
lecturership	2	5	–	3	47	1	9	–
lectureship	4	8	1	6	24	5	88	14

Table 3.63 *Lecturership* vs. *lectureship* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
lecturership	4	1	–
lectureship	7	–	–

Table 3.64 *Lecturership* vs. *lectureship* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
lecturership	99	4	1020	6	49	–
lectureship	118	52	6580	183	1910	131
	1 : 1	1 : 13	1 : 6	1 : 31	1 : 39	–

A look at the OED is once more informative, as it shows that variation between *lectureship* and *lecturership* was more common in late nineteenth-century English.¹¹⁰ While the major varieties have carved out a clear preference for *lectureship*, IndE usage has apparently retained a degree of variability which used to be normal on both sides of the Atlantic as well.

‘Inflammable’ vs. ‘flammable’. Among the near-but-not-identical words tested in this study, contemporary IndE turned out to be most different from BrE and AmE in the case of the pair *flammable/inflammable*. Both words can be used synonymously today to say that ‘something catches fire easily’. See two examples from IndE:

(139) <#>I mean how if if you don’t have any inflammable material in your house the neighbour will not be able to burn anything [IND2000 Broad Disc 02]

(140) Class B fires are those in which flammable liquids, oils, and grease are burning. [The Tribune; May 12, 2000¹¹¹]

110. *Lectureship* is the elder of the two forms, its first attestation dating back to the year 1634. *Lecturership*, labeled ‘rare’, is the more recent coinage, its first attestation in the OED quotations base going back to 1891.

111. www.tribuneindia.com/2000/2000/05/13/windows/fact.htm, online February 7, 2002.

Table 3.65 *Flammable* vs. *inflammable* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLE ST	BaE IN	BrE GU	AmE WP
flammable	1	12	6	11	22	1	9	3	2	57	36
inflammable	36	22	5	52	129	13	35	124	1	23	–

Table 3.66 *Flammable* vs. *inflammable* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
flammable	279	534	18200	1900	13400	22400
inflammable	283	86	2710	216	1460	1620

While ICE-India did not return examples of any of the two words in context, the frequencies of the two variants in the online newspapers from India, South Asia, Britain and the US (February 7, 2002) brought to light a noteworthy distribution, cf. Table 3.65. It is clear that *inflammable* is widely preferred in South Asia today, whereas *flammable* is taking the lead in BrE and AmE. A similar picture emerged when looking at the distributions of *inflammable* and *flammable* across the top-level country domains (February 7, 2002), cf. Table 3.66.

The explanation for this stylistic difference between IndE and the other varieties lies in an interesting case of language planning that has apparently resulted in divergent diachronic developments among Englishes worldwide. According to Burchfield (1996: 301), *flammable* was reintroduced and pushed in Britain and America to replace the more traditional *inflammable* only in the second half of the twentieth century, namely to avoid ambiguity which may arise from mistaking the prefix *in-* for a negative.¹¹² In that course, AmE appears to have adopted *flammable* most rigorously, with BrE following shortly behind. In IndE, no such large-scale change in preference appears to have set in so far, and *inflammable* continues to be the more common variant, at least as regards the written domain. Again, it remains to be seen whether IndE usage will ultimately converge toward the international usage standard in that particular instance. Around the turn of the millennium, South Asian English was obviously in a class by itself.¹¹³

112. This move was pushed in Britain by the British Standards Institution but so far “has met with only partial success” (Burchfield 1996: 301). Peters 1995 points out for the US and Australia that “[t]he US National Fire Protection Association adopted it [*flammable*] in the 1920s but in Australia it has only recently become standard on warning signs.” (282)

113. Note that *inflammable* apparently continues to be the normal choice in figurative language to say that a person is ‘full of emotions or violence’ (OALD). This use is not attested in the Primary Corpus and rare in the online newspaper corpora. Also see S. Upendra’s language

3.4 Americanisms and Briticisms in the IndE lexicon

Since the beginning of the second half of the twentieth century, the influence of AmE on Englishes around the world has intensified (cf., e.g., McArthur 2002: 370–6 or Svartrvik/Leech 2006: 157–9). The popularity of American literature and popular music, of American films and television programs, and the advent of the personal computer with its American-based English-language software are only some of the factors that have contributed to this process, which has resulted in the spread of AmE forms and structures around the English-speaking world. In this context, AmE has even been termed “[a] major unifying force” (Leitner 1992a: 217) for Englishes worldwide.¹¹⁴ Since AmE had repeatedly been observed to exert influence on IndE but no systematic study had been conducted on that issue (cf., e.g., Taylor/Rao 1981; Spitzbardt 1987: 130; Leitner 1991b: 222; Shekar/Hedge 1996: 57), my interest in the intrusion of AmE lexical items into IndE had been kindled. In view of the continuing prestige of the BrE model in India, no large-scale reorientation of IndE usage conventions away from its parent variety BrE toward AmE within the standard usage range was expected. It was nevertheless hypothesized that if IndE had opened up to the influence from AmE over the years, this might be reflected in a somewhat greater proportion of AmE-typical lexical items in the structured corpora around the turn of the millennium than around 1978. When scanning the Primary Corpus more systematically for ‘classic’ lexical Americanisms, however, relatively few turned up, which suggested that the postcolonial force of AmE had so far only had a mild impact on IndE in the usage domains studied. Yet in spite of the rarity of lexical Americanisms in my material, a closer look at the items that were uncovered revealed noteworthy new insights into the interaction of IndE with AmE and BrE both from a synchronic and from a diachronic viewpoint.

Following Leitner’s (1991b) suggestion that “[a] look at spelling practices is ... revealing because it can show (i) if Indian English aligns more with its parent, viz. British English, or (ii) already shows influences from American English” (222), I started out my investigation by looking at spelling, an area for which the two prestige

column *Know your English* in *The Hindu* on December 21, 1999, to which a reader, obviously concerned about how to use the word pair properly, addressed the question: “What is the difference between “flammable” and “inflammable?” Upendra’s answer was that “[t]here is no difference; the two words can be used synonymously. Both *flammable* and *inflammable* can be used to refer to things that can easily catch fire.” But Upendra also added that “[w]hen you are referring to a person’s temper, however, you cannot use the word *flammable*. You can only use the word *inflammable*.” (www.the-hindu.com/1999/12/21/stories/13210675, online February 9, 2002)

114. Cf. Modiano 1996 on the influence of AmE on English in Europe, Möbärg 1998 and Virtanen/Lindgrén 1998 on the influence of AmE on English in Scandinavia, and Collins/Peters 1988 and Taylor 2001 on the interaction between AmE and AusE.

varieties BrE and AmE have carved out noticeable differences over the centuries.¹¹⁵ Historically, BrE spelling can be said to have its roots in Dr Johnson's dictionary of 1755, whereas AmE spelling today broadly follows the norms laid out in Noah Webster's later dictionaries of 1828 and 1838 (cf. Görlach 1985: 18–9, 1995: 133–4; Venezky 1999: 25; Bauer 2002: 62–7). While both spelling systems are known in India today, it is the BrE system that is predominantly taught and employed in schools across the country, which is also the most likely reason why AmE spellings are rare in the writing compiled in Kolhapur and the Primary Corpus. But American-based spellings are certainly not unknown. See the following examples:

- (141) The Golden Bridge pottery center in Pondicherry is the only place in the south which teaches this art form but they take only seven people every year.
[IND2000 Press Rev 05]
- (142) In his defense, Nayak said any move to lodge a police complaint would only aggravate his problems.
[IND2000 Press Rep 04]
- (143) Channeling the Narmada dam overflow is the only hope.
[IND2000 Press Rep 02]

There is no sign that any of the AmE-typical spellings was used self-consciously in the examples above. The spelling <center>, used in example (141), has been standard in AmE since Noah Webster supported it in favor of <centre> in his “radical dictionary of 1806”, as Peters (1995: 640–1) informs us. BrE would usually have <centre> in this case (cf. Venezky 1999: 26). Similarly, <defense> in example (142) and <channeling> in example (143) reflect current AmE practice, whereas BrE prefers the variants <defence> and <channelling> (cf. Venezky 1999: 26). When comparing the number of American-based and British-based spellings in the structured corpora, examples of the types above were of rather exceptional status. Moreover, there was no evidence that AmE spellings had made any significant inroads into IndE journalistic writing over the years. It can therefore be established that contemporary written IndE to date shares with BrE a diachronically stable preference for the noun endings <re>, <our> and <ce> in words like *centre*, *favour* and *defence* and for the doubling of consonants in words like *channelling*, *travelling* and *modelling*. Also, IndE and BrE are alike in their use of <ce> and <se> in words like <licence>/<license> and <practice>/<practise> to mark the distinction between noun uses (marked <ce>) and verb uses (marked <se>), while AmE uses <practice> and <license> irrespective of the grammatical roles of the forms (cf. Carney 1994: 476; Venezky 1999: 26; Bauer 2002: 64–5). Moreover, both IndE and BrE use the verbal endings <ise> and <ize> side by side, whereas AmE has standardized on <ize> (cf. Carney 1994: 475; Bauer 2002: 62; cf. Appendices 3.29–3.36). Taken

115. Leitner 1991b found a “preference for the *-re*, *-our*, *-ise* and *-ll-* versions in words such as *metre*, *colour*, *realize*, and *traveller*” (222) in the Kolhapur Corpus and concluded for 1978 that “the British tradition prevail[ed]” (222).

together, those findings support Melchers/Shaw's (2003) view that "South Asian English is predominantly spelt in the British style" (139).

A closer look at the stratification of the AmE-favored spelling variants in the Kolhapur Corpus brought to light a noteworthy text-category-specific distribution. In Kolhapur, AmE spellings are found more frequently in the categories of academic and scientific writing. Of the 26 instances of <center>, <color>, <favor>, and <offense> in Kolhapur, 20 were found in text category J, which comprises specimens of learned and scientific writing. Similarly, <iz> spellings in Kolhapur cluster in that particular text category (cf. Appendix 3.37). It appears that IndE learned and scientific writing is slightly more open to adopting AmE spellings than, for example, IndE quality newspapers – a finding which corresponds to results produced by Peters (1995: 406) on AusE spelling practices. AmE-typical spellings, it can be concluded, are endorsed especially by scholars and scientists across the English-speaking world.

Though IndE at large, hence, remains oriented toward BrE spelling conventions, individual users of IndE are always free to choose between the BrE and the AmE models to meet their personal communicative needs. A look at India's quality newspapers on the Internet reveals that AmE spellings are regularly exploited for their international cachet, for instance in the worlds of sports and business. In examples (144) and (145), the spelling <color> has been appropriated for the name of a golf tournament and for the name of a business company:

- (144) NEW DELHI: Digvijay Singh, one of the most improved golfers on the Hero Honda Golf Tour, reclaimed the top spot on the Order of Merit after his sensational victory at the Color Plus Open last Friday.

[*The Tribune*; August 29, 2002¹¹⁶]

- (145) CAN there be a perfect way of betting on a Supreme Court verdict? It appears so going by the trading pattern in Color Chem Ltd (CCL) stock on the major bourses following announcement of an open offer on Monday by its foreign promoter and persons acting in concert.

[*The Hindu*; April 10, 2003¹¹⁷]

AmE spellings may, however, appear unacceptable to Indian users of English whenever unduly smacking of 'Americanization'. The following example is an excerpt from a book review in *The Hindu*. Here, the reviewer criticizes the editor of an Indian literary work for employing AmE spellings:

Better editing could have been done, eliminating obvious repetitions in many places. Cleaning up of the copy was in the main left out as one finds out, encountering error after error beginning with the "Acknowledgements" page and

116. www.tribuneindia.com/2002/2002/08/29/sports.htm, online February 8, 2003.

117. www.hinduonnet.com/businessline/2003/04/10/stories/2003041002381300.htm, online February 8, 2003.

thereafter almost on every page. The adoption of American spelling also strikes a jarring note. (*The Hindu*; April 15, 2001¹¹⁸)

While one can only guess how many users of IndE would share this critic's view, it is likely that AmE spellings are generally considered more acceptable for the purpose of international communication. In the following excerpt from an article published in *The Hindu*, its author rather pragmatically recommends the use of AmE spellings in job applications addressed to US companies, cf.

Remember, too, that spelling can differ when writing American English and U.K. English. For example, words like *colour* or *favour* are spelled *color* or *favor* in American English. See if the spell check programme on your computer allows you to check for American spelling to make your missive even more appropriate to the market you are trying to contact. (*The Hindu*; January 28, 2001¹¹⁹)

Only time can tell if AmE spelling variants will eventually be employed for more domestic uses as well and spread more evenly across text types and registers in IndE. Around the turn of the millennium, there were no signs of any large-scale orientation away from the BrE model in India's quality newspapers or in the corpus of IndE student writing.

Next to spelling, IndE and BrE were also found to share a liking for several common nouns that are rare in AmE. When looking up the word pairs *railway/railroad*, *working day/workday*, *public transport/public transportation* and *pavement/sidewalk* in the Primary Corpus and in the electronic newspaper corpora, it was always the first variant that turned out more common in IndE and BrE, whereas the second variant predominated in AmE. As the structured corpora inform us, we seem to be dealing with diachronically stable regional differences (cf. Appendices 3.38 and 3.39). See the following examples of the IndE-typical forms:

- (146) "After speaking to several other old people in the area," he says, "I came to a conclusion that he worked on the long railway that was constructed between Thailand and Myanmar by the Japanese Army." [IND2000 Press Rev 04]
- (147) The precinct has to endure the wear and tear of eight million feet every work-ing day. [IND2000 Press Rep 05]
- (148) Increase in diesel price leads to inflationary pressure, he said, pointing out that diesel was the chief fuel for public transport. [IND2000 Press Rep 02]
- (149) The morning rain keeps falling on the already wet pavement of the Birla Sabhaghar, passers-by scuttling up and down, huddled under umbrellas. [IND2000 Press Rev 04]

118. www.hinduonnet.com/2001/04/15/stories/1315017l.htm, online February 8, 2003.

119. www.hinduonnet.com/2001/01/28/stories/1328046a.htm, online February 8, 2003.

But again, although rare in IndE newspapers today, American forms can occasionally be encountered, and the online newspaper corpora are a valuable resource for tracing their inroads in more detail. One factor contributing to their spread in IndE quality broadsheets appears to be their use by international news agencies. See example (150), which illustrates how an Americanism (*sidewalk*) traveled into Chandigarh's *Tribune* via a quotation in a DPA news item:

- (150) WASHINGTON: Shel Silverstein, author of noted children's books such as "Where the Sidewalk Ends", "Giraffe and a half" and "A Light in the Attic", died on Monday ABC News reported. [The *Tribune*; May 12, 1999¹²⁰]

Lexical Americanisms can also be found in texts composed by Indians who have been exposed to AmE first hand. Example (151) has been taken from a letter to the editor, in which the author compares working conditions in India and in the US. Reporting a conversation that he had with a friend in the US the previous night, the author uses the 'American' *workday*, an unusual lexical choice by IndE standards, cf.

- (151) A friend of mine, who settled in the US a couple of years ago, called on us yesterday, and by the time talk veered around to work, I realised that we had actually been talking about his work all the time, and had just then started on mine. What has really happened to our workday? [The *Hindu*; March 19, 2001¹²¹]

In example (152), the lexical Americanism *public transportation* has been employed by an Indian who, it becomes known, used to do business in the United States:

- (152) You know the USA is a huge country, where the car is the most popular mode of transport. And Los Angeles, where I was based, is a big geographical sprawl where there is no reliable public transportation. [The *Tribune*; December 10, 2003¹²²]

Example (153) is also noteworthy in this context in that the AmE-preferred *sidewalk* has been paraphrased twice in another letter to the editor of the *Deccan Herald*, cf.

- (153) If hawking on the sidewalk (pedestrian footpath) is made legal, Bangalore would become another Chennai ... Being in the US now, I am able to appreciate the power of the individual tax payer. ... Did you know that

120. www.tribuneindia.com/1999/99may12/world.htm, online December 28, 2003.

121. www.hinduonnet.com/businessline/2001/03/19/stories/101939w1.htm, online December 28, 2003.

122. www.tribuneindia.com/2003/2003/12/10/edit.htm, online December 28, 2003.

Table 3.67 *Film* vs. *movie* vs. *picture* ‘moving pictures’ [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
film	38	71	50	23	17	16
movie	–	10	1	4	2	18
picture	–	1	3	–	4	2

I could object or call the police if my neighbour regularly parks the car in his/her driveway blocking the sidewalk (footpath) in front of his house?
[*Deccan Herald*; May 8, 2006¹²³]

It turns out that the author is a non-resident Indian in the United States, whose letter was glossed so as not to confuse an Indian readership. In the big picture, using (and paraphrasing) *sidewalk* would be unusual by IndE newspaper norms.

Another interesting set of words that is nowadays put to variable use in IndE, BrE and AmE is found in the semantic field of ‘motion pictures’. In Burchfield’s (1996: 147) words, “[i]n Britain one goes to the *cinema* or *the pictures* to see a *film*. In America one goes to the *movies* or to a *theatre* to see a *movie*. In Australia and NZ [New Zealand] one goes to the *pictures* or a *picture theatre* to see a *film*.” If we follow Nihalani et al. (1979, 2004), IndE may be even more variable than that, as Indians are said to go to *cinema halls*, *cinema houses*, *picture halls*, *picture houses*, *talkie-houses* or simply to the *pictures*, the *movies*, the *talkies* or the *bioscope*, where they enjoy *films*, *pictures*, *movies*, or *talkies* (cf. also Melchers/Shaw 2003: 143). See, first, the distribution of the three lexical items *film*, *movie* and *picture* ‘motion picture’ in the Primary Corpus (Table 3.67).

As Table 3.67 indicates, *film* and *movie* are common in Indian, British and American newspapers nowadays, whereas *picture* ‘moving picture’ was a rare choice at the turn of the millennium in all varieties, its only attestation in IndE found in the context of a company name (cf. example (156) below). The distribution in the Primary Corpus also informs us that *movie*, widely perceived as an Americanism today, appears to have spread across AmE journalistic writing only in the second half of the twentieth century and has been picked up by IndE and BrE journalists alike (cf. Appendix 3.41). See the following examples of the competing variants in context:

- (154) I just walked into a theatre in Hyderabad, told the manager I would pay him for the hall and released the film.
[IND2000 Press Rev 10]

123. www.deccanherald.com/deccanherald/images/hawkerresponse3.asp, online September 2, 2006.

Table 3.68 *Film* vs. *movie* vs. *picture* ‘moving pictures’ [Primary Corpus; broadcast].

	IndE IND2000 Broad	BrE UK1991 Broad
film	17	29
movie	2	3
picture	–	–

(155) Most of the comedies which features Kamalhassan, barring MMKR, has never got fullsome praise from the public, and ‘Tenali’ too is one such movie which has created mixed feelings among the audiences. [IND2000 Press Rev 05]

(156) A sincere effort from Talking Pictures and Hansal Mehta, the film is worth a look if black comedy is what you are looking for. [IND2000 Press Rev 04]

It is questionable whether Nihalani et al.’s (1979) statement that “‘picture’ continues to be used by many IVE [Indian variants of English] speakers in preference to the (now very widespread) BS [British Standard English] ‘film’” (139) is still valid today. In the IndE broadcast material, both *film* and *movie* were attested but *picture* was not, cf. Table 3.68.

In reference to the location where motion pictures are shown, IndE users may choose among the compound nouns *cinema hall*, *cinema house* and *movie hall*, all of which are found recurrently in Indian newspapers today but are rarely used in BrE and AmE newspapers. Here, *cinema* and *movie theater* would arguably be more common choices, cf.

(157) We bunk classes or sometimes in some cases we skip our school also and go to far off places with our friends or to cinema halls to enjoy the latest movies. [IND2000 Ess 01]

(158) Recreation centres as well as points of entertainment such as cinema houses crowd in south Bombay; some of them have however gone to the residential areas of central Bombay also. [IndE KOL J 28]

(159) One is in the movie hall. The movie has reached its climax. [The Tribune; August 12, 2000¹²⁴]

The investigation of the online newspapers (Google search dated December 30, 2003), thus, confirms Nihalani et al.’s (1979, 2004) claims that *cinema hall* “is rarely found in contemporary BS [British Standard English], though it once existed there” and that “[c]inema house’ is also sometimes heard in IVE [Indian variants of English]”

124. www.tribuneindia.com/2000/2000/08/12/ldh.htm, online December 30, 2003.

Table 3.69 *Cinema hall vs. movie hall vs. picture hall.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
cinema hall	10	710	1220	49	533	165	3	1
cinema house	2	65	8	4	213	2	1	1
movie hall	4	94	17	7	17	20	–	1
movie house	–	25	2	3	3	2	33	290
movie theatre/-er	6	107	16	13	34	37	49	1861
picture hall	–	–	–	–	5	1	–	–
picture house	–	7	5	–	3	1	34	12

Table 3.70 *Talkie vs. bioscope.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
talkie	1	166	15	4	29	855	29	42
bioscope	–	58	3	–	8	7	2	–

(1979: 49, 2004: 47). The compound *picture house*, which Nihalani et al. (1979, 2004) describe as being “of quite frequent occurrence in IVE [Indian variants of English]” (1979: 139, 2004: 139), is rare in the IndE newspaper material, though (Table 3.69).

The items *talkie* and *bioscope*, discussed in Nihalani et al. (1979: 34, 175, 2004: 32–3, 176), are also recurrent in IndE newspapers (December 30, 2003), cf. Table 3.70. But as is the case for BrE and AmE, when *talkie* or *bioscope* appear in IndE newspaper articles, it is usually in the context of discussing the history of motion pictures. See examples (160) and (161), which are typical in this regard:

- (160) After training under R. Prakash, he turned filmmaker and created history by establishing the first talkie studio in South India, “Srinivasa Cinetone” (also known as “Sound City”) in Lawder’s Gate-Vepery area.

[*The Hindu*; July 1, 2000¹²⁵]

- (161) Primitive cinematic devices like the ‘cinematograph’, ‘vitascope’ and ‘bioscope’ came into being and began to be exploited in the 1890s almost simultaneously in the USA, France, Germany and Great Britain.

[*The Tribune*; February 26, 2000¹²⁶]

125. www.hinduonnet.com/2000/07/10/stories/09100224.htm, online December 30, 2003.

126. www.tribuneindia.com/2000/2000/02/27/spectrum/index.htm, online December 30, 2003.

It is possible, of course, that *bioscope* and *talkie* are still more common in IndE than in BrE and AmE in reference to ‘cinemas’ in other text types or in spontaneous speech.¹²⁷

Another two arguably less transparent lexical similarities between IndE and BrE were uncovered in the ways the function words *towards* and *afterwards* are put to use. Nowadays, *towards* and *afterwards* are preferred to *toward* and *afterward* in India and Britain, while contemporary AmE shows the opposite preference pattern. See the competing variants in context:

- (162) First time author Vandana Majumdar has come out with a collection of short stories that not only deals with basic themes that have preoccupied women down the ages but one which also highlights male intolerance towards women in general. [IND2000 Press Rev 08]
- (163) Although the editorial notes that I as state parks director asked for a study of whether additional concessions can be granted without “violating our resources,” it seems to suggest that this department has a cavalier attitude toward parks resources. [US1992 Press Edit 06]
- (164) Soon afterwards, Shri Narayan Singh filed a civil suit in the munsif court in Munger against all three with the same charges. [IND2000 Press Rep 10]
- (165) Describing his “walk” afterward, Quayle said: “You’ll probably see it on TV. ...” [US1992 Press Rep 01]

A look at the one-million-word corpora of IndE, AmE and BrE suggests that the preference for the s-less variants in written AmE has become more pronounced over the years, especially in the case of *afterward*, cf. Table 3.71. IndE, on the other hand, was still closely aligned with BrE in this particular instance around the year 2000 (Google search dated December 28, 2003), cf. Table 3.72. But given the high recurrency rates for *toward* and *afterward*, those variants also have a foothold in BrE and IndE today as do the s-variants in AmE, which renders differences strictly quantitative (cf. Mair 1998: 144).¹²⁸

In spite of the relatively close affinity of IndE to BrE thus emerging from the corpus data, several lexical features were found to be shared by IndE and AmE but not by BrE, which goes to show that relations between and across varieties of English are never clear-cut. The compound noun *taxi stand* is common in India and the US, whereas BrE usually has the compound *taxi rank*. See the two competing variants in context, cf.

127. Bauer 2002 quotes *bioscope* as an example of lexical colonial lag in South Africa: “Lexical lag can be illustrated with the word *bioscope*, until recently the word for ‘cinema’ in South Africa, long after the word had vanished in Britain.” (6)

128. ICE-India: *towards/toward*: 273/7; *afterwards/afterward*: 32/5.

Table 3.71 *Toward(s)* and *afterward(s)* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
towards	337	316	293	64	17
toward	18	14	11	385	360
afterwards	26	76	54	14	6
afterward	2	1	-	16	21

Table 3.72 *Toward(s)* and *afterward(s)* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
towards	10100	6430	2670	14800	12600	1060	16400	807
toward	734	270	188	2840	590	43	1990	7370
afterwards	372	292	63	122	831	68	3450	297
afterward	36	23	8	24	44	1	124	1310

- (166) A taxi stand right in front of the Raikhy Cinema on the busy Cinema Road here, which is in existence for a long time now, is involved in a thick controversy with the cinema owner and the taxi operators making contradictory claims, the district and civic administration turning a deaf ear and the traffic police allegedly colluding with the taxi drivers.

[*The Tribune*; June 20, 2000¹²⁹]

- (167) It's hard to know what to expect when the lift glides towards the bottom and the doors open. But few would expect to see a Ford transit van. Two Ford transit vans in fact. A taxi rank at the centre of the Earth. This is no cramped underground workings.

[*The Guardian*; May 1, 2003¹³⁰]

Table 3.73 gives the frequency distributions of the two compounds in the online newspaper corpora from India, Britain and the United States (December 28, 2003).

Note that IndE also uses the combination *bus stand*, which may have come into being on the analogy of *taxi stand*, cf.

- (168) By the time he reached the Old Bus Stand area around 12 noon, a crowd of about 200 protesters gathered there.

[IND2000 Press Rep 11]

129. www.tribuneindia.com/2000/2000/06/21/ldh.htm, online December 28, 2003.

130. www.guardian.co.uk/life/feature/story/0,13026,946657,00.htm, online December 28, 2003.

Table 3.73 *Taxi stand* vs. *taxi rank*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
taxi stand	14	72	11	1	252	40	4	78
taxi rank	–	2	1	–	1	1	66	3

Table 3.74 *Bus stand* vs. *bus stop*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
bus stand	19	5660	2150	64	4790	250	10	8
bus stop	7	1500	490	28	616	162	940	1460

In BrE and AmE, *bus stand* is rare, *bus stop* being the more common choice (Google search dated December 28, 2003), cf. Table 3.74.

Another shared lexical feature of IndE and AmE is the preference for *garbage* over *rubbish*, the BrE-favored variant (also cf. Görlach 1995: 115–7):¹³¹

- (169) But a lot of the improvement was masked by a single-word element of the problem. Hawkers. Their litter, garbage and use of pavements for residence.
[IND2000 Press Rep 05]

- (170) The great pity is these people cannot be made to help clear the 260 tonnes of litter taken from motorways verges each week and 7,000 bags of rubbish that have to be removed from London's Underground stations.
[UK1991 Press Edit 04]

Table 3.75 (December 28, 2003) suggests that *garbage* is firmly established and preferred both in IndE and in AmE (cf. Appendix 3.42).

Likewise, IndE and AmE use the compounds *price hike* and *price rise* side by side whereas *price rise* is the preferred choice in BrE:¹³²

- (171) After all we should have been taken into confidence by the Prime Minister before the price hike.
[IND2000 Press Rep 02]
- (172) As regards external factors, oil importing developing country economies were badly hit by the 'shocks' of two sets of oil price rises, in 1973 and 1979.
[BrE FLOB J 43]

131. ICE-India: *garbage/rubbish*: 28/2.

132. ICE-India: *price rise/price hike*: 7/1.

Table 3.75 *Rubbish vs. garbage.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
rubbish disposal	1	1	–	–	–	1	26	1
garbage disposal	45	529	88	8	284	48	10	229
rubbish dump	–	18	2	2	–	5	155	4
garbage dump	7	299	64	7	283	90	56	174

Table 3.76 *Price rise vs. price hike.*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
the price rise	359	112	20	48	488	28	453	6
the price hike	330	353	40	250	585	56	46	9

See again the distribution of the two variants in the online newspapers available on the World Wide Web (December 30, 2003) for empirical evidence (Table 3.76).

A final similarity uncovered between IndE and AmE is observed in the shared liking among IndE and AmE users for the combination *full fledged*, whereas BrE users normally go for *fully fledged*, cf.¹³³

(173) This is evident from the support it has got from the Gujar's United Front that seeks the separation of the Jammu region from Kashmir and its conversion into a full-fledged Duggar and Gujar state within India and under the Union Constitution. [IND2000 Press Edit 09]

(174) As a result of cost cutting measures, some plans for further fully fledged BBC local radio stations have been abandoned in favour of providing a limited output (up to six hours per day) of local news and events within the programming of existing local stations and under their management.

[BrE FLOB G 76]

Table 3.77 provides the frequency distributions of the two variants across the online corpora from India, Britain and the US (December 30, 2003).

What exactly are the reasons behind those similarities between IndE and AmE is difficult to reconstruct in hindsight, and more diachronically sensitive research will be needed to reveal which of the observed distributions are the result of the postcolonial force of AmE (with IndE having moved away from its parent variety) and which are due to divergent developments in BrE throughout the twentieth century. Both factors

133. ICE-India: *full fledged/fully fledged*: 5/0.

Table 3.77 *Fully fledged* vs. *full-fledged*.

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
fully fledged	5	3	2	5	7	–	202	4
full fledged	412	338	124	541	718	76	40	121

Table 3.78 *Okay*, *I guess* and *gonna*.

	IndE IND2000 Broad	BrE UK1991 Broad
okay	48	–
I think	138	322
I guess	8	3
I suppose	6	20
going to	55	120
gonna	25	–

may, in fact, be responsible for creating the variational profile observable today. The very existence of such similarities is a reminder that IndE, though an offspring of BrE, does not automatically share all usage preferences with its parent.

Let me note in this context that the distributions of several other ‘classic’ Americanisms and Britishisms in the IndE corpus data defied a straightforward interpretation in terms of any clear-cut usage preference. As regards the word pairs *lorry/truck*, *flat/apartment* and *holidays/vacation*, for example, usage appears to be highly variable on a global scale today, which is suggested by the fact that neither IndE nor BrE or AmE shows any obvious preference in the online newspaper corpora – a finding which corresponds to observations made by Benson et al. (1986: 34–5) two decades ago (cf. Appendices 3.43 and 3.44). It appears that what used to be more or less pronounced regional lexical differences between BrE and AmE in those cases has been in the process of evaporating over the years, possibly under the contemporary influence from AmE on a global scale. There are also some indirect hints in the corpus data pointing to a more forceful intrusion of AmE-based forms into spontaneous IndE speech. See, for instance, the comparatively frequent use of the agreement marker *okay*, the discourse marker *I guess* and the operator *gonna* in the IndE broadcast data vis-à-vis BrE, cf. Table 3.78. But the extent to which we are witnessing the spread of AmE-typical conversational language in those particular instances will have to be pinpointed more succinctly once more comparable spoken data from AmE becomes available.

To conclude, although some mild influence from AmE is felt in IndE journalistic writing, in the broadcast sections and in the student essays today, the impact of AmE

on those particular text categories is moderate and the affinity of contemporary IndE to its parent variety BrE still noticeable, most obviously in the area of spelling. That said, the analysis has produced evidence that IndE and AmE also share several lexical features that are much rarer in BrE (for example *taxi stand*, *price hike*, *garbage*, and *full fledged*). Whether or not those differences between IndE/AmE and BrE are due to the postcolonial force of AmE on IndE or the result of divergent diachronic developments remains to be seen. Future research will probably also find more obvious traces of the AmE influence on IndE in other less formal and less traditional text types and registers of IndE, for example Netspeak or youth slang. Within the usage range investigated here, however, IndE is still broadly in keeping with the BrE model.

3.5 Chapter summary

Even though the overall proportion of variety-specific lexical features in the IndE corpus data must not be overrated, there is plenty of evidence to show that users of IndE draw freely on the possibilities of borrowing, word formation and semantic change to expand their communicative possibilities and innovate their vocabulary according to their own expressive and referential needs. The Primary Corpus has been helpful in establishing that loanwords tend to cluster in texts communicating aspects that are specific to India's socio-cultural reality. In India's quality newspapers, loanwords are found predominantly in press reviews, where they may be employed in formal and informal contexts, for example in critiques of classical concerts or in reviews of Indian movies. In such environments, markedly casual and creative uses of loanwords were slightly more common at the turn of the millenium than twenty-two years before, which might indicate an increasingly more relaxed and more self-confident attitude on the part of IndE quality newspapers toward adopting variety-specific lexical items. In IndE news reportage, on the other hand, loanwords are still relatively rare, which might reflect a special concern with supraregional (and international) intelligibility on the part of India's national newspapers and national broadcasting companies.

Complementing the Primary Corpus, the Kolhapur Corpus, ICE-India and the online corpora have been useful in describing the availability and status of numerous individual lexical items more context-sensitively than previous lexicographical research had accomplished. Evidence from the online corpora suggests that several lexical items are limited in currency to regions within India. The Hindi loanword *Vidhan Sabha* and the fairly young coinage *timepass* were found to be more common in Hindi-speaking territories around the year 2000, whereas the compound *petrol bunk* occurred in high numbers almost exclusively in newspapers from India's south. On a pan-South-Asian scale, the items *lathicharge*, *bandh*, *dharna*, *rasta roko* and the hybrid compound *road roko* as well as the coinages *speedmoney*, *hydel*, *senti*, *enthu* and *hi-fi* 'posh' were found to be more common in India's quality newspapers than in the Pakistani, Sri Lankan and Bangladeshi newspapers consulted. In contrast, the loanwords *hartal* and *gherao*, the coinages *eve-teasing*, *upgradation*, *delink*, *chargesheet*,

incharge, *undertrial*, *funda*, *sympathy wave*, *month-end*, *January-end*, *parliament session*, *telephonic conversation*, *departmental store* and uses of *meet* to denote 'political meetings' are nowadays employed in journalistic writing across South Asia and should therefore be considered lexical South Asianisms.

The analysis of the online newspaper corpora further suggests stylistic differences among India's quality newspapers as a source of variation and differentiation in written IndE. For example, while *The Hindu* (Chennai) predominantly uses the solid spelling *lathicharge* and the hybrid compound *road roko* to talk about the blocking of roads, *The Tribune* (Chandigarh) prefers the hyphenated spelling *lathi-charge* and the all-English compound *road blockade*. Given the findings presented in this chapter, *The Tribune* appears to be slightly more conservative in adopting hybrid compounds than *The Hindu*. On the other hand, *The Hindu* is less advanced than *The Tribune* (and also the other four quality newspapers consulted) in omitting the plural *-s* in currency statements like *Rs 100 lakhs* and *Rs 100 crores*, which exemplifies a dated usage by local IndE usage norms.

Numerous lexical differences between contemporary IndE, BrE and AmE discussed in this chapter are not due to IndE-internal innovation but must be ascribed to divergent diachronic developments in Britain and America. *Upliftment* 'uplift', common today in India and South Africa but virtually absent from BrE and AmE, is a case in point. According to the OED, *upliftment* used to be more frequent in the prestige varieties in the early twentieth century but has, by now, gone out of use while staying on in the Indian subcontinent and in South Africa. Similarly, *lecturership* 'lectureship', *dentenu* 'detainee', *to demit* 'to console', *untoward* 'unexpected' and *thrice* 'three times' are nowadays employed in relatively high numbers in the IndE texts investigated here while being much rarer in BrE and AmE. In this regard, the study of the function words *while/whilst*, *among/amongst*, *amid/amidst* and *till/until* has been particularly insightful for the purpose of documenting the workings of linguistic change in IndE vis-à-vis BrE and AmE. Contemporary IndE differs from the international prestige varieties in its pronounced preferences for *amidst* (over *amid*) and *till* (over *until*), both of which represent developmental stages of English that have been superseded in the prestige varieties. On the other hand, IndE has been more radical than BrE in dropping *whilst*, which is even less common in AmE than in IndE today. While AmE has also dropped *amongst*, this variant continues as a minority variant both in BrE and in IndE. At present, AmE and, to a lesser extent, BrE are diverging from IndE by replacing *inflammable* with *flammable*.

When scanning the closed corpora for uses of lexical Americanisms and Britishisms in IndE, the Indian data showed numerous similarities with BrE in the usage range analyzed. IndE shows great affinity to BrE in the use of the lexemes *pavement*, *working day*, *railroad* and *public transport* and also in the area of spelling, which has been shown to be largely in keeping with the British model to date. Current influence from AmE on IndE is felt in the Primary Corpus in the spread of contractions in India's quality newspapers, the occasional use of AmE-based spellings in writing,

especially in texts representing learned and scientific writing, and mild traces of informal AmE conversational language in speech. IndE and AmE are nowadays also alike in using *price hike* alongside *price rise*, *taxi stand*, *garbage*, *full-fledged* and *alternate* 'alternative' whereas BrE prefers *price rise*, *taxi rank*, *rubbish* and *fully-fledged* and largely adheres to the semantic distinction between *alternative* and *alternate*. While the postcolonial force of AmE is a likely factor reinforcing the use of American-based words in IndE, those lexical similarities between IndE and AmE might also have come into being as a consequence of divergent diachronic developments in BrE, with contemporary AmE acting as an additional conservative force.

Using the closed corpora and the online corpora side-by-side has also resulted in putting numerous other previous claims on the IndE lexicon into place. IndE, for example, does not automatically give preference to more formal lexical items but is stylistically as varied as other varieties, too, can be expected to be. In the usage range under inspection, IndE does not prefer *ancient* to *old*, *demise* to *death*, *bosom* to *chest*, *comely* to *pretty*, or *resplendent* to *dazzling*, and any description of IndE limiting itself to such stereotypical assessments runs danger of misrepresenting IndE usage grossly. Likewise, the confusion of near-but-not-identical words such as *tense* and *tensed* is rare and should not be considered typical of IndE usage at large. Moreover, there is no proof in the available data that IndE generally uses nominal or adjectival compounds more frequently within the standard usage range than BrE or AmE, although IndE has been shown to be capable of carving out preferences for individual compounds (e.g. *telephonic conversation*, *parliament session* or *sympathy wave*). Nor is there any convincing empirical evidence for the claim that derivation using Latinate prefixes should be considered to be any more productive in IndE than in BrE or AmE. Since those word-formation patterns have seen a rise in popularity in quality newspapers worldwide, it has, of course, been difficult to uncover potential IndE-specific usage preferences in those areas. It is only in the case of copulative compound nouns and adjectives of the types N+cum+N and adj+cum+adj that IndE can be said to differ from BrE and AmE, although the quantitative difference observed is rather marginal.

CHAPTER 4

Lexicosyntax

Next to lexis, variation within IndE and across varieties has been claimed to be remarkably high in the area of lexicosyntax, and numerous items have been listed as characteristic lexicosyntactic features of IndE in previous descriptions of the variety, in particular particle verbs and verb-complementation patterns.¹ Since the status of most of those items in contemporary IndE usage was unclear, the Primary Corpus, the Kolhapur Corpus, ICE-India and the online databases were consulted to see whether a combination of quantitative and qualitative work would illuminate the ways in which contemporary IndE was variable and possibly also innovative in those areas around the turn of the millennium. This chapter looks at combinations of verbs with the particles *up*, *down*, *out*, *away* and *off* (4.1) and puts several more hypotheses on the use of multi-item words and phrases (4.2) as well as complementation patterns (4.3) to test.

4.1 Particle verbs: Verbs combining with *up*, *down*, *out*, *away*, and *off*

From a learner's perspective, Yadurajan 2001 is certainly right by stating that "[p]ractically all the particles cause problems" (82). On the other hand, Sahgal/Agnihotri 1985 found that educated users of IndE were quite accepting toward variation in this area, which suggests that independent innovation may be particularly likely to happen in this field – all the more so as intelligibility rarely seems to be hampered significantly. According to Sahgal/Agnihotri 1985, "deviant complex constructions and deviations involving the use of particles and prepositions constitute two ends of the continuum of acceptability, the former being the least and the latter the most acceptable" (125–6). Nihalani et al. (1979, 2004) state that the particle *up* "is extremely frequently inserted after verbs by speakers of IVE [Indian variants of English] in contexts where it would be deemed superfluous in BS [British Standard English]" (1979: 186, 2004: 188). Moreover, "[in] contexts where the form without 'up' is equally as likely as

1. For references to lexicosyntactic variability in IndE cf. (in chronological order) Whitworth (1982 [1907]: 59–78), Spitzbardt (1976: 41–44), Verma (1978: 212–5, 1982: 183), Nihalani et al. (1979, 2004), Bansal (1983: 3–5), Labru (1984: 75–102), Pandharipande (1987: 150), Shastri (1988b: 24–5, 1996), Lukmani (1992: 160), Sridhar (1992: 142–5), Gramley/Pätzold (1992: 443), Leitner 1994, Parasher (1994 [1983]: 150–5), Hansen et al. (1996: 225), Yadurajan 2001, Rogers (2002: 196–7), Trudgill/Hannah (2002: 131–2), Melchers/Shaw (2003: 141), Mukherjee/Hoffmann 2006, and Hoffmann/Mukherjee 2007.

the form with it, IVE [Indian variants of English] speakers almost invariably prefer the latter" (1979: 186, 2004: 188). Lukmani 1992, too, observes that "*up* ... is frequently inserted where it is not required" (160). Beside *up*, also *down*, *out*, *away* and *off* have been said to combine with verbs rather freely in IndE (cf. Appendices 4.1–4.4). Hence, it was decided to start out the investigations into IndE lexicosyntax by looking at those five particles in context using the Primary Corpus as a starting point. Would those particles be found to combine differently with verbs in IndE than in BrE or AmE?

4.1.1 The distribution of unrecorded particle verbs in the Primary Corpus

After extracting all instances of *up*, *down*, *out*, *away* and *off* from the Primary Corpus with the help of the WordSmith tool and identifying all verb-particle combinations among those tokens manually in the next step, phrasal, prepositional and phrasal-prepositional verbs listed in the corpus-based COB dictionary were excluded from further consideration. To identify those, the corpus material was scrutinized carefully to uncover potential differences between the codified standard and contemporary IndE in form and/or function. This was not a straightforward task, as differences in meaning between verbs and their related verb-particle combinations (e.g. *find* and *find out*) were sometimes not obvious. According to my reading of the data, only thirty tokens of a total of 1,027 tokens ended up not straightforwardly complying in form or function with the codified norms. To pinpoint the status of those unrecorded combinations more succinctly, the Kolhapur Corpus and ICE-India usually turned out to be too small, and additional snapshots from the Internet were taken. Before looking at the individual unrecorded particle verbs in more detail, let us consider their distribution in the Primary Corpus.

Table 4.1 gives the token frequencies of the unrecorded verb-particle combinations identified in the Primary Corpus. Considering that only about three percent of all combinations with *up*, *down*, *away*, *off* and *out* qualify as unrecorded cases, one is tempted to say that variability in particle usage is rather a marginal feature in the usage range under investigation, and it would be misleading to speak of it as a defining characteristic feature of educated IndE at large – even more so as particle verbs, according to Schneider (2004: 234), generally appear to be less frequent in IndE than elsewhere in the English-speaking world.² In fact, no unrecorded particle

2. Using the ICE corpora representing British, Singaporean, Philippine, Indian and East African English (both Kenyan and Tanzanian), Schneider 2004 found the propensity to use intransitive particle verbs to be "very high in Singapore (at 68 occurrences per one million words), where PVs [particle verbs] are used even more frequently than in Great Britain (55), while, conversely, it is lower everywhere else and very low in Tanzania (29), India (26) and the Philippines (35), especially when compared to the other written data." (235) Interestingly, Schneider 2004 also observed for IndE and East African Englishes that particle verbs occur relatively more frequently in writing than in speech whereas the reverse is true for BrE and SgE. Cf. Schneider (2004: 235–6).

Table 4.1 Unrecorded particle verbs with *up*, *down*, *away*, *off* and *out*.*

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad News	IndE IND2000 Broad I/D	IndE IND2000 Ess
away	– (21)	– (30)	– (9)	– (8)	– (13)
down	– (39)	– (39)	– (19)	1 (10)	2 (9)
off	– (23)	– (28)	– (9)	– (3)	– (1)
out	4 (100)	6 (119)	2 (58)	4 (70)	2 (16)
up	1 (122)	– (115)	1 (104)	– (58)	7 (14)
	40 texts 5 tokens	40 texts 6 tokens	20 texts 3 tokens	20 texts 4 tokens	10 texts 11 tokens
Tokens per text	0.13	0.15	0.15	0.25	1.10
Tokens per 1,000 words	0.06	0.08	0.08	0.10	0.55

* Note: figures in parentheses are total numbers of particle verbs with respective particle

verbs with *away* and *off* were found in the data, which might primarily be due to the relatively limited size of the corpus and the comparatively low proportion of those two particles. It was only for *out*, *up* and *down* that unrecorded combinations occurred at all.³ On the other hand, the stratification of the few tokens is noteworthy. When comparing the text categories, the proportion of unrecorded combinations is highest in the category representing nonprofessional writing (student essays; 1.10 tokens per text), followed by the categories of spontaneous speech (broadcast discussions and interviews; 0.25 tokens per text), scripted speech (broadcast news; 0.15 tokens per text), and, finally, professional writing (journalistic prose; 0.15 tokens per text for the year 2000 and 0.13 tokens per text for 1978). In other words, within the usage range covered by the corpus data, IndE usage is slightly more off

3. According to Shastri (1988b:24–5), approximately one third (27.65%) of all verb-particle combinations with *up* and *down* show “peculiar features” in the Kolhapur Corpus, which suggests that IndE writing may well be more variable beyond the confines of the text category of quality newspaper prose in less heavily edited texts. However, Shastri 1988b gives no precise definition of what exactly is subsumed under the label “peculiar features,” and several of the examples listed in his paper appear to conform to international usage standards (e.g. *He beats up his children for no reason*). Shastri 1988b himself concludes on a careful note by stating that “[o]ur observations regarding the peculiarities of IndE usage may well be within the creative mechanism of the language.” (25) Also cf. Shastri (1996:72).

the codified norm in educated speech than in journalistic writing and clearly in a class by itself in the category of student writing. This suggests text type, mode and linguistic competence as possible influential factors in creating variability. While the difference between writing and speech could reflect the variant degrees of planning and editing underlying the production of the respective texts, the notably higher proportion of unrecorded verb-particle combinations in the essays points to learner difficulties in acquiring the conventions of using English phrasal, prepositional and phrasal-prepositional verbs – a particularly tricky area to master for learners of English around the globe.⁴ But independent innovation might also be a reason for unrecorded combinations to feature more prominently in less heavily edited writing than journalistic writing, which is why I decided to take a closer look at the unrecorded particle verbs in a next step to pinpoint their status and separate stable innovative forms from unstable ones.

4.1.2 Pinpointing the status of unrecorded verb-particle combinations

When cross-checking the unrecorded verb-particle combinations with the wider World Wide Web, only a few of the combinations detected in the Primary Corpus qualified as stable features of contemporary IndE and even fewer as independent South Asian innovations. Consider first examples (1) to (5), which represent ad-hoc combinations of verbs and particles that according to my reading of the data should not be considered to be distinctive features of IndE:

- (1) We have to throw our old traditions and have to make an identity of our own which depicts out qualities and values adapting with the world of today and not to blindly follow our traditions as it is granted to us by our forefathers.
[IND2000 Ess 07]
- (2) Every parents tries to fulfil the wishes of their children. They sacrifice their own wishes for the sake of their own child and spend all their income in fulfilling up the wishes of their child.
[IND2000 Ess 01]
- (3) Now it's for Dasmunshi to speak his mind out. [IND2000 Press Rep 17]
- (4) <#>uh my question to the honourable panel members<,> even one year after the Kargil operation<,> we have not yet been able to know<,> why we had to fought the war<,> <#>who are responsible for that<,> and why have no heads rolled down
[IND2000 Broad Disc 02]

4. For treatments of particle verbs in IndE textbooks cf. Mason (1978:74–82), Dhillon (1998:52–60), Wren et al. (2001:136–48).

- (5) Mr. Narayanan would agree that the Council of Ministers is the President's Council of Ministers. In the light of this, one fails to see logic in his unwarranted advice – “we should not throw away the baby with the bath-water”.
[IND2000 Press Edit 07]

In example (1), which was found in a student essay, the student probably confused the verbs *pick* and *depict*, producing the combination *depict out*, which is not common in IndE but simply a mistake of a type likely encountered in learner Englishes around the world. In example (2), the choice of the particle *up* after the verb *fulfill* could be said to function as an additional aspectual marker by adding a sense of completion to the action denoted, or it may simply have been triggered on the analogy of the prepositional verb *fill up*. In the third and fourth examples, users added the particles *out* and *down* to two set phrases, namely *speak your mind* and *heads will roll*, which is in contrast to the usage conventions codified in COB. In example (5), which is taken from a press editorial, an inappropriate particle has been used, the standard expression in IndE continuing to be *to throw out the baby with the bath-water*. Since the phrase occurs in a passage of quoted speech, it is possible that we are dealing with a written representation of a user's slip of the tongue. Note that none of the collocations has so far had any significant impact on general IndE usage, at least not in writing. This is suggested by a look at the IndE online corpora, which provide no evidence that the combinations in question are used recurrently in published writing today (cf. Appendix 4.5.). It can therefore be concluded that the examples above, more than anything, reflect the insecurity on the part of individual users concerning the conventions of English particle-verb usage.

Generally speaking, the difficulty of differentiating semantically between lexical verbs on the one hand and related particle verbs on the other is at the core of many of the unrecorded uses in the Primary Corpus. For instance, the lexical verbs *dress*, *lift*, *grow* and *take* would arguably have been more appropriate in the next set of examples than the verb-particle combinations employed, although it must be acknowledged that differences in meaning between the two choices are sometimes only minute, cf.

- (6) Our traditions include our religion, our customs, beliefs, the way we dress up, the way we speak, respect our elders and treat small children.
[IND2000 Ess 07]
- (7) Why is it that the bans imposed on India after Nuclear Tests are lifted up but are still over Pakistan?
[IND2000 Ess 07]
- (8) Another factor which contributes towards corruption is that every person wants his work done in a very short span. And the opponent get [sic] the

work done in stipulated period in exchange of a few amount of money [sic]. This way the precious time and unnecessary bureaucratic hurdle that come [sic] in the way of life is solved. Consequently, the corruption grows up. [IND2000 Ess 03]

- (9) He's alone responsible for every decision he takes up in life.

[IND2000 Ess 01]

Nihalani et al. (1979:80, 2004:79–80) observe a similar blurring of semantic distinctions in the use of the verb *find* 'to see someone or something and learn where they are' and its companion *find out* 'to learn something you did not already know, especially by making a deliberate effort'. The results of my investigations suggest that the two forms are indeed conflated occasionally in IndE writing and speech, cf.

- (10) Firstly, people can not [sic] find out enough time for physical relaxation. The time table, or the daily schedule has been prepared in such a way, that there is not enough time to think about anything. [IND2000 Ess 10]

- (11) "So, not only Indian Olympic Association [IOA] but districts and state associations should also try to involve the youths in sports as much as they can. They should conduct proper coaching camps to find out the talents," she said. [IND2000 Press Rep 11]

- (12) <\$F><#>see basically<,> the problem today is<,> that the<,> citizens of Sri Lanka of Tamil origin<,> they have been subjected to a lot of uh problems in their own country <#>they are treated as a second-class citizens<,> <#>so<,> a solution has to be found out for that [IND2000 Broad Disc 10]

This has also been observed by Schneider 2004, who noticed *find out* to be common "in Singapore and spoken English in India" (238) in ICE. But it is important to keep in mind that IndE maintains the semantic difference between *find* and *find out* far more often than it does not, at least in educated writing. This is suggested when looking at India's national newspapers on the Internet. The distributions of the collocations *find (out) a solution* across India's electronic quality newspapers (Google search dated June 12, 2002) signal that we might be dealing with a recurrent but overall rare phenomenon in IndE, cf. Table 4.2. Note that the conflation of the two forms is not restricted to IndE alone, but common also across South Asia (June 12, 2002), cf. Table 4.3. Yet still, when comparing the frequencies of the two collocations across the top-level country domains (June 12, 2002), it is striking to see that *find out* collocates with the noun *solution* more frequently under the 'in' domain than elsewhere, cf. Table 4.4.

More research will be necessary to explain whether the collocation *find (out) a solution* is representative of uses of *find* and *find out* at large and, if that should be the case, why the conflation of the two forms is more recurrent in IndE than elsewhere. In light of the evidence above, it can be established that the feature described in Nihalani

Table 4.2 *Find (out) a solution* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
find a solution	268	289	87	71	514	118	277	134
find out a solution	9	7	2	–	23	3	–	–

Table 4.3 *Find (out) a solution* [Internet: PakE, BaE and SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
find a solution	358	27	41
find out a solution	36	4	1

Table 4.4 *Find (out) a solution* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
find a solution	952	690	111000	6650	25900	11600
find out a solution	58	2	39	3	27	11
(out) : out	16 : 1	345 : 1	2846 : 1	2217 : 1	959 : 1	1055 : 1

et al. (1979:80, 2004:79–80) to date remains a minority variant in the usage range under investigation.

Another such recurrent but rare collocation is *lower down*, which has been pointed out by IndE style critics before. According to Nihalani et al. (1979:117, 2004:117), the combination *lower down* is rare in BrE and AmE but common in IndE, where it is said to collocate with nouns referring to abstract concepts and ideas. In the set of unrecorded verb-particle combinations from the Primary Corpus, one such example can be found, cf.

- (13) Under the influence of this vulgar western culture boys go to the pubs and discos. They dance with the girls, drink most expensive wines, commit various crimes and then intercede their parents to give money for their release from the jail. These debasing habits not only lower down their moral but also their parents [sic] moral. [IND2000 Ess 04]

A look at the online corpora from India (June 11, 2002) again helps put Nihalani et al.'s (1979, 2004) claim into perspective. While this use may be more common in less heavily edited written IndE or in speech, it is rare in contemporary journalistic writing, for example in collocations of the type *lowering (down) the N*, cf. Table 4.5. The feature is also occasionally found in PakE journalistic writing. On the other hand, it is

Table 4.5 *Lower (down)* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
lowering the N	111	144	25	35	337	70	200	288
lowering down the N	1	–	–	–	6	–	–	–

Table 4.6 *Lower (down)* [Internet: PakE, BaE and SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
lowering the N	259	19	8
lowering down the N	6	–	–

not attested in the two corpora from Bangladesh and Sri Lanka, which are comparatively small (June 11, 2002), cf. Table 4.6. When browsing the Internet (June 11, 2002), *lower down* turns out to be slightly more recurrent in the South and South East Asian regions than elsewhere, but the conflation of the two competing variants appears to be not as common in IndE as that for *find (out)*, cf. Table 4.7. However, even if IndE and SgE are indeed more variable than the other varieties in that particular instance, the numerical evidence from the World Wide Web remains difficult to interpret precisely because we are dealing with a marginal feature.⁵

Of the thirty unrecorded verb-particle tokens uncovered in the Primary Corpus, only three combinations qualify as stable particle verbs used in IndE in a characteristic fashion vis-à-vis BrE and AmE, namely *chart out*, *chalk out* and *take out*. Consider, first, examples (14), (15) and (16) of *chart out* ‘outline’ from IndE press texts:⁶

- (14) The steering groups used to chart out their activities within the framework of policy guidelines provided by the apex Indo-U.S. Defence Policy Group [DPG]. [IND2000 Press Rep 01]
- (15) An I&B ministry paper has charted out regulations and the three possible options if DTH is cleared. [IND2000 Press Rep 02]
- (16) Regardless of whatever decision is taken or not taken – by the Nuclear Regulatory Commission in Washington on the supply of fuel for Tarapur, New Delhi will have to reappraise its atomic policy and chart out a clear course. [IND1978 Press Rep 05]

5. ICE-India: *lower/lower down*: 23/0.

6. Neither *chart* ‘outline’ nor *chart out* ‘outline’ is attested in ICE-India.

Table 4.7 Lower (down) [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
lowering the N	3460	663	89500	9740	43700	44900
lowering down the N	14	3	26	–	9	2
(down) : down	247 : 1	221 : 1	3442 : 1	–	4856 : 1	22450 : 1

The OED informs us that *chart out* ‘delineate, outline’ is probably not an independent IndE innovation, but a combination originating in the nineteenth century, the first attestation of its figurative use dating back to 1842. It appears that *chart out* has gone out of general use in Britain and America, where the shorter *chart* is nowadays more common. Accordingly, *chart out* cannot be found in contemporary dictionaries of English, for example COB. In IndE, on the other hand, the combination has stayed on, which is illustrated in Table 4.8 (Google search dated August 3, 2002). The particle verb *chart out* is also available in South Asian journalistic prose (August 3, 2002), but the figures returned from the Internet are too small to allow for any more precise insights into possible regional differences, cf. Table 4.9. That *chart out* has a particularly strong foothold in South Asia today is also suggested when examining the use of the combination across the top-level country domains on the World Wide Web (August 3, 2002), cf. Table 4.10. When comparing the ratios in Table 4.10, South and South East Asian usage is in a class of its own, with writing published under the ‘.in’ and ‘.sg’ domains employing *chart out* more recurrently than happens in the other varieties. In this light, the phrasal verb qualifies as a South Asianism, but only in quantitative terms and from a synchronic viewpoint.

The combination *chalk out*, which is synonymous with *chart out* and also no longer recorded in the major dictionaries, falls in the same category. See the following two examples from IndE news broadcast:

- (17) <\$A><#>the party urged the government to chalk out an action plan to modernise all the ammunition depots in the country [IND2000 Broad News 07]
- (18) <\$A><#> The Congress-I spokesman Mr Pranab Mukherjee said that the Working Committee of the party will meet this morning to chalk out the funeral arrangements [IndE ICE-India S2B020]

Chalk out is also found recurrently in IndE writing, for example in press reportage (example (19)), business letters (example (20)) or academic writing (example (21)):

- (19) The whole show has been managed in such a way that a senior board member even chalked out who should initiate the discussion asking for leniency towards the players. [IND2000 Press Rep 17]

Table 4.8 *Chart out* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
charted out	48	58	10	5	65	23	–	–
charted the N	7	11	2	1	14	2	51	34
charted out the N	6	3	2	–	3	3	–	–
outlined	375	551	96	98	600	175	3390	983

Table 4.9 *Chart out* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
charted out	13	2	–
charted the N	3	3	–
charted out the N	–	–	–
outlined	656	134	73

Table 4.10 *Chart out* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
charted the N	44	76	6830	214	733	714
charted out the N	8	5	20	1	1	10
(out) : out	6 : 1	15 : 1	342 : 1	214 : 1	733 : 1	71 : 1

- (20) The Kendra could chalk out a course of training young researchers through direct participation, in these historic languages and scripts.

[IndE ICE-India W1B019]

- (21) Thoughts which would hitherto have been considered blasphemous, such as a free market economy, are now being openly, unabashedly and forcefully expressed and elaborate programmes chalked out for their speedy attainment.

[IndE ICE-India W2B012]

The OED informs us that *chalk out* is of Early Modern English origin, its first attestation in the OED dating back to the sixteenth century. Frequent in IndE quality broadsheets today, *chalk out* is conspicuously absent from that text category in BrE and AmE (Google search dated August 3, 2002), cf. Table 4.11. Next to IndE, PakE and BaE also use this combination, which, on the basis of this evidence, is best treated as a South Asianism (August 3, 2002), cf. Table 4.12. In the rest of the English-speaking world, *chalk out* appears to be on the wane (August 3, 2002), cf. Table 4.13. Note in passing that when *chalk out* is found outside the ‘in’ domain, it is usually in contexts

Table 4.11 *Chalk out* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
chalked out a plan	28	17	2	7	110	15	–	–
devised a plan	5	4	1	3	14	2	25	24

Table 4.12 *Chalk out* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
chalked out a plan	117	1	–
devised a plan	6	2	3

Table 4.13 *Chalk out* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
chalked out a plan	23	–	8	1	2	1
devised a plan	13	34	786	183	567	516
chalk out : devise	2 : 1	–	1 : 98	1 : 183	1 : 284	1 : 516

discussing India, which is further evidence that this particle verb is enjoying an active retirement only in the Indian subcontinent, cf.

- (22) India chalked out a plan for her economic growth in a protective manner.
[AmE⁷]
- (23) In the meanwhile, the dacoit leader called his men and chalked out a plan.
[AusE⁸]
- (24) At this Parma Nand, the prime minister of Ajmer Chand, chalked out a plan.
[BrE⁹]

Hence, *chalk out* stands as a contemporary South Asianism – but it is most likely not an independent Asian innovation.

Only when it comes to the verb-particle combination *take out* do we finally observe an innovation that has most probably originated in the Indian subcontinent. It is a semantic rather than lexicosyntactic innovation, though. Used in international

7. www.desperdes.us/culture3.html, online August 3, 2002.

8. www.mandala.com.au/chaitanya_bhagavat/ant5.html, online August 3, 2002.

9. www.sikh-history.co.uk/anandpur_sahib.html, online August 3, 2002.

English to say that someone leads someone else to a restaurant or theatre after inviting them, *take out* can be employed in IndE also to say that someone leads a group of people in a gathering, for example in a demonstration, procession or rally. See the following examples from the Primary Corpus:

- (25) According to the press release of the ABVP, about 100 university students took out a procession on the campus demanding the immediate and unconditional release of DUSU leaders. [IND1978 Press Rep 13]
- (26) Protesting against the lathicharge, the Congress MLAs – Mr. K. R. Suresh Reddy [Balkonda] and Mr. Bajireddi Goverdhan [Armoor] – took out a rally and marched to the Mamidipalli crossroads where they began a sit-in. [IND2000 Press Rep 11]
- (27) <#>in Pakistan an explosion at a rally of Islamic militants in southern Karachi<,> has killed three people and injured thirty<,> <#>the rally was taken out by the militants' group Lashkar-e-Taiba [IND2000 Broad News 04]

ICE-India, too, includes four tokens of *take out* in that particular sense, and a Google-based search query conducted on the Internet (July 17, 2002) returned numerous further examples of the collocation *took out a rally* from IndE journalism. These findings point to a fairly high degree of integration of this new extended sense in present-day IndE, cf. Table 4.14. While the collocation is not found in the BrE and AmE data, it is attested for PakE journalistic prose (also cf. Baumgardner 1993b: 47), but not for BaE or SrLE (July 17, 2002), cf. Table 4.15. That the use of *take out* in reference to rallies is nowadays limited to South Asia is further suggested by the absence of the collocation outside the 'in' domain on the Internet (July 17, 2002), cf. Table 4.16. Hence, the collocation *take out (a rally, a demonstration etc.)* is a strong contender for the status of a contemporary innovative semantic South Asianism.

To conclude, of the thirty unrecorded verb-particle combinations with *up*, *down*, *out*, *away* and *off*, only three can, from a synchronic point of view, be considered stable characteristic features of the IndE usage norm, which puts a question mark on any description of IndE seeing variability in the use of verb particles as a defining element of IndE usage at large. The corpus data show a more diversified picture of the linguistic situation in IndE: while IndE differs only moderately from international usage conventions in qualitative terms within the usage range studied (one example of a South Asian innovation being the use of *take out* in reference to political gatherings), quantitative differences between IndE and other varieties, especially the prestige varieties, are noticeable across text types and modes (e.g. *find out* 'find', *lower down* 'lower'). Divergent diachronic development, finally, appears to be the most probable reason for the relatively frequent use of *chart out* and *chalk out* in South Asian Englishes vis-à-vis BrE and AmE.

Table 4.14 *Take out* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
took out a rally	39	81	33	5	207	15	–	–

Table 4.15 *Take out* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
took out a rally	158	–	–

Table 4.16 *Take out* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
took out a rally	3	–	–	–	–	–

4.2 More variation in the use of particle verbs

This chapter zooms in on a further set of particle and multi-item verbs and phrases that have been pointed out as potential lexicosyntactic Indianisms. At the outset of the analysis, I compiled a list of 97 combinations of multi-item verbs and phrases which had been discussed in reference to IndE before (Whitworth 1982 [1907]; Nihalani et al. 1979; Yadurajan 2001) but which were not listed in COB (cf. Appendix 4.6). On scanning the Primary Corpus for examples of those combinations in use, only seventeen turned up in the database – a result to be expected considering the modest size of the text collection and one which was in line with my earlier investigations into verb-particle combinations with *up*, *down*, *out*, *away* and *off* (cf. Chapter 4.1). To test whether any of the attested combinations had gone down the path of integration in IndE over the years and whether they would qualify as South Asianisms or even Indianisms, the double-layered approach of using the structured and the online corpora side-by-side was employed again to describe the status and the currency of those combinations.

4.2.1 Distribution in the Primary Corpus

Let us first take a look at the stratification of the seventeen marked-up ‘different’ combinations in the Primary Corpus. Table 4.17 gives the token frequencies of the individual combinations across the text categories represented in the Primary Corpus.

Table 4.17 ‘Different’ particles [Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess
aspire for NP	–	–	1	–
be admitted in NP	–	–	1	1
get/seek admission in NP	–	–	–	1
take the law in your hands	–	–	1	–
translate NP in NP	–	1	–	–
call NP as	1	–	–	3
comprise of NP (active)	–	1	1	1
discuss about NP	–	–	3	–
emphasise on/upon NP	–	1	–	–
investigate into NP	–	–	–	1
rename NP as	1	–	1	–
request for NP	–	–	1	–
stress on NP	–	4	–	–
term NP as	–	2	–	–
urge for NP	–	1	–	–
avail NP	–	1	–	–
bunk NP	–	–	1	–
	40 texts	40 texts	40 texts	10 texts
	2 tokens	11 tokens	10 tokens	7 tokens
	2 types	7 types	8 types	5 types
Tokens per text	0.05	0.28	0.25	0.70
Tokens per 1,000 words	0.03	0.14	0.13	0.35
Types per text	0.05	0.18	0.20	0.50
Types per 1,000 words	0.03	0.09	0.10	0.25

Although the token frequencies in Table 4.17 are low and must be interpreted with caution, the distribution shows a similar patterning of unrecorded combinations across the sub-sections of the Primary Corpus as Table 4.1 did (cf. Chapter 4.1). Again, it is the category of student writing that is characterized by the highest token ratios of unrecorded combinations, which suggests variant degrees of planning and editing as well as variant degrees of competence as possible determining factors for the extent to which contemporary IndE may be variable. Since the use of particle verbs is oftentimes simply a matter of convention, the comparatively high degree of variability in the student essays is not unexpected, as learners of English are likely to be less familiar with the conventions governing the use of multi-item verbs and phrases. While there is no dramatic difference between educated speech and educated writing around the year 2000, it is interesting that the press material from the year 2000 is once again slightly more variable than the material from 1978. Although one

must hesitate to speak of a sweeping diachronic development in this instance, the observation is in line with research findings presented elsewhere in this book. It appears that Indian press English has been opening up to forms not part of standard English over the years.¹⁰

What kinds of ‘unrecorded’ uses are we confronted with exactly? Of the seventeen combinations that had come to my notice, five differed from the codified standard in that users of IndE had chosen different particles than those listed in COB. In ten cases, IndE featured particles after verbs that would usually take none according to COB. The remaining two cases involved the use of verbs without particles in contexts that would usually have particles by international standards. Next, the available data are examined in more detail to see whether those uses are typical of contemporary IndE.

4.2.2 Use of ‘different’ particles

Consider first examples (28) to (32), which illustrate choices of prepositions that must be considered unusual by standard English:

- (28) <\$B><#>an early summer this year has led to a rise in cases of gastro enteritis in Ahmedabad<,> <#>so far one thousand three hundred and thirty-five cases have been reported<,> with an average of thirty-five to forty cases<,> being admitted in hospitals in the city each day [IND2000 Broad News 16]
- (29) But, the life of an adult is full of tensions like getting admission in schools, afterwards burden of examinations etc. [IND2000 Ess 01]
- (30) He credits much of what he has achieved to his gentle, soft-spoken wife, Shamasree Lal, grand-daughter of Ramananda Chatterjee who was the editor of the much famed and fiery Prabasi, Modern Review, and daughter of noted historian Kalidas Nag. A constant companion to him since his college days, she has co-translated in English with her husband the Last Poem of Tagore. [IND2000 Press Rev 03]
- (31) <#>after the break the elephant menace in villages near Bangalore where farmers are taking the law in their hands <#>stay with us [IND2000 Broad News 17]
- (32) <#>Mr Maran said the government aspired for an annual trade target<,> of at least twenty-five billion dollars in the next five years [IND2000 Broad News 04]

10. Cf. the results on informal loan vocabulary (Chapter 3.1.1), unrecorded count uses of non-count nouns (Chapter 5.2), nonstandard past-perfect forms (Chapter 5.4) and nonstandard direct and indirect questions (Chapter 5.7).

In examples (28) and (29), standard English would usually have *to* in place of *in* after the verb *admit* and the noun *admission*.¹¹ In example (30), the verb *co-translate* would be followed by *into*,¹² and *into* would also be used in the idiomatic phrase *to take the law into one's (own) hand* in example (31).¹³ Moreover, in example (32), *to* would be regarded as the standard particle following the verb *aspire*.¹⁴ When looking up those five combinations in the electronic newspaper corpora in selected collocations (October 16, 2002), only one turned out to be a common choice in IndE journalism (*aspire for*), three were recurrent but rather rare (*translate in*; *admit in*; *take the law in one's hand*), and one did not appear at all (*admission in*). See the distributions of the collocations collected in Table 4.18 in detail.

Since the collocations chosen for this Internet search query do not represent the entire functional range of the combinations discussed, the explanatory potential of the distributions above is limited. However, it is fair to say that IndE journalistic writing is slightly more variable in its choice of prepositions than BrE and AmE in those instances. While BrE and AmE newspapers, too, show some variability going beyond the norms codified in COB, the four nonstandard instances from the BrE and AmE newspaper corpora appear to be exceptions representing the fuzzy edges of natural language use when weighed against the well over 200 standard instances found in the online *Guardian* and *Washington Post*. In the IndE newspapers, on the other hand, the combination *aspire for* appears to be firmly established, possibly on the analogy of the noun-particle collocation *aspiration for* and the synonymous verb *hope for*, which are common the English-speaking world over. The combinations *to translate in*, *to take the law in one's (own) hand*, *to be admitted in*, and *admission in* are established to lesser degrees in IndE journalistic writing than *aspire for* but are also found repeatedly across the corpora.

Another look at the Internet informs us that none of the 'different' combinations is used exclusively in IndE nowadays. All of them are, for instance, also attested in Pakistan's quality broadsheet *The Dawn* (cf. Appendix 4.16). And when searching the Internet more widely across the top-level country domains (October 16, 2002), the collocations are found elsewhere in the English-speaking world, too. Table 4.19 gives the recurrency rates of the six collocations across the top-level country domains (also cf. Appendix 4.8 for the absolute frequencies of the collocations).

11. According to COB, *admit* is typically used in the patterns *admit N*, *admit that*, *admit to N/ing*, *admit -ing*, and *admit* plus quotation. The expressions *to admit s.o. to hospital* and *to admit s.o. to an organization, group* are said usually to occur in the passive. The pattern given for *admission* is *admission to N*.

12. The patterns recorded in COB are *to translate N from/into N* and *to translate as* (ergative use).

13. The only pattern recorded in COB is *to take the law into one's own hands*.

14. The patterns recorded in COB are *to aspire to N/ing*; *to aspire to-inf*.

Table 4.18 'Different' particles [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
admitted to hospital	76	127	26	44	470	101	228	5
admitted in hospital	2	3	4	9	22	6	–	–
admitted into hospital	1	–	3	–	–	1	2	1
admission to unversity	7	2	1	1	–	1	2	1
admission in university	–	–	–	–	–	–	–	–
admission into university	2	–	–	–	5	–	–	–
was translated into (+ lg.)	4	15	6	1	33	2	24	10
was translated in (+ lg.)	2	–	1	–	3	–	–	–
took the law into	5	8	–	–	9	2	4	5
took the law in	1	–	–	–	6	1	–	–
aspired to the	4	2	–	–	3	–	18	13
aspired for the	4	7	1	–	8	1	–	–

Table 4.19 'Different' particles [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
aspired to the :	77:1	3:1	46:1	16:1	91:1	31:1
aspired for the						
was translated into :	6:1	5:1	25:1	25:1	29:1	20:1
was translated in						
admission to :	3:1	54:1	28:1	–	–	–
admission in						
took the law into :	2:1	20:1	107:1	8:1	31:1	7:1
took the law in						
admitted to :	1:1	–	2278:1	114:1	1783 :1	82:1
admitted in hospital						

Although it is problematic to compare and interpret the figures in Table 4.19, as they were computed from absolute token frequencies varying greatly in size (which again is due to the tremendous differences in size among the individual top-level country domains), the search results do confirm what has already been shown for the register of newspaper journalism: written IndE usage is slightly more variable than BrE and AmE today. Four of the five collocations recur more frequently in IndE than in BrE or AmE. But note that IndE usage under the 'in' domain is, above all, characterized vis-à-vis the other territories by a comparatively high proportion of the collocation *admitted in (hospital)* – a collocation that was found to be rare in India's quality newspapers. On the other hand, the proportion of the collocation *aspired for the*, which is frequent in Indian journalistic writing, is comparatively low under

Table 4.20 *Admit, admission and translate.*

	IndE ICE-IND writing	IndE ICE-IND speech
admit N to (standard)	10	2
admit N in/into/for (nonstandard)	3	11
admission to (standard)	9	1
admission in (nonstandard)	5	8
translate (lg.) into (lg.) (standard)	7	10
translate (lg.) to/in (lg.) (nonstandard)	4	3

the ‘in’ domain. It thus appears that text-type and register-specific constraints are at work in IndE allowing or sanctioning the use of unrecorded particle verbs in complex ways, and no monolithic explanation for IndE usage at large suggests itself. There is evidence coming from ICE-India that nonstandard combinations may nowadays be more common in less heavily edited forms of speech and writing. This is, for example, the case for the combinations *to admit in/into* and *admission in*, which outnumber their standard English counterparts *to admit to* and *admission to* in spoken, but not in written language. On the other hand, the combinations *translate in* and *translate to* appear to be rare in both modes, cf. Table 4.20.

In view of this evidence, more register-specific research across registers is called for to illuminate the nature of lexicosyntactic variability in IndE in the use of verb particles more precisely. It would also be interesting to know how acceptable those combinations have become with educated users of IndE over the years, as this would help predict the future path of lexicosyntactic nativization in IndE. A future attitude survey might consider taking those variables into account.

4.2.3 Use of ‘additional’ particles

Sometimes, IndE uses particles after lexical verbs which take no particles in standard English. While, for example, standard English employs the verbs *discuss*, *emphasise/emphasize*, *stress*, *investigate*, *request* and *urge* without prepositions following them, IndE has been observed also to use particles with those verbs,¹⁵ and the examples from the Primary Corpus below are in line with that claim at first sight. Consider:

15. Cf. Bansal (1983: 3), Leitner (1992a: 224) and Parasher (1994 [1983]: 154) on *discuss about* in IndE. *Discuss about* was also one of the least acceptable items in Sheorey’s (2006: 192) ‘error gravity’ survey among 59 English college teachers from India, which suggests a low degree of institutionalization of this combination.

- (33) <\$C><#>well joining us now in the studio to discuss a little more about what's happening in Sri Lanka and in Tamil Nadu<,> is Aditi Phadnis who is with the Business Standard and also someone who's been closely following the conflict in Sri Lanka for a while<,> [IND2000 Broad Int 10]
- (34) The film emphasizes on the humanity of both sexes, about how what is supposedly right for the man is right for the woman too and the other way round. [IND2000 Press Rev 09]
- (35) The recent tour of South Africa to India, disclosed about some of these "worms" in the international cricket. South African captain, Hansie Cronje agreed of taking [sic] money after the Delhi police investigated into the matter. [IND2000 Ess 04]
- (36) <\$A><#>he also<,> successfully requested for Sharif and his six co-accused<,> to be present in the court when it meets tomorrow [IND2000 Broad News 02]
- (37) Majumdar particularly stresses on the emerging option of marriage and the various facets of love and divorce. [IND2000 Press Rev 08]
- (38) He urged for corrective steps as the situation is going out of proportion. [IND2000 Press Rep 02]

The highlighted combinations have in common that they were employed in an unself-conscious manner, probably coming into being on the analogy of the related standard English noun-particle combinations *discussion about*, *emphasis on*, *investigation into*, *request for*, *stress on*, and *urge for*. A look at the larger ICE-India suggests that those combinations are indeed recurrent in contemporary IndE, but they are not generally preferred. Note the relatively high frequency of occurrence of *stress (up)on* and *request for* in spoken IndE, cf. Table 4.21.

A similar picture emerged when consulting the Internet corpora, which revealed that the verb-particle combinations in question have established themselves to variant degrees in the English-speaking world in the domain of published writing.¹⁶ This is suggested by a comparison of the recurrency rates across the top-level country domains in collocations of the types Verb-ING PART THE and Verb-ING THE (Google search dated October 18, 2002), cf. Table 4.22.

It turns out that *urge for*, *investigate into* and *discuss about* are comparatively rare under the 'in' domain, whereas *emphasis/ze on*, *stress on* and *request for* appear to have a much stronger foothold. What is more, both Indian writing and Singaporean writing under the 'in' and 'sg' domains show similar degrees of openness for the unrecorded forms to enter written usage. What may have caused this likeness of the two

16. Cf. Appendix 4.9 to Appendix 4.26 for the token frequencies of the collocations in South Asia's online newspapers and across the top-level country domains.

Table 4.21 *Discuss, emphasise, stress, urge, request, investigate* [ICE-IND].

	IndE ICE-IND writing	IndE ICE-IND speech
discuss N (standard)	31	109
discuss about N (nonstandard)	4	8
emphasis/ze N (standard)	7	18
emphas/zise (up)on N (nonstandard)	–	–
investigate N (standard)	19	7
investigate into N (nonstandard)	–	1
request N (standard)	69	55
request for/to N (nonstandard)	–	6
stress N (standard)	11	7
stress (up)on N (nonstandard)	1	6
urge N (standard)	4	26
urge (up)on N (nonstandard)	1	1

Table 4.22 *Discuss, emphasise, stress, urge, request, investigate*
[Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
urging for the :	1:113	1:146	1:643	1:87	1:525	1:624
urging the						
investigating into the :	1:82	1:45	1:3272	1:3871	1:2047	1:2085
investigating the						
discussing about the :	1:64	1:17	1:860	1:467	1:1323	1:1191
discussing the						
emphasis/zine on the :	1:10	1:13	1:361	1:207	1:775	1:741
emphasis/zine the						
stressing on the :	1:6	1:7	1:434	1:249	1:340	1:653
stressing the						
requesting for the :	1:6	1:4	1:326	1:223	1:596	1:424
requesting the						

varieties remains open for debate. We seem to be dealing with a pan-Asian phenomenon deserving further research.

One group of verbs taking additional prepositions in the Primary Corpus is particularly noteworthy as interference from Indian languages has been claimed to be responsible for creating variability. Consider examples (39), (40) and (41):

(39) And yet Mr. Jain calls them as “essentially superficial differences.”

[IND1978 Press Edit 09]

- (40) The area around the Shahdara flyover, including near Shyam Lal College and the flyover's descent towards Rajender Nagar in UP, has been officially termed as one of the black spots of Delhi. [IND2000 Press Rep 14]
- (41) <\$D><#>no this commission Sir should be renamed as National Population Association [IND2000 Broad Disc 07]

IndE usage does not comply with the conventions of standard English in the examples above, which would likely not have the particle *as* after *call*, *rename* and *term*. According to Sridhar 1992, the use of *as* in such contexts represents a “pervasive characteristic of Dravidian English ... probably motivated by the fact that, in Dravidian languages, speech material such as technical terms, names, and quotations have to be marked by the quotative particle ...” (142–3). As regards the examples above, it is plausible to claim that *as* functions as a quotative marker, since *as* in examples (39) and (40) precedes direct and indirect quotations and in example (41) a name. But since the two examples from writing appeared in the Delhi-based *Times of India*, the general validity of Sridhar's (1992) hypothesis remains questionable, as the phenomenon appears to go beyond those subvarieties of IndE that have Dravidian languages as background languages. Are we dealing with a stable expansion of the functional load of the preposition *as* on a pan-Indian scale? Addressing this issue is not a straightforward task as standard English itself uses *as* quite flexibly in complex transitive complementation after verb and direct object (cf. Quirk et al. 1985: 1200). This makes it difficult to interpret IndE variability only within the interpretative framework of substrate influence. In standard English, some verbs, e.g. *describe*, always take *as* after the direct object (as in example (42)), while others, e.g. *consider*, take the preposition optionally (as in examples (43) and (44)). See the following examples from BrE and AmE:

- (42) Al-Qaddissiya, Iraq's defense ministry daily, described Bush as a “cursed criminal,” British Prime Minister John Major as “worthless” and French President Francois Mitterrand as “the mean old man.” [US1992 Press Rep 05]
- (43) I'd like every woman to consider politics as an option. [BrE FLOB F 13]
- (44) “There is one especial test of respectability in plate,” he remarked; “we seldom find it but, when we do, we consider it the most correct thing and the best guarantee of solid prosperity that anything in plate can give.” [BrE FLOB G15]

For yet other verbs, the presence or absence of *as* involves a distinct change in meaning, for example in the case of *name* ‘give a name to’ (no *as*) and *name as* ‘identify, choose’. See examples (45) and (46) from IndE:

- (45) I don't know why they named her Pandora. [IndE KOL R 01]

Table 4.23 *Call/rename/term (as)* [KOL, Primary Corpus, ICE-IND].

	IndE Kol 1978	IndE IND 1978 Press	IndE IND 2000 Press	IndE IND 2000 Broad	IndE IND 2000 Ess	IndE ICE-IND writing	IndE ICE-IND speech
call N N	325	10	16	14	5	129	336
call N as N	16	1	–	–	3	16	101
rename N N	1	1	–	–	–	–	–
rename N as N	2	1	1	1	–	–	1
term N N	19	2	2	–	–	3	3
term N as NP	16	–	2	–	–	4	13

- (46) The DSP and the Police authorities said that they are ready get to the cooperation of the different industries to save their factories, but they said that ‘they regarded the police as their personal servants; they feel we should arrest anyone they name as a trouble-maker.’ [IndE KOL H 16]

In all those instances, the preposition *as* can be said to designate “a copular relation, particularly in specifying a role or status associated with the direct object” (Quirk et al. 1985: 1200). When taking the inherent variability of standard English into consideration, can IndE still be said to use *as* in an extended way as quotative marker?

A look at the Kolhapur Corpus, the Primary Corpus and ICE-India brings to light that contemporary IndE is, in fact, more variable than BrE and AmE when it comes to the use of *as* in complex-transitive complementation involving the verbs *call*, *rename* and *term* – but the preposition is not equally likely to combine with all those verbs in all contexts. Consider the distribution of the competing complementation patterns in the Kolhapur Corpus, the Primary Corpus and ICE-India (Table 4.23). The construction *call N as N* is a recurrent variant in present-day IndE today, although it is not preferred in any text category. Rare in journalistic writing (column 2) and in the broadcast material (column 3), the pattern appears to be more of a second choice in non-public writing (represented by the student essays; column 4) and in other speech settings (as represented by ICE-India; column 7). On the other hand, while the constructions *rename N as N* and *term N as N* are found less frequently in the corpus data, they appear to be quantitatively on a par with their internationally established standard variants. In fact, the proportion of combinations with *as* is higher in the younger datasets (Primary Corpus; ICE-India) than in the Kolhapur Corpus from 1978, which might be an indication that their acceptability among educated users has been on the rise over the past two decades. The claim that uses of *as* with *term* and *rename* are fully established in present-day IndE today is also supported by the results of a Google-based search query across the six Indian online newspaper corpora (October 13, 2002), which are summarized in Table 4.24.

Table 4.24 *Call/rename/term (as)* [Internet: IndE, BrE, AmE newspapers].*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
called him the	12	11	1	5	42	9	78	87
called him as the	–	–	–	–	–	–	–	–
was called the	42	38	18	13	113	17	177	166
was called as the	–	–	1	–	–	–	–	–
renamed it N	3	3	2	1	8	4	25	21
renamed it as N	1	1	–	–	8	–	–	–
was renamed the	9	6	–	–	8	4	13	9
was renamed as the	2	1	1	1	5	–	–	–
termed it the	4	2	–	–	6	–	–	4
termed it as the	3	7	1	3	22	–	–	–
was termed N	11	23	6	4	58	5	23	15
was termed as N	21	12	2	–	66	1	–	–

* Excluded from the count were free combinations such as *the Army was called as the district police 'rebelled'*.

In light of those findings, Yadurajan's (2001) statement that "[a] very common construction in IE is the use of *as* with *call*" (30) can be modified. The usage described by Yadurajan 2001 is found repeatedly in contemporary IndE speech but is still notably absent from journalistic writing. On the other hand, the use of *as* with the verbs *rename* and *term* has become fully established even in India's quality broadsheets. It is possible that since the high-frequency word *call* is used in many contexts without a 'naming' sense, constructions without *as* seem a more natural choice also in cases when *call* is used to denote that 'someone gives someone or something else a name or title'. *Rename* 'to change a name into a new name' and *term* 'to say what other people call something', by contrast, are basically limited in use to their naming and quoting functions, which may make the additional use of *as* as a quotative marker more likely, possibly due to interference from languages like Telugu (cf. Sridhar 1992: 142–3). But note that PakE journalistic writing, too, shows similar degrees of variability, which suggests that the use of *as* with *rename* and *term* is not restricted to areas where Dravidian languages are spoken (Google search dated October 13, 2002), cf. Table 4.25. When adopting a more international perspective (October 13, 2002), South Asian usage stands out by frequently using *as* after *rename* and *term*, but not after *call*, cf. Table 4.26.

Within the standard usage range covered by this study, then, *as* is particularly likely to occur across registers, modes and text types with the two verbs used exclusively for naming or quoting, namely *rename* and *term*. In those instances, L1 interference might well be a factor reinforcing a dimension of variability already inherent in English. After *call*, on the other hand, the use of *as* is recurrent only in speech, but rare in educated writing. It is conceivable that scholastic mediation and prescriptive

Table 4.25 *Call/rename/term (as)* [Internet: PakE, BaE and SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
called him the	14	–	–
called him as the	–	–	–
renamed it N	6	–	1
renamed it as N	5	–	2
termed it the	31	1	–
termed it as the	10	2	–

Table 4.26 *Call/rename/term (as)* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
called him the	74	123	5310	176	719	1660
called him as the	1	–	1	–	–	–
renamed it the	8	14	801	37	219	296
renamed it as the	4	2	13	–	7	3
termed it the	6	7	132	35	56	42
termed it as the	7	–	2	–	2	–

pressure should have had a greater impact on IndE usage in the case of *call*, since this high-frequency verb has been brought to users' attention repeatedly by language critics (cf. Yadurajan 2001:30) and might therefore be avoided in nonstandard patterns in more heavily edited language use, for example in national newspapers. In general, more regionally sensitive research is needed for all three verbs to assess the force of substrate influence across linguistic areas more succinctly.

Another unrecorded use found in the Primary Corpus that is to be given attention to in this context is *comprise of*, which Algeo 2006 refers to as "one of the shibboleths of prescriptive usage guides" (242). According to Yadurajan 2001, "[t]he use of a preposition with *comprise* (*comprise of*) is incorrect" (40). A look at the corpus data reveals that usage in IndE is once again more complex than Yadurajan's (2001) generalization makes one believe. In fact, the combination *comprise of* is interesting as standard English itself is quite variable in putting this verb-particle combination to use both grammatically and semantically. Labeled formal in COB, the verb *comprise* is usually employed without preposition to say that something has a number of things or persons as its parts or members. See the following representative example from the Primary Corpus:

- (47) The series comprised the recitals of the young singers, Abhishek Raghuram, Randhini and Roshini, Balamuralikrishna, Sriram Parthasarathy, Sangeetha Krishnamoorthy and V. Sankaranarayanan. [IND2000 Press Rev 02]

In example (47), *comprise* means ‘include’ or ‘contain’. The verb can also be used in the passive, in which case the preposition *of* is normally added. In example (48) from the Kolhapur Corpus, *comprised of* means ‘made up of’, cf.

- (48) His patta or holding is comprised of wastelands and wetlands.

[IndE KOL F 06]

There is another construction with *comprise* that became established only in the twentieth century. Peters 1995 describes it as “the mirror-image of the traditional use: it begins with the parts that make up the whole, rather than the whole which consists of certain parts” (155). Here, *comprise* has the meaning ‘combine to make up’. For IndE it can be established that this latest use of *comprise* had arrived in IndE by 1978, as example (49) from the Kolhapur Corpus proves, cf.

- (49) Take, for instance, the resettlement colony of Shakurpur. Here, about 10,000 families, comprising a population of about 50,000, have been resettled.

[IndE KOL G 13]

While the three uses above must be considered standard, IndE uses yet another construction with *comprise*. See the following three examples that were uncovered when scanning the Primary Corpus:

- (50) The Doon School is a boarding school comprising of five hundred boys from all over the country.

[IND2000 Ess 09]

- (51) <#>thirty-seven per cent of Chhattisgarh’s population comprises of tribals

[IND2000 Broad News 18]

- (52) Another part of the exhibition comprises of the Hands and Feet Series.

[IND2000 Press Rev 10]

In IndE, it is also possible to insert the preposition *of* when *comprise* is used in the active voice to mean ‘consist of’. When searching for more examples of this construction, ICE-India returned eight tokens of *comprise of* (vis-à-vis 33 standard uses of *comprise*) from a wide range of text categories, for example private letters (example (53)), academic writing (example (54)) and broadcast talks (example (55)). This finding suggests a fairly highly degree of integration of this combination:

- (53) Now that everything is back to normal, Raju is relaxed & carrying out with his routine work. Invariably I & Kuku are with him in the Engine Control Room. Kuku has become an expert seafarer. He does his homework regularly; I don’t have to pester him to open his books. As far as his food is concerned, he’s not fussy about anything. (His diet comprises of milk, eggs, fruits & bread.)

[IndE ICE-India W1B-014]

- (54) It is an integrated system comprising of a low power X-ray generator (25 W) to excite the samples and a high resolution (160 eV at 5.9 keV) detector system coupled to a microprocessor controlled multichannel analyser.

[IndE ICE-India W2A-032]

- (55) <\$A><#>The Council will be presided over by the Prime Minister<,> and will comprise of the Defence External Affairs<,> Finance<,> and Home ministers

[IndE ICE-India S2B-004]

A Google-based search query conducted on the World Wide Web (July 24, 2002), using the collocations *comprising the* and *comprising of the*, revealed that the more elaborate active form is a recurrent minority variant in IndE journalistic writing today, but absent from the two BrE and AmE quality newspapers, cf. Table 4.27. When adopting a broader perspective, however, it becomes obvious that the active use of *comprise of* is not restricted to IndE alone. It is, for example, also more widely available across South Asia (Google search dated July 24, 2002); cf. Table 4.28. And when testing the currency of the unrecorded function of *comprise* on the Internet across top-level country domains (July 24, 2002), even more regional differences surfaced. Consider Table 4.29. Although the function has so far not been recognized in the major dictionaries of contemporary English, *comprise of* ‘consist of’ is used throughout the English-speaking world today. It is strikingly rare in AmE, though. We have, therefore, not come across a qualitative difference between IndE and other varieties of English. Rather, what distinguishes South Asian English from other varieties is the spread of the construction into the domain of quality newspaper prose, a category from which the feature has been shown to be noticeably absent in the BrE and AmE material consulted. In view of the comparatively high proportion of *comprising of the* for BrE, it is possible that this construction started out as a more recent innovation in BrE. Note that the OED quotations base features one example of *comprise of* used in the active voice. It is from a British source dating from 1973.¹⁷

4.2.4 ‘No’ particles

While it is comparatively easy to search corpora for ‘different’ uses of particles, it is impossible to run automatic search queries for cases of omitted particles unless a corpus has been tagged manually for this particular feature. Though there is little reason to believe that the deletion of particles should be any more innovative a force in IndE than the adding or changing of particles, this claim cannot be backed up empirically with the material at hand. Consider first examples (56) and (57) from spoken IndE, in which particles were apparently dropped at the spur of the moment:

17. This view is supported by findings reported by Algeo 2006, according to whom “CIC [Cambridge International Corpus] British texts have 0.2 iptmw [instances per ten million words] of *comprise of*; American texts have none.” (242)

Table 4.27 *Comprise (of)* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
comprising of the	15	7	5	3	39	–	–	–
comprising the	240	919	52	62	935	105	77	29
of : (of)	1 : 16	1 : 131	1 : 10	1 : 21	1 : 24	–	–	–

Table 4.28 *Comprise (of)* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
comprising of the	6	5	–
comprising the	458	45	31
of : (of)	1 : 76	1 : 9	–

Table 4.29 *Comprise (of)* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
comprising of the	634	137	8960	583	949	118
comprising the	7550	4380	138000	15900	91000	51500
of : (of)	1 : 12	1 : 32	1 : 15	1 : 27	1 : 96	1 : 436

- (56) <\$A><#>the Samata Party<,> an important partner in the ruling National Democratic Alliance government at the Centre<,> has urged the government to take sugar and cerosene out of the price-control regime<,> <#>party president Jaya Jaitly told reporters yesterday<,> that a consensus on the issue was arrived at the two-day national executive meeting<,> of the party which concluded yesterday [IND2000 Broad News 07]
- (57) <\$A><#><[2>very true</[2></[2>and very angry reactions coming in Mr Baig from the Indian Cricket Board chief uh as well<,> <#>uh Mr A-C Muthaiah is known to be<,> very restrained in all that he's uh said all this while <#>today he's almost lashed out Mr Ali Bacher <#>we'd like to take that up with you as well [IND2000 Broad Int 07]

While in the first excerpt, standard English would arrive *at* a consensus, the speaker in example (56) used the verb without the preposition, most probably because he was hesitant to use *at* twice in a row. Note that the temporal adverbial added after the verb starts with the preposition *at*. In example (57), a similar online production problem might have led to the dropping of the preposition *at*, by which *lash out* is usually followed. It is important to note that neither usage is representative of the patterns in

which IndE normally puts those two verbs to use. We are dealing with nothing more but unstable adhocisms.

On the other hand, the verb *bunk* 'to run away from school or work', which has been pointed out in reference to IndE by Nihalani et al. (1979: 39, 2004: 38) and Hankin (2003: 64), is used regularly without a particle in IndE whereas BrE uses the phrasal verb *bunk off* for the same meaning. The OED informs us that *bunk* arrived in the English lexicon in the second half of the nineteenth century in two uses, namely an intransitive use meaning 'to be off' and a transitive use 'to expel from school.' Today, those uses are apparently no longer common, as a look at the major dictionaries of contemporary English reveals. According to COB, informal British usage continues to have the phrasal verb *bunk off* to say that someone stays away from school or university without permission. In IndE, it is usually *bunk* that is used that way (cf. Crowther 1996: 1440). Standard English would most probably have the synonymous verb *skip* in such a context. The following is an example from an IndE broadcast interview:

- (58) <\$B><#>uhm<,> I think it was a relief<,> <#>I mean I went to college because I wanted to be the first Kapoor to join<,> college<,> to be I mean to go to college <#>I went to Sophia <#>I was there<,> just for a couple of months <#>and then I sat like my mom down and then we I told my entire family<,> that what am I doing here I'm just bunking classes sitting around in the canteen going to watch movies [IND2000 Broad Int 03]

Five similar examples from spoken IndE are also found in ICE-India. Furthermore, with the help of the online corpora, it can be established that this usage is not uncommon in writing either. In India's online newspapers, *bunk* collocates more frequently with *school* or *classes* than *bunk off* or *skip* do (Google search dated July 24, 2002); cf. Table 4.30. It is occasionally also used in South Asian journalistic writing outside India (Google search dated July 24, 2002), cf. Table 4.31. When looking up the collocations *bunk school* and *bunk off school* more widely on the Internet (July 24, 2002), the former collocation also turned up on numerous British and South African websites. We are, thus, not dealing with a feature restricted to IndE alone but one that is of larger currency in the English-speaking world (cf. Table 4.32). It appears that *bunk* is notably frequent in South Africa but not uncommon in BrE either. On account of that, *bunk* is most likely not a variety-induced independent innovation of IndE but, in fact, a historical 'predecessor' of the particle verb *bunk off*. It appears that the independent history of IndE has resulted in the retention of another lexical feature that has gone out of use elsewhere in the English-speaking world – but apparently not everywhere.

The same is true for the ways in which IndE employs the often-quoted item *to avail* (cf. Spitzbardt 1976: 41; Nihalani et al. 1979: 27, 2004: 25; Parasher 1991: 171, 1994 [1983]: 150; Yadurajan 2001: 18). Contemporary international English uses the expression *to avail oneself of* 'to take advantage of' predominantly in formal contexts, as in example (59):

Table 4.30 *Bunk (off)* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
bunk classes	5	20	1	7	18	8	–	–
bunk off classes	–	–	–	–	–	–	1	–
bunk school	4	3	1	2	2	3	1	–
bunk off school	–	–	–	–	–	–	8	–
skip classes	1	3	1	–	18	3	–	2
skip school	–	2	1	–	10	1	11	14

Table 4.31 *Bunk (off)* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
bunk classes	1	–	–
bunk off classes	–	–	–
bunk school	–	–	–
bunk off school	1	–	–
skip classes	1	–	1
skip school	3	–	–

Table 4.32 *Bunk (off)* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
bunk school	1	–	184	58	–	–
bunk off school	–	–	747	4	3	–

- (59) Availing themselves of their position as lessees or owners of the land, they told their ryots that the rents of all those who refused to sow indigo would be raised. [IndE KOL F23]

IndE also uses *avail of* in such contexts, cf.

- (60) Those producers would not have availed of her services if her English pronunciation was far from perfect. [IndE KOL B16]
- (61) There is serious dissatisfaction with the existing model of medical and health care services with its emphasis on hospitals, specialities and super specialities and highly trained doctors which gets limited in practice mostly to urban areas and which is availed of mainly by the well-to-do classes.

[IndE KOL H24]

This pattern, according to the OED, goes all the way back to Shakespeare and was still common in BrE throughout the nineteenth century. It has been singled out as incorrect usage by Indian style pundits (cf. Yadurajan 2001: 18), as has been the following third variant:

- (62) SMS could be availed through ordinary cell phones. [IND2000 Press Rep 15]
- (63) ... whenever the problems arise in front of her husband the educated wife would readily avail him with her fantastic ideas. [IND2000 Ess 08]

According to Spitzbardt 1976, this use exemplifies “[a]n outstanding grammatical feature of IE [IndE] verbal constructions” (41), namely the dropping of prepositions. More recently, Yadurajan 2001 has characterized *avail* as a verb “frequently used in Indian English ... without a preposition” (18). What is not mentioned by Spitzbardt 1976 and Yadurajan 2001 is that the transitive use of *avail* without *of* is, again, most likely not an independent development starting out in IndE but a pattern that, according to the OED, used to be more common in English in previous centuries. What therefore characterizes IndE today is once again the freedom taken by its users in retaining a degree of variability that has been shed elsewhere, and labeling such uses incorrect runs danger of denying IndE its independent history. Consider the distribution of the three compatible forms in written English, which shows the verb being more frequent in IndE around 1978 than in BrE and AmE, cf. Table 4.33. ICE-India returned a similar distribution both for writing and for speech, cf. Table 4.34. All three patterns are also recurrent in India’s quality newspapers. Table 4.35 gives the distribution of the variants in the online corpora (July 26, 2002). In fact, newspapers in India, Pakistan, and Bangladesh appear to be largely homogeneous today in terms of giving preference to the shortest of the three competing forms (cf. Baumgardner 1993b: 47), whereas none of the forms was found to occur in the Sri Lankan online newspaper consulted (July 26, 2002), cf. Table 4.36.

The distribution across top-level country domains supports the reading of mono-transitive *avail* without particle being a quantitative variety differentiator for South Asia from a contemporary perspective, although the pattern is occasionally also found elsewhere in the English-speaking world (Google search dated July 26, 2002).¹⁸ Consider Table 4.37. Strictly speaking, IndE can be claimed to have gone independent innovative ways here only insofar as the use of *avail* has become common in informal contexts as well. Note that Crowther (1996: 1435) even labels the verb informal.

18. Cf. Mukherjee/Hoffmann (2006: 165), who report uses of *avail* in this particular pattern in the British National Corpus.

Table 4.33 *Avail (of)* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
avail oneself of NP	4	2	1	4	–
avail of NP	9	–	–	–	–
avail NP	2	–	–	–	–

Table 4.34 *Avail (of)* [ICE-IND].

	IndE ICE-IND writing	IndE ICE-IND speech
avail oneself of NP	2	1
avail of NP	2	3
avail NP	3	4

Table 4.35 *Avail (of)* [Internet: IndE, BrE, AmE newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
avail themselves of the opportunity	–	2	–	–	16	–	2	1
avail of the opportunity	2	3	–	3	22	8	–	–
avail the opportunity	6	5	2	1	23	3	–	–

Table 4.36 *Avail (of)* [Internet: PakE, BaE, SrLE newspapers].

	PakE DA	BaE IN	SrLE ST
avail themselves of the opportunity	10	4	–
avail of the opportunity	13	5	–
avail the opportunity	60	8	–

Table 4.37 *Avail (of)* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
avail themselves of the opportunity	11	19	289	52	340	248
avail of the opportunity	36	–	127	3	16	1
avail the opportunity	153	1	10	–	8	7

4.3 Even more variation in IndE lexicosyntax? Case studies

As the combined approach of using closed and online corpora side by side had proved successful in describing the status of multi-item verbs and phrases across registers, text categories, and regions, it was decided to go ahead and test the validity of several more hypotheses that had been found in previous descriptions of IndE. For additional diachronic information, the OED and its quotations base were again consulted. What was ultimately uncovered was more evidence showing that variation in IndE cannot be described appropriately only in reference to contemporary exonormative usage standards. Both variety-internal innovative tendencies as well as the independent history of English in the Indian subcontinent must be taken into account to interpret the variational profiles of IndE in the data.

4.3.1 *Enjoy*

Enjoy ‘to find pleasure in doing or experiencing something’ is a verb for which IndE usage has frequently been said to differ erroneously from the codified norms of standard English. However, a closer inspection of the corpus data suggests a different reading of the situation in contemporary IndE. Consider first examples (64) and (65) from the Primary Corpus:

- (64) It is the only period when a child can enjoy to his full extent.

[IND2000 Ess 01]

- (65) <\$A><#>but did it also mean that you missed out on some of the carefree fun and some of the irresponsibilities that teenagers automatically indulge in

<\$B><#>uh I guess when I see my sister today<,> I do feel that you know she’s uh been able to enjoy a lot more than I did<,> <#>but uh you know I was so focussed and so determined to do to do what I wanted to do<,> to be an actress<,> that I don’t regret not going through all that

[IND2000 Broad Int 03]

Both examples are slightly off the norms of standard English as codified in COB in that the direct object after the verb *enjoy* is missing. In standard English, reflexive pronouns would arguably be more likely choices in both cases. It is for this reason that Nihalani et al. (1979: 74, 2004: 73) warn Indian users of English against employing *enjoy* intransitively in contexts like *come and enjoy* and *yes, I enjoyed very much* and recommend using the monotransitive pattern instead. However, the intransitive use of *enjoy* by itself should not be considered a structural deviation, as *enjoy* is used intransitively elsewhere in the English-speaking world, too – for example as an imperative form meaning ‘Bon appétit’, predominantly so in AmE according to Quirk et al. (1985: 358). A look at the Primary Corpus suggests that intransitive *enjoy* ‘to take pleasure in a situation, to have a good time’, while being absent from the BrE and

Table 4.38 *Enjoy* [Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess
enjoy N/-ing	6	10	3	19
enjoy pron-refl	–	–	–	–
enjoy (intransitive)	–	–	2	1

Table 4.39 *Enjoy* [ICE-IND].

	IndE ICE-IND writing	IndE ICE-IND speech
enjoy N/-ing	44	92
enjoy pron-refl	2	2
enjoy (intransitive)	2	28

AmE data, appears to fill a communicative need in present-day IndE – but only in less formal contexts. In the Primary Corpus, it is found in the speech section and in the text category of student writing. It is absent from the press material, cf. Table 4.38. Another look at the larger ICE-India corpus supports the view that this usage has greater foothold in IndE in less heavily edited and more informal contexts, cf. Table 4.39. From a quantitative viewpoint, the transitive pattern *enjoy N* is the most common pattern in IndE, as is the case in BrE and AmE. Intransitive *enjoy* is more common in IndE than the reflexive variant *enjoy oneself*, which has been given a separate entry in COB. See more examples of the usage in question from another IndE student essay (example (66)), from a piece of non-academic writing (example (67)), and from a private conversation (example (68)):

- (66) They think that when mother and father return home they can enjoy with them. [IndE ICE-India W1A-004]
- (67) Our friends are dead and we ourselves that once enjoyed with them and now sorrow for them must go too! [IndE ICE-India W2B-003]
- (68) <\$><#>so all the departments can come together we can just enjoy have fun [IndE ICE-India S1A-062]

Interestingly, intransitive *enjoy* was among the more acceptable features in Sahgal/Agnihotri's (1985:124) acceptability study more than two decades ago, for which the authors had used the sentence *we have a party tonight – why don't you come and enjoy*. Rejected by only about seven percent of their informants, this sentence was accepted by a great majority of users as good (34.09%) or informal (59.09%) English. It remains to be shown whether IndE will see any further spread of the intransitive

pattern into more formal language in the future. If we follow Parasher 1991, the functional expansion of *enjoy* might even reflect a larger trend in IndE lexicosyntax, namely “a tendency in IE [IndE] to use certain transitive verbs intransitively,” which Parasher 1991 labels “[o]ne of the major deviations found in the VP [verb phrase]” (172). This hypothesis, too, will deserve more in-depth research with the help of larger corpora.

4.3.2 *Contest*

Another verb that has been singled out for being used wrongly as an intransitive verb in IndE is *contest* (cf. Whitworth 1982 [1907]:120), a verb which, according to the COB, is normally used as a monotransitive verb to say that someone ‘takes part in’ an election or competition, especially in BrE. But as is the case for intransitive *enjoy*, a closer look at the Primary Corpus suggests that IndE has evolved stable IndE-specific uses of this verb that are recurrent and should not be regarded as errors. Consider examples (69) and (70) from IndE, which are in line with international usage conventions in that *contest* is followed by a direct object:

- (69) <\$F><#>the Allahabad police on Monday caught a gang with fifty-three extremely sophisticated fire-arms<,> and a large quantity of ammunition<,> <#>the guns were brought from Burhanpur in Madhya Pradesh<,> and were reportedly meant for those contesting civic elections in U-P<,> scheduled for the twentieth of this month [IND2000 Broad News 12]
- (70) <\$D><#>there was consensus though on some other issues<,> the setting up of a delimitation commission to ensure more proportionate constituencies<,> allowing a person to contest a Rajya Sabha election from a state<,> even if the person doesn’t reside there<,> debarring candidates from contesting elections<,> only if they are convicted [IND2000 Broad News 15]

The pattern *contest N from N* is also common in IndE journalistic writing to specify the region in which a candidate is running for office. Consider:

- (71) Take Chief Minister Om Prakash Chautala, of the Indian National Lok Dal (INLD) who is contesting the elections from both Narwana and Rori constituencies. In Rori, which he represented last time, he is pitted against his own brother ... [IND2000 Press Rep 13]
- (72) Whether it is last-minute switching of constituencies when things get too hot for them in their home areas, or contesting the elections simultaneously from two constituencies, they are doing it all. [IND2000 Press Rep 13]

More research on the wider World Wide Web revealed that this pattern meets no communicative need outside South Asia today (cf. Appendix 4.27). If it occurs at all in non-

Asian sources under the top-level country domains, it appears to be in texts discussing aspects of South Asia, as in the following excerpt from an academic paper on India's educational system dated August 2000 that was published under the '.uk' domain:

- (73) It has now been made very clear by legal experts and through several judgments of the High Court and the Supreme Court of India that non-teachers and non-graduates are equally eligible for contesting elections from teacher and graduate constituencies respectively. [BrE¹⁹]

What is relevant about *contest N from N* for the lexicosyntactic description of the verb is the fact that the direct object may be left out in this particular environment, which has given rise to the verb-particle combination *contest from* 'to run for office (in an election) in a particular constituency or area or party'. See examples (74) and (75):

- (74) The BJP, an alliance partner of the INLD, whose leaders are not known for contesting from two constituencies, has also succumbed to this affliction in Haryana. [IND2000 Press Rep 13]
- (75) Former Chief Minister Bhajan Lal's son Chander Mohan is contesting from Kalka as the Congress candidate. [IND2000 Press Rep 13]

This combination is also restricted in availability to South Asia today, where it is common in published writing. The combination can be found outside the '.in' top-level country domain as well, but, interestingly, if it does, it usually refers to South Asian contexts (cf. Appendix 4.27). See example (76) from a BBC news item dated December 11, 2002:

- (76) Although born in northern Gujarat, Mr Modi is contesting from a constituency in Ahmedabad – a strong support base of Hindu voters. [BrE; BBC News²⁰]

Contemporary IndE finally also uses *contest* intransitively to say that someone runs for office, as in examples (77) and (78):

- (77) In deference to the wishes of Bhattarai who wants younger leaders to take over, former prime minister Surya Bahadur Deuba contested securing 43 votes against Koirala's 69. [IND2000 Press Edit 02]
- (78) We will finalise parameters to ascertain the ground realities well in advance. The best possible candidate will be offered a ticket to contest, irrespective of the party. [IND2000 Press Rep 13]

19. www.economics.ox.ac.uk/Members/geeta.kingdon/PublishedPapers/epw-polecon-paper.pdf, online February 2, 2003.

20. www.news.bbc.co.uk/1/hi/world/south_asia/2562727.stm, online February 2, 2003.

Since further examples of intransitive *contest* are relatively easy to come by when browsing India's online newspaper corpora, the intransitive pattern qualifies as a stable feature of contemporary IndE. Are we dealing with a grammatical innovation of IndE, though? Interestingly, the intransitive use of *contest* 'to strive or contend (in a general sense)' is also found in the OED, its first attestation dating from 1618. Common in nineteenth-century English, the pattern appears to have gone out of general use in the major varieties, which is probably why none of the major dictionaries of contemporary English makes mention of it. IndE, by contrast, seems to have retained this possibility, which therefore qualifies as a quantitative variety differentiator for IndE from a synchronic perspective – but probably not as a lexicogrammatical innovation.

4.3.3 *Cater*

Nihalani et al. (1979, 2004) warn against uses of the verb *cater* 'to provide' without a preposition, as in *we can cater all types of parties* (Nihalani et al. 1979: 44, 2004: 42), and encourage readers to employ the prepositional verbs *cater for* and *cater to* instead. Yadurajan (2001: 19) even goes one step further and labels such uses incorrect. A look at the major dictionaries of current English informs us that there is nothing much awkward or incorrect about using *cater* transitively without a preposition; this use has, for example, been recognized in the COB dictionary as an acceptable grammatical variant without any regional or stylistic label.²¹ See the following three examples from IndE, BrE and AmE which were found on the Internet. In each, *cater* has been used to say that 'someone or something provides for a group of people, an organization or an event':

- (79) It caters the whole of Malabar. [IndE²²]
- (80) PJ Murray is a widely experienced practitioner in the field of commercial finance broking. He caters all kinds of businesses from SME to larger corporate bodies, offering an extensive knowledge of what is possible financially. [BrE²³]
- (81) Chartwells is San Juan College's food contractor and caters all events on campus. [AmE²⁴]

21. This puts Algeo's (2006: 242) description of BrE and AmE varietal differences into perspective, in which the author labels the transitive use of *cater* an AmE-specific form.

22. www.calicutmedicalcollege.ac.in/Transfusion%20Medicine/index.html, online March 9, 2002.

23. www.murrayco.co.uk/aboutus.htm, online March 9, 2002.

24. www.sjc.cc.nm.us/pages/186.asp, online March 9, 2002.

See the alternative prepositional verbs *cater for* and *cater to* in context:

- (82) The biggest bank in the world, the Bank of America, has a one-man branch in Rocky Mountains. This essentially caters for hunters. [IndE KOL A 04]
- (83) A theatre catering to a large and varied audience has necessarily to be entertaining. [IndE KOL G 50]

Cater is occasionally also used without a preposition in Englishes worldwide to say that ‘someone or something takes something into account’. In such contexts, it frequently collocates with the noun *need*, cf.

- (84) Jagadhri railway station, which caters the needs of millions in Haryana, Himachal Pradesh and parts of Uttar Pradesh, is a neglected station where the Northern Railway has not shown any inclination to improve facilities. [IndE; *The Tribune*; April 2, 2001²⁵]
- (85) CAPA’s product line caters the need of hospitality properties with PC-based property management, central reservation and call accounting solutions. [BrE²⁶]
- (86) Our virtual and reseller hosting services are cheap and sufficient to cater the needs of any personal or a fairly small-corporate brand. [AmE²⁷]

Although this use has not been recorded in any major dictionary, it can be found recurrently across the top-level country domains. The relative rarity of the pattern might, however, point to its rather low degree of integration at present. More common in this environment are uses of *cater for* or *cater to*, cf.

- (87) The term “social welfare-services” denotes, “services intended to cater for the special needs of persons and groups who, by reason of some handicap – social, economic or mental – are unable to avail themselves of, or are traditionally denied, the amenities and services provided by the community. [IndE KOL F 22]
- (88) The commercial banks can cater to the economic needs of such slum dwellers under their various schemes for lending to weaker sections of the society ... [IndE KOL A 39]

When comparing the distribution of the three patterns with third-person-singular *caters* in collocations of the type *caters (for/to) the needs* across the online newspaper

25. www.tribuneindia.com/2001/2001/04/02/haryana.htm, online March 9, 2002.

26. www.carhирental.co.uk/direcory.php?browse/Business/Hospitality/Software/property_Management_Systems, online March 9, 2002.

27. www.optimiz.us/content/39/505/, online March 9, 2002.

Table 4.40 Complementation of *cater*.

	.in	.sg	.uk	.za	.au	.us
caters the needs	60	–	50	1	16	3
caters for the needs	31	35	17100	1370	9960	7
caters to the needs	770	242	1470	223	640	495

corpora and across the top-level country domains (Google search dated March 9, 2002), a noteworthy regional distribution surfaced. Consider first the frequencies of the three collocations across the top-level country domains, which are summarized in Table 4.40.²⁸

The fact that the collocation *caters the needs* comes in second in number only under the ‘.in’ domain (whereas it is outnumbered by both prepositional variants elsewhere) could be a sign that *cater N* ‘to take account of’ has become established to a slightly higher degree of integration in IndE than in other varieties today. But even if that should be the case, it is a minority variant and likely to remain one considering the prescriptive pressure from the available style guides. Moreover, Table 4.40 illustrates that both IndE and SgE show a remarkable affinity to AmE in giving preference to the prepositional verb *cater to*. BrE, on the other hand, uses *cater for* more often, which is also more common a choice in SAfrE and AusE. This regional difference also surfaces when comparing the frequencies of the prepositional verbs *cater to* and *cater for* (used in the third-person singular) across top-level country domains (March 9, 2002); cf. Table 4.41. And it is also observed when comparing India’s quality newspapers to Britain’s online *Guardian* and the American *Washington Post* (March 9, 2002), cf. Table 4.42. A look at the closed corpora lends further weight to the claim that *cater to* is more common a choice in IndE and AmE than in BrE today. The pattern *cater N*, on the other hand, is too rare to be attested here.²⁹ Cf. Table 4.43. Similarly, only *cater to* is attested in ICE-India (seven tokens), but not *cater for* or *cater N*.

Thus, Yadurajan’s (2001) recommendation that *cater to* should be used only “in the transferred sense” (19) ‘to take account of’ does not do justice to IndE usage preferences. Evidence from the online corpora and the closed corpora indicates that IndE users nowadays prefer *cater to* in all semantic environments, as do users elsewhere in the English-speaking world, for example in Singapore and in the United States. More diachronically sensitive research will be necessary to answer the question of how this similarity between South Asian Englishes and AmE has come about, though. According to the OED, *cater for* is the older of the two variants, its first attestation dating

28. Cf. Appendix 4.28 for the distribution of the three collocations in the online newspaper corpora.

29. Algeo 2006 reports that “[i]n CIC [Cambridge International Corpus], *cater for* is more than 100 times as frequent in British texts as in American; *cater to* is 3 times as frequent in American texts as in British.” (242)

Table 4.41 *Cater for* vs. *cater to* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
caters for	193	641	771000	119000	303000	879
caters to	24700	14700	72300	9870	70600	24600
for : to	1 : 128	1 : 23	11 : 1	13 : 1	4 : 1	1 : 28

Table 4.42 *Cater for* vs. *cater to* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
caters for	3	346	4	2	14	3	121	4
caters to	101	531	133	103	545	136	72	267

Table 4.43 *Cater for* vs. *cater to* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
cater N	–	–	–	–	–
cater for N	6	7	12	–	–
cater to N	13	–	2	6	3

back to the year 1600. The combination *cater to*, on the other hand, appears to be a nineteenth-century innovation, its first attestation in the OED dating from 1840. It is therefore not unlikely that *cater to* became established in AmE in the second half of the nineteenth century and spread into South and South East Asian Englishes afterwards, possibly at an accelerated speed in the second half of the twentieth century. If that were true, written IndE and SgE under the top-level country domains would have been more open to influences from AmE than, for example, SAfrA or AusE in this particular instance.

4.3.4 *Protest*

Like *cater*, the verb *protest* ‘object to’ may be complemented with direct or prepositional objects in contemporary English, which is contrary to what we find in Yadurajan 2001, who lists under his “examples of verbs wrongly used in IE [IndE] without a preposition” (19) the sentence *we must protest the rise in prices*. Corpus evidence shows that uses of *protest* ‘object to’ with a direct object have become established in IndE, BrE and AmE to variant degrees of integration. See an example of the pattern *protest N* ‘object to’ from IndE:

- (89) <\$E><#>Another villager was killed on Tuesday after police opened fire on a group of locals who had gathered to protest the previous night's killings
[IND2000 Broad News 16]

Alternatively, *protest* can be followed by a prepositional object:

- (90) Sir, – When the public protested against the shooting of Water at Kashi, the director, Ms. Deepa Mehta, talked about her rights guaranteed in the Constitution.
[IND2000 Press Edit 07]
- (91) Both did not question or protest at the old woman being left out of most family outings and visits.
[IndE KOL K 47]
- (92) John was the favourite of both parents, and when William protested about this, he was told that he would eventually “beg his bread” at John’s door.
[BrE FLOB G 16]

While those prepositional variants and the pattern *protest N* can be used more or less interchangeably, slight semantic differences seem to be involved. While *protest against N* often implies an organized protest, the combinations *protest at* and *protest about* imply a protest started by individuals (cf. Peters 1995:618). Historically, *protest N* and *protest against N* are the older forms, their respective first attestations in the OED dating back to 1599 and 1612. The combinations *protest at* and *protest about*, on the other hand, are twentieth-century creations dating back to 1922 and 1958 respectively. What makes the verb *protest* particularly interesting for a cross-varietal analysis is the observation that the pattern *protest N* appears to have seen a rise in popularity in AmE writing in the second half of the twentieth century, cf. Table 4.44.

The figures in Table 4.44, low as they may be, also suggest that IndE around the year 1978 was slightly more open than BrE in the early 1990s toward using the pattern *protest N* as a second variant.³⁰ This view is confirmed when comparing uses of *protest* in the IndE and BrE online newspaper corpora around the year 2000. The results of a Google-based search (dated November 1, 2002) reveal that *protest N*, even if it is not the preferred variant, is more common in five (of six) Indian newspapers than in Britain’s online *Guardian*. In the *Washington Post*, by contrast, the pattern *protest N* is clearly preferred. Cf. Table 4.45 for the exact distribution of the two competing variants.

When zooming further out, *protest N* is the most common pattern under the ‘in’ domain, whereas it is much rarer in BrE (Google search dated November 1, 2002), cf. Table 4.46.

From a synchronic point of view, AmE and BrE constitute the two ends of a preference scale, along which Englishes worldwide might be placed, with AmE represent-

30. In ICE-India, *protest N* is attested once and *protest against N* 14 times. No other prepositional variants occur here. Cf. Appendices 4.29, 4.30 and 4.31 for the distribution of *protest N* and its prepositional variants in the Primary Corpus and in Kolhapur, LOB, FLOB, BROWN and FROWN.

Table 4.44 Complementation of *protest* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
protest NP	3	–	1	3	9
protest PREP NP	13	11	14	3	2

Table 4.45 Complementation of *protest* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
protested the	73	244	32	31	137	120	86	334
protested ...								
... against the	93	357	85	42	999	147	134	51
... at the	6	16	1	8	29	7	124	17
... about the	2	2	1	–	6	1	44	–
(PR) : PR	1 : 1.4	1 : 1.5	1 : 2.7	1 : 1.6	1 : 8	1 : 1.3	1 : 3.5	4.9 : 1

Table 4.46 Complementation of *protest* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
protested the	403	161	14200	457	9870	23800
protested ...						
... against the	199	122	23800	447	10500	633
... at the	26	10	9910	233	392	400
... about the	4	14	575	67	163	53
(PR) : PR	1.8 : 1	1.1 : 1	1 : 2.4	1 : 1.6	1 : 1.1	22 : 1

ing a clear-cut preference for the pattern *protest N* and BrE representing the opposite preference (if not as pronounced). Along this scale, South and South East Asian Englishes are nearer the AmE pole and AusE and SAfrE nearer the BrE pole. This, in fact, corresponds to the distribution found for the choice between *cater to* and *cater for* (cf. Chapter 4.3.3), which has been interpreted in terms of postcolonial influence from AmE. Could it be that the contemporary force of AmE is generally greater in South and South East Asia than in Australia or South Africa in the area of verb grammar? More evidence will have to be collected to illuminate the impact of AmE after World War II in those particular instances.³¹

31. The Internet data are in line with Bauer's (2002) description: "While US English tends to prefer the construction *We protested the decision*, British English is more likely to use *We protested against the decision* (with the possibility of using *at* or *about* instead of *against*). Australian and New Zealand Englishes allow both." (54)

4.3.5 *Want*

The use of *that*-clause complements with the verb *want* has received considerable attention from language critics of IndE (cf. Nihalani et al. 1979:190, 2004:192) and researchers (cf. Shastri 1996:77) alike and has repeatedly been quoted as a lexicogrammatical peculiarity of IndE.³² See the following examples of the pattern in question, which were found in the Primary Corpus:

- (93) “I want that women belonging to OBCs be also given their due share in the women’s reservation Bill and I should be allowed to move the necessary amendment in Parliament,” she said. [IND2000 Press Rep 05]
- (94) As we are Indians we also want to adopt western [sic] culture but all of them do not want it that western [sic] culture should come to our country ... [IND2000 Ess 05]

In example (93) (from a newspaper report), *want* has been used monotonitively with a *that*-clause, in example (94) (from a student essay) in a ditransitive pattern with an object and a *that*-clause. Both constructions are not grammatical according to the norms laid out in Quirk et al. (1985:1193), as *want* belongs to a small group of verbs taking complementation by a *to*-infinitive clause with a subject in such contexts. Sahgal/Agnihotri 1985 report that the ‘*want* plus *that*-clause’ construction type was among the more acceptable features in their attitude survey and regarded as ‘good’ or ‘informal’ English by 72% of their informants.³³ This led them to speculate in the mid-1980s that “this construction ha[d] become a part of the normative behaviour of the educated speakers of English in India” (Sahgal/Agnihotri 1985:122). But there is no evidence from around the turn of the millennium that this construction has had any major impact on IndE usage over the years, although it is, indeed, a recurrent phenomenon.³⁴ See the exact distribution of the available complementation types with *want* in the Primary Corpus first (Table 4.47).

When we compare the frequencies of the ‘*want* plus *that*-clause’ pattern (line 2 in Table 4.42) to that of the pattern ‘*want* plus *to*-infinitive with subject’ (line 6) across text categories, the student essays section (with a ratio of 1 : 1) turns out to be slight-

32. Cf. Smith-Pearse (1968 [1934]:15), Verma (1978:214, 1982:183), Nihalani et al. (1979:176, 190), Bansal (1983:4), Gramley/Pätzold (1992:443), Hansen et al. (1996:225–6), George C (1998:19), Trudgill/Hannah (2002:134), Viereck et al. (2002:210), Nihalani et al. (2004:177, 192).

33. The sentence used by Sahgal/Agnihotri 1985 to elicit users’ attitudes was *Mohan wants that he should win the election*. It was considered ‘good’ English by 44.19%, ‘informal usage’ by 27.91%, and ‘wrong English’ by another 27.91% of the informants.

34. There is no evidence that the grammatical behavior of *want* in IndE should differ significantly from BrE or AmE in any other way within the usage range looked at. For a comparison of IndE with BrE and AmE cf. Appendices 4.32–4.34.

Table 4.47 Complementation of *want* [Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess
want (verb)	24	53	118	58
want (NP) that	–	1	2	1
want (standard uses)	24	52	116	57
of those:				
want NP (N/adj)	4	13	33	4
want NP to-inf	5	7	15	1
want NP -ed	1	1	–	1
want NP -ing	–	–	2	–
want to-inf	14	31	66	51
wanna	–	–	–	–

ly more variable than the IndE broadcast texts (2 : 15) and the IndE press material around the year 2000 (1 : 7), which suggests learner difficulties and variant degrees of planning and editing as possible influential factors in creating variation. Note that no example of *want* plus *that*-clause is found in the 1978 press texts, which is another finding suggesting a mild increase in variability in IndE quality newspapers between 1978 and 2000. But it is clear that, on the whole, IndE users remain largely adamant about employing '*want* plus *that*-clause' constructions vis-à-vis the standard pattern. The same picture emerged when looking at the other available databases. In ICE-India, the proportion of the '*want* plus *that*-clause' pattern is relatively high in text category S1A representing private conversations (with 6 instances as against 22 instances of the pattern '*want* plus *to*-infinitive with subject'). It is much rarer elsewhere in the corpus. On the World Wide Web, *that*-clause complements after *want* also appear to be rare. See, for example, the distribution of the two collocations *he wants that they* and *he wants them to* in the online newspaper corpora (Google search dated October 19, 2002) in Table 4.48. A similar result was produced when zooming further out and looking up the same two collocations across the top-level country domains (October 19, 2002); cf. Table 4.49. In view of this evidence, the degree of integration of the pattern '*want* plus *that*-clause' is rather low in IndE – a situation which is probably encouraged by 'remedial' teaching and exonormative pressure.

4.3.6 The *to*-infinitive vs. N-*ing* divide

Yet another area for which IndE lexicogrammar has long been said to display considerable variation is the infinitive vs. V-*ing* divide (cf. Whitworth 1982 [1907], Nihalani et al. (1979, 2004), Parasher (1991, 1994), Shastri 1996). Previous descriptions of IndE have been contradictory in this regard. A century ago, Whitworth (1982 [1907]: 140)

Table 4.48 Complementation of *want* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
he wants that they	–	–	–	–	–	–	–	–
he wants them to	8	16	2	3	17	5	28	75

Table 4.49 Complementation of *want* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
he wants that they	1	–	–	–	1	–
he wants them to	30	197	12000	248	646	1850

claimed that IndE displayed a tendency toward overusing *V-ing* forms at the expense of infinitive constructions. Decades later, Nihalani et al. (1979: 116, 2004: 116–7) stated the opposite by pointing out that “[t]he tendency to prefer the *to* infinitive form to the gerund (with or without a preposition) is a major characteristic of IVE [Indian variants of English].” None of those hypotheses can be given a firm empirical backing on the grounds of the available data, and IndE, at least as far as the usage range under investigation is concerned, should, again, be considered broadly in line with international conventions. After looking at 43 lexical items previously mentioned in reference to IndE in this respect, only three showed unusual complementation patterns in the Primary Corpus (cf. Appendix 4.35 for the complete list of variables). Consider the following examples:

- (95) <\$C><#>they are looking forward to go to their own homeland to their homes<,> where they have been living for centuries and centuries
[IND2000 Broad Disc 06]
- (96) In fact, her remark describing the Congress as “a large family” where “we listen to and accommodate various views” is a subtle remark that could be intended for those who fall within the Congress ranks but nurse grievances, as much as for those who have broken away and could contemplate to return if circumstances are right.
[IND2000 Press Edit 03]
- (97) <\$A><#>the relief director S-N Devi says that relief works in the affected areas of Saurashtra<,> Kutch<,> and North Gujarat was being taken up on war-footing<,> with a view to provide employment in those areas
[IND2000 Broad News 13]

Examples (95), (96) and (97) have in common that in each case an infinitive construction was chosen in a context that would demand a gerund to be in accordance with standard English. The speaker in example (95) complemented the multi-word verb *look forward to* with an infinitive instead of a gerund – a usage quoted by Nihalani et

al. (1979: 116, 2004: 116–7) as evidence for their trend hypothesis. In example (96), the use of the infinitive after *contemplate*, which is another item discussed in Nihalani et al. (1979: 54, 2004: 54), is unusual by international usage norms as *contemplate* belongs to a category of verbs taking noun or N-*ing* complements only. In the example above, *contemplate* has been treated on the analogy of verbs such as *start* and *begin*, for which both complementation patterns are available without implying any significant difference in meaning. In the last example, example (97), a gerund would be the international standard choice after *with a view to*, which is a multi-item combination that standard English has come to treat as a complex preposition. Before addressing the issue of how common the attested usages are in IndE today, see the following examples from the Primary Corpus, which illustrate that the internationally established variants are also available to users of IndE:

- (98) The Japan Open triumph could not have been better-timed for the Lee-Hesh duo, coming as it is after four failures on the trot. “I am pleased with the way things are shaping up and if we can keep this up, we can look forward to doing something good in the world doubles championship which comes up in December,” said Leander Paes, the Indian ace who is in homecity on a two-day break. [IND2000 Press Rep 08]
- (99) Indian Muslims have never contemplated being disloyal to India. [IND2000 Press Edit 09]
- (100) With a view to reviving listener interest in this phase, Mudhra organised a three-day festival of pallavis at Bharatiya Vidya Bhava Mini Hall in Mylapore from June 11. [IND2000 Press Rev 02]

Since the closed corpora were too small to allow for any more meaningful insights, the online corpora were tapped to look at the two variants in a number of representative collocations (Google search dated December 28, 2002). But even though the number of attestations for the verbs *look forward to* and *contemplate* were low in those corpora as well, which did not offer much leeway for interpretation, it is fair to say that the infinitive with *look forward to* and *contemplate* is rare in IndE today, at least with regard to the category of published writing, which puts a question mark on Nihalani et al.’s (1979, 2004) trend reading of their data (cf. Appendix 4.36). On the other hand, the infinitive after *with a view to* is a common, albeit not preferred choice, cf. Table 4.50.

Similar results were obtained when consulting the top-level country domains (December 28, 2002). While there is no evidence that *look forward to* or *contemplate* take the infinitive any more often in IndE writing than in the other varieties, the choice of the bare infinitive after *with a view to* stands out as a peculiarity of IndE (cf. Appendix 4.37). Note that although the infinitival variant is available across all six top-level country domains, it is only for IndE that it qualifies as preferred variant, at least as far as the collocation *with a view to provide* is concerned, cf. Table 4.51.

Table 4.50 *Look forward, contemplate, with a view* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
looked forward to seeing	1	3	–	1	3	2	17	17
looked forward to see	–	2	–	–	–	–	–	–
contemplates going	–	–	–	–	1	–	–	1
contemplates to go	–	–	–	–	–	–	–	–
with a view to providing	4	149	18	13	220	6	5	2
with a view to provide	3	40	16	4	82	3	–	–

Table 4.51 *Look forward, contemplate, with a view* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
looked forward to seeing	11	16	11100	108	866	993
looked forward to see	1	10	115	6	30	25
contemplates going	1	3	177	4	87	6010
contemplates to go	1	–	–	–	–	–
with a view to providing	799	75	40600	538	22000	189
with a view to provide	907	23	257	68	211	14

Table 4.52 *With a view to* plus infinitive/V-ing [OED quotations base].

	N	with a view to + V-ing	V
< 1800	11	1	6
1800–1850	17	5	15
1850–1900	34	19	15
>1900	13	43	–

How can the distribution for *with a view to* in Table 4.51 be explained? A look at the OED quotations base reveals that the development of *with a view to* into a multi-word preposition in BrE and AmE might have taken place only in the nineteenth century, cf. Table 4.52. Since the tide of change has been with the V-ing pattern, IndE can be considered least advanced among world Englishes today. Note in this context that *with a view to* is one of the forms that was also discussed by Parasher (1994: 150–1). Including the variable in a small-scale survey among IndE, BrE and AmE speakers conducted in the late 1970s, Parasher 1994 noted that the infinitival variant was unacceptable to his British and American informants but acceptable to users of IndE. On explaining his findings, he argued that “[s]ince in most Indian educational institutions several eighteenth and nineteenth century texts are prescribed for English studies, the

IndE speaker does not find the dated expressions odd" (Parasher 1994 [1983]: 151). On account of the search results presented above, it seems not unlikely that exposure to seventeenth, eighteenth and nineteenth-century literature has contributed to prolonging the life of the pattern *with a view* plus *to*-infinitive in the Indian subcontinent. Around the turn of the millennium, the infinitive was a viable grammatical variant in IndE in that particular environment (also cf. Parasher 1991: 172–3).

4.4 Chapter summary

Lexicosyntactic differences between IndE and other varieties of English are rarely of a qualitative, more often of a quantitative and register-specific kind. Among the verb-particle combinations with *up*, *down*, *off*, *away* and *out* in the Primary Corpus, for example, only about three percent were found to differ in use from the norms of standard English, the majority of which appeared in the category of student writing and represented unstable learner difficulties rather than stable variety-specific uses. However, when zooming in on the individual items in context, using the Primary Corpus and the online databases, several stable IndE innovations could be identified. It can be established that IndE journalism uses the particle verb *contest from* innovatively to say that 'someone runs for office in a particular constituency'. In the case of the particle verb *take out*, IndE has gone innovative ways by extending its meaning: *take out* can nowadays be used in IndE to say that someone 'leads' a protest demonstration. IndE is also slightly more flexible in using the verb *enjoy* intransitively to say that 'someone has a good time', although this use appears to be restricted to less carefully planned language use. Interference from India's background languages might be a factor reinforcing the use of *as* after verbs of naming and quoting, for example *term* and *rename*. The verb *avail*, dated by the norms of standard English, has survived in the Indian subcontinent, where it is nowadays used innovatively in markedly informal contexts in the sense 'to use, to take advantage of'.

Sometimes, differences between IndE and other varieties were found to be restricted to specific domains of use. This appears to be the case for active uses of *comprise of*, which have spread into India's quality newspapers. This has apparently not yet happened in BrE or AmE, though the pattern is used recurrently in those two varieties in non-professional writing. The combined use of the Primary Corpus alongside the online corpora also produced evidence for the conflation of the verbs *find out* and *find* and of *lower down* and *lower* to be more common in IndE than elsewhere in the English-speaking world. Likewise, the verb-particle combinations *request for*, *stress on*, *emphasize on*, *discuss about*, *investigate into* and *urge for*, each resembling well-established standard English noun-particle combinations, are more frequent in IndE today than in the other varieties considered. Those forms have become integrated to variant degrees across text categories, where they are usually employed in an unself-conscious manner as minor variants alongside their standard English 'preposition-free' companions.

In several other instances, IndE was found to use items and patterns that are no longer employed in the prestige varieties. Those particular features most likely are not independent innovations but remnants of nineteenth-century English. The verb *bunk* (as in *to bunk school*) and the transitive use of *avail* 'to take advantage of' (as in *to avail an opportunity*) are two cases in point, the first item also being common in South Africa. The particle verbs *chart out* and *chalk out*, too, are found more frequently in South Asian Englishes today than in BrE and AmE. Another interesting case is provided by the combination *with a view to*, which BrE and AmE have developed into a multi-item preposition followed by a noun phrase or gerund. In IndE, by contrast, the bare infinitive after *with a view to* is still in use in relatively high numbers, which, according to the OED, used to be the case also in Britain and America in the nineteenth century.

In two case studies, IndE showed a remarkable affinity to AmE, namely in the preference for *cater to* (as opposed to the BrE-favored *cater for*) and in the comparatively high proportion of the monotransitive pattern *protest N* 'object to'. More diachronically sensitive research will be necessary to clarify whether we are dealing with cases of divergent diachronic developments in BrE or with IndE-internal changes away from BrE toward AmE under the postcolonial influence from AmE.

Several usage patterns which previous descriptions of IndE have held up as characteristic features of IndE were rarely ever found. The patterns *want* plus *that*-clause, *contemplate* plus *to*-infinitive and *look forward* plus *to*-infinitive were infrequent both in the closed and in the online corpora. While those patterns might well be more common in less heavily edited forms of speech and writing, they appear to be unacceptable to users of IndE in more formal and more professional language use.

Morphosyntax and grammar at the sentence level

The topics discussed in this chapter cover several areas in syntax and grammar for which IndE has been said to go innovative ways, namely article usage (5.1), the noun phrase along the count and number divides (5.2), the past perfect (5.4), tense with *since* and temporal adverbials of the past (5.5), question formation (5.7) and focus syntax (5.8). In addition, two syntactic variables for which Englishes worldwide were showing clear signs of change in the twentieth century are discussed, collective nouns (5.3) and the mandative subjunctive (5.6).

5.1 Noun phrase: Articles

Variation in the use of articles has repeatedly been brought up as a characteristic feature of new Englishes in general and of IndE in particular, and hypotheses as to why IndE should vary in this highly complex linguistic field abound.¹ It has, for example, been noted that there is a “widespread Indian tendency to disregard the articles, both definite and indefinite” (Spitzbardt 1976:40) due to interference from Hindi and other major Indian languages (for example Tamil) which do not have a comparable article system (cf. Leitner 1987:341, 1988:388; Parasher 1991:169; Ramamurthi 1991:123–6; Shekar/Hedge 1996:60).² But when looking at the corpus data, numerous usages were found which differed from the codified norms of standard English not only in that articles were ‘missing’, but also in that articles had been added in places where standard English would have the zero article. In yet other instances, users had apparently replaced surface articles. In other words, the substrate hypothesis alone was not likely to account straightforwardly for the entire spectrum of variation in the data. Agnihotri et al. 1984, in a large empirical study on the use of articles in IndE conducted within an error-analytical framework, found a similar degree of variability in their

1. Cf., e.g., Spitzbardt (1976:40, 1987:130), Daswani (1978:121), Mehrotra (1982:158), Kachru (1986:40), Schmied (1989b:92–3), Lukmani (1992:161), Hansen et al. (1996:224), Shekar/Hedge (1996:60), Hickey (2004:545), Kirkpatrick (2007:94).

2. Hindi, for example, has no definite article but marks the definiteness of a noun phrase by word order. The numeral *one* can be used for specific indefinite references. Cf., e.g., Sharma (2005:537).

material.³ Noticing a “habit of dropping articles” (127) alongside “indiscriminate insertions, sometimes correct and sometimes incorrect” (127), the authors argued that the English article system was rather difficult to acquire for Indian learners because of its context-sensitivity, with the choice among definite, indefinite and zero articles determined largely by situational, pragmatic and stylistic factors. But by concluding that the situation in contemporary IndE was basically a reflection of “learning strategies and our teaching (or its absence) over the years” (127), Agnihotri et al. 1984 ignored the possibility that IndE might have carved out variety-specific usage patterns in certain linguistic environments that should be treated as stable variants rather than errors. More recently, in the mold of Platt et al.’s (1984: 52–9) comparative work on Englishes worldwide, the possibility of an emergent variety-specific semantic-pragmatic article system has been brought into the discussion for IndE and other new varieties of English (cf. Leitner 1987: 341, 1988: 388). Sharma 2005, for example, sees “an economical, disambiguating principle to the use of overt articles” (535) at work in IndE, according to which speakers are more likely to omit articles in noun phrases whose referents can be identified unambiguously from the context. Noun phrases introducing new referents, on the other hand, are said to be less likely to omit articles.⁴ “Is there a systematic use of the articles in South Asian English?” Kachru (2005: 47) asks in his latest survey of the structural characteristics of IndE – and comes to the conclusion that “[t]he present research does not provide a definitive answer” (48). What answers does the Primary Corpus hold in store? This chapter presents the findings.

5.1.1 Quantitative analysis

To get an overview of the availability of *a(n)* and *the*, the Primary Corpus was consulted first for the distribution of definite and indefinite articles in IndE, AmE and BrE. Since the corpora were untagged, there was no straightforward way of relating the absolute number of surface articles to the total number of noun phrases in the corpus or to the number of zero articles and other determiners, which was bound to limit the explanatory potential of this quantitative approach (also cf. Schmied 1994: 224–5). It was nonetheless hypothesized that if IndE was indeed showing a trend toward article deletion under the influence of Indian languages, this might be reflected in a significantly lower number of definite and indefinite articles in the IndE sections of the

3. Agnihotri et al. 1984 examined the use of articles by 366 Indian university students who were asked to read a text from which definite and indefinite articles had been deleted and add articles where felt necessary. A parallel study was conducted with Ethiopian and Yemeni students (cf. Agnihotri 1992).

4. Sharma 2005 examined the use of articles in conversational language recorded in interviews with 12 first-generation Indians living in the United States, who “acquired English to varying degrees in India, emigrated as adults to the United States, and, for the most part, maintained their multilingual repertoires” (541).

Table 5.1 Definite articles per 1,000 words [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
the	80.3	79.0	67.5	64.0	71.4	60.6

Primary Corpus than in AmE and BrE. But the frequency distributions were more complex than that and defied any one-dimensional reading of the data along those lines. Let us briefly consider them in more detail.

Table 5.1 gives the frequencies of the definite article *the* per 1,000 words across the IndE, BrE and AmE press texts.⁵ According to Table 5.1, the number of definite articles per 1,000 words is higher in the IndE press material than in BrE or AmE. This runs counter to what one might expect when subscribing to the substrate hypothesis, according to which we would expect fewer, rather than more, definite articles in IndE. Table 5.1 informs us that between 1978 and 2000, IndE journalistic writing was diachronically stable in having more definite articles than BrE and AmE journalism during the three decades between the early 1960s and early 1990s. This observation corresponds to results reported by Sand (2004:287), who found that the IndE press texts in ICE-India also included a higher number of definite articles (88.1 per 1,000 words) than their BrE counterparts (72.0 per 1,000 words).⁶ Interestingly, among the varieties of English studied by Sand 2004, IndE is characterized by the highest number of definite articles in the press genre.⁷ Since the comparatively high number of definite articles in IndE suggests a high number of noun phrases in IndE writing, we might have come across quantitative evidence for what Parasher 1991 describes as a tendency “to use nominal forms where native speakers preferred verbal forms” (186), for example when choosing between variants like *acknowledge receiving this letter* and *acknowledge the receipt of this letter*. This hypothesis on IndE syntactic style variation will merit further research once a grammatically tagged corpus of IndE is available. This will permit systematic comparison of the number of noun phrases across varieties and text types.

Table 5.1 is interesting in yet another respect. It reveals that BrE and AmE journalistic writing saw a noticeable decline in the number of definite articles between the early 1960s and early 1990s, if slightly more pronounced in AmE. This development

5. Cf. Appendix 5.1 for the token frequencies of the definite and the indefinite article in the Primary Corpus.

6. The fact that the frequencies in ICE are slightly higher than in the Primary Corpus could be due to differences in sample size and corpus design.

7. Sand 2004 compares British, New Zealand, Jamaican, East African (Kenyan), Indian, and Singaporean English on the basis of ICE.

has not gone unobserved in previous research and has been related to stylistic trends in American and British newspaper English affecting the proportion of definite articles. Quirk et al. 1985, for example, mention as an influential factor the spread of “Time-style” (267) in English-language journalism, according to which definite articles are optional in longer premodified noun phrases (e.g. *Gurdarpur District Police Chief J. P. Birdi*, qtd. in Leitner 1988:391). Likewise, the spread of descriptive appositives without definite articles in preposed or postposed positions (Jucker 1992:207–50; Bauer 2002:56) has been observed to lower the proportion of surface articles in newspaper prose. Considering the frequencies in Table 5.1, IndE journalistic writing may not have followed those developments full force. This is also suggested by Sand’s (2004:90) case study on the various types of noun-phrase appositions with *president*, *minister* and *chairman* in IndE, for which she found BrE-typical appositive constructions with a definite article (e.g. *the Union minister for commerce, Mr Arun Nehru, has emphasized the need for raising exports*) to be more frequent than their competing AmE-typical constructions without article, which were found to be common in SgE as well.⁸ As Sand 2004 points out, “contact varieties have developed independent stylistic preferences which influence the occurrence of the definite articles” (290). More examples of such preferences in IndE will be provided in the qualitative investigation below.

Let us turn to the speech sections of the Primary Corpus. Again, there is no straightforward quantitative evidence that would suggest that the definite article is ‘underused’ in IndE. A look at the absolute token frequencies in Table 5.2 reveals that IndE, as compared with BrE, has more definite articles in the broadcast news section and approximately the same number in the categories of spontaneous speech. The comparatively high frequency of *the* in IndE news broadcast corresponds to the distribution in ICE-India for spoken IndE monologues in general (78.7 per 1,000 words in categories S2A and S2B subsuming speeches, lectures, broadcast news etc.; cf. Sand 2004:287). On the other hand, the frequency of *the* in the interviews and discussions is lower in the Primary Corpus than in ICE-India, for which Sand reports 69.9 definite articles per 1,000 words (2004:287). This might be due to differences in corpus design, as the category of ‘public dialogues’ in ICE-India comprises a larger spectrum of texts ranging from broadcast discussions all the way to classroom lessons. More importantly, Table 5.2 also reveals that IndE and BrE are alike in having fewer definite articles in spontaneous speech than in scripted speech or journalistic writing. This might best be explained in terms of the differences in linguistic complexity that are encountered when moving from predominantly unplanned texts (e.g. spontaneous speech) to carefully planned texts (e.g. newspaper prose).

It is only in the student essays that IndE finally appears to have noticeably fewer definite articles than BrE. However, even this generalization is problematic. While the Primary Corpus returned 47.8 tokens of *the* per 1,000 words for the student essays

8. The alternative constructions compared by Sand 2004 are of the types *Mr. P. J. Patterson*, *(the) prime minister of Jamaica* and *(the) Jamaican Prime Minister(,)* *P. J. Patterson*; cf. Sand (2004:290).

Table 5.2 Definite articles per 1,000 words [Primary Corpus; broadcast and essays].

	IndE IND2000 Broad News	BrE UK1991 Broad News	IndE IND2000 Broad I/D	BrE UK1991 Broad I/D	IndE IND2000 Ess
the	83.4	75.9	55.2	57.7	47.8

section, Sand (2004:287) found 62.8 tokens per 1,000 words in ICE-India and 73.9 tokens per 1,000 words in ICE-GB. In other words, while IndE student writing does seem to have fewer definite articles than BrE, the gap between BrE and IndE in ICE is not as wide as that ‘inside IndE’ between the Primary Corpus and ICE-India. One reason for this could be a difference in levels of competence between the two groups of students contributing to the respective corpora. The Primary Corpus includes essays by secondary-school students in their final school year. Although the corpus contributors had received seven years of English-medium education by the year 2000, a closer look at the essays reveals that their writing skills and linguistic competence varied notably. The student essays in ICE-India, on the other hand, appear more homogeneous and more elaborate. Written by university students with an experience in using the English language that probably exceeded that of the contributors to the Primary Corpus, those essays are more complex in structure and language, which might have resulted in using more noun phrases and, consequently, more definite articles.

When zooming in on the indefinite articles in the Primary Corpus, the absolute token frequencies turned out even more difficult to interpret in a coherent fashion. In the press texts, IndE journalists appear to use slightly fewer indefinite articles than their BrE or AmE colleagues. However, the gap in absolute numbers has narrowed over the years, and there is no statistically significant difference to be found between IndE, BrE and AmE at the turn of the millennium; cf. Table 5.3. For IndE and BrE, Sand (2004:293) uncovered similar frequencies in the press sections of ICE-India (22.8 definite articles per 1,000 words) and ICE-GB (25.7), which suggests that the category of journalistic writing is remarkably uniform across varieties as far as the frequency of the indefinite article is concerned.

In the broadcast news category, too, IndE behaved as was to be expected given the results from ICE. See Table 5.4. While the distribution of the indefinite article in the IndE broadcast news section is almost identical with that in ICE-India (24.7 indefinite articles per 1,000 words; cf. Sand 2004:293), differences are somewhat more pronounced in the categories of spontaneous speech (ICE-India: 20.5; ICE-GB: 19.8), which might again simply be due to the differences in corpus design.

In the category of student writing, IndE is again most strikingly different from BrE. While the student essays in the Primary Corpus feature 29.4 indefinite articles per 1,000 words, Sand (2004:292) found 23.9 tokens per 1,000 words in ICE-GB and only 13.5 tokens per 1,000 words in ICE-India. Given those figures, the number of indefinite

Table 5.3 Indefinite articles per 1,000 words [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
a/an	25.0	27.0	28.5	27.3	28.8	28.3

Table 5.4 Indefinite articles per 1,000 words [Primary Corpus; broadcast and essays].

	IndE IND2000 Broad News	IndE IND2000 Broad I/D	IndE IND2000 Ess	BrE UK1991 Broad News	BrE UK1991 Broad I/D
a/an	24.0	25.1	29.4	27.4	28.0

articles in the student essays is unusually high, which deserves a closer look. The qualitative investigation will return to this issue below.

What conclusions can be drawn from this exercise? First and foremost, no monolithic explanation for IndE article usage suggests itself on the basis of the frequency counts alone. IndE does certainly not generally ‘overuse’ or ‘underuse’ surface articles, but various factors such as text type, mode and style appear to interact in a highly complex manner to produce the variational profiles showing in the data. If we assume L1-interference to have an impact on IndE, this influence has not affected the distribution of the definite or the indefinite articles in IndE at large in any numerically graspable sense, and a more qualitative, microscopic approach is called for to arrive at a meaningful interpretation of the data. However, it is clear from the empirical evidence produced above that statements like Shekar/Hedge’s (1996), who claim as “a striking example of the grammar of IE [IndE]” that “IE speakers generally tend to delete the English definite/indefinite articles or use them inconsistently” (60), distort the situation in IndE grossly. Descriptions like this do not match the data investigated here and should therefore be taken with a pinch of salt.

5.1.2 Qualitative investigations: Article usage in the student essays

To shed more light on the finer nuances of IndE article usage, a closer look was taken at the student essays, the least professionally edited form of writing represented in the Primary Corpus. It was assumed that the essays would be particularly interesting to scrutinize, as they represented advanced learners’ writing which was likely to display a wide spectrum of ‘different’ uses ranging from unstable learner errors all the way to possible stable IndE-specific innovations – data that could be compared to more heavily edited forms of speech and writing elsewhere in the Primary Corpus and on the World Wide Web. The student essays were therefore reread closely and particular attention was paid to those instances of article use that could be said to differ in

Table 5.5 'Different' uses of articles [Primary Corpus; essays].

found ... expected ...	a(n)	the	zero	
a(n)	–	4 (2.5%)	38 (23.5%)	42 (25.9%)
the	5 (3.1%)	–	63 (38.9%)	68 (42.0%)
zero	8 (4.9%)	44 (27.2%)	–	52 (32.1%)
	13 (8.0%)	48 (29.6%)	101 (62.3%)	

function from the norms laid out in Quirk et al. (1985:265–87), which were used as a grammatical yardstick. All in all, 162 instances in the 20,000-word sub-corpus were uncovered that differed from the codified norms in that definite or indefinite articles had been added, exchanged or left out in ways not straightforwardly compatible with Quirk et al. 1985. This also meant that the overwhelming majority of articles in the student essays conformed to the norms of standard English, which was an important first finding as it suggested a fairly close adherence of Indian advanced learners' writing to the norms of standard English at the micro-level in spite of the differences in token frequencies that had been uncovered in the Primary Corpus for *the* and *a(n)*. Table 5.5 shows the distribution of the three types of 'different' uses in the IndE student essays.

Although the marked-up cases are relatively few, the distribution of the types of 'different' uses is noteworthy. The majority of 'different' uses in the student essays (63.4% of all marked-up uses) involves the omission of a surface article, with the dropping of the definite article being slightly more common (38.9%) than the dropping of the indefinite article (23.5%).⁹ It is possible that influence from Hindi was an influential factor here, as all essays were written by students speaking Hindi as their mother tongue. However, examples of articles used in places where standard English would arguably have none come in second in number, making up one third of all marked-up uses (32.1%). It is in those cases that the 'deletion' of articles alone fails to offer a convincing explanation, and other interpretations must be given.¹⁰ Evidence

9. In Agnihotri et al.'s (1984:117) study, the zero article used in place of surface articles was also the most common 'different' phenomenon. In 58.10% (55.87%; 42.30%) of all cases where the authors had expected the indefinite article *a* (the definite article *the*, the indefinite article *an*) to occur, the zero article had been used.

10. In Agnihotri et al.'s (1984:117) study, *the* was found in 7.80%, *a* in 3.67% and *an* in 0.71% of all cases where zero had been expected.

for the indefinite article used in place of zero or *the* is scarce, amounting to only 8% of all marked-up uses. This, in turn, suggests stylistic and pragmatic differences, rather than structural innovation, as the most probable reason for the comparatively high number of indefinite articles in the Primary Corpus student essays vis-à-vis ICE-India, which has been observed above (cf. Table 5.4). In comparison to the deletion and addition of articles, the exchange of surface articles is a minor phenomenon in the student essays (5.6%).¹¹ Next, let us take a closer look at the marked-up cases and discuss the ways in which IndE article usage is ‘different’ from the norms outlined in Quirk et al. 1985.

5.1.2.1 Zero article for definite article

The discussion starts out by looking at the most frequently occurring ‘different’ type of article use in the essays, namely that of the zero article used in place of the definite article. According to Quirk et al. 1985, the main function of the definite article *the* in standard English is to mark a noun phrase as definite, that is “as referring to something which can be identified uniquely in the contextual or general knowledge shared by speaker and hearer” (265). Interestingly, when looking at the IndE student essays, the zero article was found to occur in all (rather than just a selection of) semantic and pragmatic environments outlined in Quirk et al. (1985: 265–72) for the use of *the* with specific reference (i.e. reference to a “particular specimen” (265) in the real world) – a finding which corresponds to Leitner’s (1987) small-scale qualitative analysis of a private letter belonging to he terms the “lower mesolect” (342). Consider first examples (1) and (2):¹²

- (1) In this practical world of today an individual has no time left for following his old traditions and customs. If one continuous [sic] to go by it he will be left behind in the race of success. Earlier due to scarce means of communication [THE] world was divided into small parts ... [IND2000 Ess 07]
- (2) They hesitate to do the simplest thing and easiest task. They cannot face any difficulty. They can’t face any danger. Being always shut up within the four walls of the home, they have practically no contact with the outside world. They have no friends to be proud of. They have no knowledge of men and things around them. They sometimes feel sad and lonely ... However they

11. In Agnihotri et al.’s (1984: 117) study, *the* was found in 12.08% (10.24%) of all cases where *a* (*an*) had been expected; *a* (*an*) was found in 5.03% (2.16%) of all cases where *the* had been expected. Sand (1998: 125) uncovered a similar distribution of ‘unusual’ articles in her press and broadcast data from Jamaica, which showed “a tendency for zero-article instead of the definite article, followed by articles where none is expected. Very rarely do we find the two surface articles exchanged.”

12. The symbols [THE], [A] and [AN] stand for uses of zero articles in places where definite or indefinite articles were expected to occur.

wish to move with [THE] outside world and satisfy their craving for knowledge of the outside world, they can't. [IND2000 Ess 01]

In standard English, the definite article in noun phrases such as *the world*, *the cosmos* or *the earth* adds a sense of uniqueness; as Quirk et al. 1985 put it, “there is the presupposition that, in our experience or fields of interest, there is only one such object” (266). In two noun phrases in examples (1) and (2), students apparently did not feel the need to mark the uniqueness of their referents in the surface structure of their utterances, leaving the task of identifying the reference as unique to the addressee, the reader. This does not pose any serious threat to intelligibility, as the referents can be recovered straightforwardly from the given contexts. In fact, in both essays, the relevant referents have already been introduced via the underlined nouns phrases, which happen to conform to the norms of standard English. Examples (1) and (2), thus, resemble closely those presented by Sharma 2005 to illustrate that surface articles in her data were particularly likely to be dropped in noun phrases “with unambiguous discourse reference” (557), whereas noun phrases introducing new or inferable referents normally had surface articles.¹³ This led her to hypothesize that “a systematic use of overt articles mainly for purposes of discourse disambiguation might be operative for these Indian English speakers” (Sharma 2005:556). As regards examples (1) and (2), it is plausible to follow Sharma 2005 and claim that the students considered further definite articles superfluous, as in both instances the uniqueness of reference to *the world* had been established previously in discourse. Note, however, that the author of example (2) did not use the zero article consistently in the final sentence but returned to the definite article in the last underlined noun phrase, which could reflect some deeper-running insecurity in coping with the conventions of standard English article use.¹⁴

In example (3), on the other hand, zero articles occur more consistently, their use possibly triggered by the relative familiarity of the noun-phrase referents within the given context:

- (3) Man first lived in a part of land, that land became his home, his home made a village, [THE] village made a city and [THE] city made a country. [IND2000 Ess 08]

In the next set of examples, the choice of the zero article is somewhat more difficult to account for in terms of discourse familiarity, as it becomes clear from reading the entire essays that the zero articles occur in noun phrases presenting new information. This leaves the task of recovering the uniqueness of the noun-phrase referents

13. Sharma 2005 categorized the participants in her interviews into three groups of variant proficiency. In all three groups, surface articles were found to be omitted significantly more frequently in noun phrases with higher familiarity status.

14. IndE is variable in its choice between zero and definite article in such environments also on the Internet, cf. Appendix 5.3 for the distribution of the collocations *to (the) outside world*.

to readers' general knowledge. In examples (4) and (5), the article-free noun phrases refer to a uniquely identifiable period of time (the 1990s) and to a uniquely identifiable language (the English language).¹⁵

- (4) This has been going around since [THE] early 90's but the real proof has come out now. [IND2000 Ess 04]
- (5) These movies in [THE] English language tell us about various incidents that have occurred in those countries. [IND2000 Ess 06]

Likewise, to understand the uniqueness of reference in examples (6), (7) and (8) below, knowledge about India's political and economic system and India's literary heritage is necessary.¹⁶

- (6) Political corruption can be tackled only at [THE] political level and the battle against it must start at the election platform. [IND2000 Ess 03]
- (7) India has a mixed economy i.e. [THE] private sector as well as [THE] public sector dominates [sic]. [IND2000 Ess 06]
- (8) Rantilya says in [THE] famous Arthashastra, "Just as it is impossible not to taste the honey at the lip of one's tongue so it is impossible for a government servant not to eat up at least a bit of the king's revenue." [IND2000 Ess 03]

Note that each of the nonstandard noun phrases in examples (4) through (8) has been premodified by adjectives. This, in fact, may not be coincidental. Sharma 2005, too, found that modified and quantified noun phrases were more likely to have zero articles than 'bare' nouns, which she interpreted as further evidence for the emergence of an innovative pragmatic system of article use in IndE:

[W]hereas bare nouns require overt articles, modified nouns are more likely to be associated with omission of the article, and quantified nouns actually favor null articles. If modification restricts the potential real-world referents of a NP [noun phrase] token (e.g., *the cold day*) and quantification is understood as uniquely identifying the referents (e.g., *the first day*), then we can argue that exactly the same principle of disambiguation that was found for relative givenness lies behind that pattern as well ... Thus, two of the most robust factors for omission of articles in the present data – NP modification and discourse givenness – both point to a single discourse pragmatic system of article use primarily for the purpose of disambiguation. (Sharma 2005: 558)

15. Cf. Appendix 5.4 for the distribution of the collocations *since (the) early 1990s* on the Internet.

16. Cf. Appendix 5.5 for the distribution of the collocations *at (the) political level* on the Internet.

Consider in this light examples (9) and (10), in which zero articles precede two modified nouns. While standard English would use *the* for ‘logical’ reasons in such contexts (cf. Quirk et al. 1985:270), the interpretation of the superlative adjectives *greatest* and *highest* strongly suggests unique referents and may, for that reason, have made the use of definite articles appear superfluous, cf.

- (9) [THE] Greatest advantage which this golden period of childhood enjoy [sic] is that ... [IND2000 Ess 02]
- (10) But nowadays since these women have adopted the western [sic] culture we can watch them on the path of Progress [sic]. We can see them being doctors, engineers, scientists etc. which has not only given them [THE] highest rank in the society but also respected them for being in side by side of men in every field of life. [IND2000 Ess 06]

Examples (11) through (14) illustrate that zero articles are also found in noun phrases whose referential meaning is specified by means of postmodification.

- (11) And, I strongly believe that when I come up with [THE] evil consequences of betting and it's [sic] being legalised, you all will also agree with what I believe. [IND2000 Ess 03]
- (12) He is not aware of [THE] hardship of life. [IND2000 Ess 06]
- (13) In my view [THE] adopting of western culture is not a sign of advancement. [IND2000 Ess 06]
- (14) It is not able to face [THE] problem regarding population, food etc. [IND2000 Ess 06]

Examples (15) and (16) below are noteworthy in that capital spellings were used to highlight the uniqueness of reference of two noun phrases graphically – a strategy that is usually reserved in writing for proper nouns, which are commonly used without surface articles. Hence, the workings of analogy may have triggered the use of zero articles here. In example (15), the reference is to a uniquely identifiable sports team (the national cricket team of South Africa), in example (16) to specific nuclear tests conducted by India and Pakistan at the beginning of the new millennium (around the time when the essays were composed):

- (15) Let's not go very far but we have all just gone through a stage when all of us were simply left shocked when we came to know that Hansie Cronje, the ex-captain of [THE] South African Cricket Team [sic], was involved in match fixing. [IND2000 Ess 03]
- (16) Why is it that the bans imposed on India after [THE] Nuclear Tests [sic] are lifted up but are still over Pakistan? [IND2000 Ess 07]

Example (17) deserves a closer look in this regard, too, as the zero article has been used with a name for which standard English has come to use the definite article, *the (United) States*.¹⁷

- (17) A major problem in our country is the problem of being divided in different regions. By retiring our Indian cricketers would be able to encourage and find tallent [sic] in [THE] States. [IND2000 Ess 04]

The student who wrote that sentence may not have been aware of this usage convention.¹⁸

Next to using definite articles for specific reference, standard English employs definite articles also for what Quirk et al. 1985 call sporadic references, i.e. references made to an institution, to aspects of mass communication or to modern transport and communication “which may be observed recurrently at various places and times” (269). In example (18) from another IndE student essay, the zero article has been used in such a context with the noun *computer*. It is possible that the zero article with *computer* has been triggered on the analogy of *television*, which may be used with the zero article by standard English norms:

- (18) ... Every child wants to spend his/her time with friends, spends maximum time on [THE] computer or television ... [IND2000 Ess 05]

In standard English, *the* is also used frequently with body parts instead of possessive pronouns (cf. Quirk et al. 1985: 270). In example (19) from an IndE student essay, it is not *the* but the zero article that was used in such a context:

- (19) On the other hand, men cannot weep and thus they are much more prone to health problems most probably related to [THE] heart. [IND2000 Ess 09]

The may finally also function generically in standard English, although this function “is rather limited.” (Quirk et al. 1985: 282). According to Quirk et al. 1985, plurals like *the Romans* or *the Carthaginians* refer collectively to a group of people as a whole, the function of *the* in such contexts thus being ‘generic’ rather than ‘specific’. Without the definite article, a sense of ‘generalization’ would be lost. Consider in this regard example (20), in which the zero article has been used where *the* would arguably have been the more usual choice, cf.

17. Cf. Appendix 5.6 for the distribution of the collocations *live in (the) United States* on the Internet.

18. Cf. Sand 1998 for similar examples of *United States* preceded by zero from Jamaican English, for which a number of possible explanations are offered that also apply to the IndE data presented here: “a certain insecurity in article usage, or a tendency to hypercorrect usage, mirroring other common country names or perhaps an interference of the quasi-synonymous term *America*.” (127)

- (20) In ancient times, India was a main centre of business. Then came the rule of [THE] Britishers. [IND2000 Ess 07]

Although this is the only such example in the student writing, its occurrence suggests that in IndE the zero article may also be intruding on the domain occupied in standard English by the definite article used for generic reference.

In the light of the findings presented so far, it appears that IndE students use the zero article in all environments in which standard English would usually have the definite article. But as the quantitative analysis has shown, the likelihood for the zero article to occur in such contexts is rather low in this particular text category, which makes it difficult to weigh the overall influence of individual factors such as pragmatic innovativeness, interference from the students' mother tongue Hindi, individual competence and editorial carefulness. Moreover, while the zero article in specific, sporadic or generic reference is recurrent but rare in the student essays, it appears to be considerably less common in professional writing such as newspaper reportage, which is suggested by even the most cursory glance at the IndE press sections of the Primary Corpus. This observation corresponds to Kachru's (2003) in a microcorpus of letters to the editor from India and three other regions (Nigeria, Singapore and the USA), which did not produce evidence for any major usage differences to exist across varieties in this text genre. But there is always the possibility of certain 'article-free' collocations going down the path of integration and spreading into more heavily edited writing. An interesting case in point is illustrated in example (21), cf.

- (21) As [THE] majority of our neighbouring asian [sic] countries have adopted western [sic] culture so [sic] we cannot keep going on with our typical Indian culture. [IND2000 Ess 06]

Evidence from the closed and the online corpora suggests that IndE is currently seeing the combination *majority of* developing into a phrasal quantifier along the lines of combinations such as *plenty of* and *lots of*, which also consist of a quantity noun followed by *of*. *Majority of* is found once in the press section of the Primary Corpus, cf.

- (22) The extent of the injury could be gauged from the fact that [THE] majority of Megha's first serves (approximately 80 per cent) found the net. [IND2000 Press Rep 18]

In ICE-India, the combination *majority of* is particularly frequent in the text categories representing spontaneous spoken language, cf. Table 5.6.

When browsing India's quality newspapers on the Internet (February 11, 2003), *majority of* is found to have a foothold there, too, for example in collocations of the type *in (a/the) majority of*; cf. Table 5.7. In the online *Guardian* and in the online *Washington Post*, on the other hand, the use of *in majority of* is rare (February 11, 2003). The only example of its kind in the electronic *Washington Post* corpus was

Table 5.6 (*A/the*) majority of [ICE-IND].

	IndE ICE-IND writing	IndE ICE-IND speech
majority of	5	8
the majority of	11	3
a majority of	7	–

Table 5.7 (*A/the*) majority of [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
in majority of	–	100	21	7	246	2	–	1
in the majority of	–	108	7	4	97	4	61	199
in a majority of	11	204	21	7	380	8	9	8

Table 5.8 (*A/the*) majority of [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
in majority of	779	82	560	95	832	573
in the majority of	653	670	574000	24200	161000	49200
in a majority of	628	105	24900	1170	20900	15300

encountered in a headline, i.e. in a context in which articles can readily be dropped even by international usage standards, cf.

- (23) Placebos Improve Mood, Change Brain Chemistry in Majority of Trials of Antidepressants
[*Washington Post*; May 7, 2002¹⁹]

A final look at the top-level country domains confirms the view that the combination *majority of* is relatively more frequent in IndE today than in other varieties (Google search dated February 11, 2003), cf. Table 5.8.

Within the standard usage range, Indian users, it thus appears, are going innovative ways today in this particular instance, not in terms of restructuring the article system as such, but in terms of turning a noun phrase into a phrasal quantifier – a development which might have been accelerated by the fact that surface articles in IndE appear to be generally more optional in the context of quantified noun phrases (cf. Sharma 2005:558). But when considering the relatively high number of usage

19. www.washingtonpost.com/ac2/wp-dyn/A42930-2002May6?language=printer, online February 11, 2003.

differences among other varieties of English in this area (cf. Sand 2003: 427–8; Algeo 2006: 43–56), IndE usage is not entirely novel but falls within a dimension of variability that is likely exploited throughout the English-speaking world.

5.1.2.2 *Zero article for indefinite article*

The student essays returned 38 examples of the zero article used in places where standard English would likely have an indefinite article, which equals 23.5% of all marked-up uses. Spread evenly across the various contexts of use usually associated with *a(n)*, the zero article turned out to be a recurrent but overall rare second option.

According to Quirk et al. 1985, the major function of the indefinite article is that of an unmarked article “in the sense that it is used for singular count nouns where the conditions for the use of *the* do not obtain” (272). In such environments, *a* and *an* are used when the reference to an entity cannot be recovered uniquely from the knowledge shared by speakers and hearers, writers and readers, most typically when a referent that has not been mentioned before is introduced in discourse. See examples (24) and (25) from the student essays, in which the zero article was found in such an environment:

- (24) ... it should be clear that those who are weak only have [AN] inferiority complex ... [IND2000 Ess 08]
- (25) So to do so, to develop to live [A] luxurious life people in India are adopting western culture in hope they would become advance [sic] like other developed country. [IND2000 Ess 06]

The indefinite article also takes over generic functions when it is used to pick out “any representative member of the class” (Quirk et al. 1985: 281). In examples (26) through (29), it is the zero article functioning this way, cf.

- (26) Presently the roads and houses of India are full of people dressed in Western style and speaking [A] western language, that is English. [IND2000 Ess 05]
- (27) The golden period of [A] human’s life is the childhood period. [IND2000 Ess 02]
- (28) In Childhood [sic] [A] child has all sorts of comforts and luxuries. [IND2000 Ess 01]
- (29) It is believed that [A] child is a true image of God. [IND2000 Ess 02]

The examples above resemble those provided by Platt et al. (1984: 54–5) to back up their claim that many new varieties of English, including IndE, do not employ articles to make a distinction between definite and indefinite reference but to communicate a distinction between specific and non-specific information (also cf. Sand 2003: 417; Sharma 2005: 538–9; Schneider 2007: 83). This is explained by the authors in terms of influence from languages that make a specific-non-specific distinction, for example

Table 5.9 (A) *lot of* [Primary Corpus, ICE-IND].

	IndE KOL 1978	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess	IndE ICE-IND writing	IndE ICE-IND speech
a lot of	82	2	10	32	4	33	127
lots of	12	–	–	7	–	10	25
lot of	4	–	–	5	4	5	113

Chinese.²⁰ A ‘non-specific’ reference according to Platt et al. 1984 is one to “persons, things, etc. [which] are not particular ones but belong to a group, type or species” or, more generally, to items “unknown to the speaker or writer”, “irrelevant to the issue” being discussed or “thought to be obvious” (54). While it cannot be ruled out that individual users of IndE may follow those rather general pragmatic principles, the rarity of such uses in the available corpus material makes it seem rather impossible that IndE at large is currently in the process of innovating its article system on a grand scale along those lines. It is, for instance, more likely that the zero article with *child* in examples (28) and (29) was used on the analogy of similar uses of *man* and *woman*, which may be used generically without articles (cf. Quirk et al. 1985:282).

Finally, standard English also uses the indefinite article in several set phrases, one of which is the quantifying expression *a lot of*. In IndE, the indefinite article is dropped regularly in this environment (cf. Parasher 1991:169). Consider example (30), in which the indefinite article has been omitted:

- (30) These foods are very nutritious and have [A] lot of calories and fats.
[IND2000 Ess 06]

Outside the student essays section, *lot of* is notably more frequent in the speech section of the Primary Corpus, whereas it is relatively rare in India’s quality newspapers. In ICE-India, too, *lot of* is quantitatively on a par with *a lot of* in the speech sections, but it is a minority variant in writing, cf. Table 5.9.

In published writing on the wider World Wide Web, however, the combination *lot of* is recurrent. From a comparative viewpoint, it is interesting to see that the omission of the article appears to be more frequent under the ‘in’ domain than under the other five domains consulted, at least in the collocation *with (a) lot of*. See the snapshot from the Internet (February 11, 2003) in Table 5.10.

Since the indefinite article no longer carries any communicative load within the confines of the quantifying expression *a lot of*, its omission does not severely hamper intelligibility, which might be one reason why the combination *with lot of* is common

20. Platt et al. (1984:55) give the following two examples from IndE: *It looks like cat* (It belongs to the species cat) and *Everyone has car* (No particular car or cars are being discussed).

Table 5.10 (A) *lot of* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
with a lot of	858	768	665000	33600	266000	126000
with lots of	704	11600	1510000	50200	358000	140000
with lot of	1140	310	12500	606	662	3680

(cf. Sharma 2005: 558). An additional internal reason for *lot of* to have spread extensively in IndE could be the simplification of the consonant cluster *ts* in *lots of*, which is not uncommon in IndE speech and might be carried over into the written domain in less heavily edited forms of writing. The relatively high proportion of *lot of* in the IndE speech sections of the Primary Corpus and in ICE-India points in this direction.

5.1.2.3 Definite article for zero article

Let us turn to those marked-up usages in which the definite article was found in contexts in which the zero article had been expected to occur. Making up approximately one third (27.2%) of all ‘different’ uses that had been tagged, this type is particularly interesting as its very occurrence provides a strong argument against the substrate hypothesis as sole explanatory framework for variation in IndE article usage.

In standard English, the zero article precedes noncount abstract nouns when they are used generically (cf. Quirk et al. 1985: 282, 286–7). In the IndE student essays, the definite article is also occasionally encountered in such environments. In examples (31) and (32), students employed the definite article before the noncount nouns *childhood* and *humanity* in generic use:

- (31) The childhood is that period of time when child learn more by experience than by precept ... [IND2000 Ess 01]
- (32) Women can work as nurses, engineers and surgeons and can prove to be a great help to the humanity. [IND2000 Ess 08]

Elsewhere in the Primary Corpus, both *childhood* and *humanity* show the regular properties of noncount abstract nouns, which makes examples (31) and (32) look rather exceptional in the big picture and suggests learner difficulties in distinguishing between generic and nongeneric uses of those nouns as the main factor behind the variability of IndE observed here.

The zero article is also found in standard English with plural nouns in generic use to identify classes of persons, things, or abstractions that are considered an “undifferentiated whole” (Quirk et al. 1985: 282). Thus, while the statement *cigarettes are bad for your health* would be interpreted as a statement about the dangerous consequences of smoking cigarettes in general, the use of the definite article in a statement like *the cigarettes are bad for your health* would add a sense of specificity that can fully be interpreted only when taking the wider context of the statement into account. See the

next set of examples, in which the definite article, according to my reading of the data, was used generically in the IndE student essays:

- (33) In old times the people being narrow minded imposed some evils in the society like caste restriction. [IND2000 Ess 05]
- (34) Under the influence of this vulgar western culture boys go to the pubs and discos. They dance with the girls, drink most expensive wines ... [IND2000 Ess 04]
- (35) Boys find it embarrassing to wear the traditional kurta-pyjamas and go after funky jeans ... [IND2000 Ess 07]
- (36) ... India has produced a number of sound spinners in the recent years. [IND2000 Ess 04]

Note that with the choice between *the* and *zero* primarily motivated pragmatically, it is not a straightforward task to explain the unusualness of those examples out of the larger contexts in which they originally occurred. Example (33) was found in a discussion essay on whether the adoption of western culture was a sign of advancement. Since the student's aim was to compare present developments in Indian society to past developments, it would arguably have been more effective to leave out the definite article before *people*, as the definite article adds a sense of specificity to a statement that is meant to describe a general fact about Indian society. On the other hand, the definite article may be said to function pragmatically as a softener avoiding the directness and harshness of stating a de-contextualized fact, especially since the noun phrase *the people* introduces new information within the given discourse context – a situation which, according to Sharma 2005, may trigger uses of surface articles more readily than other contexts of use for the purpose of disambiguation (cf. Chapter 5.1.2.1). Whatever pragmatic motivation the student may have felt when writing the essay cannot be reconstructed in hindsight. In examples (34) and (35), too, the choice of the definite article adds a specifying sense that according to my reading of the data does not fit the given contexts. *The* precedes plural nouns in statements that were made about classes of places (*pubs and discos*), people (*girls*) and objects (*kurta-pyjamas*) as undifferentiated wholes, for which the zero article would arguably have been the standard choice. But again, even if the choices made by the authors of the essays must be considered unusual by written usage standards, they are grammatically not incorrect. Note that in examples (34) and (35) both students vacillate freely between the definite and the zero article within one sentence, which may be indicative of some general insecurity in using definite articles on the part of the individual users. Example (36) is arguably the most unusual example in this context, as the phrase *in recent years* has risen to the status of a set phrase that has already become codified in major dictionaries of contemporary English (e.g. in COB), which makes *in the recent years* sound rather unfamiliar from the viewpoint of standard English.

Table 5.11 Definite/zero article with *police* [Primary Corpus; press].*

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
the police	18	21	6	6	1	4
ZERO police	3	7	8	9	16	8

* Excluded from the count were uses of the noun *police* in proper names and compounds (e.g. *police sources*) as well as premodified or postmodified uses of the noun (e.g. *another police*, *the best police*, *the police of Delhi*).

An interesting example of IndE displaying a stable variety-specific profile vis-à-vis BrE and AmE is found with the noun *police*. When used in a collective sense, *police* ‘the police force’ may nowadays take both the zero article and the definite article by standard English norms. Interestingly, however, IndE today is more open toward employing the definite article than the prestige varieties are. See first example (37) of *police* used in a collective sense and preceded by *the*, which was found in a student essay:

- (37) South African captain Hansie Cronje agreed of taking [sic] money after the Delhi police investigated into the matter. [IND2000 Ess 04]

In the press and broadcast sections of the Primary Corpus, uses of both the definite and the zero article with collective *police* are attested, cf.

- (38) <\$A><#>the police fired thirty-two thirty-five rounds at four places on Monday<,> in which one woman died and three others sustained injuries [IND2000 Broad News 03]
- (39) Police fired eight rounds and lobbed more than a dozen tear-gas shells to scatter the 1000-strong mob, which gathered outside the Yamuna Pushta police post on Friday night. [IND2000 Press Rep 11]

On comparing the IndE, BrE and AmE press texts, noteworthy distribution of the two variants emerged, cf. Table 5.11. Although the absolute numbers are small, it is clear that IndE journalistic writing displays a diachronically stable preference for the definite article with *police*, whereas the zero variant outnumbers its alternative in BrE and AmE newspapers.²¹ This difference is also reflected in the broadcast data, which suggests that we have come across a mild usage difference between IndE and the major varieties that goes beyond the confines of Indian press English, cf. Table 5.12. This

21. According to Sand (1998: 130), Jamaican English, too, favors the definite article with *police* ‘police force’ by a ratio of 30:10 in her press material and by a ratio of 30:2 in her radio broadcast material.

Table 5.12 Definite/zero article with *police* [Primary Corpus; broadcast].

	IndE IND2000 Broad	BrE UK1991 Broad
the police	21	4
ZERO police	11	47

Table 5.13 Definite/zero article with *police* [ICE-IND].

	IndE ICE-IND writing	IndE ICE-IND speech
the police	43	113
ZERO police	14	31

Table 5.14 Definite/zero article with *police* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
arrested by THE police	123	382	83	41	19700	59	106	9
arrested by police	24	59	67	42	37	42	335	107

view was confirmed when consulting ICE-India. The definite article with *police* 'police force' is nowadays clearly preferred in IndE across registers and text categories, cf. Table 5.13.

When looking up the two collocations *arrested by police* and *arrested by the police* in the online newspaper corpora from India, Britain and the US (February 12, 2003), even more evidence was produced suggesting that the definite article with *police* was more common a choice in IndE than in BrE or AmE, cf. Table 5.14.

Across the top-level country domains, the definite article with collective *police* is particularly common in IndE, but also in AmE (February 12, 2003), cf. Table 5.15.

While the use of the definite article with *police* (used in a collective sense), hence, does not constitute a categorical difference among the varieties looked at, contemporary IndE does stand out in displaying a comparatively strong liking for the definite article. Small as it is, this finding is noteworthy as it represents a stable IndE usage preference for a surface article in a context where BrE and AmE would prefer zero. But rather than reflecting an independent innovative trend in IndE, the observed difference may be a repercussion of a more general development in Englishes worldwide concerning the grammatical properties of the noun *police*. As shown elsewhere in this study (cf. Chapter 5.3.2), the major varieties BrE and AmE have moved *police* 'the po-

Table 5.15 Definite/zero article with *police* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
arrested by THE police	141	110	14800	390	526	383
arrested by police	123	133	24700	449	766	336

lice force' toward the status of an unmarked plural noun only over the past 150 years. Nowadays, *police* almost exclusively takes plural verbal and pronominal concord. In the course of this change, the zero article with collective *police* appears to have become integrated in BrE and AmE to a comparatively high degree to talk about police authorities as an undifferentiated whole. IndE does not seem to have followed the same path and, to date, remains more variable in using *police*. As the evidence from the corpora clearly shows, IndE users today employ both singular and plural verbal and pronominal concord with collective *police* in notably high numbers. This inherent flexibility in putting *police* to use might also be a reason why the definite article is still more of a grammatical option in IndE.

Another noun for which IndE article usage is slightly more variable today than BrE or AmE usage is *society*, which usually takes the zero article in standard English when used generically to denote 'people in general, thought of as a large organized group'. Consider the following three examples. In each, *society* has been used in a generic sense but has been preceded by the definite article.

- (40) He starts following the trend of the society and his innocence gets lost forever.
[IND2000 Ess 03]
- (41) Singing western songs means that a person is very litrate [sic] and has high standard in the society ...
[IND2000 Ess 06]
- (42) There have been cases when a deserving boy has not been made a prefect, for example, and in his place a less deserving boy has been appointed who probably exercises more influence in the society.
[IND2000 Ess 10]

A look at the broadcast section of the Primary Corpus shows that the zero article is also available in IndE, cf.

- (43) <\$B><#>no but fair is not I mean the other models obviously uh you know society likes people who are not just fair
[IND2000 Broad Disc 09]

When analyzing the Primary Corpus more closely for the choice of articles with *society* denoting 'people in general', IndE turned out to be more variable than BrE and AmE in this instance as well, although the figures returned from the corpus were too small to draw any definitive conclusions. While *the* was not attested for BrE and AmE, it was a recurrent second variant in the IndE data from 1978 and 2000; cf. Table 5.16.

To see whether or not the distribution in the small Primary Corpus was coincidental, the frequencies of the collocations *in (the) Indian society* and *in (the) British*

Table 5.16 Definite/zero article with *society* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
the society	3	3	–	–	–	–
ZERO society	16	5	–	5	3	9

Table 5.17 Definite/zero article with *society* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
in Indian society	56	236	17	43	134	65	6	8
in the Indian society	16	27	7	15	54	4	–	–
in British society	6	15	2	5	8	6	210	15
in the British society	–	–	4	2	6	–	–	–

society in the online newspapers (Google search dated February 15, 2003) were determined. Again, IndE had more definite articles than BrE and AmE, cf. Table 5.17. While no attestation was found for the definite article in the online editions of the *Guardian* and the *Washington Post*, both the zero and the definite article were recurrent variants in the Indian newspapers, with the zero article being the more common choice throughout.

A similar picture emerged when browsing the top-level country domains on the Web (February 15, 2003), cf. Table 5.18. Interestingly, Table 5.18 indicates that the collocation *in the British society* is particularly rare in BrE. To BrE users, the combination *British society* is inherently definite and apparently needs no further surface marker of definiteness. The definite article seems to be more of a choice in BrE when using *society* to refer to (and specify) foreign territories. IndE users, by contrast, are less rigorous in their use of the zero article when employing *society* to refer to their own country and people, which implies that the combination *Indian society* is not as inherently definite in IndE as *British society* is in BrE. Hence, we have come across another subtle, but distinct pragmatic difference seeing the definite article used slightly more frequently in IndE than in BrE and AmE.

From a more global perspective, it is clear that the use of *the* with *society* is not limited to IndE alone. According to Sand 2004, who looked at various sub-corpora of ICE, contemporary Jamaican English, Kenyan English and Singaporean English show similar degrees of variability:

In the category of generically used collective nouns, *society* is of special interest because the usage with and without article is almost evenly distributed in ICE-JAM [Jamaica] (zero: 86/*the*: 86), ICE-IND (zero: 114/*the*: 90), and ICE-EA(K)

Table 5.18 Definite/zero article with *society* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
in Indian society	495	19	669	69	204	323
in the Indian society	167	2	85	6	41	35
	3 : 1	10 : 1	8 : 1	12 : 1	5 : 1	9 : 1
in British society	26	7	44100	120	246	345
in the British society	3	1	218	–	8	34
	9 : 1	7 : 1	202 : 1	–	31 : 1	10 : 1

[East Africa: Kenya] (zero: 151/*the*: 135). Even in ICE-SIN [Singapore], one-fifth of the occurrences of *society* with generic meaning are preceded by a definite article. The situation is less conclusive for the other lexemes under analysis.

(Sand 2004: 292)

On interpreting why varieties of English, in particular contact varieties, should be quite variable in this case, Sand 2004 refers to the “strong link in linguistic typology between the concept of definiteness and animacy and humanness” (295), which might make nouns referring to human beings and their activities particularly likely to be seen as definite in new varieties of English and, hence, be preceded by a definite article (in contrast to inanimate nouns). But variability in article use in phrases like *Indian society* or *British society* also touches on a dimension of variability inherent in the standard English, namely the distinction between definite descriptions and names. Quirk et al. (1985:294) distinguish between what they call definite descriptions (*the Oxford road* ‘the road to Oxford’), which take an initial definite article, and names (*Oxford Road*), which do not. Since the boundary between the two categories is not clear-cut, as Quirk et al. (1985:295) admit, variation across the *the* vs. zero divide is a common feature of standard English, and the question of whether the zero or the definite article is used in such contexts becomes primarily a matter of pragmatic convention. Evidence from the student essays suggests that IndE users might generally be more likely to use the definite article in such contexts than AmE or BrE users. Consider first examples (44) and (45), which illustrate the use of the zero article and the definite article in reference to one of India’s most famous boarding schools, (*the*) *Doon School*, cf.

- (44) For example the Doon School only caters to boys, gives no chance to girls.

[IND2000 Ess 09]

- (45) At present this school - I am referring to Doon School - is an all boys school.

[IND2000 Ess 09]

Variation of this type can also be observed in the IndE broadcast material. Both *the* and zero may precede *Breach Candy Hospital*, one of India’s most prestigious hospitals, cf.:

Table 5.19 Definite/zero article with names [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
the + name	4	8	4	1	3	–
ZERO + name	5	9	6	5	23	4

- (46) <\$B><#>doctors who constantly monitored the progress of Mr Vajpayee at the Breach Candy Hospital in Mumbai<,> were satisfied with the good recovery Mr Vajpayee has made<,> and found him fully fit to be discharged from the hospital today [IND2000 Broad News 04]

- (47) <\$A><#>a fortnight after his successful knee-replacement surgery<,> Prime Minister Atal Bihari Vajpayee was discharged today from Breach Candy Hospital [IND2000 Broad News 04]

To put those observations in a wider context, all proper nouns denoting hospitals and educational facilities in the press sections of the Primary Corpus were identified, with *school*, *university*, *hospital* being the key words in the search queries. In this sample, IndE has a slightly higher proportion of the definite article than BrE and AmE, cf. Table 5.19. The results from the broadcast section point in the same direction, speaking of a more pronounced liking for the definite article in IndE than BrE, cf. Table 5.20.

If the use of names for hospitals, schools and universities is representative for the treatment of names in IndE in general, IndE users can be said to employ the definite article more often with proper names than their BrE and AmE counterparts. Note that variability in IndE article usage along the gradient from definite descriptions to names also extends to names which have become fully integrated in BrE or AmE, such as *Harvard University*, *Cambridge University* or *Buckingham Palace*, cf.

- (48) In the study they have observed and analysed over a long time the health of 11130 men who matriculated at the Harvard University between 1916–1950. [The Tribune; December 3, 2000²²]
- (49) Dr. Koshy is now pursuing a Ph.D course in geriatrics at the Cambridge University on scholarship. [The Hindu; January 25 2002²³]

22. www.tribuneindia.com/2000/2000/12/03/spectrum/fitness.htm, online February 8, 2003.

23. www.hinduonnet.com/2002/01/26/stories/2002012603130300.htm, online February 8, 2003.

Table 5.20 Definite/zero article with names [Primary Corpus; broadcast].

	IndE IND2000 Broad	IndE UK1991 Broad
the + name	7	1
ZERO + name	1	3

Table 5.21 Definite/zero article with names [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
to the Breach Candy Hospital	2	15	1	–	8	2	–	–
to Breach Candy Hospital	9	11	–	5	17	8	–	–
at the Harvard University	1	8	2	1	4	5	–	1
at Harvard University	25	55	12	31	56	22	159	372
at the Cambridge University	–	4	3	4	4	3	–	–
at Cambridge University	14	68	14	9	43	18	620	67
at the Buckingham Palace	–	17	–	1	8	2	1	–
at Buckingham Palace	12	34	15	54	33	27	716	132

- (50) [Caption:] Queen Elizabeth II greets England captain David Beckham and manager Sven-Goran Eriksson during a reception for the Football Association at the Buckingham Palace Tuesday. [*The Telegraph*; November 21, 2002²⁴]

See the distribution in Table 5.21 for empirical evidence (Google search dated February 8, 2003).

An interesting reading of this phenomenon is offered by Chesterman 1993, who suggests treating variation between *the* and zero article in proper names within a discourse-pragmatic framework in terms of users' variant degrees of familiarity with the places referred to. According to Chesterman 1993,

[w]e say *Buckingham Palace* but *the Mariinsky Palace* ... This appears to be because the English building is more familiar to English people, who are therefore "insiders" with respect to it. In some sense, then, familiar buildings are felt to have clearer conceptual outlines; and herein lies the key to the difference between null and *the*. *The* gives a concept bounds, limits; it concretizes, defines. But if a concept is already familiar enough, clear enough in and of itself, it does not need such extraboundedness. (Chesterman 1993: 20)

24. www.telegraphindia.com/1021121/asp/sports/story_1405916.asp, online February 8, 2003.

It is therefore possible, Chesterman 1993 argues, that users of English as a second or foreign language should feel a particularly strong communicative need to specify their referents. In Chesterman's (1993) words, "non-native speakers may feel somehow less confident that their referents will be understood correctly and therefore tend to over-specify, to over-emphasize that their referents are indeed "familiar"" (Chesterman 1993:20). Though further cross-varietal research will be necessary to address this claim more systematically, the data presented in this chapter lend quantitative support to his view for the case of IndE and provide another argument against the general validity of the 'underuse' hypothesis for IndE article use.²⁵

While the relatively frequent use of the definite article with *police*, *society* and with proper names thus distinguishes IndE mildly from both BrE and AmE, contemporary IndE must be said to stand between the two major varieties in the case of another noun denoting "'institutions' of human life and society" (Quirk et al 1985:277), *hospital*. In standard English, *hospital* may be preceded by the zero or by the definite article when used in a 'medical treatment' context. AmE has been noted to show a preference for the definite article, whereas BrE has been claimed to prefer zero (Quirk et al. 1985:277). Interestingly, IndE shows a high degree of variability in the choice between zero and definite article without displaying any clear-cut preference for the definite article, as might be expected from the findings on *police* and *society*, for which the definite article was more of an option in IndE than in BrE or AmE. In the essays section, only one example of *hospital* used in a 'medical treatment' sense is attested, which is preceded by *the*, cf.:

- (51) A patient is admitted in the hospital so that he may come back home fit and fine. [IND2000 Ess 09]

In the IndE broadcast section of the Primary Corpus, both variants are encountered, cf.

- (52) <\$A><#>the Home Minister Mr L-K Advani stayed in the hospital throughout the morning [IND2000 Broad News 03]
 (53) <\$L><#>Mr Maran has been in hospital for the past eleven days [IND2000 Broad News 12]

A look at the press sections of the Primary Corpus suggests that while BrE and AmE users have opposing preferences, IndE usage is divided in that particular instance, cf. Table 5.22.²⁶ A similar picture emerges for the broadcast data, cf. Table 5.23. In ICE-India, the preference is clearly with the definite article, cf. Table 5.24. On the Internet,

25. The recurrent use of the definite article with proper names has also been observed for other varieties of English, cf. Mair (1992b:88–9) and Sand (2004:291) on Jamaican English.

26. Excluded from the count were proper names, references to buildings and institutions, compounds (*hospital sources*) and premodified nouns (*another hospital*, *the best hospital*).

Table 5.22 Definite/zero article with *hospital* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
the hospital	2	1	–	–	5	1
ZERO hospital	1	1	4	5	–	–

Table 5.23 Definite/zero article with *hospital* [Primary Corpus; broadcast].

	IndE IND2000 Broad	BrE UK1991 Broad
the hospital	2	–
ZERO hospital	5	2

Table 5.24 Definite/zero article with *hospital* [ICE-IND].

	IndE ICE-IND writing	IndE ICE-IND speech
the hospital	4	30
ZERO hospital	1	7

however, IndE vacillates freely between *the* and zero article. See, for example, the distribution of the collocations *admitted to (THE) hospital* across India's online newspaper corpora (February 12, 2003); cf. Table 5.25. Under the top-level country domains, IndE usage falls in between BrE and AmE (February 12, 2003), cf. Table 5.26.

Interestingly, according to Table 5.26, SgE displays a behavior similar to that observed for IndE. In AusE and SAfrE, on the other hand, the preference for zero is much more pronounced, much like in BrE. Whether this distribution is the result of more recently exerted influence from AmE on IndE (and other South and South East Asian varieties of English) or the result of Asian-specific pragmatic conventions not sanctioning the use of the definite article with nouns having generic meaning as rigorously as BrE does cannot be reconstructed on the basis of the available evidence, of course. What is clear is that South and South East Asian English usage must be positioned between BrE and AmE.²⁷

Let me conclude by pointing out an idiomatic context in which IndE has come to prefer *the* to the zero article in contradistinction to the major varieties. Of the proverbs

27. Cf. Sand (1998: 132), who observed in Jamaican press and broadcast English “a clear adherence to British usage” in the case of *hospital*.

Table 5.25 Definite/zero article with *hospital* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
admitted to the hospital	63	234	55	40	627	163	41	91
admitted to hospital	61	253	63	47	436	144	298	24

Table 5.26 Definite/zero article with *hospital* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
admitted to the hospital	126	221	15200	360	887	19700
admitted to hospital	208	374	207000	11500	71600	554
the : (the)	1 : 2	1 : 2	1 : 14	1 : 32	1 : 81	36 : 1

and sayings listed in Yadurajan (2001: 80–2) as examples of deviant article use in IndE, the expression *last but not the least* has evolved as a preferred variant, whereas AmE and BrE prefer *last but not least*. See a representative example from IndE:

- (54) In the pages of that history, the names of Siraj-ud-Doula and Mohanlal of Bengal, Haider Ali and Tipu Sultan and Vellu Thampi of South India, Appa Sahib Bhonsle and Peshwa Baji Rao of Maharashtra, the Begums of Audh, Sardar Shyam Singh Atariwala of Punjab and last but not the least Rani Lakshmibai of Jhansi, Tantia Tope, Maharaj Kanwar Singh of Dumraon, Nana Sahib are for ever engraved in letters of gold ... [IndE KOL G 31]

In spite of prescriptive warnings, the expression has become firmly established in IndE journalistic writing (Google search dated February 12, 2003), cf. Table 5.27. But a look at the top-level country domains (February 12, 2003) reminds us that differences among world Englishes must not to be considered absolute, cf. Table 5.28. However, IndE does stand out with its noticeable preference for the lengthier variant. It is possible that we have come across an IndE innovation that might have come into being by means of analogy, as the ‘logical’ use of the definite article for specific reference shows great affinity to superlative forms (cf. examples (9) and (10) above). In this regard, the variant *last but not the least* also resembles standard English set phrases such as *to say the least* or *in the least*.²⁸

5.1.2.4 Indefinite article for zero article

The discussion now turns to the rare cases in the IndE student essays in which the indefinite article was used in places where the zero article had been expected to

²⁸ ICE-India returned one example of *last but not the least* (text W2D-001) but no example of *last but not least*.

Table 5.27 *Last but not (the) least* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
last but not least	17	19	4	1	30	1	39	31
last but not the least	33	41	18	10	138	16	–	–

Table 5.28 *Last but not (the) least* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
last but not least	582	13700	245000	21000	67600	32400
last but not the least	822	355	818	204	237	746

occur. Additional indefinite articles for the zero article were not common among the marked-up cases (4.9%), and the few instances that were found in the student essays represented little more than learner difficulties. In example (55), a student employed indefinite articles in the context of an enumeration:

- (55) A human being has to go through three important phases of life such as baby, an adult and lastly an old age. [IND2000 Ess 01]

As the nouns *baby*, *adult*, and *old age* function as names given to the stages of human life, they would not be preceded by the indefinite article in standard English. Similarly, the indefinite article in the next example is superfluous, as the student refers to the lexical item *woman* as a vocabulary word, cf.

- (56) We all know the fact that when we add WO to MAN then you get a woman ... [IND2000 Ess 08]

In the next set of examples, the zero article would arguably be more likely a choice by the standards of international English, as the nouns in question are used in noncount senses.

- (57) Our traditions stop inter-caste marriages because two people following different traditions cannot live together peacefully but this cause divides humans and brings an ill feeling for others. [IND2000 Ess 07]
- (58) Men are the weaker sex. History is a proof to the above said sentence. [IND2000 Ess 08]
- (59) Mind you, this wasn't for an official work but for a mere holiday. [IND2000 Ess 10]

5.1.2.5 Exchange of surface articles

Similarly, the exchange of surface articles is a minor phenomenon in the essays. Rather than indicating any variety-specific usage patterns, the marked-up examples speak of a certain sense of insecurity in article usage that is likely found in student writing around the English-speaking world. In five instances (or 3.1% of all marked-up forms), *a* was used for *the*, as in example (60):

- (60) “Childhood is the golden period of life. All of us are aware of this fact. “Childhood” is a first stage of life which is full of joy & glory.

[IND2000 Ess 01]

The would arguably be more appropriate a choice here as the concept of ‘first stage of life’ refers to a uniquely identifiable phase in life, which makes the use of the indefinite article sound rather awkward in this particular context by the norms of standard English. In contrast to this, the indefinite article would have been more appropriate in the next example, as it becomes clear from reading the entire essay that the student’s intention was to communicate the idea that drinking alcohol is one of many signs of ‘smartness and advancement’ and not the only (or most important) one, cf.

- (61) Today, drinking wine is considered to be the sign of smartness and advancement but we know how advanced they are ...

[IND2000 Ess 05]

Example (61) is one of only four instances in the essays (totalling 2.5% of all marked-up forms) in which *the* was used for *a(n)*, which contrasts with Agnihotri et al.’s (1984: 118) claim that “[t]he tendency to use the definite article in place of the indefinite article is relatively strong.” Moreover, the rules governing the use of *a* and *an* were violated in an additional six cases. As compared with 592 standard uses of the indefinite article, those deviations from the norms of standard English represent foremost a common learner difficulty:

- (62) ‘Life is like an photograph so always smile.’

[IND2000 Ess 01]

- (63) People who want to study and go in for competitive exams should be given a option about playing sports and studying during that time.

[IND2000 Ess 10]

Interestingly, nowhere in the essays has *one* been found to replace the indefinite article *a(n)* (cf. Taylor/Rao 1981: 141; Spitzbardt 1987: 130; Leitner 1988: 390). Hence, substituting *one* for *a*, which has been said to be reinforced by the influence from Indian languages (see, for example, Ramamurthi 1991 on interference from the Tamil numeral ‘oru’), does not qualify as a characteristic feature of IndE at large, although it may be more common in less heavily edited forms of speech and writing. A similar conclusion was drawn by Sharma 2005, whose investigations point in the same direction:

the common claim that Indian speakers use *one* in place of *a* with specific indefinite NPs [noun phrases] was not supported. Although individuals showed a noticeably higher rate of use of *one* in standard (emphatic, specific) reference than native English speakers might have, very few instances were clearly non-standard ... (Sharma 2005: 551)

5.1.3 Summary

How systematic is the use of articles in South Asian English? Evidence from the corpora suggests that contemporary IndE, within the range of texts covered by this study, broadly conforms to the rules of standard English article use as formulated by Quirk et al. 1985. This also means that IndE cannot be said to generally overuse or underuse the definite or the indefinite article, which puts a question mark on the validity of any description of IndE that seeks to capture the nature of article usage in IndE in any such simplistic, one-dimensional terms. In the IndE student essays, the use of zero articles in place of surface articles was the most frequently observed usage difference vis-à-vis standard English. The zero article was found to occur in all environments that are usually reserved for the definite article and the indefinite article in standard English. But rather than pointing to a single influential factor, the data showed several forces to be responsible for creating variation simultaneously, for example innovative tendencies at the discourse-pragmatic level and general learner difficulties in dealing with the standard English article system. Interestingly, there are several contexts of use in which IndE nowadays consistently employs the definite article more frequently than BrE and AmE do, most noticeably when using proper names, nouns denoting institutions of human life (e.g. *police*, *society*) and the lexicalized phrase *last but not the least*. This puts any theory into place that would predict the non-occurrence of surface articles to be a general characteristic feature of IndE due to the influence from substrate languages such as Hindi or Tamil. What the corpus data show instead is that users of IndE exploit the inherent possibilities of the English article system freely to meet their specific communicative needs, which has resulted in stable modifications to the codified norms of the standard English article system in ways not uncommon for other varieties of English either (cf. Sand 2003, 2004). But there is no evidence in the data that would support the view that IndE is currently undergoing any fundamental restructuring of its article system.

5.2 Noun phrase: Countability and number

Next to variation in article use, variability along the count-noncount and singular-plural divides has been claimed to be a peculiarity of the IndE noun phrase. IndE has been said to treat noncount nouns as count nouns (e.g. *equipment*; *equipments*), introduce morphological number contrast for unmarked plurals (e.g. *aircraft*; *two*

aircrafts), use plural nouns in the singular (e.g. summation plurals; *a pant, a trouser, a pajama*), and treat singular nouns as variable nouns (e.g. *news*). One of the earliest comments on IndE variability is found in Whitworth 1982 [1907], who remarked that “[n]o grammatical mistakes are made by Indian writers in the use of nouns in English, except in respect of number” (85). In his view, “it is a comparatively common mistake to put into the plural nouns which, according to native usage, have no plural in the particular sense in question” (85). Decades later, Daswani 1978 noted “the wrong use of plural -s in *disbursements* and *representations*” (121). When Nihalani et al. 1979 published the first edition of their usage handbook, they spoke of “the propensity of many speakers of IVE [Indian variants of English] to prefer plural forms where most BS [British Standard English] speakers would select the singular” (68) – a view also shared by Sahgal/Agnihotri 1985, who stated that “Indian speakers have a tendency to pluralize the nouns in all the contexts” (125). According to Parasher 1991, too, “[o]ne of the major sources of deviation in the noun phrase ... is that certain non-count nouns such as *equipment, evidence, bread, advice* are treated as count and therefore pluralised to convey the plural meaning” (167).²⁹ Given this plethora of comments over the years, there was little doubt that IndE was variable across the count-noncount and singular-plural divides, not least because other varieties of English, including the prestige varieties BrE and AmE, had been shown to vary considerably in this area of the language system (cf. Platt et al. 1984: 46–65; Algeo 2006: 43–104; Schneider 2007: 85). Furthermore, since Sahgal/Agnihotri 1985 had uncovered that variation in the noun phrase regarding countability and number was among the “somewhat acceptable” (cf. Sahgal/Agnihotri 1985: 126–7) grammatical features, it appeared all the more reasonable to hypothesize that many of the forms that had come to the attention of researchers over the past one hundred years might have turned into stable usage patterns by the turn of the millennium.³⁰

29. Also cf. Parasher (1994: 148), Hansen et al. (1996: 224), George C. (1998: 16), Yadurajan (2001: 109–11).

30. The following items were tested for acceptability by Sahgal/Agnihotri 1985: “He performed many charities in the course of his career” (Place in implicational hierarchy: 9; Good: 48.84% - Informal usage: 20.93% - Wrong English: 30.23%); “She’s a tease, she loves to pull your legs” (Place in implicational hierarchy: 12; Good: 23.26% - Informal usage: 53.49% - Wrong English: 23.26%); “In his hearts of hearts he knows he shouldn’t have done it” (Place in implicational hierarchy: 7; Good: 54.55%; Informal usage: 27.27%; Wrong English: 18.18%). The results by Sahgal/Agnihotri 1985 are borne out by Sheorey’s (2006: 191) ‘error gravity survey’, according to which variability across the singular/plural divide is one of the more acceptable grammatical phenomena. The following sentences were tested for acceptability by Sheorey 2006: “She is always finding faults with me”, “Please don’t pull my legs” and “They sell old equipments.”

5.2.1 Quantitative analysis

At the outset of this case study, a list of variables was compiled containing nouns that had been discussed in reference to IndE in the two most recently published reference handbooks on IndE (Nihalani et al. 1979, 2004 and Yadurajan 2001) as displaying unusual variation along the count-noncount and singular-plural divides. After adding thirteen more nouns and expressions from Quirk et al. 1985 and Parasher 1991, the final list included 78 items whose behavior was then studied in the available corpora.³¹ Using the corpus-based COB dictionary as benchmark for established count, non-count, singular and plural uses in standard English, I scanned the Primary Corpus and the Kolhapur Corpus for authentic IndE examples of unrecorded noun uses. As the majority of nouns belonged to the category of dual-class membership nouns, for which both count and noncount uses were available in standard English, a close reading of the corpus material was necessary at this stage to uncover possible grammatical and/or semantic differences between IndE and the codified standard. All in all, 23 of the 78 items that had been mentioned as Indianisms displayed the nonstandard behavior ascribed to them at least once. To assess the status of those attested unrecorded uses in contemporary IndE more succinctly, the Internet and ICE-India were consulted in a next step. For additional diachronic information, the OED quotations base was tapped.

Consider first the stratification of those 23 nouns and expressions that were encountered in the data in patterns not codified in COB. A 'plus' in Table 5.29 indicates at least one 'different' (that is to say, non-codified) use of the noun or expression in question (e.g. *a traffic, homeworks, aircrafts*) in the respective sub-corpus, a 'minus' a lack thereof.

As Table 5.29 illustrates, the majority of unusual noun uses in the data not complying with the COB is found in the one-million-word Kolhapur Corpus, the Primary Corpus being much too small to allow for any major insights into the availability of the 'different' noun uses in question. When considering further that even in the Kolhapur Corpus the token frequencies of the 'different' uses are always lower than those of the standard uses, one is left with little interpretative leeway apart from stating that contemporary IndE largely follows the codified conventions. Nonetheless, the stratification of the few unrecorded cases across the Kolhapur Corpus and the Primary Corpus deserves a closer look. While no substantial difference can be posited to exist between speech and writing or between learner English and more professional uses of English, it turns out that Indian press English around the year 2000 was slightly more open toward unrecorded noun uses than Indian press English twenty-two years before. While not a single nonstandard noun use is found in the press material from 1978, six such examples can be found in the material dating back to the year 2000. Although this finding is beyond statistical significance, it corresponds to the distribution of other uncoded uses in IndE quality newspapers, for example that of

31. Cf. Appendix 5.7 for a complete list of variables.

Table 5.29 Unrecorded count uses [Kolhapur Corpus, Primary Corpus].

	IndE KOL 1978	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess
content	+	–	+	–	–
employment	+	–	+	–	–
equipment	+	–	+	–	–
legislation	+	–	+	–	–
agitation	+	–	–	–	–
aircraft	+	–	–	–	–
anger	+	–	–	–	–
bread	+	–	–	–	–
co(-)ordination	+	–	–	–	–
damage	+	–	–	–	–
furniture	+	–	–	–	–
homework	+	–	–	–	–
in all probability	+	–	–	–	–
in detail	+	–	–	+	–
information	+	–	–	–	–
machinery	+	–	–	–	–
time(s) immemorial	+	–	–	–	–
traffic	+	–	–	–	–
training	+	–	–	–	–
tuition	+	–	–	–	+
advice	–	–	+	–	–
resistance	–	–	+	–	–
at all cost(s)	–	–	–	+	–
	20 types	–	6 types	2 types	1 type
	500 texts	40 texts	40 texts	40 texts	10 texts
Types per text	0.04	–	0.15	0.03	0.10
Types per 1,000 words	0.02	–	0.08	0.02	0.05

unrecorded verb-particle combinations (cf. Chapters 4.1 and 4.2).³² Hence, Table 5.29 provides more evidence for the claim that the grip of standard English on Indian quality newspapers was loosening between 1978 and 2000. Let us take a closer look at the ‘different’ noun uses next.

32. Cf. the results on informal loan vocabulary (Chapter 3.1.1), unrecorded verb-particle combinations (Chapters 4.1 and 4.2), nonstandard past-perfect forms (Chapter 5.4) and nonstandard direct and indirect questions (Chapter 5.7).

5.2.2 Qualitative investigations

Quirk et al. 1985 label nouns “denoting an undifferentiated mass or a continuum” (246) noncount nouns. From a grammatical point of view, noncount nouns cannot be pluralized and are normally not used with the indefinite article. The majority of nouns differing from the codified norms in the IndE material are classified ‘noncount’ in COB, but also display the grammatical behavior of count nouns in that they are either preceded with an indefinite article or pluralized. Consider first the following set of examples:

- (64) A huge anger filled me. [IndE KOL K 46]
- (65) Let us therefore comprehend and grasp truly the divine purpose and by a coordination of all the tattvas in our lives forge them into a Unity and harmony with the original purpose of the Divine. [IndE KOL D 08]
- (66) Because of his vast knowledge of the West and having had a scientific training Coomaraswamy was eminently equipped to make, as reflected through his works, the “fusion of an Oriental intuition with the scientific discipline of the West.” [IndE KOL G 36]
- (67) His style was unique. It has, however, been a one-way traffic. [IndE KOL B 16]
- (68) Children in particular found her quite helpful in regard to crises arising from undone homeworks or the ill humour of the pandits of the primary school. [IndE KOL K 23]
- (69) All informations gathered from the agent’s report incorporating the surveyor’s findings will be important when liability is under consideration. [IndE KOL E 29]
- (70) The spice and its oil are used in baked goods like breades, cakes and pies in the Scandinavian countries. [IndE KOL E 20]
- (71) The provisions of this rule shall apply to the employees in scheduled employments other than agricultural employment. [IndE KOL H 32]
- (72) After the advent of the Portuguese, the talent of wood carving in Goa, was turned to making quality furnitures copying the Western style and retaining the Indian influences, Goan furniture became a unique blend of Eastern and Western culture. [IndE KOL E 02]

In examples (64) to (67) the nouns *anger*, *coordination*, *training* and *traffic* are preceded with indefinite articles, which makes them similar in function to count nouns. According to Quirk et al. 1985, the use of the indefinite article in the given contexts must be considered exceptional but not impossible by standard English norms, which allow for the use of *a/an* with noncount nouns when “the noun refers to a quality or

other abstraction which is attributed to a person" (287) or when "the noun is pre-modified and/or postmodified" (287). Since both criteria apply in examples (64) to (67), IndE usage falls within what seems a general (albeit quite rare) possibility in standard English. The noncount nouns *homework*, *information*, *bread*, *employment* and *furniture* in examples (68) to (72), on the other hand, are pluralized, which would be unusual in standard English in the given contexts. How 'typically IndE' are such uses, though? Since the closed corpora were too small to allow further insights into their degrees of integration, the top-level country domains on the Web were consulted in search of more such examples to test the hypothesis whether or not IndE had semantically reclassified the nouns in question. However, none of the 'different' noun uses appeared exclusively in IndE in the collocations selected for those Google-based search queries, nor was any one such use notably more frequent under the 'in' domain than under the other top-level country domains consulted (Google search dated May 3, 2001); cf. Table 5.30.

Table 5.30, hence, offers no quantitative evidence that IndE usage should be considered to differ fundamentally from international usage conventions. In all of the chosen collocations, the 'different' uses attested in the structured corpora are recurrent in contemporary IndE, but their token frequencies are rather low when measured against the frequencies of their standard English alternatives. Moreover, since the 'different' variants are also found in other varieties, they seem to represent little more than the fuzzy edges of natural language use. On the basis of the available empirical evidence, one should therefore hesitate to claim that IndE has undergone any large-scale "semantic reclassification" (Platt et al. 1984: 50–1) of the noncount nouns in question. It is well possible, though, that more qualitative research taking into account the entire functional range of those nouns will reveal further stable IndE-specific usage patterns in the future.

Other nouns do display a more characteristically IndE variational profile on the Internet. *Legislation* 'a law or laws passed by a government', which has been discussed in reference to IndE noun grammar before (cf. Nihalani et al. 1979: 113, 199–200; 2004: 113), is a case in point. Labeled noncount in COB, *legislation* shows the characteristic features of a count noun in examples (73) and (74):

- (73) While the first one was an enabling legislation for fixing minimum standards of equipment, physical infrastructure and manpower in all medical institutions, the second was a legislation that would require all doctors to re-register with the medical council of India once in every five years, after undertaking a prescribed module of continuing medical education.

[IND2000 Press Rep 06]

- (74) In the matter of Harijans or the tribes or women constitutional safeguards, legislations and social organisations alone may not work well so long as there is hiatus between the word and the deed between the heart and the hand.

[IndE KOL B 24]

Table 5.30 Count and noncount nouns: selected collocations
[Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
in great anger	13	2	217	8	139	64
in a great anger	–	–	68	4	15	23
in great angers	–	–	–	–	–	–
	–	–	3 : 1	2 : 1	9 : 1	3 : 1
project coordination	369	259	12200	318	18600	26500
project coordinations	–	–	4	–	2	–
	–	–	3050 : 1	–	9300 : 1	–
research training	10700	604	411000	15900	370000	25200
research trainings	6	2	166	5	10	29
	1783 : 1	302 : 1	2476 : 1	3180 : 1	37000 : 1	869 : 1
sports training	1220	424	140000	552	24900	19600
sports trainings	1	13	2	–	78	1
	1220 : 1	33 : 1	120000 : 1	–	319 : 1	19600 : 1
one way traffic is	2	4	117	10	30	157
a one way traffic is	–	–	–	–	–	–
	–	–	–	–	–	–
my homework	725	1600	58400	1650	15900	72700
my homeworks	1	8	286	4	95	133
	725 : 1	200 : 1	204 : 1	413 : 1	167 : 1	547 : 1
gave information	126	49	27600	410	11300	9580
gave informations	–	–	7	–	1	–
	–	–	3943 : 1	–	11300 : 1	–
have bread	40	132	636	211	1020	759
have breads	1	2	27	1	7	2
	40 : 1	66 : 1	24 : 1	211 : 1	146 : 1	380 : 1
give employment to	212	4	820	109	167	191
give employments to	–	–	–	–	–	–
	–	–	–	–	–	–
furniture	143000	130000	11600000	479000	2830000	2350000
furnitures	487	2200	28100	1200	9270	16200
	294 : 1	59 : 1	413 : 1	399 : 1	305 : 1	145 : 1

While example (73) must be considered exceptional but acceptable according to standard English conventions (cf. Quirk et al. 1985:287), the plural *legislations* in example (74) is unusual. Further research on the Web across the top-level country domains brought to light that the plural *legislations* is a recurrent variant in contemporary IndE while being comparatively rare in BrE, in AmE and elsewhere in the English-speaking world, even though it is not unheard of there either. Consider the total token frequencies of the items *legislation* and *legislations* across top-level country domains (May 3,

Table 5.31 *Legislation(s)*.

	.in	.sg	.uk	.za	.au	.us
legislation	212000	151000	20600000	1360000	12800000	7910000
legislations	40700	650	48300	12800	84200	617
	5:1	232:1	4265:1	106:1	152:1	12820:1
enabling legislation	217	42	28200	9990	90700	130000
enabling legislations	20	1	7	21	120	33
	11:1	42:1	4029:1	476:1	756:1	3939:1
passed legislation	240	62	21400	546	45100	128000
passed legislations	13	–	1	–	2	3
	18:1	–	21400:1	–	22550:1	42666:1

2001) as well as the distribution of those two competing variants in the collocations *enabling legislation(s)* and *passed legislation(s)* in Table 5.31.

Even when bearing in mind that the distributions of the chosen collocations provide only a snapshot of the entire functional range of the noun *legislation*, Table 5.31 suggests quite clearly that the plural *legislations* is nowadays more common in IndE than in the other varieties. Why this should be the case remains open for debate. While independent IndE innovation by means of semantic reclassification is an ever-present possibility, it is also interesting to see that the OED quotations base contains two examples of *legislations* dating back to nineteenth-century British sources, which suggests divergent diachronic developments as an alternative interpretation for the observed regional differences. We might, in fact, be dealing with another usage feature that used to be more common in Englishes worldwide but has been shed in the major varieties of English.³³

Equipment ‘things which are used for a particular purpose’ falls in the same category as *legislation* in that the noun is labeled noncount in COB but is recurrently found in count uses in contemporary IndE. Consider examples (75) and (76) from the Primary Corpus and Kolhapur:

- (75) The CBI stumbled upon a sophisticated satellite communication equipment when it last week raided the premises of an NRI and thus unearthed a racket enabling people making international trunk calls at much cheaper rates than charged by the official telecom department. [IND2000 Press Rep 13]
- (76) Domestic articles or equipments like agricultural implements, a few earthen vessels for preparation and storing of rice-beer, cooking utensils, bamboo baskets, brooms made of a type of grass called Jono, one or two cots made of wood and rope of Sawai grass used as beds as well as for sitting, wooden

33. The two examples of *legislations* in the OED quotations base date back to 1828 and 1869. There is another example of *legislation* preceded with the indefinite article from 1838.

measures like Poila Topa mats, or Jati made of a type of grass called Kita etc., are found both in the houses of agriculturists as well as the industrial workers.
[IndE Kol H 11]

In example (75), *equipment* occurs with the indefinite article. Since the noun has been premodified extensively, such a use would probably be acceptable internationally (cf. Quirk et al. 1985:287). The plural *equipments* in example (76) is arguably more unusual. In fact, the distribution of *equipments* across the top-level country domains signals that although the form is by no means used exclusively in IndE today, IndE makes more frequent use of it than other varieties, at least as regards the collocations selected for the Internet search query (dated May 3, 2001), cf. Table 5.32.

In light of Table 5.32, *equipments*, like *legislations*, appears to be more of a choice in South and South-East Asia, with IndE and SgE showing a greater likelihood of using the plural form across the top-level country domains than BrE, AmE, SAfrE and AusE. Again, a look at the OED is informative as it suggests that the plural *equipments* might have been more common in nineteenth-century English. In the OED, count uses for *equipment* are attested for as late as 1873. Hence, rather than being a more recent independent Asian innovation proper, the count use of *equipment* could be another archaic form that has gone out of general use in the major varieties of English but has stayed on more firmly in the Asian region. Note, too, that more than two decades ago, Nihalani et al. (1979:75) observed for *equipment* “a tendency to use the plural form ... in Britain too.” Given the evidence presented in Table 5.32, *equipments* indeed seems slightly more common in BrE than in AmE today, but it is clearly less common than in the Asian region, particularly in IndE.³⁴

Like *legislation* and *equipment*, the noun *agitation* has been discussed by Nihalani et al. (1979:19, 2004:17) as an alleged peculiarity of IndE noun grammar. Labeled a noncount noun in COB, the item was found to co-occur with the indefinite article and to take the regular plural-s in the Indian corpus material, cf. examples (77) and (78):

- (77) The Congress Working Committee (CWC) member, Mr. Rajesh Pilot, today warned the Centre of a nation-wide agitation to highlight the plight of farmers.
[IND2000 Press Rep 19]
- (78) While conceding that public property should be protected, the opposition members felt that the enactment of this legislation would serve as a cover for suppressing all the agitations of the working class.
[IndE KOL A 15]

34. Parasher 1991 found *equipments* to be acceptable among his Indian informants, whereas it was rejected by his British and American informants. He concluded on the note that “[s]ince *equipments* appeared 5 times in the corpus its use seems to be rising.” (167) In Sheorey’s (2006:191) ‘error gravity’ survey, too, plural *equipments* was one of the more acceptable items differing from the codified norms of standard English.

Table 5.32 *Equipment(s)*.

	.in	.sg	.uk	.za	.au	.us
equipment	887000	557000	35200000	1760000	9580000	11000000
equipments	486000	27100	219000	12600	58100	30000
	2 : 1	21 : 1	161 : 1	140 : 1	1649 : 1	367 : 1
medical +						
equipment	957	16300	473000	24700	278000	263000
equipments	509	526	1890	44	563	778
	2 : 1	31 : 1	250 : 1	561 : 1	494 : 1	338 : 1
communication +						
equipment	682	700	67100	910	48800	74000
equipments	374	204	387	15	478	121
	2 : 1	3 : 1	173 : 1	61 : 1	103 : 1	612 : 1
sports +						
equipment	5910	4870	1680000	12000	116000	66300
equipments	257	834	21300	10	68	80
	23 : 1	6 : 1	79 : 1	1200 : 1	1706 : 1	829 : 1

Table 5.33 *Agitation(s)*.

	.in	.sg	.uk	.za	.au	.us
agitation	24900	771	254000	16400	153000	105000
agitations	13600	13	861	110	648	735
	2 : 1	59 : 1	295 : 1	149 : 1	236 : 1	143 : 1

According to Nihalani et al. 2004, “[t]he plural form of the noun does not occur in BS [British Standard English]” if the noun is to denote a ‘protest’ or a ‘rally’” (17). This is borne out by a look at the online corpora (May 4, 2001). Here, such plural uses are rare in BrE, AmE, AusE and SAfrE today while being much more frequent in SgE and especially in IndE, cf. Table 5.33.

Have we come across another echo of the past rather than an independent innovation? When looking up the plural *agitations* in the OED quotations base, count uses of *agitation* were indeed found to be common in nineteenth-century English, with the last attestation for pluralized *agitation* in the OED dating from 1880.

Two more nouns labeled noncount in COB are put to count uses more frequently in IndE today than in other varieties, namely *advice* ‘council’ and *tuition* ‘private lesson’. See examples (79) and (80) from the Primary Corpus:

(79) Before we begin, an advice. Buy this book.

[IND2000 Press Rev 01]

Table 5.34 *Advice(s)* and *tuition(s)*.

	.in	.sg	.uk	.za	.au	.us
advice	385000	379000	102000000	1720000	16200000	4270000
advices	16900	642	131000	872	171000	20300
	23 : 1	590 : 1	779 : 1	1973 : 1	95 : 1	210 : 1
gave advice	51	75	35000	322	13000	1140
gave an advice	2	3	9	–	15	1
	26 : 1	25 : 1	3889 : 1	–	867 : 1	1140 : 1
tuition	63000	64200	5890000	127000	1360000	1710000
tuitions	343	42	9500	19	200	19300
	184 : 1	1529 : 1	620 : 1	6684 : 1	6800 : 1	88 : 1

- (80) But that is not enough we also skip our tuitions and rome [sic] here and there with our friends and fool around. [IND2000 Ess 01]

The distributions across the top-level country domains (May 2, 2001) suggest that the pluralization of both nouns appears to be more common a practice in IndE than elsewhere on the Internet today, cf. Table 5.34.

Note that the plural form *advices* is used recurrently throughout the English-speaking world in the context of business English, where it denotes ‘formal communications from a distance’. The plural is much rarer with *advice* when it denotes ‘council’ or ‘opinions given’. In South Asia, *advices* ‘council, opinions given’ is apparently quite common but remains a minority variant to date. The form *tuitions* is noticeably frequent in contemporary AmE under the ‘us’ top-level country domain, too. But a closer look at the attestations for *tuitions* in IndE and AmE reveals that this plural form is nowadays used in distinctly different senses in IndE and AmE, with *tuitions* usually denoting ‘private lessons’ in the former but ‘tuition fees’ in the latter variety. Consider the following representative example from an AmE legal document:

- (81) In connection with the issuance of the bonds authorized by section 15–1483, or in order to secure the payment of such bonds and interest thereon, the board shall have power by resolution to: 1. Fix and maintain tuitions, fees, rentals and other charges from students, faculty members and others using or being served by, or having the right to use or the right to be served by, any project. ... [AmE³⁵]

It is the AmE variant that is also found in other varieties of English, although it has not been recorded in the major dictionaries yet. The use of *tuitions* meaning ‘private lessons’

35. www.azleg.state.az.us/FormatDocument.asp?inDoc=/ars/15/01484.htm&Title=15&DocType=ARS, online May 2, 2001.

appears to be a peculiarity of IndE, which suggests that we might be dealing with two independent developments in AmE and IndE taking different semantic directions.

Variability in the next set of examples from IndE revolves around nouns that may have count and noncount functions in standard English. Labeled “dual-class membership” nouns by Quirk et al. (1985:247), those nouns often show distinct differences in meaning across the count-noncount divide. It is such differences that have been blurred in the following instances:

- (82) The Project Director of Women’s Development Programme (WDP), Ms. Chitra Rathore, said the women running the station maintain a computerised record of the milk collected, test the fat contents and issue computerised dairy receipts. [IND2000 Press Rep 01]
- (83) Lohia points out the damages caste system is doing to the Indian society both in political and economic field. [IndE KOL G 19]
- (84) More generally, it is in the family that we face the deepest resistances to our freedom, it is in the family that our morals are most policed, it is in the family that our oppression is directly enforced in the false name of tradition. [IND2000 Press Rev 10]

When referring to ‘the amount or proportion of something that a substance contains,’ *content* is normally used in the singular. In example (82), the IndE user opted for a plural form, which according to COB is common only if the reference is to ‘things inside a container’ or, more generally, to the ‘subject of a book, a speech, or television program’. Similarly, the plural *damages* is a common choice in standard English when the reference is to ‘a sum of money someone has to pay because he or she has damaged another person’s reputation’. When denoting ‘a physical harm’ or ‘an unpleasant effect’, as in example (83), *damage* is usually noncount. Likewise, standard English employs *resistance* as a count noun in the context of physics and electrical engineering, whereas it usually functions as a noncount noun when denoting ‘a refusal to accept something’ or ‘an attack consisting of fighting back against the people who have attacked you’. This difference was not maintained in example (84).

Again: how characteristic is such blurring of semantic distinctions for contemporary IndE at large? To place the examples from the structured corpora in a wider context, the currency of the singular and plural forms of *content*, *damage* and *resistance* was tested in various semantic environments to see whether or not IndE was on its way toward shedding the distinctions described above. But no convincing evidence was found that usages of the types exemplified in (83), (84) and (85) are any more common in IndE today than in the rest of the English-speaking world. Consider the search results (dated April 29, 2001) in Table 5.35.

If IndE shows any variety-specific behavior in Table 5.35, it is in the recurrent singular use of *damage* to denote ‘a sum of money to be repaid to someone’ (as in *pay damage to*), a usage warned against by Nihalani et al. (1979:62, 2004:61). But in the big

Table 5.35 *Content(s), damage(s), resistance(s).*

	.in	.sg	.uk	.za	.au	.us
fat content	956	545	77800	10400	39700	16200
fat contents	47	18	657	68	479	445
	20 : 1	30 : 1	118 : 1	153 : 1	83 : 1	36 : 1
gold content	105	15	765	120	531	138
gold contents	3	–	46	1	73	17
	35 : 1	–	17 : 1	120 : 1	7 : 1	8 : 1
table of content	784	438	22600	769	60800	14700
table of contents	177000	65500	5830000	186000	4080000	7420000
	1 : 226	1 : 150	1 : 258	1 : 242	1 : 67	1 : 505
do damage to	72	78	11000	345	617	828
do damages to	–	1	9	–	–	2
	–	78 : 1	1222 : 1	–	–	414 : 1
a lot of damage	257	182	41100	1270	21200	775
a lot of damages	5	3	36	5	5	32
	51 : 1	61 : 1	1142 : 1	254 : 1	4240 : 1	24 : 1
pay damage to	8	–	8	–	–	5
pay damages to	38	66	690	95	11100	524
	1 : 5	–	1 : 86	–	–	1 : 105
a lot of resistance	41	26	664	448	538	708
a lot of resistances	–	1	2	–	1	–
	–	26 : 1	332 : 1	–	538 : 1	–

picture, IndE appears to mark the different senses of *content*, *damage* and *resistance* in accordance with standard English usage norms, and no substantial difference can be said to exist between IndE and the other varieties, at least as far as published writing on the Internet is concerned. While it is possible that the blurring of semantic distinctions of the types illustrated in examples (82) to (84) is more common in less heavily edited writing or in speech, concrete empirical evidence for this is yet to be produced.

Two more items that have been discussed in reference to the singular-plural distinction in IndE are *machinery* and *aircraft*. While COB lists under *machinery* non-count and singular uses depending on whether the reference is to ‘machines in general’ (noncount) or ‘the system and all the procedures that a government uses to deal with things’ (singular), plural uses of *machinery* used in the latter particularized sense are also common in IndE today, cf.

- (85) In the context of Rolling Plan commencing in April, 1978, timely availability of data on the economic activity in the non-agricultural sector covered by the Economic Census conducted by the Government of India throughout

the country in collaboration with the State Government machineries will go a long way in fixing the reasonable targets and allocation of resources.

[IndE KOL B 24]

Considering the communicative necessity of referring to the multitude of India's state and local governments and their administrative bodies, it is not surprising to find that the plural form should fill a special communicative need in IndE. But the plural is not unusual elsewhere in the English-speaking world either, as a look at the World Wide Web readily suggests (Google search dated April 29, 2001). Even so, it appears to be notably rarer there, especially in BrE and AmE, which supports Nihalani et al.'s (1979: 118, 2004: 118) statement that "[t]his noun is normally treated as non-count in BS [British Standard English] and is rarely, if ever, found in the plural". Table 5.36 indicates that the plural, which is common in reference to India's political system, appears to be less common when *machinery* is used in more technical contexts to denote 'machines in general'. This is suggested by the fairly low frequency of the compound *farm machineries*. Since the plural *machineries* is attested in the OED for nineteenth-century English, the restriction of *machinery* to noncount and singular contexts might again be a more recent development in the major varieties which has not been followed in IndE full-force (also cf. Platt et. al 1984: 51).³⁶

A similar explanation can be offered for *aircraft* 'airplane'. A count noun in standard English, *aircraft* belongs to the small class of nouns taking the zero plural. IndE uses both zero plural and s-plural side by side, cf.

(86) ... the VIP aircraft that fly the President and the Prime Minister are of the same vintage ...
[IND2000 Press Edit 03]

(87) Para 4 of this Resolution had asked Member States to take the following measures, separately or collectively, in conformity with the Charter to bring about the abandonment of apartheid policies of the South African Government: ... refusing landing and passage facilities to all aircrafts belonging to the Government of South Africa and companies registered under the laws of South Africa.
[IndE KOL F 02]

The plural form *aircrafts* has been singled out by Nihalani et al. (1979: 20, 2004: 18) as an Indianism and is also mentioned by Platt et al. (1984: 52). Evidence from the World Wide Web (April 29, 2001) suggests that *aircrafts* is indeed more common in present-day IndE than in BrE, AmE, AusE, SAfrE or SgE, cf. Table 5.37. Though the plural *aircrafts* is nowadays in use on a global scale, the proportion of s-plural forms is particularly high in IndE, where the item is found to occur alongside the synonymous forms *airplanes* and *aeroplanes*, which themselves take the s-plural. Hence, analogy might have a hand in supporting the more regular s-plural also in *aircrafts*. But a look

36. The last use of *machineries* in the OED quotations base is found in a quotation dating from 1866.

Table 5.36 *Machinery/-ies*.

	.in	.sg	.uk	.za	.au	.us
government +						
machinery	661	145	800	238	409	116
machineries	42	2	8	3	8	–
	16 : 1	73 : 1	100 : 1	79 : 1	51 : 1	–
farm +						
machinery	10200	9	117000	189	120000	75300
machineries	99	1	5	–	3	–
	103 : 1	9 : 1	23400 : 1	–	40000 : 1	–

Table 5.37 *Aircraft(s)* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
aircraft	103000	71100	4940000	2010000	1860000	1320000
aircrafts	13700	950	49000	617	20800	9970
	8 : 1	75 : 1	101 : 1	3258 : 1	89 : 1	133 : 1

Table 5.38 *Aircraft(s)* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
aircrafts	101	258	238	114	204	36	46	47
airplanes	56	342	71	101	133	52	373	13600
aeroplanes	21	251	56	23	142	57	579	13

at the history of English is once more informative. The OED quotations base reveals that the plural *aircrafts* was more common in the nineteenth century to refer to different kinds of flying-machines such as balloons and airships. It appears that only from the early 1930s on did the use of *aircraft* become restricted to denoting an airplane or airplanes in a collective sense. In the course of this development, the zero plural established itself as regular choice in BrE and AmE. Given the available data, IndE does not seem to have latched on to this development, which started in the first half of the twentieth century. Interestingly, in four of six India's online newspapers the plural *aircrafts* turned out to be more common than *airplanes* or *aeroplanes* (May 5, 2001), which is further evidence for the continuing high degree of integration of the regular plural in IndE today, cf. Table 5.38. *Aircrafts*, we can conclude, stands as a contemporary Indianism from a quantitative viewpoint, but it is most likely not an independent innovation.³⁷

37. Also cf. Platt et al. (1984:52).

Let us round off this survey of 'different' noun uses in IndE by pointing out yet another area of variability along the singular-plural divide observed in the IndE material. In the idiomatic phrases *in detail*, *from time immemorial*, and *in all probability*, plural nouns are occasionally chosen in IndE though it is the singular form that constitutes the norm in standard English, cf.:

- (88a) <\$A><#>the two-day meeting discussed organisational matters in detail
[IND2000 Broad News 06]
- (88b) <\$C><#>the decisions<,> which involve the people's interest<,> should be discussed in the allies parties' meetings in details
[IND2000 Broad News 05]
- (89a) The deft handwork of the folk craftsmen with the brush, chisel and needle completes the products which have delighted the world, from time immemorial.
[IndE KOL E 01]
- (89b) From times immemorial, the musicologist has accepted the existence of two kinds of music ...
[IndE KOL E 11]
- (90a) They were in all probability issued by Murshid Quli Khan as Diwan or Subadar along with the imperial orders of appointment as token of his assurance and support to the Sadar Kanungoe.
[IndE KOL H 29]
- (90b) In all probabilities the candidates set up by them will be regarded by the voters as hardly different from the Independent candidates.
[IndE KOL A 31]

While none of the second variants (88b – 90b) is preferred in IndE, all are used currently, possibly due to the force of hypercorrection and in spite of the fact that their use has been branded by style critics (cf. Yadurajan 2001: 109–11). In comparison with other varieties, IndE writing seems to be slightly more variable in this area (Google search dated August 2, 2001), cf. Table 5.39.

On the other hand, the plural-s is frequently dropped in IndE when the phrase *at all costs* is used, cf.

- (91a) <\$D><#>two things<,> one<,> we want that the territorial integrity and sovereignty of Sri Lanka must be maintained at all cost
[IND2000 Broad Disc 04]
- (91b) "Anargha" refers to Rama's keen interest in maintaining Dharma at all costs.
[IND2000 Press Rev 01]

This seems to happen more often in IndE and SgE than in the rest of the English-speaking world (Google search dated August 7, 2001), cf. Table 5.40.

Here, the similarity to the related expression *at any cost* may be responsible for creating the variational profile showing in the data. Another influential factor could be transference from speech: we might, in fact, be dealing with a written repercussion of

Table 5.39 Count and noncount uses of nouns: lexicalized phrases.

	.in	.sg	.uk	.za	.au	.us
in detail	121000	42300	5410000	156000	1420000	1200000
in details	46900	646	215000	26100	189000	38100
	3 : 1	65 : 1	25 : 1	6 : 1	8 : 1	31 : 1
from time immemorial	547	242	23700	818	13700	594
from times immemorial	127	3	135	7	35	7
	4 : 1	81 : 1	176 : 1	117 : 1	391 : 1	85 : 1
in all probability	543	427	109000	15200	54900	37600
in all probabilities	54	1	119	62	35	16
	10 : 1	427 : 1	916 : 1	245 : 1	1569 : 1	2350 : 1

Table 5.40 *At all cost(s)*.

	.in	.sg	.uk	.za	.au	.us
at all cost	189	100	24800	608	9470	526
at all costs	619	275	305000	26500	99000	31200
	1 : 3	1 : 3	1 : 12	1 : 44	1 : 10	1 : 59

the simplification of the consonant cluster *ts* that is frequently heard in spoken IndE (cf. Chapter 5.1.2.2 for a discussion of the high proportion of *lot of* vis-à-vis *lots of*).

5.2.3 Summary

Evidence from the closed corpora and the online corpora suggests that IndE differs less heavily from standard English in quantitative terms within the standard usage range than might be expected when looking at the available feature-list descriptions. Of the 78 claimed Indianisms chosen for investigation in this chapter, only about one third (23) displayed the behavior in the data that had previously been pointed out as being characteristic of IndE. None of the attested ‘unrecorded’ uses was generally preferred in IndE, nor was any such use found to be restricted to IndE alone. *Legislation*, *equipment* and *agitation*, which are classified noncount in COB, as well as *machinery*, which is a singular noun by COB norms, regularly show the morphological properties of count nouns in contemporary IndE. Historically, count uses of those nouns appear to have been more common in English in the nineteenth century, which suggests that the quantitative differences nowadays observable between IndE and other Englishes worldwide might reflect divergent diachronic development rather than independent South Asian innovation. Similarly, the count noun *aircraft* frequently takes the s-plural in contemporary IndE, whereas the prestige varieties BrE and AmE have grammaticalized the unmarked plural. Semantic innovation can be shown to have affected IndE in

the use of the plural *tutions*, which has acquired an IndE-specific sense, that of ‘private lessons’. Generally speaking, the comparison of the press material in the Primary Corpus across time has revealed that variability seems to be on the rise in IndE quality newspapers in this area of the language system, albeit slowly.

5.3 Nouns, pronouns and verbs: Concord with collective nouns

This chapter looks at a subset of nouns for which a notable degree of variation has been observed in English for centuries (cf. Bauer 2002: 50, 88): collective nouns. Collective nouns are common nouns or proper nouns that are characterized by taking verbs and pronominal coreferents either in the singular or plural without displaying a change in number in the noun. In example (92), the collective noun *government* is followed by a verb in the singular, in example (93) by a verb in the plural:

- (92) <\$C><#>well as of the moment<,> we are not aware of any formal approach that uh Sri Lankan government has made to Pakistan
[IND2000 Broad Disc 03]
- (93) Government have to work out a mechanism whereby social expenditure is effectively targeted to the poor; will India’s “pro-poor” parties ever stop playing games?
[IND2000 Press Edit 02]

In example (94), *government* is coreferential with the singular pronoun *it*, in example (95) with the plural pronoun *they*, cf.

- (94) <\$A><#>after the shooting<,> the Israeli government said it would deploy special army units<,> trained in guerilla warfare [IND2000 Broad News 18]
- (95) <\$D><#>but government should move in and take decisive action<,> which they’re not doing
[IND2000 Broad Disc 05]

Examples (92) and (94) illustrate what has been termed ‘syntactic agreement’ or ‘grammatical concord’, examples (93) and (95) ‘semantic agreement’ or ‘notional concord’. While the choice between grammatical and notional concord, according to Quirk et al. 1985, often shows “a difference in point of view” (316), with grammatical concord stressing “the nonpersonal collectivity of the group” (316) and notional concord emphasizing “the personal individuality within the group” (316), users of English today do not oscillate freely between those two possibilities. Some collective nouns almost always take the singular pattern whereas others almost always take the plural, with yet other collective nouns showing more or less pronounced leanings toward either end of the scale that need not be identical for verbal and pronominal concord (cf. Siemund 1995: 365–9). Moreover, there are regional, stylistic and diachronic dimensions to the variation at stake, which have been described most extensively for BrE and AmE to date. Let us briefly review those.

In general, differences among varieties of English are more pronounced for verbal concord than for pronominal concord. While plural concord is a likely option with pronouns in many varieties today (cf. Hundt 1998a: 84; Levin 1998a: 197), plural verbal concord with collective nouns is rare in contemporary AmE but quite frequent in BrE (cf. Siemund 1995: 365–9; Hundt 1998a: 80–9; Levin 1998a, 1998b). However, corpus-based research has also brought to light that singular verbal concord was gaining ground in BrE writing throughout the twentieth century, which has been ascribed to the growing influence from AmE on a global scale.³⁸ Bauer (1994: 61–6), using a compilation of BrE press editorials taken from *The Times* covering the years 1900 to 1980, noted a statistically significant trend toward singular verbal concord with collective nouns. Similarly, Siemund (1995: 365–9), using the press sections of the LOB and FLOB databases, detected an increase in the proportion of nouns taking singular verbal concord in BrE between 1961 and 1991. When Hundt 1998a studied the behavior of a selection of collective nouns in three quality newspapers from Britain, the United States and New Zealand, she found her New Zealand English data to be more closely related to AmE than to BrE and concluded that “variable verbal concord does not seem to be deeply rooted in the grammatical system but rather a highly specific British affectation characteristic of late 19th and early 20th century standard BrE” (88).

While it was not to be expected that IndE would behave in vastly different ways from the prestige varieties in this area, Nihalani et al.’s (1979: 141, 2004: 141) comment on the frequent use of singular concord with *police* ‘police force’ aroused my interest in that it suggested that IndE-specific usage preferences might exist in this traditionally quite variable field of English grammar as well. In the light of the diachronic interpretations offered by Bauer 1994, Siemund 1995 and Hundt 1998a, this seemed all the more promising an area to go into as it offered a possibility of gauging the distance between IndE and BrE at the grammatical level in a diachronically sensitive fashion: would IndE turn out to have followed a “highly specific British affectation” (Hundt 1998a: 88) a century ago or not?

5.3.1 Quantitative analysis

To address this issue, a list of 37 collective nouns was put together for which comparable data from BrE and AmE were available (cf. Siemund 1995; Levin 1998a, 1998b; Hundt 1998a). After scanning the IndE sections of the Primary Corpus for all occurrences of those nouns, tokens with concord marks were separated from those without

38. According to Bauer 2002, the argument of AmE being an influential external factor on other varieties “is one of those cases that is hard to prove, since variation between the two forms has persisted at all times in British English (see for example Visser 1963: §77), and we could just be seeing a process of gradual drift.” (88)

ones by following the methodological procedure outlined in Levin (1998a: 195).³⁹ For cross-varietal comparison, Siemund's (1995) study of verbal concord in BrE newspaper prose, which had been based on the press sections of the LOB and FLOB corpora and allowed for direct comparison, was consulted next. As Siemund 1995 had only dealt with verbal concord in press texts, the BrE sections of the Primary Corpus were scanned for comparable data on singular and plural verbal and pronominal concord in BrE speech and writing. To compare concord patterns in IndE to those in AmE, the studies by Hundt 1998a and Levin (1998a, 1998b) were consulted and the Internet corpora tapped for further supplementary data.

5.3.1.1 *Verbal concord*

Let us start out by looking at verbal concord in IndE. Table 5.41 gives the distribution of singular and plural verbal concord with 37 collective nouns in the IndE press and broadcast sections of the Primary Corpus.⁴⁰

Since the size of the Primary Corpus is modest, the absolute frequencies of the relevant concord patterns are too low to study the behavior of individual nouns in sufficient detail. Of the 37 nouns on the list, four (*company, crowd, staff* and *youth*) are not attested at all in suitable grammatical environments (i.e. environments showing concord marks), and a further ten (*audience, cast, elite, federation, firm, gang, generation, press, public* and *university*) are represented less than three times. Nonetheless, it is clear from the information in Table 5.41 that contemporary IndE, like BrE and AmE, does not vacillate randomly between singular and plural verbal concord but shows distinct preferences for either singular or plural verbs at the level of the individual lexical item. Moreover, when zooming in on the proportions of nouns taking only singular, only plural and variable (i.e. singular and plural) verbal concord, a striking regional difference between IndE and its parent variety BrE surfaces. Consider the proportion of collective nouns taking singular, plural and variable concord in the press texts first (cf. Table 5.42).⁴¹

Table 5.42 reveals that the pull toward the singular pattern is clearly stronger in the IndE press material than in the BrE material. While more than two thirds of the collective nouns in the IndE press texts take singular verbal concord only (1978: 68.2%, 2000: 79.2%), less than half of the nouns do so in BrE (1961: 36.1%; 1991: 46.9%). In this light, IndE displays a behavior close to what has been said to be characteristic of AmE usage: an obvious leaning toward the singular pattern in writing (cf. Hundt

39. In Levin (1998a: 195), concord for each token of a noun was counted only once except for cases when both singular and plural concord were found to agree with the same noun ('mixed concord').

40. For a comparison with BrE see Siemund 1995.

41. The distribution for BrE has been taken from Siemund 1995, who based his counts on the complete press sections of LOB and FLOB. This explains why the absolute type (and token) frequencies are slightly higher for BrE than for IndE.

Table 5.41 Singular and plural verbal concord with collective nouns.

	IndE IND1978 Press sg : pl	IndE IND2000 Press sg : pl	IndE IND2000 Broad sg : pl	Total sg : pl
army	–	1 : 0	11 : 0	12 : 0
association	2 : 0	1 : 0	1 : 0	4 : 0
audience	–	1 : 0	–	1 : 0
board	–	3 : 0	4 : 0	7 : 0
cast	–	1 : 0	–	1 : 0
commission	4 : 0	5 : 0	6 : 0	15 : 0
committee	4 : 0	2 : 0	3 : 1	9 : 1
community	3 : 0	–	4 : 1	7 : 1
company	–	–	–	–
corporation	1 : 0	1 : 0	2 : 0	4 : 0
council	–	3 : 0	1 : 0	4 : 0
couple	0 : 1	0 : 1	0 : 1	0 : 3
crew	–	0 : 1	2 : 0	2 : 1
crowd	–	–	–	–
data	1 : 2	–	3 : 0	4 : 2
department	5 : 0	1 : 0	4 : 0	10 : 0
elite	–	1 : 0	–	1 : 0
faculty	–	–	–	–
family	–	–	3 : 0	3 : 0
federation	0 : 1	–	–	0 : 1
firm	1 : 0	–	–	1 : 0
gang	–	1 : 0	–	1 : 0
generation	1 : 0	–	1 : 0	2 : 0
government	29 : 0	22 : 3	101 : 3	156 : 6
group	3 : 1	1 : 1	8 : 1	12 : 3
institute	1 : 0	2 : 0	–	3 : 0
majority	0 : 1	0 : 1	1 : 1	1 : 3
opposition	1 : 0	2 : 0	6 : 0	9 : 0
party	7 : 0	9 : 0	21 : 0	37 : 0
police	0 : 2	1 : 1	2 : 10	3 : 13
population	2 : 0	1 : 0	6 : 0	9 : 0
press	–	–	1 : 0	1 : 0
public	1 : 0	1 : 0	–	2 : 0
staff	–	–	–	–
team	0 : 1	7 : 0	5 : 2	12 : 3
university	1 : 0	–	–	1 : 0
youth	–	–	–	–

Table 5.42 Verbal concord with collective nouns [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press
# (types)	22 (100%)	24 (100%)	36 (100%)	32 (100%)
# (singular concord only)	15 (68.2%)	19 (79.2%)	13 (36.1%)	15 (46.9%)
# (variable concord)	2 (9.1%)	2 (8.3%)	18 (50%)	10 (31.3%)
# (plural concord only)	5 (22.7%)	3 (12.5%)	5 (13.9%)	7 (21.9%)

1998a: 83; Levin 1998a: 196). In addition, Table 5.39 suggests that variable verbal concord is nowadays less common in IndE than in BrE, at least within the domain of journalistic writing. Both in 1978 and in 2000, collective nouns displaying variable concord were rare in IndE (1978: 9.1%; 2000: 8.3%). By contrast, half of the collective nouns found in the 1961 press material from BrE showed this behavior (1961: 50%). A language generation later, the proportion of variable cases had decreased significantly in BrE (1991: 31.3%), but was still considerably higher than in IndE around the turn of the millennium. Siemund (1995: 336) suggests that the lower proportion of nouns taking variable concord in BrE around the year 1991 might be an indication that BrE was in the process of grammaticalizing verbal concord at an accelerated pace at the close of the twentieth century. But unlike in IndE, no obvious preference for the singular pattern had emerged in BrE by then. In fact, the proportion of nouns taking plural verbal concord in BrE went up from 13.9% in 1961 to 21.9% in 1991. All in all, the findings in Table 5.39 confirm what Hundt's (1998a) analysis of New Zealand English hints at: BrE is in a rather exceptional position among world Englishes today in showing a special liking for variable verbal concord with collective nouns. If the comparatively high proportion of variable verbal concord is interpreted as an echo of "a highly specific British affectation characteristic of late 19th and early 20th century standard BrE" (Hundt 1998a: 86), it can be concluded that IndE quality newspapers today do not show any such obvious traces of late Victorian English.

When turning to the broadcast material, a similar distribution surfaced, thus indicating that usage differences between IndE and BrE in the treatment of collective nouns are nowadays not limited to the newspaper genre alone, cf. Table 5.43. As is the case for the newspaper material, singular verbal concord is more frequent in the IndE data than in BrE. But it is also interesting to see that the proportion of nouns taking variable concord is significantly higher in the IndE broadcast material (2000: 30.4%) than in the IndE newspaper texts (1978: 9.1%; 2000: 8.3%), which suggests speech being a somewhat more likely context for variable verbal concord to occur than writing – a finding which corresponds to an observation found in Quirk et al. 1985, according to whom "the plural is more popular in speech, whereas in the more inhibited medium of writing the singular is probably preferred" (758).

Table 5.43 Verbal concord with collective nouns [Primary Corpus; broadcast].

	IndE IND2000 Broad	BrE UK1991 Broad
# (types)	23 (100%)	22 (100%)
# (singular concord only)	15 (65.2%)	9 (40.9%)
# (variable concord)	7 (30.4%)	8 (36.4%)
# (plural concord only)	1 (4.3%)	5 (22.7%)

In summary, if we hypothesize the Primary Corpus to be indicative of trends in IndE at large and assume further that the general development in world Englishes throughout the twentieth century was toward singular verbal concord with collective nouns, the press and broadcast data suggest that IndE, like New Zealand English, is more advanced than BrE in this process, possibly under the influence from contemporary AmE. From a contemporary perspective, the continuing influence of AmE on press Englishes worldwide is likely keep the proportion of collective nouns with plural verbal concord rather low, though further research will be necessary to describe the interdependencies between AmE and the rest of the English-speaking world in this particular instance in more detail. From a diachronic viewpoint, yet another interpretation suggests itself. If we follow Hundt 1998a and interpret the comparatively high proportion of variable verbal concord with collective nouns in present-day BrE as a repercussion of a BrE ‘affectation’ typical of the late nineteenth and early twentieth centuries, it appears that IndE (and, for that matter, New Zealand English and AmE) never followed BrE in adopting and integrating this practice to the same extent. This, in turn, would suggest a relatively high degree of independence of IndE from its parent variety BrE around the turn of the nineteenth and twentieth centuries.

5.3.1.2 *Pronominal concord*

Differences between IndE and BrE are more difficult to interpret in the case of pronominal concord. Table 5.44 gives the distribution for singular and plural pronominal concord in IndE and BrE for the 37 nouns dealt with in this study. In addition to considering singular and plural personal pronouns, this case study follows Levin (1998a: 193–4) by counting the relative pronoun *which* as displaying singular pronominal concord and relative *who* as displaying plural pronominal concord.⁴² With ten of the 37 nouns not co-occurring with any personal pronoun or relative pronoun

42. Levin 1998a refers to Jespersen (1949 III:120) and Juul (1975:96), according to whom “relative *which* is used when a collective is thought of as a unit and *who* when it is thought of as a collection of individuals” (Levin 1998a: 193). As was the case for verbal concord, concord for each token of a noun was counted only once except for cases when both singular and plural concord agreed with the same noun.

Table 5.44 Singular and plural pronominal concord with collective nouns.

	IndE IND1978 Press sg : pl	IndE IND2000 Press sg : pl	IndE IND2000 Broad sg : pl	BrE UK1961 Press sg : pl	BrE UK1991 Press sg : pl	BrE UK1991 Broad sg : pl
army	–	1 : 0	1 : 1	1 : 0	1 : 0	2 : 0
association	–	–	–	2 : 1	2 : 0	–
audience	1 : 0	0 : 1	–	–	0 : 3	2 : 0
board	–	–	–	2 : 1	0 : 1	–
cast	0 : 1	–	–	1 : 0	1 : 0	–
commission	7 : 0	–	0 : 1	1 : 0	2 : 0	–
committee	4 : 0	2 : 0	6 : 1	3 : 0	3 : 0	–
community	1 : 1	2 : 0	–	–	0 : 1	2 : 1
company	1 : 0	–	1 : 0	1 : 0	4 : 0	2 : 1
corporation	3 : 0	1 : 0	–	–	1 : 0	–
council	–	–	1 : 1	5 : 0	3 : 0	1 : 0
couple	–	–	–	–	–	0 : 4
crew	–	–	0 : 1	–	–	0 : 1
crowd	2 : 2	1 : 1	–	0 : 1	–	–
data	–	–	–	–	–	–
department	2 : 0	1 : 0	–	1 : 0	3 : 0	2 : 1
elite	1 : 1	–	–	–	–	–
faculty	–	0 : 1	–	–	–	–
family	–	1 : 1	1 : 0	0 : 1	–	2 : 0
federation	–	–	–	–	1 : 0	–
firm	–	–	–	0 : 1	2 : 0	–
gang	–	1 : 0	–	–	–	–
generation	–	0 : 1	0 : 1	0 : 1	–	–
government	21 : 0	9 : 0	17 : 1	14 : 3	12 : 0	2 : 3
group	8 : 3	0 : 2	5 : 2	–	3 : 1	1 : 2
institute	–	–	–	1 : 0	–	–
majority	–	–	–	–	–	0 : 1
opposition	–	1 : 0	2 : 2	–	0 : 1	–
party	1 : 0	5 : 0	1 : 0	4 : 1	4 : 0	1 : 3
police	1 : 0	0 : 2	0 : 2	0 : 1	0 : 1	0 : 4
population	1 : 0	1 : 0	–	0 : 2	–	–
press	–	–	–	–	–	–
public	1 : 0	0 : 1	–	1 : 2	0 : 1	–
staff	–	0 : 1	0 : 1	0 : 1	0 : 1	0 : 1
team	2 : 1	3 : 1	3 : 3	0 : 3	–	0 : 1
university	–	–	–	–	–	–
youth	0 : 2	0 : 1	–	–	–	–

Table 5.45 Pronominal concord with collective nouns [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press
# (types)	18 (100%)	21 (100%)	21 (100%)	21 (100%)
# (singular concord only)	11 (61.1%)	8 (38.1%)	8 (38.1%)	13 (61.9%)
# (variable concord)	5 (27.8%)	3 (14.3%)	5 (23.8%)	1 (4.8%)
# (plural concord only)	2 (11.1%)	10 (47.6%)	8 (38.1%)	7 (33.3%)

at all, the explanatory potential of the available evidence is again rather limited when it comes to the individual behavior of the nouns on the list. But taken together, the data are noteworthy in that the proportions of nouns taking singular, plural and verbal pronominal concord are indicative of further interesting usage differences between contemporary IndE and BrE journalistic writing, cf. Table 5.45.

Although the figures in Table 5.45 are small and do not allow for sweeping generalization, it is striking to see that IndE and BrE journalistic writing apparently moved in opposite directions in the second half of the twentieth century. While the proportion of nouns taking variable or plural pronominal concord went up over the years in the IndE material (from 38.9% in 1978 to 61.9% in 2000), it is the proportion of nouns taking only singular concord that rose significantly in BrE (from 38.1% in 1961 to 61.9% in 1991). In other words, plural pronominal concord had become slightly more of an option in IndE quality newspapers by the turn of the millennium, whereas singular pronominal concord had become more common in BrE newspapers. Those findings can, again, be set in relation to the conclusions drawn by Hundt 1998a on analyzing written BrE, AmE and New Zealand English usage. When Hundt 1998a compared quality newspapers from New Zealand, the United States and Great Britain, she also discovered that “AmE, while generally showing a clear tendency towards singular concord with verbs, ha[d] (with the exception of *team*) the highest figures for plural pronominal concord” (84–5). Why should that be the case? It is possible that since notional concord is less often realized by verbal concord in IndE and AmE writing, the need to use plural pronominal concord might be greater in those two varieties than in BrE writing, where notional concord is frequently realized by verbal concord.⁴³ Moreover, if plural pronominal concord should be more of an option in

43. Such a near-complementary distribution of grammatical and notional concord would go against the Agreement Hierarchy proposed by Corbett (1983: 10), according to which the probability of occurrence for notional (plural) concord should be expected to increase when moving down the hierarchy from premodifiers to predicates to relative pronouns all the way to personal pronouns. Since both grammatical and notional concord are possible options for predicates in BrE, Corbett’s model would predict notional concord to be more likely than grammatical concord further down the hierarchy, i.e. with relative and personal pronouns. This is apparently not true for the BrE data investigated here, although the model fits the IndE data (with

Table 5.46 Pronominal concord with collective nouns [Primary Corpus; broadcast].

	IndE IND2000 Broad	BrE UK1991 Broad
# (types)	15 (100%)	16 (100%)
# (singular concord only)	3 (20.0%)	4 (25.0%)
# (variable concord)	7 (46.7%)	6 (37.5%)
# (plural concord only)	5 (33.3%)	6 (37.5%)

AmE, IndE journalism might have been under the increased influence from AmE after 1978. More corpus-based research into the relationship between IndE and AmE is called for to substantiate this hypothesis across varieties, registers and text types.

In the broadcast sections, regional differences in pronominal concord are not as pronounced as in writing. In fact, IndE and BrE behave in remarkably similar ways, cf. Table 5.46. It appears that the differences in pronominal concord which are observable in the IndE and BrE newspaper material are not characteristic of IndE and BrE usage at large, and one should be careful not to overgeneralize. It is possible, of course, that more text-type and register-specific investigations using larger text collections will uncover further usage differences between IndE and other Englishes in the future.

5.3.2 Case study: *Police*

Since the Primary Corpus was too small to yield satisfactory insights into the patterning of the individual collective nouns, it was decided to supplement the findings in the Primary Corpus with further investigations on one particular noun that had been discussed in reference to IndE before: *police*. In BrE and AmE, *police* occupies a rather unusual position within the set of collective nouns by almost exclusively taking plural verbal and pronominal concord (cf. Hundt 1998a: 82–3). On the other hand, *police* was pointed out by Nihalani et al. (1979: 141; 2004: 141) as a grammatical Indianism in terms of taking singular concord. To see whether IndE usage was indeed significantly different from BrE and AmE, the first 100 occurrences of verbal and pronominal concord with *police* in the online edition of the *Indian Express* (date of analysis: November 4, 2001) were sampled and analyzed. Those findings were then compared with the results on BrE, AmE and New Zealand English reported by Hundt (1998a: 82–3). Table 5.47 gives the distribution of singular and plural verbal and pro-

the exception that distributional differences between relative and personal pronouns were not considered due to the overall low token frequencies in the Primary Corpus). Note that Levin's (1998a: 197) findings based on the *New York Times* and *The Independent* are not strictly in line with the Agreement Hierarchy either, which raises doubt as to the model's general validity for varieties of English at large.

Table 5.47 Concord with *police*.*

	IndE IE	BrE GU	AmE MH	NZE DOM/EVP
verbal concord (sg : pl)	21 : 79	1 : 99	0 : 100	1 : 99
pronominal concord (sg : pl)	19 : 81	3 : 97	1 : 99	0 : 100

* IE vs. GU: $p < 0.001$; IE vs. MH: $p < 0.001$; IE vs. DOM/EVP: $p < 0.001$.

nominal concord in the in India's *Indian Express*, Britain's *Guardian*, America's *Miami Herald* and New Zealand's *Dominion/Evening Post*.⁴⁴

While BrE, AmE and NZE do not show any statistically significant differences in their treatment of *police*, IndE usage is different in that singular concord with *police* appears in significantly higher numbers both for verbs and pronouns, although the singular pattern is not preferred in the *Indian Express*. It thus emerges that in spite of the prescriptive pressure from India's style critics (cf. Yadurajan 2001: 110), singular concord with *police* was recurrent practice in IndE journalistic writing around the year 2000.⁴⁵ See the following illustrative examples of singular and plural verbal concord with *police* from the online edition of the *Indian Express*:

- (96) For the first time in its history, the Mumbai police was handling a case as mammoth as this, with ramifications that touched Pakistan's Inter Services Intelligence (ISI) in such a tangible manner, nailing underworld don Dawood Ibrahim with the most deadly conspiracy he had ever hatched.

[*Indian Express*; September 18, 2001⁴⁶]

- (97) The Mumbai police were tipped off on the possible involvement of Mumbai-based ISI operatives in the hijack four days after the incident.

[*Indian Express*; October 2, 2001⁴⁷]

Examples (98a) and (98b) display singular and plural concord with relative pronouns, examples (99a) and (99b) with personal pronouns:

44. The figures for BrE, AmE and NZE are taken from Hundt (1998a: 82–3).

45. Singular verbal concord with *police* appears to be a diachronically stable second option in IndE. In the 1978 Kolhapur Corpus, plural verbal concord outnumbered singular verbal concord by a ratio of 10:5 in the press texts and by a ratio of 6:3 in non-journalistic writing. Singular pronominal concord was rarer, plural pronominal concord outnumbering singular concord by a ratio of 9:1 in the press categories and by a ratio of 8:0 in the other categories.

46. www.indianexpress.com/story/14464.html, online November 4, 2001.

47. www.indianexpress.com/story/16924.html, online November 4, 2001.

- (98a) For the Mumbai police which has arrested Bollywood financier Bharat Shah for his alleged links with mafia don Chhota Shakeel, this must come as music.
[*Indian Express*; February 4, 2001⁴⁸]
- (98b) Rashid's ordeal started on July 30, when he was dragged out of his cousin's apartment in Vasant Kunj by the Delhi police who suspected him to be a militant.
[*Indian Express*; September 5, 2001⁴⁹]
- (99a) Even as the police began its search, Kundal released the child on Monday in front of a telephone booth in Mirzapur area.
[*Indian Express*; January 4, 2001⁵⁰]
- (99b) The police wanted to know whether he printed the wedding cards too. They found, however, that he did not.
[*Indian Express*; February 3, 1999⁵¹]

Why should IndE differ from the prestige varieties in this particular instance? A look at the OED quotations base suggests that the grammaticalization of the plural verbal concord pattern with *police* 'police force' apparently took place in BrE and AmE only in the second half of the nineteenth century. When looking up the two collocations *police is* and *police are* (with *police* meaning 'police force'), the distribution shown in Table 5.48 emerged. If the grammaticalization of the plural pattern with *police* in BrE (and AmE) is regarded as the result of a development starting in the nineteenth century, South and South East Asian Englishes might not have followed the major varieties in this development full force. According to this interpretation, IndE usage would have to be considered more conservative than BrE and AmE in this particular instance in terms of maintaining a degree of variability that by now has largely been shed in the prestige varieties.⁵² Singular verbal concord is nowadays common in IndE not only in the *Indian Express* but in other national quality broadsheets as well, which suggests a fairly high degree of integration of the pattern in question nationwide. See Table 5.49, which gives the distributions of the collocations *if/when (the) police is* and *if/when (the) police are* in the online editions of the *Indian Express* (IE), the *Hindu* (HI), the *Deccan Herald* (DH), the *Hindustan Times* (HT), the *Tribune* (TR) and the *Telegraph* (TE) as well as in London's online *Guardian* (GU) and in the *Washington Post* (WP) (Google search dated November 7, 2001). Interestingly, Chandigarh's *Tribune* is characterized by a comparatively high

48. www.indianexpress.com/sunday/story/9588.html, online November 4, 2001.

49. www.indianexpress.com/res/web/pIe/archive_full_story.php?content_id=14019, online November 4, 2001.

50. www.indianexpress.com/story/3196.html, online November 4, 2001.

51. www.indianexpress.com//story/1052.html, online November 4, 2001.

52. That IndE treats *police* more variably than BrE or AmE is also shown by the fact that IndE is more likely to use the definite article with *police* in various contexts, cf. chapter 5.1.

Table 5.48 Verbal concord with *police* [OED quotations base].

	POLICE + is	are
< 1800	3	–
1800–1850	1	1
1850–1900	–	6
>1900	2	30

Table 5.49 Verbal concord with *police* [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
if police is	–	–	1	–	1	–	–	–
if police are	–	–	–	–	–	3	5	9
when police is	–	–	–	–	–	–	–	–
when police are	–	3	–	–	1	3	4	7
if the police is	–	–	1	–	15	–	–	–
if the police are		25	3	3	3	8	17	6
when the police is	1	2	–	–	13	–	–	–
when the police are	–	9	2	–	1	2	3	1
POLICE + SG. (total)	1	2	2	–	29	–	–	–
POLICE + PL. (total)	–	37	5	3	4	16	29	23

proportion of tokens of the singular variant, which is rarer in the other five Indian publications and conspicuously absent from Britain's online *Guardian* and from the *Washington Post*. From a diachronic view, usage in *The Tribune* is least advanced in terms of adhering to internationally established contemporary usage conventions. Note that *The Tribune* was also found to be slightly more careful in adopting new coinages and hybrid compounds than the other five newspapers.⁵³ Hence, among the six IndE quality newspapers looked at, *The Tribune* qualifies as the stylistically most conservative publication in lexis and grammar.

When looking at IndE usage under the 'in' top-level country domain (November 7, 2001), on the other hand, the preference in Chandigarh's *Tribune* looks less

53. Concord patterns with collective nouns have been observed to be affected by house-style conventions in other varieties of English as well, see for examples Siemund (1995: 368-9) on specific usage patterns of *government* in *The Times* (BrE), Hundt (1998a: 86) on the use of *government* and *team* in New Zealand's *Dominion* and *Evening Post*, and Levin (1998a: 197), who interprets quantitative differences in concord-marking between the AmE *New York Times* and the BrE *Independent* in reference to the possibility of the presence or absence of style guides.

Table 5.50 Verbal concord with *police* [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
if police is	1	–	4	–	–	8
if police are	2	–	236	26	468	90
when police is	–	–	1	–	–	–
when police are	1		239	25	409	116
if the police is	8	1	16	1	1	2
if the police are	4	3	16900	98	536	276
when the police is	6	1	2	2	3	1
when the police are	2	2	755	43	242	137
POLICE + SG. (total)	15	2	23	3	4	11
POLICE + PL. (total)	9	5	18130	192	1655	619

exceptional, cf. Table 5.50. While the forces of analogy may well trigger occasional uses of the singular pattern in BrE and AmE, singular verbal concord is rare in the prestige varieties within the confines of the chosen collocations. In AusE and SAfrE, too, the singular pattern appears to be uncommon. In South and South East Asia, by contrast, the linguistic situation is different. Though the absolute figures are quite small for IndE and SgE, it is obvious that those two varieties are characterized by a comparatively high proportion of singular verbal concord in non-journalistic writing. Under the ‘in’ domain, instances of singular verbal concord even outnumber those for plural concord. In this light, singular concord with *police* qualifies for the status of quantitative grammatical South Asianism. Time will tell whether IndE in general and Chandigarh’s *Tribune* in particular will maintain this usage difference vis-à-vis the prestige varieties. Ongoing prescriptive pressure from India’s style critics on that matter might encourage change in the opposite direction (cf. Yadurajan 2001: 110).

5.3.3 Summary

The comparison of concord patterns with collective nouns in contemporary IndE and BrE has revealed that IndE features a lower proportion of collective nouns showing variable and plural verbal concord than BrE, for which plural verbal concord was more natural a choice around the turn of the millennium. In this regard, IndE resembles what has been described as being typical of contemporary AmE usage: a pronounced preference for the singular pattern with the majority of collective nouns. This finding throws an interesting light on the nature of linguistic change in IndE during the past one hundred years. If the popularity of plural concord is seen as a BrE-internal phenomenon characteristic of standard BrE around the turn of the nineteenth and twentieth centuries, IndE was independent enough around that time not to be affected by those developments. The same is valid for AmE, which nowadays acts as a conservative force for singular verbal concord. As the analysis of the corpus material also shows, no-

tional concord is realized in IndE quality newspapers today predominantly by means of plural pronominal concord, in BrE publications by plural verbal concord. Again, IndE approximates more closely what has been said to be the case for AmE. However, IndE usage differs from both AmE and BrE at the level of the individual word in its treatment of *police* ‘police force’. *Police* is used recurrently with singular verbal and pronominal concord by Indian users, especially in Chandigarh’s *Tribune*, whereas contemporary AmE and BrE have come to treat this noun as an unmarked plural noun. In reference to diachronic evidence from the OED quotations base, it has been argued that this synchronic usage difference has resulted from divergent diachronic developments in IndE, BrE and AmE over the past two hundred years.

5.4 Verb phrase: The past perfect

Contradictory claims have been made regarding the functions of the past perfect in IndE. In standard English, the most common function of the past perfect is to denote an event, state or habit “anterior to a time of orientation in the past” (Quirk et al. 1985: 196), cf.

- (100) Third, even though the Pakistanis had an opportunity to cripple the naval vessel INS Vikrant’s strike capability by destroying the naval squadron, they could not do so. That is because we had anticipated the attack and had prepared ourselves for any eventuality. [IND2000 Press Edit 09]
- (101) “We knew it was Fazal. And we kept our heads together and attacked him,” said Zaatini. When the attack did come, late in the fourth set, the Indians were certainly taken by surprise. Until then, the home team hadn’t faced a single breakpoint on serve – for a full two hours and more – and Paes himself was up 40–15 on serve in the eighth game. [IND2000 Press Rep 07]

Next to its ‘past-before-the-past’ meaning, the past perfect can also be employed to express hypothetical meaning:

- (102) How I wish this sensible thought had occurred to the party before it toppled the BJP Government at the Centre and forced a general election on the people. [IND2000 Press Edit 07]
- (103) If religion alone had been the basis of the nation-state, Pakistan would not have splintered to create Bangladesh. [IND2000 Press Edit 05]

And it can be used in indirect speech constructions to indicate “backshift into the more remote past” (Quirk et al. 1985: 197):

- (104) <\$A><#>briefing newsmen later Mr Sayeed said<,> Mr Atal Bihari Vajpayee had assured him of all possible help [IND2000 Broad News 01]

- (105) <\$B><#>well it's a<,,> very very<,,> great loss to Congress party<,,> and I think to the country<,,> and as you said that he had<,,> in a short time<,,> carved for himself <,,> a special space in national life<,,> <#>he was a politician true politician [IND2000 Broad Int 01]

In addition to those standard functions, IndE has been claimed to use the past perfect more freely also in contexts where standard English would use other tense forms, for example the simple past or the present perfect (cf. Comrie 1985; Kachru 1987; Agnihotri et al. 1988; Leitner 1989a, 1989b; Sharma 2001; Rogers 2002). Comrie 1985 states that the past perfect is 'overused' in IndE because of interference from languages like Hindi and Urdu:

In Hindu-Urdu and Armenian, the form that most closely corresponds formally to the English pluperfect seems to have two meanings, related but distinguishable: on the one hand, it can indicate a situation located prior to a past reference point, and when used in this way it corresponds to the English pluperfect. It can also be used, however, for a remote past event even when there is no intervening reference point, and this clearly differs from the English use of the pluperfect ... (Comrie 1985: 69)

The IndE past perfect, Comrie 1985 concludes, has acquired an additional function, namely that of signaling "remote pastness," which he considers "one of the salient characteristics of many varieties of Indian English, under the influence of the broader range of uses of the formally similar form in Hindi and some other languages of the sub-continent" (69).

Like Comrie 1985, Sharma 2001 reports that the IndE past perfect is recurrently associated with preterite and present perfect meanings, which she interprets in terms of "an emergent anterior, completive (perfective, rather than perfect) marking and a concomitant loss of the necessity for a distinct reference point in the context" (370). Sharma 2001 provides empirical evidence from the Kolhapur Corpus for such uses of the past perfect to cluster in texts representing regional bureaucratic writing, which is seen as a linguistic reflection of the complex sociocultural realities of using English in India:

Unlike the national press, the national administrative or civil services extend to all corners of the country and across classes, thereby integrating a very disparate population into a single institutional language code. For this reason, while government documents in native English-speaking nations may represent a high level of formality, the Indian bureaucratic register in fact reflects a potentially more indigenized standard than, say, that of the national English language media. (Sharma 2001: 361)

In the press texts from 1978 (taken from text category A of the Kolhapur Corpus), Sharma 2001 finds innovative uses of the past perfect to be significantly less common

than in bureaucratic writing, which she explains in terms of a greater orientation toward internationally established usage conventions within the English-language national print media in India.

According to Leitner (1989a, 1989b), IndE also shows innovative tendencies at the discourse-pragmatic level. Drawing on observations made by Kachru 1987 in a passage of expository writing translated from Hindi into English, Leitner 1989a states that “the simple past and past perfect have a backgrounding function and do not only serve to mark temporal sequence as such” (176). Therefore, the vacillation between the past perfect and other tenses carries a communicative load by foregrounding or backgrounding text passages in their relation to one another. This, in effect, leads to an ‘overuse’ of the past perfect in contemporary IndE. As Leitner 1989b puts it:

Die Vertextung ist demnach durch makrodiskursive Faktoren gesteuert. Sie verursachen die Divergenz zwischen dem indischen und dem muttersprachlichen Englisch und haben, neben der ‘Überverwendung’ des past perfect, vor allem eine wesentliche Folge für das Sprachsystem. Sie überdecken die Erfordernisse der relativ ikonischen Enkodierung objektiver Ereignisfolgen, und Ereignissequenzen sind nicht mehr aufgrund des Tempusgebrauchs, sondern aufgrund des Allgemeinwissens und der Ereignislogik rekonstruierbar. (Leitner 1989b: 49)

Analyzing the English of Indian advanced learners, Agnihotri et al. 1988, on the other hand, find that the past perfect is less, rather than more, common in their IndE material than expected. Agnihotri et al. 1988 observe that the past perfect is frequently replaced by the simple past, which they explain in an error-analytical framework:

Our study suggests that the students find the use of the past perfect more difficult than the use of the present perfect. Our experience with the undergraduate students substantiates the results ... The use of the past perfect involves viewing a completed event in the past form from the past time perspective. A large number of students tended to use the simple past instead of the past perfect.

(Agnihotri et al. 1988: 94)

While the authors do not attempt to give their data a functional reading, the results of their analysis are worth taking note of, as their ‘substitution’ reading runs counter to the ‘overuse’ hypothesis underlying the work of Comrie 1985, Kachru 1987, Leitner (1989a, 1989b) and Sharma 2001.

L1-interference, pragmatic innovativeness, and learner difficulties – given the spectrum of hypotheses on IndE, it was a natural decision to consult the Primary Corpus and take a closer look at the IndE past perfect in order to test Comrie’s (1985), Kachru’s (1987), Leitner’s (1989a, 1989b), Sharma’s (2001) and Agnihotri et al.’s (1988) hypotheses systematically on the grounds of written and spoken corpus data.

5.4.1 Quantitative analysis

To address the issue of whether the past perfect was overused or underused in IndE vis-à-vis the major varieties, the Primary Corpus was consulted first. The absolute numbers of past-perfect forms in the three varieties under investigation were computed by means of extracting all tokens of *had* (including contractions) and identifying the past-perfect forms manually. Since the Primary Corpus was untagged, there was no straightforward way of comparing the proportion of past-perfect forms to those of other tense forms in IndE and in the other varieties, which would have been desirable in a next step to describe the position of the IndE past perfect in relation to other tense forms in IndE more exactly. Bearing this limitation in mind, I hypothesized that if the past perfect was used more frequently in IndE and had acquired new pragmatic functions there, possibly due to the influence from Indian languages, this might show in a significantly higher number of past-perfect forms in IndE than in BrE or AmE across text types and registers. If, on the other hand, the learner context was a strong influential factor, this might be reflected in a significantly lower number of past-perfect forms in the IndE material. However, when comparing the absolute frequencies of the past perfect in IndE, BrE and AmE, a much more complex picture emerged that did not suggest any such monolithic interpretation of the data, but pointed to sharp differences among the text categories represented.

Consider first the distribution of the past-perfect forms in the press material. As Table 5.51 illustrates, the past perfect is, indeed, more common in the IndE press texts than in comparable texts from BrE and AmE, which appears to support the ‘overuse’ hypothesis at first sight. But it is important to note that the difference in token frequencies between IndE and BrE quality newspapers is statistically significant only when comparing the older datasets from 1978 (IndE) and 1961 (BrE). The difference between IndE and BrE is not statistically significant for the younger datasets from 2000 (IndE) and 1991 (BrE). Hence, the claim that the past perfect is generally ‘overused’ in IndE as compared with BrE can no longer be given a statistical backing. In fact, IndE and BrE, from a quantitative perspective, appear to have moved more closely together over the years, at least as regards the domain of quality newspapers. In this respect, the findings in Table 5.51 raise doubt as to the representativeness of the 1978 Kolhapur Corpus for more contemporary written IndE and, hence, to the validity of any description of IndE based solely on this text collection, for example Sharma 2001.

Table 5.51 also illustrates that while the quantitative difference between IndE and BrE quality newspapers had diminished over the years, AmE quality newspapers were in a class by themselves throughout the second half of the twentieth century in using significantly fewer past-perfect forms than the other two varieties. This research finding corresponds to observations made in previous comparative studies dealing with differences between BrE and AmE (cf. Algeo 1988: 19, 2006: 27; Harris 1996: 38; Sharma 2001: 355–60). As Algeo (2006: 27) sums up, “British [English] is especially more likely to use the past perfect where it is logically called for, to denote an action or

Table 5.51 Past perfect [Primary Corpus; press].*

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
tokens	206	176	148	166	112	118
tokens per text	5.2	4.4	3.7	4.2	2.8	3.0
per 1,000 words	2.6	2.2	1.9	2.1	1.4	1.5

* IND1978 vs. UK1961: $p < 0.005$; IND1978 vs. US1961: $p < 0.001$; IND2000 vs. US1992: $p < 0.001$; UK1961 vs. US1961: $p < 0.05$; UK1991 vs. US1992: $p < 0.005$.

state that existed prior to some other past action or state.”⁵⁴ To substantiate this claim further with the help of the Primary Corpus, one particular environment in which this regional difference had been said to surface often was inspected more closely: indirect speech (cf. Algeo 1988: 19; Sharma 2001: 354, 359). Generally speaking, verb forms that are used in the past tense in direct speech are normally backshifted into the past perfect in indirect speech when the reporting verb is in the past.⁵⁵ While BrE has been characterized in terms of a “tendency to sequence tenses” (Algeo 2006: 29), AmE has been said to suspend backshift more frequently than other varieties. This is borne out when zooming in on cases of reported speech after the reporting verb *say* in the Primary Corpus. Consider Table 5.52, which also shows that IndE takes an in-between position.⁵⁶

While American journalists suspended backshift in about half of the cases of reported speech after *say* (US1961: 54.4%; US1992: 48.2%), Indian journalists only did so in about one third of the cases (IND1978: 30.1%; IND2000: 34.8%). In the BrE press texts, by contrast, the suspension of backshift is less common and appears to have become even rarer over the years, with the proportion of instances of suspended

54. According to Biber et al. (1999:462), the perfect aspect is more common in BrE than in AmE by a ratio of 4:3.

55. According to Quirk et al. 1985, backshift in reported speech may be suspended “if the relative time reference clear” or “when the time-reference of the original utterance is valid at the time of the reported utterance.” (1027)

56. Table 5.52 gives the token frequencies of finite verb forms in the context of reported speech after past-tense and past-perfect forms of *say* in the *inquit* phrase. Excluded were cases of reported speech with present-tense and present-perfect forms of *say* in the matrix clause as they do not trigger backshift. The category ‘completed’ includes all forms where no further backshift was possible, the category ‘suspended’ those where backshift, according to my reading of the data, would have been possible without altering the meaning of the utterances in question.

Table 5.52 Backshift in reported speech after past-tense *said*.

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
# said + ind. speech	73 (100%)	155 (100%)	122 (100%)	112 (100%)	149 (100%)	168 (100%)
completed backshift	51 (69.9%)	101 (65.2%)	85 (69.7%)	96 (85.7%)	68 (45.6%)	87 (51.8%)
suspended backshift	22 (30.1%)	54 (34.8%)	37 (30.3%)	16 (14.3%)	81 (54.4%)	81 (48.2%)

Table 5.53 Suspended backshift.

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
suspended backshift	22 (100%)	54 (100%)	37 (100%)	16 (100%)	81 (100%)	81 (100%)
past-tense	17 (77.3%)	14 (25.9%)	14 (37.8%)	5 (31.3%)	29 (35.8%)	21 (25.9%)
other	5 (22.7%)	40 (74.1%)	23 (62.2%)	11 (68.7%)	52 (64.2%)	59 (74.1%)

backshift after *say* dropping significantly from 30.3% (UK1961) to 14.3% (UK1991).⁵⁷ That the suspension of the backshift rule is a contributing factor (though a minor one) in keeping the number of past-perfect forms relatively low in AmE vis-à-vis IndE and BrE becomes clear when looking at the absolute numbers of unshifted past-tense forms in the three varieties, which is higher in AmE (US1961: 29; US1992: 21) than in IndE (IND1978: 17; IND2000: 14) and in BrE (UK1961: 14; UK1992: 5). See Table 5.53 for empirical evidence.

Table 5.53 illustrates further that IndE journalism had caught up with AmE and BrE by the year 2000 in recurrently suspending backshift also in other temporal contexts, for example in present-tense or future contexts. This used to happen rarely in India's quality newspapers around 1978 (making up only 22.7% of all cases of sus-

57. Quality newspapers in India, Britain and the US also saw a significant rise in the use of direct speech after the *inquit* verb *say* in the second half of the twentieth century, cf. Appendix 5.8. According to the Primary Corpus, IndE journalistic writing has been participating in the international trend toward colloquializing the written norms of journalistic writing over the past two decades, which is why Sharma's (2001) claim that "IndE appears to paraphrase direct quotes much more than either AmE or BrE" (350) is no longer valid.

Table 5.54 Choice of tense with *after*.

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
after + past tense	10	24	2	21	24	26
after + past perfect	–	–	7	2	3	3

pending backshift after *said*) but was noticeably more common a phenomenon twenty-two years later (making up approximately three-fourth of all cases of suspended backshift after *said*). In this regard, too, IndE journalistic writing can be said to have converged toward, rather than diverged from, the norms of standard English.

Another environment in which AmE has been said to make less use of the past perfect than BrE is in the context of time adverbials and temporal clauses (cf. Biber 1999 et al.:469), for example in clauses introduced by the conjunction *after*. See the two competing variants in examples (106) and (107) first:

- (106) <\$A><#>Aditi uh<,> before we go on to the domestic dimension of what's happening if we just go back to that report<,> that the Tigers have now launched<,> two attacks against Sri Lankan army<,> after a nine-day lull<,> <#>is this something that we were going to expect<,> after the the the cease-fire offer had been rejected [IND2000 Broad Int 10]

- (107) If it required the intervention of Graham Greene, who had not known Narayan earlier, to persuade Hanmish Hamilton of London to publish "Swami and Friends" after it was rejected by many other publishers, it could be just an instance of the intervention of destiny for setting matters right. [IND2000 Press Rev 01]

According to Quirk et al. (1985:196), both the past perfect and the simple past can be used in clauses introduced by *after*. While the use of the past perfect in example (106) places the time of the 'cease-fire offer' clearly before the time at which 'we were going to expect' a terrorist attack, the temporal relation must be reconstructed by the reader in example (107) by means of interpreting the meaning of the conjunction *after*. When looking at the press sections of the Primary Corpus, quantitative differences between AmE, BrE and IndE quality newspapers were found to be only marginal in this particular environment, cf. Table 5.54.

According to Table 5.54, quality newspapers in India, the US and Britain nowadays clearly prefer the past-tense to the past-perfect variant. Interestingly, BrE quality newspapers apparently underwent a slight change in preference between 1961 and

Table 5.55 Past perfect [Primary Corpus; broadcast and essays].*

	IndE IND2000 Broad News	BrE UK1991 Broad News	IndE IND2000 Broad I/D	BrE UK1991 Broad I/D	IndE IND2000 Ess	BrE ICE-GB Ess
tokens	73	124	35	77	16	34
tokens per text	3.7	6.2	1.8	3.9	1.6	3.4
per 1,000 words	1.8	3.1	0.9	1.9	0.8	1.7

* IND2000(Broad News) vs. UK1991(Broad News): $p < 0.025$; IND200(Broad I/D) vs. UK1991(Broad I/D): $p < 0.005$.

1991, with the past tense having become more of an option in journalistic writing over the years. This might have been triggered by the external pressure from AmE journalistic writing, although considerations of space favoring the use of the shorter of the two variants might also have played a role. IndE journalistic writing, on the other hand, is characterized by a diachronically stable preference for the past-tense variant in the available data, which again goes to say that IndE journalists cannot generally be said to ‘overuse’ the past perfect.⁵⁸

More striking quantitative differences between IndE and BrE in the frequency of the past perfect were observed in the Primary Corpus when looking at the broadcast registers and the student essays. In contrast to the newspaper categories, the number of past-perfect forms in the speech data and in the essays was found to be only about half of that in the BrE corpora. Consider Table 5.55.

From a quantitative perspective, the claim that the past perfect is overused in IndE is clearly refuted at this point. If we follow Comrie 1985 and Leitner (1989a, 1989b) and expect L1 interference and pragmatic innovativeness to result in a higher number of past-perfect forms, we cannot build a case for this with respect to the broadcast registers and the essays. The comparatively low number of past-perfect forms in the IndE student essays (as compared with the student essays compiled for ICE-GB), on the other hand, might have resulted from avoiding the past perfect in the ways described by Agnihotri et al. 1988 for advanced learners’ English, but more qualitative research into the distribution of the past tense forms and other tense forms is needed to back up this claim (cf. Chapter 5.4.2). Note that the differences between spoken IndE and spoken BrE, both in scripted and spontaneous speech, are statistically significant, with IndE employing the past perfect conspicuously less often than BrE. In this regard, IndE speech is more similar to the student essays than the press material. It is possible that we have come across an area in English

58. Cf. Appendix 5.9 for the distribution of both variants in IndE and BrE speech, which is in line with the distributions in Table 5.48.

grammar for which educated IndE speech carries particularly strong traces of its acquisitional context.

5.4.2 Qualitative investigations

While the quantitative approach to the IndE past perfect raises interesting questions about the motivations behind the observed distributional patterns, the hypothesis that functional differences exist between IndE and the major varieties can only be tackled by means of a qualitative analysis. It was therefore decided to scan the Indian sections of the Primary Corpus carefully for uses of the past perfect that, according to my personal reading of the data, could not be given a straightforward 'standard English' reading as outlined in Quirk et al. (1985:195–7). Since the interpretation of the past perfect is heavily context-dependent, the task of identifying the intended meaning of the past-perfect forms found in the data was not always straightforward. Particular attention was paid to examples of the remote-pastness and discourse functions ascribed to IndE by Comrie 1985, Kachru 1987, Leitner (1989a, 1989b) and Sharma 2001.⁵⁹ See, first of all, the proportion of 'different' uses of the past perfect in IndE that could be identified unambiguously in the corpus material across the text categories (Table 5.56).

While the majority of past-perfect forms in the corpus conform to the rules formulated in Quirk et al. (1985:195–7), an interesting text-category-specific distribution of the marked-up uses surfaces. Table 5.56 strongly suggests that 'different' uses of the past perfect do not stratify evenly across registers and text types today but show an increasing likelihood of occurrence when moving from more to less heavily planned language use. According to Table 5.56, the proportion of 'different' uses is clearly the highest in the essays (proportion of unusual past-perfect uses: 87.5%), followed by the categories of spontaneous speech (Broadcast Interviews and Discussions, 51.4%), scripted speech (Broadcast News, 11%), and journalistic writing, for which the proportion of marked-up usages went up slightly between 1978 and 2000 (from 5.8% in the 1978 data to 10.2% in 2000). The student writing is in a class by itself, with 14 out of 16 occurrences of the past-perfect not complying with the rules outlined by Quirk et al. (1985:195–7). This supports Agnihotri et al.'s (1988) findings and suggests that the acquisitional context should be considered a decisive determining factor in creating variability in that particular area of IndE grammar. Moreover, Table 5.56 indicates that spontaneous speech is more variable than scripted speech. The category of newspaper prose, on the other hand, approximates most closely the usage standard in Quirk et al. 1985. This could reflect a willingness of India's quality newspapers to conform to internationally established usage conventions and, in doing so, reach out to national as well as international readerships. However, the fact that variability in IndE

59. Cf. Appendix 5.10 for the distribution of the standard functions of the past perfect in IndE.

Table 5.56 ‘Different’ uses of the past-perfect in IndE.

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad News	IndE IND2000 Broad I/D	IndE IND2000 Ess
tokens	206 (100%)	176 (100%)	73 (100%)	35 (100%)	16 (100%)
of those: ‘different’ uses	12 (5.8%)	18 (10.2%)	8 (11.0%)	18 (51.4%)	14 (87.5%)

quality newspapers was slightly higher around the year 2000 than a language generation before is another indication that India’s quality broadsheets are slowly opening up to forms that would be unusual in standard English. In other words, the press texts in the Kolhapur Corpus no longer represent contemporary IndE usage.⁶⁰

What kinds of ‘different’ uses of the past perfect are encountered in the corpus and how ‘typically IndE’ are they? When looking more closely at the marked-up forms, a complex network of past-perfect uses going beyond the codified standard was uncovered that rendered any further neat categorization difficult. While some instances allowed for a straightforward ‘remote pastness’ reading along the lines of Comrie 1985, others came closer to Leitner’s (1989a, 1989b) description of the innovative discourse-pragmatic functions of the IndE past perfect. Yet other examples seemed to be little more than learner errors of the types discussed by Agnihotri et al. 1988 or, more profanely, repercussions of the circumstances of online production. Let us look at this spectrum of usages in more detail.

Consider first a set of examples from the Primary Corpus, for which a ‘remote pastness’ reading along the lines of Comrie 1985 is plausible:

(108) ... But this is the universal truth that women are much more superior to men.
God had bestowed women with a more beautiful and attractive countenance.
They are more winsome than men ...

[IND2000 Ess 09]

(109) Today if we look back around hundred years ago we will find that there had been a lot of change in our Indian society. Or in other words we can say that

60. Sharma (2001: 361–2) found a similar cline of ‘nonstandardness’ in the use of the past perfect in the Kolhapur Corpus. Of all uses of the past perfect (except for those with hypothetical meanings) in text categories A and H, Sharma 2001 arrived at the following proportions of non-standard uses per text category: National Press 14%, National Bureaucracy 23%, Regional Press 27%, and Regional Bureaucracy 50%. Likewise, Rogers 2002 found the proportion of ‘different’ uses of the past (and present) perfect in IndE to be particularly high in e-mail correspondence, “a register that approximates some registers of spoken language in its degree of informality” (Rogers 2002: 195), whereas ‘nonstandard’ uses were rare in more heavily edited forms of writing such as news reporting or editorials.

there had been a lot of developments in our Indian society with the adoption of western culture. As the time had passed western culture had given us lot of knowledge and ways to develop our culture.

[IND2000 Ess 05]

- (110) <\$C><#>Mr Bandopadhyay both political parties are blaming each other <#>but the last time there had been such violence in uh Midnapore<,> the political parties had agreed that their workers would cooperate with the administration that they would<,> work together to try and bring peace in the area<,> <#>but yet<,> just a day after the violent incident again you are attacking each other

[IND2000 Broad Int 09]

In example (108), which has been taken from an essay on the role of women in Indian society, a student (speaking Hindi as the first language) refers to the remotest of pasts – the creation of the world – and uses the past perfect where standard English would likely have the simple past. The use of two present-tense forms preceding and following the past-perfect form supports the interpretation of the past perfect signaling ‘remote pastness’, as no obvious point of orientation in the past is recoverable from the context that would indicate an intended ‘past before the past’ meaning. Similarly, in examples (109) and (110), past-perfect forms are used after future-tense and present-tense forms respectively to refer to completed actions in the past. Again, the simple past would be a more likely option in those contexts according to my reading of the data. In all three examples, the past perfect does not hamper intelligibility and allows rather unproblematically for a reconstruction of the temporal order of the events referred to. It is possible, of course, that the past-perfect forms in (108) and (109) were triggered by means of interference from the students’ mother tongue, Hindi. But even if that should be the case, it is important to note that such uses of the past perfect are not unheard of in other varieties either. Algeo 2006, for example, lists several quotations from BrE sources in which the past perfect is used although “the past perfect is not clearly appropriate by the usual interpretation that it signals an action or state anterior to some other action or state” (27).⁶¹ Explaining his observation, Algeo 2006 states that “[t]he British preference for the past perfect appears to have produced it even when the context does not suggest it” (27). Sharma 2001, too, reports several instances of past-perfect forms with preterite meaning in BrE and concludes that “some ambiguity in pluperfect meaning can arise in native varieties” (356). If interference from languages like Hindi is therefore taken to be an influential factor in the examples above, it can only be claimed to reinforce an ambiguity inherent in the meaning of the standard English past perfect due to its high pragmatic sensitivity.

61. See the following example from Algeo (2006: 27): “Amy came in and stared at me until I had noticed the dirty sweater and holed jeans she had exchanged for her earlier get-up.” (BrE; from *The Green Man* by Kingsley Amis)

Hence, one should hesitate to speak of a structural divergence of IndE from standard English in that particular instance, but rather more carefully of an extension of the pragmatic possibilities of the standard English past perfect.

While examples (108), (109) and (110) correspond to Comrie 1985, the next set of examples shows that the past perfect, next to its temporal function, may also carry specific discourse-pragmatic functions in IndE, as has been suggested by Kachru 1987 and Leitner (1989a, 1989b). Not only do the past-perfect forms here refer to remote points in the past (and, as such, could be replaced by the simple past); they are also employed at specific turning points in discourse to highlight a macrotextual switch in temporal orientation from present to past. Consider example (111), which was found in a student essay:

- (111) It is the leading point that women in modern India are good administrators. Women rulers like Chand Bibi, Elizabeth and Rani of Shansi had proved to the world their solid worth. They ruled more efficiently than men and brought glory to their nations. Women are more intelligent and hard-working.

[IND2000 Ess 08]

Here, the shift from present tense to past perfect corresponds to a shift from 'argumentation' to 'narration'. While the past perfect (*Women rulers like Chand Bibi, Elizabeth and Rani of Shansi had proved to the world their solid worth*) situates the temporal point of reference of the passage firmly in the past, the simple past is used for elaboration once past-time orientation has been established (*They ruled more efficiently than men and brought glory to their nations*). Examples of this type of tense sequencing are not exclusive to student writing, but can also be found recurrently in the categories of broadcast speech and journalistic writing. Note that in examples (112), (113) and (114), the past-perfect forms co-occur with temporal expressions of the past (highlighted in bold print):

- (112) <\$A><#>okay<,> now<,> you know **a few days ago<,>** the Minister for Information and Broadcasting Mr Arun Jaitley he had appeared on this programme<,> <#>he was sitting where you are sitting<,> <#>he said one of the reasons why<,> the there has been a delay<,> in the uh introduction of the new Broadcast Bill<,> is because<,> after the Prime Minister decided to set up this task force on convergence uh<,> the various ministries Telecommunications Information and Broadcasting and your ministry Information Technology<,> they're all looking at a consolidated law<,> which will take into account the rapid changes in<,> technological developments that have taken place and<,> that is one reason why the Broadcast Bill as well as the new Telecom Bill as well as the Information Technology Bill has been slightly delayed ...

[IND2000 Broad Int 04]

- (113) One way of cutting down pollution and making sure that commuters aren't fleeced is by adopting a simple system. A clearly visible colour coding system

can be devised by which taxis can be identified, depending on whether they run on petrol, diesel or CNG. Accordingly, they should have different rate cards. The public will naturally prefer those running on CNG, which is far cheaper, and this in turn will encourage more taxi owners to switch over to CNG. Pollution will be reduced even as consumer interests are protected. Something similar had happened a few years ago, when rear engine autorickshaws were being manufactured. People preferred to take them because they were far more comfortable and less noisy. As a result front engine autorickshaws were soon phased out. [IND2000 Press Edit 09]

- (114) The judicial system hard pressed by mounting arrears needs to revitalise itself to meet the challenging scenario in the new millennium. Unless it opts for a progressive methodology to deal with the obsolete laws, the justice administration system might continue to remain static, a silent listener to the public outcry against injustice. **Not many weeks ago**, the Supreme Court had delivered a landmark judgment holding that the government would be liable to pay compensation to a woman on whom its doctors had performed an unsuccessful operation for birth control. It said that the government would pay for the upkeep of woman Santara's child born in spite of such the operation till the age of 18 years. [IND2000 Press Rep 05]

In the Primary Corpus, the past perfect was also found to occur in text- and turn-initial positions to place the actions, states and events reported on firmly in the past. Note again the co-occurrence of the past-perfect forms with past-time adverbials in examples (115) and (117):

- (115) Sir, – My wife and I had invested in a fixed deposit (FDR 61673, date of maturity, **22 January 1998**) in DCM Financial Services ... [IND2000 Press Edit 08]
- (116) <\$A><#>Zee News had invited Mothers' Day messages from people all over India<,> <#>the response has been overwhelming<,> <#>for the ten best messages watch our English bulletin at eleven tonight<,> [IND2000 Broad News 20]
- (117) <\$D><#>Sonal now you're actually you're just seventeen years old so you're still actually in the junior circuit which you have been playing<,> <#>but now you're playing the ladies' circuit as well <#>how do you manage to uh yah to juggle between the two circuits<,>
<\$E><#>uh I'd been playing the junior circuit **last year**<,> <#>and uh this year is my uh I'll be turning eighteen on the ninth of June <#>so I'll get to play the juniors only till uh December two thousand <#>so I decided maybe<,> after that I have<,> as it is I have to play the women circuit so maybe start<,>

you know a year earlier so I get more advantage by playing uh good tournaments

<\$D><#>Sonal you'd played the Australian Open in **January this year** where you reached the second round <#>you were in the main <{><[>draw</><#>uh what are the plans for the other Grand Slams later this year

<\$E><#><[>right</><#></>

<\$E><#>uh I have got<,> uh entry in the qualifying for the French and I'm sure to get into Wimbledon's <#>so<,> uh that's the junior tournaments' line-up for the year two thousand [IND2000 Broad Int 01]

In example (115), the past perfect is employed at the beginning of a letter to the editor, in (116) at the beginning of a broadcast news report. Similarly, in example (117), two past-perfect forms are used in spontaneous speech in turn-initial position. The function of the past perfect in those cases can, in my opinion, be described adequately in the words of Leitner 1989a, according to whom the choice of the past perfect in such environments "precludes the expression of such notions as 'immediacy', 'on-the-spot-ness'" (175).

In example (118), similar to example (111), a past-perfect form is found in what can be termed 'narrative-initial' position, cf.

- (118) <\$C><#>see<,> the point that that that Pamela made<,> was<,> I would like to uh first answer that<,> <#>in Mathura there were three incidents<,> <#>one of them was between the uh the head of the institution and the ex-D-G-C and the president and uhm one Communist Party member<,> <#>both were guardians of students who were studying there<,> <#>suddenly they had hiked up their fees to around five hundred rupees<,> and there was a lot of uh uh argument and that uh culminated into some kind of a quarrel<,> <#><{>it had nothing to do <#>then comes the Mathura Mathura incident where the priest was<,> struck down<,> by <#>now that was a typically tribal I don't know there there is a uh a ex uh erstwhile criminal tribe

[IND2000 Broad Disc 05]

The past tense would arguably be a more likely tense form by standard English usage norms in the example above. From a discourse perspective, on the other hand, the past perfect can be said to fulfill an additional macrotextual function by signaling the onset of narrative action. After the scene of the narrative has been set by providing information on setting and characters in the past tense, the telling of the action sets in (<#>suddenly they had hiked up their fees to around five hundred rupees<,> and there was a lot of uh uh argument and that uh culminated into some kind of a quarrel<,>). It is at this junction that a new reference point in the past is introduced by means of using the past perfect (i.e. the moment after the dramatic rise in fees). This switch from past tense to past perfect is therefore not haphazardous but is related to the switch

from 'showing' mode to 'telling' mode, functioning meaningfully as a pragmatic focusing device advancing the plot of the narrative.

In contrast to (118), the next example (119) suggests that, within longer stretches of narrative, the function of the past perfect in IndE may sometimes be reduced to that of backgrounding information in relation to its context, with the past perfect shedding its temporal referential function completely, cf.

- (119) <\$B><#>yeah in fact I used to take part in all the school plays debates everything elocution<,> <#>I used to be the first one really to to recite my poetry because I like mugged it up like four days on end <#>and<,> I remember that play <#>in fact I even played uh Jesus Christ<,> and I'd worn yeah and I'd worn my mother's yoga dress<,> like her kurtah a white kurtah as a robe<,> <#>I'd taken it to wear it and a dupatta from her <#>and<,> I acted in a lot of school plays [IND2000 Broad Int 03]

While recounting an episode from her childhood using the simple past, an actress remembers several details about her costume in a school performance. This information is added to the narrative by means of using the past perfect. The information provided by the speaker in *and I'd worn yeah and I'd worn my mother's yoga dress* is not essential to an understanding of the storyline and, in this regard, can be claimed to provide 'background' information. By standard English rules, the past progressive would arguably be a more likely choice to describe that particular situation in the past. Only the third past-perfect form in the passage above (*I'd taken it to wear it*) can be given a past-before-the-past reading, as the narrated incident of the speaker taking a piece of clothing from her mother's wardrobe happened prior to the act of performing in the school play.

Examples (120) and (121) finally illustrate cases for which it becomes increasingly difficult to give the choice of the past perfect a coherent interpretation even if we take into consideration the possibilities that the past perfect may refer to the remote past and, at the discourse level, be used to highlight or background information. In these examples, the present perfect would be a more natural choice by the rules of standard English. See example (120) first:

- (120) Adoption of western culture has helped a lot in increasing the standard of women in India. Womens [sic] are seeing the british [sic] women and see them working at offices. They get inspired from these british [sic] women and to have a better standard in society they had started working. Now womens [sic] are in no field behind from [sic] men. They have gained encouragement after seeing british [sic] women and had started working. This has proved to be better and had helped in advancement of women society [sic] ... [IND2000 Ess 06]

Starting out with an unproblematic present-perfect form (*has helped*), example (120) continues in the second sentence with a description of the status and the behavior of Indian women in society, for which the present progressive (*is seeing*) alongside the simple present (*see*) is used. The statement *they had started working*, which occurs twice in the following two sentences, can hardly be given a remote-pastness reading, as this would contradict the choice of the present tense in *they get inspired* and that of the present perfect in *they have gained encouragement*. Nor does the past perfect seem to have a discursive function here. In fact, the change from present perfect to past perfect in the last sentence seems haphazardous. It appears that, above all, learner difficulties in mastering the English tense and aspect system were involved in creating the mix of forms observable here. The co-occurrence of several linguistic features that are clearly nonstandard, for example the plural *womens*, the spelling <british> and the compound *women society*, supports this interpretation.

Example (121), from another student essay, displays a similar mix of forms:

- (121) Wherever the question concerning the development of Indian Society is raised, there is always turmoil and confusion among the people and the society. Somebody says that the western culture had contributed a lot in the development of Indian society while other says that this is not so. ... Today, the western [sic] culture had influenced the mind of a teenager so much so, that if he doesn't perform any western [sic] task he becomes restless. Children are learning all debasing habits. [IND2000 Ess 04]

Here, only the first past-perfect form *had contributed* can be given a 'remost pastness' reading along the lines of Comrie 1985, as the temporal reference to *the development of Indian society* is possibly meant to be a reference to the remote past. The second form *had influenced* cannot be interpreted in this way, as the context makes it clear that the temporal reference is to a period of time lasting up until the present (*today*). From a discourse perspective, however, it would be possible to argue that the statement *the western culture had influenced the mind of a teenager so much so* provides background information to the main statement that *if he [the child] doesn't perform any western task he becomes restless*, which itself paves the way for the final conclusion that *[c]hildren are learning all debasing habits*.

Although uses of the past perfect of the type illustrated in (120) and (121) are also found in the Primary Corpus outside the student essays section, they are conspicuously rare there, which raises some doubt as to their general acceptability in professional writing and speech. In the next three examples, unusual past perfect forms were found to occur in spontaneous speech. Again, the present perfect would arguably be the more likely choice, cf.

- (122) <\$B><#>what the Chief Minister<,> Mr Karunanidhi has said<,> during his birthday<,> was blown out of proportion <#>certain portions have been taken out<,> and they have blown out of proportion<,> <#>he had made

some suggestion this time<,> <#>he did not say that he is against the decision of the government [IND2000 Broad News 01]

- (123) <\$A><#>in sports news this morning<,> a last-minute penalty allowed Botafogo<,> to escape with a two-all draw in the first leg of their Brazilian cup tie<,> against Internacional<,> <#>Botafogo will host the return leg<,> and their two away-goals had given them the upper hand in the second-round tie [IND2000 Broad News 05]

- (124) <\$G><#>this is you're planting the wrong trees <#>where you do plant you plant eucalyptus which feeds further desertification because it is very water-demanding <#>and you're destroying the indigenous species the treasury that the first chipko three hundred years ago<,> saved and on the basis of this the entirely traditional religion had thrived in the desert regions<,> <#>you have destroyed all your conservation culture and your conservation technology<,> including the conservation<,> of water<,> on small-scale<,> where it falls [IND2000 Broad News 06]

But it is important to note that similar uses are occasionally also found in other varieties of English, including the prestige varieties, and therefore cannot be considered features limited to IndE alone. When analyzing the past perfect in the press sections of the BROWN, LOB, and Kolhapur corpora, Sharma 2001, for example, came across several past-perfect forms ambiguous in meaning between the present perfect and the past perfect in her BrE data, which led her to the conclusion that variation in IndE should be seen as “extending out of an ambiguous usage already present in native standard English” (366). The following excerpt exemplifies another such occurrence of the past perfect with present-perfect meaning found in the IndE newspaper section of the Primary Corpus:

- (125) However, the job of balancing the singles and doubles commitments is a ticklish one, and the star himself was candid enough to admit it. “Doubles is my bread-and-butter as it had been our priority for the last two years. It's not easy for the body for take such a beating day in and out, especially when you are 27 and had been in the circuit for so long. Frankly, even I don't have the answer on how to strike a balance between the two,” he trailed off. [IND2000 Press Rep 08]

Interestingly, the two unusual past perfect forms in example (125) made their way into the press corpus in a passage of quoted speech, which renders the excerpt similar in style to examples (122), (123) and (124) from IndE speech. From a bird's-eye view, example (125) must be considered an exceptional instance in the category of quality newspaper prose. Both quantitatively and qualitatively, IndE national newspapers are broadly in line with the codified norms of standard English.

5.4.3 Summary

As is the case for the use of articles in IndE, the analysis of the past perfect has revealed that one should hesitate to characterize IndE grammar stereotypically in terms of an 'overuse' or 'underuse' of standard English forms. While it is possible to delineate contexts in IndE in which the past perfect is nowadays used in variety-specific ways when taking Quirk et al. 1985 as a yardstick, there is no monolithic explanation for the nature of variability at large. The qualitative investigation of the corpus material shows that IndE users have appropriated the pragmatic sensitivity of the standard English past perfect to their own communicative ends. The IndE past perfect is regularly used with past-time reference in contexts without an obvious point of orientation prior to the time of writing or speaking, which might possibly be reinforced by L1 interference. The past perfect can also be found with present-perfect meaning. But since such uses can be found in BrE and AmE as well, they should not be regarded as independent IndE innovations proper but as extensions of the pragmatic possibilities already inherent in the standard English past perfect. Corpus evidence also shows that the use of the past perfect with past-tense meaning in IndE has given rise to the recurrent use of past-perfect forms in text- and turn-initial positions and as a discourse-pragmatic highlighting device which can be employed in various ways to foreground or background the information value of stretches of discourse in relation to their contexts. The quantification of the 'pragmatic innovativeness' of IndE has proved difficult since IndE remains embedded in processes of variation and change affecting the use of the past perfect in Englishes worldwide (e.g. in the context of suspending backshift in reported speech). There is evidence, though, that nonstandard uses of the past perfect are not distributed evenly in IndE across registers and text categories today but are nowadays more common in less heavily edited forms of speech and writing, for example in broadcast interviews and discussions and in student essays. India's quality newspapers, on the other hand, adhere largely to the norms of standard English. Whether nonstandard uses of the past perfect will spread further in the future is a question of whether or not users will be willing to employ those forms in higher numbers. Prescriptive pressure is likely to remain a stumbling block for further innovation (cf. Nihalani et al. 1979: 95, 2004: 94–5; Yadurajan 2001: 117–8). See in this respect Yadurajan's (2001) critical comment:

The verbal group with *had* (*had received/written* etc.) is frequently misused in Indian English. There is a belief that an event in the distant past should be expressed with this type of structure (called the "past perfect"). Thus: *Twenty years ago, I had lived in Delhi*. Another mistaken use of this form is seen in a sentence like: *Last week I had been to Mysore*. What is wrong with: *Last week I was in Mysore*? The use of the past perfect has nothing to do with the distant past ... Nor is there any need for the past perfect to convey the meaning that you were in Mysore last week and have now come back. Your friend can see that you are back from Mysore. (Yadurajan 2001: 117–8)

5.5 More variation in tense and aspect? Two observations

This chapter reports two more observations made in the Primary Corpus on alleged syntactic Indianisms: the use of the present tense with temporal *since*-elements (5.5.1) and the use of the present perfect with temporal adverbials of the past (5.5.2). Since both features had been listed as grammatical Indianisms without providing information on their frequency, stability and acceptability, the corpus material was turned to for empirical evidence.

5.5.1 *Since*

In standard English, *since* can be used to introduce prepositional phrases or clauses specifying backward time spans “up to the ‘now’ of primary concern to the speaker” (Quirk et al. 1985:537). In such cases the perfect aspect is the regular choice in the matrix clause accompanying the *since*-element. See the following two illustrative examples:

- (126) NDA allies, while refraining from criticising the President’s speech, felt that there could be no better occasion to begin a debate on the issue keeping in mind the changes that the country has undergone in the five decades since independence. [IND2000 Press Rep 10]
- (127) And it because it has been sponsoring an award, the Chameli Devi Award for an Outstanding Woman Journalist, since 1982, it was but natural that they bring out a volume on the progress women have made in the profession ever since they first ventured into it. [IND2000 Press Rev 08]

IndE has repeatedly been said to use the present tense in the matrix clause in such contexts.⁶² According to Trudgill/Hannah 2002, “the use of the present tense with durational phrases (indicating a period from past to present)” (133) must even be considered a characteristic feature of IndE, although the authors admit that it is “unusual in more educated IndEng [Indian English]” (133). However, the situation in IndE is likely to be somewhat more complex, as nonperfective forms in the matrix clause are by no means uncommon elsewhere in the English-speaking world either. According to Quirk et al. (1985:539, 1015–6), the simple present or the present progressive are, for example, common in informal AmE “where the clause in which the *since*-adjunct operates refers to the present” (539). In informal BrE, too, those forms are said to be on the rise. Moreover, nonperfective forms are possible variants in the matrix clause when its predication is durative or when its reference is to habitual actions (cf. Quirk et al. 1985: 1016–7). This leaves one wondering how ‘different’ contemporary IndE usage is.

62. Cf. Verma (1978:213); Parasher (1991:177); Gramley/Pätzold (1992:443); Hansen et al. (1996:225); Trudgill/Hannah (2002: 133).

Table 5.57 Temporal *since* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
since	18	35	25	27	26	11
since + perfect	17	33	22	24	26	11
since + present	1	2	3	3	–	–

When looking up all occurrences of *since* plus ‘time unit’ in the Primary Corpus, uses of the simple present and the present progressive were the exception rather than the norm in IndE. For the press texts, the distribution shown in Table 5.57 emerged.

Table 5.57 shows no quantitative signs that IndE journalistic writing should be expected to differ notably from BrE or AmE writing. Interestingly, the present tense appears to be slightly more of an option in BrE newspapers than in AmE newspapers. If it is true that the present tense is more frequent a choice in informal AmE, this has apparently not been carried over into quality newspaper prose to any measurable extent in the admittedly small Primary Corpus (cf. Quirk et al. 1985: 539). On a closer look, the three IndE examples of *since* elements that are used with the present tense conform to the norms of standard English in that they feature stative uses of *be* to indicate the duration of states – a well-established possibility in standard English according to Quirk et al. (1985: 1016). Cf.:

- (128) In Finland, conditionally suspended sentence and conditional release is an established practice since 1972. [IND1978 Press Rep 18]
- (129) Some people entered the post and burnt and ransacked it. They also whisked away Ishaque Mollah, who is since absconding. [IND2000 Press Rep 11]
- (130) IT is now over a fortnight since the severe drought conditions affecting four states were highlighted by the media and the Prime Minister asked the nation to help deal with the calamity. [IND2000 Press Edit 02]

See two similar examples from BrE quality newspapers:

- (131) It is now four years since teachers’ pay negotiating rights were withdrawn by Kenneth Baker after two years of disruption in the schools – plus even more years of inter-union disagreements the Burnham Committee. [UK1991 Press Edit 02]
- (132) THIS is Adrian Noble’s first production since he took over the RSC’s orb and sceptre, and it is one which suggests that, whatever the company may lack during his reign, it will not be intelligence, subtlety or feeling for language. [UK1991 Press Rev 05]

Table 5.58 Temporal *since* [Primary Corpus; broadcast and essays].

	IndE IND2000 Broad News	IndE IND2000 Broad I/D	IndE IND2000 Ess	BrE UK1991 Broad News	BrE UK1991 Broad I/D
since	15	9	4	14	7
since + perfect	12	6	2	12	7
since + present	3	3	2	2	–

In the IndE material, the proportion of present-tense forms with *since* increases when moving from scripted speech (Broadcast News: 20%) across spontaneous speech (Broadcast Interviews and Discussions: 33%) into the category of student writing (Student Essays: 50%). This suggests that present-tense forms might be more of an option in more unplanned and more informal language, while they are rejected in more professional language use (for example quality newspaper prose or English-language television news broadcast). This would confirm Trudgill/Hannah's (2002: 131) observation, cf. Table 5.58.

When looking at the attestations in context, however, no further qualitative difference between IndE and the codified norms of standard English emerged. The following example from a broadcast news report involves the use of the present tense with stative *be* to express the duration of a state, which is a possibility in standard English, too:

- (133) <\$G><#>the Karnataka legislature is in session for the first time<,> since Dr Rajkumar was kidnapped by Veerappan<,> <#>and there was some support<,> from the opposition on the issue<,> for the Karnataka Chief Minister S-M Krishna
[IND2000 Broad News 11]

See a similar example from a BrE news broadcast:

- (134) <#E>It's now thirty-two hours since she began her shift
[UK1991 Broad News 11]

As Quirk et al. 1985 put it, “[t]he simple present or the present progressive is sometimes used in the matrix clause when that clause has habitual reference” (1017). The next three examples from IndE can be subsumed under this category:

- (135) <\$A><#>Donald who's currently playing county cricket in England said <#>my whole career with him just keeps flashing past me since I found out <#>it's incredible<,> <#>it's just greed really that's made him do this
[IND2000 Broad News 09]
- (136) Since the civilization of the Indians, the people here live in joint families.
[IND2000 Ess 07]

(137) Since independence our country is trying to develop ... [IND2000 Ess 06]

Example (138), found in spontaneous speech, is arguably the most unusual example in the Primary Corpus in that a speaker uses *since* elements twice and switches spontaneously from present perfect progressive to present progressive within one turn, cf.

(138) <\$B><#>and we we honour great musicians by Hafiz Ali Khan Award uh
<#>this has been happening since uh you know my father passed away in
nineteen seventy-two<,> and since then we are honouring all the great musi-
cians of India and abroad [IND2000 Broad Int 05]

Interestingly, uses going further beyond the norms outlined in Quirk et al. 1985 are absent from the Primary Corpus. Hence, as far as the usage range covered by the Primary Corpus is concerned, IndE should be considered broadly in keeping with standard English.

Note that the investigation of the corpus material also brought to light that the semantic distinction between *since* and *for*, which according to Nihalani et al. (1979, 2004) is often confused by “non-native speakers of English all over the world” (1979: 162; 2004: 163), is basically adhered to in the IndE material. There were only two instances in the entire Primary Corpus of *since* replacing *for*, cf.

(139) A flurry of caste parties waiting to be wooed by the DMK and the AIADMK represents a defeat for the Dravidian brand of politics practised by these two parties. Especially since all these years they had managed to cut across caste barriers, especially among the non-Brahmins. [IND2000 Press Rep 03]

(140) Based on its own evaluation report, questions can definitely be raised about the effectiveness of the World Bank in poverty alleviation since the last 50 years. [IND2000 Press Edit 05]

Since the conflation of *since* and *for* has long been pointed out as an error in school grammars and usage handbooks (cf. Wren et al. 2001: 157–9), it is unlikely that uses of the types represented in examples (139) and (140) should become any more common in professional writing or speech in the future (cf. Sheorey 2006: 190). The situation may well look different in less heavily planned or less professional uses of IndE, of course.

5.5.2 Adverbials of the past

The use of the present perfect with adverbials of the past is another phenomenon that has been pointed out as a characteristic feature of IndE.⁶³ Parasher 1991 even observes

63. Cf. Verma (1978: 213), Verma (1982: 183), Bansal (1983: 4), Gramley/Pätzold (1992: 443), Hansen et al. (1996: 225), Shekar/Hedge (1996: 60), Rogers (2002: 194–5), Hickey (2004: 546),

“a strong tendency in IE [IndE] to use the present perfect with adverbials of past time” (175), a claim which he illustrates by examples such as *they’ve done that three years ago*. On the other hand, the present perfect with *ago* was found to be one of the least acceptable patterns in Sahgal/Agnihotri’s (1985) and Sheorey’s (2006) attitude surveys. Judging the acceptability of the sentence *I have read this book yesterday*, 31.82% of Sahgal/Agnihotri’s (1985) informants labeled it ‘good English’, 9.09% ‘informal usage’, and 59.09% ‘wrong English’. On finding the usage in question at position 20 (on a list of 27 items ranging from ‘acceptable’ to ‘unacceptable’) in their implicational hierarchy, Sahgal/Agnihotri 1985 concluded that it “will be accepted as correct only by those informants who accept almost all deviant constructions tested as ‘good English’” (124). A similar result was produced by Sheorey (2006: 191), whose informants judged the sentences *he has purchased the land last year* and *we have sent them two reminders last month* to be seriously erroneous. Not surprisingly, the pattern in question is also rejected by Yadurajan (2001:7) in his usage guide and is pointed out as incorrect usage in school textbooks like Wren et. al. 2001:

Note that the Present Perfect is never used with adverbs of past time. We should not say, for example, ‘He *has gone* to Calcutta yesterday’. In such cases the Simple Past should be used (‘He *went* to Calcutta yesterday’). (Wren et al. 2001: 83)

To tackle the question whether IndE could still be said to be distinctly more variable than BrE or AmE in its choice of tenses with past-time adverbials, all occurrences of the adverbials *yesterday* and *ago* were looked up in the Primary Corpus and ICE-India and analyzed for the tense forms accompanying them. No evidence was found, however, that would support Parasher’s (1991) claim. While the BrE and AmE corpora did not return a single example of the present perfect, only two examples were found in the IndE section of the Primary Corpus vis-à-vis several dozen standard occurrences, which raises some doubt as to the validity of Parasher’s (1991: 175) trend hypothesis for IndE at large (cf. Table 5.59).

See examples (141) and (142) from IndE, the only two attestations of *ago* used with the present perfect in the Primary Corpus, cf.

(141) <\$D><#>Sonal of all these uh tournaments that you’ve played obviously the biggest one was the ten-thousand-dollar event that you’ve played in Mumbai a couple of weeks ago <{><[>where you lost</> in the final to Sai Jayalakshmy <#>was that your best showing this season [IND2000 Broad Int 01]

(142) Having tackled the problems of life and attained celestial bliss in different literary veins, these poet-saints have long ago achieved the status of transcendental philosophers whom even agnostics may read with profit. [IND1978 Press Rev 10]

Table 5.59 Tense with *ago* and *yesterday* [Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Ess
ago + past tense	22	24	25	1
ago + present perfect	1	–	1	–
yesterday + past tense	12	5	49	–
yesterday + present perfect	–	–	–	–

Table 5.60 Tense with *ago* and *yesterday* [ICE-IND].

	IndE ICE-IND writing	IndE ICE-IND speech
ago + past tense	66	51
ago + present perfect	1	4
yesterday + past tense	16	211
yesterday + present perfect	2	15

A look at the larger ICE-India shows that the present perfect with the past-time adverbials *yesterday* and *ago* is, indeed, a recurrent phenomenon – but it remains a relatively rare choice in contemporary IndE, cf. Table 5.60. Of the 22 occurrences of the present perfect with *ago* and *yesterday* identified in ICE-India, 19 were found in the speech section of the corpus. Two of the three occurrences in the written section were found in student essays and one in a private letter. On the other hand, the present perfect is noticeably absent from more carefully edited forms of writing, for example newspaper reportage, which makes example (142) look like an exception and suggests a rather low degree of acceptability of this feature among users of IndE at present (cf. Sahgal/Agnihotri 1985: 124; Sheorey 2006: 191).

Note in this context that several descriptions of IndE have also reported an unusually high degree of variability in the lexical choice between *ago*, *back* and *before*. But although IndE has been said to replace *ago* with *back* frequently (cf. Yadurajan 2001: 7; Nihalani et al. 1979: 28, 2004: 26) or *before* (cf. Nihalani et al. 1979: 33, 2004: 31), a look at the Primary Corpus reveals that this is not true for IndE journalistic writing, cf. Table 5.61. As is the case for the present perfect with *yesterday* and *ago*, *back* ‘ago’ appears to be slightly more common a choice in IndE in less professional writing and in speech. On the other hand, uses of *before* ‘ago’ are rare throughout, with the only attestation of its kind found in a student essay. It appears that we are dealing with a typical learner mistake (cf. Table 5.62).

While IndE usage may well differ more radically from the norms outlined in Quirk et al. 1985 in other text types and registers not covered here, this is not true for the

Table 5.61 *Ago, back and before* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
ago	22	24	25	47	23	25
back 'ago'	2	–	1	1	–	1
before 'ago'	–	–	–	–	–	–

Table 5.62 *Ago, back and before* [Primary Corpus; broadcast and essays].

	IndE IND2000 Broad	IndE IND2000 Ess	BrE UK1991 Broad
ago	27	1	43
back 'ago'	4	2	–
before 'ago'	–	1	–

material studied, and any description of IndE holding up those particular features as characteristically IndE phenomena runs danger of oversimplifying the linguistic situation in the Indian subcontinent unjustifiably.

5.6 Verb phrase: The mandative subjunctive

The mandative subjunctive has been in the limelight among corpus linguists working on ongoing language change in present-day English for quite some time, and interesting regional differences have been uncovered that have helped characterize varieties of English in their own right as well as in terms of their openness toward ongoing language change worldwide. Mandative subjunctive forms occur in subordinate *that*-clauses that are “introduced by an expression of demand, recommendation, proposal, resolution, intention, etc.” (Quirk et al. 1985: 156). See example (143) from IndE:

- (143) <\$A>The protesters demanded that the Minister give an assurance to apprise the Chief Minister, Mr. N. Chandrababu Naidu, and ensure rolling back of power tariff. [IND2000 Broad News 01]

The mandative subjunctive is formally distinct only when a singular subject is followed by a present-tense verb in the third-person (as in example (144)) or when the subject is followed by *be* (as in example (145)):

- (144) Back from a self-imposed political exile, BJP MP Uma Bharati on Sunday threatened to raise yet another banner of revolt, demanding that the BJP agree to a special quota for women belonging to other backward classes (OBCs) in the Women's Reservation Bill. [IND2000 Press Rep 05]
- (145) By midnight a crowd had gathered in front of the police post, demanding that Ishaque be released. [IND2000 Press Rep 11]

On the other hand, if the verb is not in the third person or if the subject is a collective noun, e.g. *government*, it cannot be distinguished straightforwardly from the indicative:

- (146) <\$D><#>but with<,> with our safeguards Mr Amar Singh<,> we will not allow you<,> whole choice no no Mr Amar Singh has to agree with the safeguards in this bill and I recommend you go through them [IND2000 Broad Disc 08]
- (147) <\$A><#>the party has also demanded<,> that the government implement all the recommendations of the court of inquiry<,> ordered into the Bharatpur depot fire<,> [IND2000 Broad News 07]

Note that the mandative subjunctive competes with periphrastic constructions, most commonly the putative *should* variant, cf.

- (148) <\$B><#>the Chief Minister of Tamil Nadu has again suggested<,> that Sri Lanka should be divided into separate countries<,> and this is in spite of the B-J-P asking him not to say so [IND2000 Broad News 06]
- (149) The committee had recommended that Mr. Basu should not be allowed to continue as the Controller of Examinations or in any other post in the University [IND2000 Press Rep 17]

But other modal variants are available, too, each showing slight differences in meaning. As example (150) illustrates, the mandative force of a modal alternative like *could* sometimes becomes clear only when studying occurrences carefully in context:

- (150) <\$D><#>he is not a Prime Minister of the B-J-P<,> <#>he is the Prime Minister of India<,> <#>so when the Prime Minister of India gets up and says ... there should be an end to conversions you should there should be a debate on conversions<,> what is all this ...
- <\$G><#>I don't know how neutral I can be in this because I also am terribly disappointed with the way the Prime Minister of this country has responded to what I think is a crisis of confidence<,> besides which I think the<,> Chief Minister of the largest state in the country<,> which is<,> U-P<,> where many of these attacks have taken place has not moved uh a finger it seems

<#>it he hasn't come out and made a public statement <#>he hasn't<,> suggested<,> that there could be a confidence-building measure

[IND2000 Broad Disc 05]

Finally, next to periphrastic constructions, the mandative subjunctive also competes with the indicative:

(151) While wisdom would demand that the demand is conceded, it is doubtful if the Punjabi feudocracy will oblige. [IND2000 Press Edit 04]

(152) <\$C><#>no <#>uh this is uh very powerful for the citizens charter needs to make the citizen<,> empowered make him aware of what he is expected to get to the government department<,> and insist that he gets it<,> and that the government comes upfront and says it will also be available corruption-free<,> [IND2000 Broad Int 01]

Previous research has shown BrE and AmE to differ substantially in their choice between subjunctive, modal and indicative options.⁶⁴ As Algeo 2006 sums up the situation, "British English uses all three of the options; American uses primarily the subjunctive but accepts the modal. The indicative option is characteristically British." (263) But corpus evidence also suggests a revival of the mandative form in BrE and throughout the English-speaking world in the second half of the twentieth century, possibly under the growing influence of AmE on world Englishes after World War II (cf. Övergaard 1995: 51, 54; Hundt 1998b: 159). This makes the mandative subjunctive a particularly interesting variable for a description of IndE, as it offers another possibility to gauge the influence from BrE and AmE in the Indian subcontinent in a diachronically sensitive fashion. One of the earliest available corpus-based studies of IndE, in fact, investigated the use of the mandative subjunctive in IndE writing. In a paper modeled on Johansson/Norheim's (1988) seminal corpus study of the usage differences between BrE and AmE revolving around the mandative subjunctive, Sayder 1989 compared the 1978 Kolhapur Corpus to the 1961 BROWN and LOB corpora and found that IndE users employed the form significantly less frequently vis-à-vis the periphrastic *should* variant than AmE but significantly more frequently than BrE users (also cf. Leitner 1991b: 222–4; Schneider 2000: 124–30). Relating his results to differences in "textual norms and textual or discourse strategies" (Sayder 1989: 65) across text types and registers, Sayder 1989 argued that the mandative subjunctive had "an 'intermediate' overall position" (65) in IndE in relation to AmE and BrE and concluded that the observed differences between IndE and AmE and BrE were "of a quantitative, not a systemic nature" (65). But on considering the time difference of seventeen years between the compilation of the Kolhapur Corpus (1978) and the

64. Several elicitation experiments have been conducted to pinpoint differences between BrE and AmE in using mandative constructions; cf. Johansson 1979, Turner 1980, Nichols 1987, Algeo 1992.

LOB and BROWN corpora (1961) as well as the fact that written AmE and BrE had undergone change during that timespan, Sayder’s (1989) study called for an update. Had IndE been affected by the reintroduction of the mandative subjunctive over the years and, if so, to what extent?

Table 5.63 brings together the available quantitative findings on the distribution of the mandative subjunctive vis-à-vis the modal *should* variant in world Englishes based on the parallel one-million-word corpora of IndE, BrE, AmE, AusE and NZE. The indicative variant, “the most characteristically British form” (Algeo 2006: 264), and other modal variants were not considered in those studies:⁶⁵

Table 5.63 Mandative subjunctive [LOB-BROWN corpus family].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992	AusE ACE 1980s	NZE WC 1980s
subjunctive	35	14	44	116	94	78	70
should	81	97	67	19	11	29	35

Table 5.63 shows that written AmE was characterized by a diachronically stable preference for the mandative subjunctive vis-à-vis the periphrastic *should* variant in the second half of the twentieth century. The preference in the BrE data, on the other hand, went with the *should* variant, although it is clear that the mandative subjunctive became more of an option also in BrE in the three decades between 1961 and 1991, which has been accounted for in terms of the rising global influence from AmE after World War II. According to Hundt 1998b, the significant narrowing of the gap in numbers between subjunctive and periphrastic constructions between 1961 and 1991 points in this direction. Such a reading of the data is in line with the results produced by Övergaard 1995, who also found a sharp increase in mandative subjunctive forms in BrE after 1960 when analyzing a wide range of written material from twentieth-century BrE and AmE sampled at 20- to 30-year intervals. In contrast to the distribution in Övergaard’s (1995) text corpus, however, FLOB saw the *should* variant outnumber the mandative subjunctive, which Hundt 1998b explained in terms of differences in corpus design. More importantly, the fact that periphrastic *should* was still a viable second variant in AmE in the early 1990s was interpreted by Hundt 1998b as an indicator that the reintroduction of the mandative subjunctive, historically the older variant, had lost some of its force on a global scale at the close of the twentieth century. In view of the fact that late 1980s’ NZE and AusE were more similar to AmE than BrE in their preference for the subjunctive construction, Hundt 1998b concluded that AmE

65. The figures are taken from Schneider (2000: 128), who revisited the Kolhapur Corpus and corrected several inconsistencies in Sayder’s (1989) count.

was “the most advanced variety” (165) in the revival of the mandative subjunctive “followed by AusE with NZE close at its heels” (165).

In light of those results, Sayder’s (1989) interpretation that the mandative subjunctive has “an ‘intermediate’ overall position” (65) in IndE as compared to BrE and AmE must be reconsidered. What is a statistically significant difference between IndE and BrE when consulting Kolhapur (1978) and LOB (1961) vanishes when comparing Kolhapur (1978) to FLOB (1991). Here, no statistically significant difference between IndE and BrE written usage emerges, and the question of where IndE should nowadays be positioned vis-à-vis BrE is open again. To address this issue, more recent data from IndE were needed. The online corpora were therefore consulted and the frequency of the mandative subjunctive was compared to that of the periphrastic *should* variant across the top-level country domains and across India’s online newspapers in a selection of commonly recurring collocations. As the verbs *demand*, *insist*, *recommend* and *suggest* had been the four most common suasive verbs in the one-million-word corpora, it was hypothesized that their behavior would also be a fair representation of the position of the mandative subjunctive vis-à-vis the modal *should* variant in contemporary IndE at large. See, first, the behavior of the four verbs *demand*, *insist*, *recommend* and *suggest* in the Kolhapur Corpus and its parallel corpora from Britain, the US, Australia and New Zealand, cf. Table 5.64.

Table 5.65 gives the frequencies of the mandative construction and the putative *should* variant with the same verbs in collocations of the type *V-ed that it (should) be* across the top-level country domains on the World Wide Web (January 8, 2003). Although a Google-run search for selected collocations can, at best, only provide a glimpse of the actual range of functions of mandative subjunctive forms in the six varieties under investigation, it is interesting to see that the figures returned by Google for AmE and BrE in this admittedly limited usage context (cf. Table 5.65) mirror the distributions found in the closed corpora (cf. Table 5.64). As is the case in BROWN and FROWN, AmE usage under the ‘us’ domain displays a clear preference for the mandative subjunctive after all four verbs. In contrast, the mandative subjunctive is less of an option in BrE. Interestingly, when zooming in on the ‘uk’ domain, the mandative subjunctive is more common than the periphrastic *should* variant after the verbs *demand*, *recommend* and *insist*. Since those verbs had been followed predominantly by periphrastic *should* in LOB and FLOB, this finding might reflect a continuation of the spread of the mandative subjunctive into written BrE after 1991 with verbs of higher volitional force, possibly due to the global pressure exerted by AmE. Under the ‘uk’ domain, the mandative subjunctive remains a minor variant in the chosen collocations only after *suggest*, for which BrE usage turns out to be furthest removed from AmE. AusE, by contrast, is more advanced than BrE in preferring the mandative subjunctive with all verbs irrespective of their volitional force, another finding which is in line with the results produced by Hundt (1998a, 1998b) and Peters 1998. Among the six varieties represented in Table 5.65, IndE is clearly most reluctant in employing the mandative subjunctive, with periphrastic *should* being more common a choice after all four verbs and the mandative subjunctive being a second variant. SgE shows slightly

Table 5.64 Subjunctive vs. *should* with selected verbs [LOB-BROWN corpus family].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992	AusE ACE 1980s	NZE WC 1980s
demand +							
subjunctive	8	2	8	19	10	12	6
should	12	3	6	–	–	–	1
recommend +							
subjunctive	6	1	6	10	6	12	15
should	24	13	20	1	3	6	5
insist +							
subjunctive	3	–	3	9	10	10	10
should	2	8	3	2	2	2	3
suggest +							
subjunctive	12	2	4	12	23	7	17
should	26	34	13	7	5	8	18

Table 5.65 Subjunctive vs. *should* with selected verbs [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
demanded that it be	7	5	149	13	90	69
demanded that it should be	21	–	38	–	19	6
	1 : 3	–	4 : 1	–	5 : 1	12 : 1
recomeded that it be	28	5	19300	175	21700	20900
recommended that it should be	53	6	9700	28	268	71
	1 : 2	1 : 1	2 : 1	6 : 1	81 : 1	294 : 1
insisted that it be	5	2	126	13	163	75
insisted that it should be	8	8	96	5	32	15
	1 : 2	1 : 4	1 : 1	3 : 1	5 : 1	5 : 1
suggested that it be	12	13	10500	89	765	19100
suggested that it should be	92	11	18700	37	545	294
	1 : 8	1 : 1	1 : 2	2 : 1	1 : 1	65 : 1

greater openness to adopting mandative forms than IndE but less strongly than BrE. The distribution for S AfrE, on the other hand, falls between BrE and AusE.

What can be concluded from those observations for the position of IndE in a global context? The comparison of Table 5.64 and Table 5.65 suggests that if the general tide in twentieth-century written English was with reintroducing the mandative subjunctive, contemporary IndE has been rather more conservative in this process than the other five varieties, which modifies Sayder's (1989) static reading of IndE taking an intermediate position in relation to BrE and AmE. What emerges from the

Internet snapshot is that the mandative subjunctive is nowadays slightly less common in IndE than in BrE, at least as regards the domain of published writing on the Internet. SgE usage shows some similarities to IndE, especially when it comes to the behavior of the two verbs *recommend* and *insist*, which suggests a certain affinity among South and South East Asian Englishes meriting further investigation. In contrast, SAfrE and AusE usage is more advanced than BrE in adopting mandative subjunctive forms, thus displaying a behavior approximating more closely the AmE end of the scale. Note that the affinity of the Antipodean varieties AusE and NZE to AmE has previously been interpreted in terms of the increased external influence from AmE after 1945 (cf. Hundt 1998a:94; Peters 1998:89). If the postcolonial force of AmE is seen as the main factor behind the global distribution of the mandative subjunctive, both IndE and SgE would have been affected by this development to a lesser extent. A variety-internal reason for the mandative subjunctive to be less common in IndE than in BrE or AmE might lie in the fact that the mandative subjunctive is a rather peripheral phenomenon in English verb grammar which is given little coverage in English language teaching. In the latest edition of Wren et al.'s (2001) classic *English Grammar and Composition*, for instance, which is still one of the most popular textbooks of English in India's English-medium schools, the mandative subjunctive finds mention only in the appendix in a chapter on differences between British and American English. Here, students are provided with the information that "[t]he British often use *should* after *demand*, *insist*, *recommend* etc. The use of the subjunctive is rather formal in Britain. Americans normally use the subjunctive. They rarely use *should*." (Wren et al. 2001:472) The modal construction, on the other hand, is treated at length. Besides, prescriptive warnings in IndE style guides of the type found in Yadurajan (2001: 147–8), who refers to the subjunctive in BrE as being 'rather formal' and warns against its use outside the context of 'stock expressions', might contribute to keeping the proportion of subjunctive forms rather low in this variety.

But even if IndE, in the big picture, is using the mandative subjunctive less frequently at present than other varieties, it is certainly not a rare form, and varietal differences are mainly quantitative in nature. When checking on the distribution of the mandative subjunctive and periphrastic *should* in India's quality online newspapers (January 8, 2003), Chennai's *Hindu* displayed a remarkable preference for the subjunctive, whereas the preference in Chandigarh's *Tribune* was with the periphrastic *should* variant. This suggests the existence of house-internal stylistic conventions favoring one variant over another, with the style in *The Hindu* being more progressive than that of *The Tribune* – a finding which appears to correlate with earlier observations on lexical and syntactic style differences between those two major quality newspapers (cf. Chapters 3.1 and 5.3). The figures returned for the other four Indian quality newspapers were too low to be indicative of any clear-cut usage preferences, cf. Table 5.66.

That stylistic considerations may generally be a factor favoring or sanctioning the use of mandative subjunctive forms was also pointed out by Sayder 1989, who found that the mandative subjunctive forms in the Kolhapur Corpus predominated in rather formal registers and text types and co-occurred with several linguistic

Table 5.66 Subjunctive vs. *should* with selected verbs [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
demanded that it +								
be	–	28	–	1	23	2	11	5
should be	–	5	2	–	48	2	–	–
recommended that it +								
be	–	4	–	–	3	–	1	4
should be	–	6	–	1	10	–	1	–
insisted that it +								
be	–	3	1	–	1	–	6	5
should be	1	6	–	–	3	1	1	–
suggested that it +								
be	–	12	1	–	4	5	–	6
should be	–	4	2	–	22	2	2	–
SUBJUNCTIVE (total)	0	47	2	1	31	7	18	20
SHOULD (total)	1	21	4	1	83	4	4	0

markers of formal English, for example with the passive voice. More recent corpus-based research on spoken BrE, using the British National Corpus, has shown that the mandative subjunctive might currently be in the process of losing its formal connotations in Britain, though (Hundt 1998b). The following example from the IndE speech section of the Primary Corpus suggests that IndE, too, is slowly beginning to employ the mandative subjunctive in co-occurrence with stylistically more informal features such as *that* deletion:

- (153) <\$B><#>the rebels who continue holding thirty-one hostages including the deposed Prime Minister<,> Mahendra Chaudhry<,> have warned that negotiations for their release<,> could fail if the military has plans to restore Fiji’s nineteen ninety-seven constitution<,> which they claim gives too much power to ethnic Indians<,> <#>they are instead demanding the country’s ethnic Indian minority be stripped of political power
- [IND2000 Broad News 09]

Special attention in previous linguistic research has also been given to the negative mandative subjunctive (cf. Schneider 2000: 129). See an example from IndE (the only instance of its kind in the Kolhapur Corpus):

- (154) His moral sense demanded that he not allow an innocent person to suffer for him.
- [IndE KOL K 02]

According to Bauer (2002: 88), this construction came into being only in the twentieth century, possibly in AmE, from where it appears to have spread across the

English-speaking world. In Bauer's (2002) words, "[w]hile the history of this particular construction is rather obscure, it does seem to be one minor case where the syntax of a colonial variety has triumphed over the home construction" (88).⁶⁶ It can be established that the construction had arrived in IndE by 1978, as example (154) from the Kolhapur Corpus illustrates.

To conclude, the comparison of the distributions of the mandative subjunctive and the periphrastic *should* variant in the Kolhapur Corpus and across the World Wide Web suggests that IndE writing has been relatively reluctant to reintroduce the mandative subjunctive in the twentieth century. As compared with the major varieties BrE and AmE, with AusE and NZE and with SAfrE, IndE appears to be lagging behind. A variety-internal explanation for this phenomenon has been seen in the continuing prescriptive pressure against uses of the subjunctive in India plus the fact that the mandative subjunctive receives little coverage in the context of teaching English as a second language. It has also been shown that some publications may embrace the mandative subjunctive more readily than others, the preference in *The Hindu* being a case in point.

5.7 Sentence level: Inversion in direct and embedded *wh*-questions

The formation of direct and embedded *wh*-questions has been another point of interest in many feature-list descriptions of IndE.⁶⁷ In standard English direct *wh*-questions are formed with a *wh*-element at the beginning (i.e. *who*, *whom*, *what*, *which*, *where*, *when*, *why* and *how*; cf. Quirk et al. 1985:78) followed by the operator in pre-subject position and the subject "in all cases except when the *wh*-element is subject" (Quirk et al. 1985:818). Examples (155) and (156) illustrate this:

- (155) What does one say of a system that holds a corrupt act to be legitimate, only provided it takes place as part of a parliamentary function?

[IND2000 Press Edit 04]

66. Schneider (2000:129) reports six instances of the negative subjunctive for BROWN and none for LOB, which supports the view that AmE may be the origin of this construction.

67. For references to IndE cf. Spitzbardt (1976:43–4), Daswani (1978:122), Pattanayak (1978:199), Verma (1978:217), Nihalani et al. (1979:98), Das (1982:147), Verma (1982:181–2), Sahgal/Agnihotri (1985:123–4), Schmied (1989b:82–4), Leitner (1989a:173–4), Parasher (1991:170), Parasher (1994:149), Schmied (1994:223), Gramley/Pätzold (1994:442–3), Kachru (1994:520–1), Hansen et al. (1996:224–5), Shekar/Hedge (1996:60), George C (1998:18–9), Bhatt (2000:74–5), Trudgill/Hannah (2002:14), Bhatt (2004:1019–21), Hickey (2004:545–6) and Gargesh (2006:104). For references to other varieties cf. Platt et al. (1984:127) and Schneider (2007:85).

- (156) And how can one forget Mr Justice Ranganath Misra's report on the massacre of Sikhs in 1984 – a document used by Rajiv Gandhi to exonerate loyalist murderers? [IND2000 Press Edit 08]

On the other hand, standard English usually has no subject-operator inversion in embedded *wh*-questions:

- (157) The minister of state for external affairs attributed his faster return to some urgent pending files. Asked how it felt to be back in the ministry, Mr Panja said: "It's fine. I am back in my office and busy doing the work." [IND2000 Press Rep 02]
- (158) "I wonder why the same people did not come to me earlier to ask how I was preparing for the games with so little money," the iron lady quipped. [IND2000 Press Rep 08]

In contrast to standard English, IndE, like many other varieties of English (cf. Kortmann/Szmrecsanyi 2004: 1154), has been claimed to suspend inversion also in direct questions. See the following examples:

- Where you went yesterday?* (Daswani 1978: 122)
Where you are working now? (Sahgal/Agnihotri 1985: 123)
When you would like to come? (Kachru 1994: 520)

At the same time, IndE has been said to have inversion in embedded questions (cf. Kortmann/Szmrecsanyi 2004: 1179):

- Please advise why did the trustees file an application* (Parasher 1991: 240)
Subash knows who is your best friend (Bhatt 2000: 95)
I wonder where is he (Trudgill/Hannah 2002: 134)

Several interpretations have been offered as to why IndE should show variation in this field, with the spectrum of explanations ranging from general learner difficulties,⁶⁸ the influence from Indian languages⁶⁹ all the way to syntactic

68. According to Parasher (1991: 170; 1994 [1983]: 149), "[t]he source of deviation here [i.e. in indirect questions] is interference from the inverted word order of direct questions."

69. According to Pandharipande (1987: 151), embedded inversion in an utterance like *he asked me that what was I doing must be considered* can be explained in terms of transfer from Hindi. Sridhar (1992: 144–5), too, states that L1 influence is a relevant factor triggering variation. Sridhar 1992 discusses the impact of L1 interference on 'Dravidian English' as follows: "This characteristic may appear at first to be an instance of overgeneralization based on the overtaught though not overlearned rule of inversion in regular questions. However, another explanation is possible: Dravidian languages do not make a clear distinction between direct and reported speech, and there are no sequence of tense phenomena. In contexts where a reported speech

innovation.⁷⁰ But while there is agreement among linguists today that variability does exist in IndE question formation, inconclusive statements have been made regarding the actual frequency of those structures in IndE and their acceptability. A language generation ago, Daswani 1978 remarked that non-inverted direct questions and embedded inversion were restricted to “some subvarieties” (122) of IndE – but he failed to specify those any further. According to Nihalani et al. (1979, 2004), variation in the word order of direct and embedded *wh*-questions is to be considered “a common feature of the rapid, informal speech of many Indian speakers of English” (1979: 98, 2004: 98). Bhatt (2000: 74–7, 2004: 1019–21), on the other hand, argues that both embedded inversion and non-inverted direct questions are stable characteristic features of what he calls vernacular IndE. In addition, a small-scale attitude survey conducted by Bhatt 2000 points to a high degree of awareness among users of IndE at the turn of the millenium that those patterns are part of a local, rather than international, English usage norm. In an elicitation test with 27 speakers of IndE, Bhatt (2000: 73–4) found that more than 80% of his informants considered variability in question formation a typical feature of IndE (rather than being typical of BrE or neither of the two varieties). Fifteen years earlier, variation in question formation had been one of the least acceptable grammatical features among the educated users of IndE interviewed by Sahgal/Agnihotri (1985: 123–4). Judging the acceptability of the construction *Where you are working now?*, 30.23% of Sahgal/Agnihotri’s (1985) informants had considered this particular usage as an example of good English, 22.73% good enough for informal use, but 43.18% wrong English. Embedded inversion was also judged “deviant by all the three groups of informants” in Parasher’s (1991: 170) survey among users of IndE, AmE and BrE, and it was among the more unacceptable grammatical phenomena in

sentence would be used by a native speaker of English, Dravidian speakers typically use a verbatim quote (Lehmann 1989). Furthermore, the quotative particle in Dravidian, which functions generally as a complementizer, can embed not only speech verbs such as *say*, but also a variety of verbs of cognition/cogitation, including *know* (Ramanujan 1964; Sridhar 1990: 2). Given these facts, it is arguable that the inversion in embedded questions is at least reinforced by the substratum if not a direct product of it. This is an example of what one might call an ‘acquisitional conspiracy,’ that is, a situation where two apparently distinct strategies have identical surface effect.” (14–5)

70. According to Bhatt (2000: 74–7, 2004: 119–21), there is more to the formation of direct and embedded questions than the mere confusion of two similar structures. In his view, the formation of direct *wh*-questions without inversion resembles the fronting of constituents in declarative sentences, which itself has been said to be a frequent unmarked option in spontaneous IndE speech (also cf. Leitner 1989: 173–4). According to Bhatt (2000, 2004), direct *wh*-questions without inversion are formed by moving the *wh*-element in the matrix clause to sentence-initial position on the analogy of the fronting of constituents. The following constructions are compared: *Only fashionable girls these boys like* vs. *These boys like only fashionable girls* and *What he has eaten?* vs. *He has eaten what?*

Sheorey's (2006:191) error-gravity survey, which was conducted among 59 English college teachers in India.

Given this spectrum of attitudes, the question arose how common the patterns really were in contemporary IndE writing and speech. To provide a corpus-based answer, all *wh*-clauses in the Indian sections of the Primary Corpus were identified with the help of the WordSmith concordance program, using the key words *who*, *whom*, *whose*, *what*, *which*, *when*, *where*, *how* and *why* (cf. Quirk et al. 1985:817). In a next step, all irrelevant tokens were filtered out manually, i.e. relative clauses, direct *wh*-questions eliciting a subject (since they do not take auxiliary inversion) and other instances of finite and non-finite *wh*-interrogatives conforming to standard English usage conventions (e.g. *wh*-questions introduced by *how come*). The remaining tokens were categorized into direct and embedded interrogatives, which was sometimes not a straightforward task as punctuation and intonation occasionally failed to help disambiguate occurrences satisfactorily. The working hypothesis was that if the interrogative patterns in question had moved further down the path of integration over the past two decades, they might be more common in the press material dating from 2000 than in the texts from 1978. Also, if the patterns were more common in less heavily edited forms of speech and writing, the student essays and the broadcast material might return more examples of direct and embedded questions not complying with the codified norms of standard English than the press texts.

5.7.1 Direct *wh*-questions

Let us first look at direct *wh*-questions in the corpus data. Table 5.67 gives the proportion of inverted and non-inverted direct *wh*-questions that were identified in the Primary Corpus.⁷¹

Though the number of tokens not conforming to standard English is low, Table 5.67 shows a noteworthy text-type-specific distribution of the 'unusual' constructions in the data. While IndE quality newspapers and IndE national broadcast closely adhered to standard English around the year 2000 in quantitative terms, non-inverted direct questions were recurrent in all three text categories, with 4 of 39 direct *wh*-question in the press texts from 2000 (11.4% of all tokens in that category), 8 of 114 in the broadcast material (7.6%) and 3 of 5 in the student essays (60%) showing no inversion. Interestingly, the press corpora show the familiar stratification of 'unusual' forms along the time axis: it appears that usage around the turn of the millennium was slightly more variable than usage in 1978, for which no single occurrence of non-inverted direct questions was returned. Within the broadcast category, *wh*-questions without inversion occur exclusively in passages of spontaneous unscripted speech, which supports Nihalani et al.'s (1979:98; 2004:98) claim on rapid speech being the main locus of the pattern in question. The student essays, on the other hand, are once

71. Cf. Appendices 5.11–5.14 for the token frequencies of the individual *wh*-questions.

Table 5.67 Inversion in direct *wh*-questions.

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Essays
# direct questions	44 (100%)	39 (100%)	114 (100%)	5 (100%)
inversion	44 (100%)	35 (88.6%)	106 (92.4%)	2 (40%)
no inversion	–	4 (11.4%)	8 (7.6%)	3 (60%)

again in a class by themselves, with the proportion of noninverted interrogatives being much higher here than elsewhere in the corpus, although the overall token frequency of interrogatives is rather low.

A closer look at the non-inverted direct questions in context revealed more insights into the nature of the variation at stake. Not only are such direct questions rare in India's quality newspapers; when the construction appears, it is usually not employed consistently but co-occurs with direct questions conforming to the codified rules of standard English. See examples (159) and (160):

- (159) Before we begin, an advice. Buy this book. If not, borrow it. Better still, steal. Serious. How many shapes our country has? How much does it change every time we blink? How much of it is always the same? How is it breathing right now? Or, how is that which we haven't ever seen, ever heard, ever thought about, how is that faring and passing us by?
Don't die without telling me where you are going, India.

[IND2000 Press Rev 06]

- (160) The AIFF was keen on starting the league before the election, but that may not happen now as IPFA is talking of dealing with officials after the election and till then it wants the league postponed. Dasmunshi and Mour will have to answer quite a few uncomfortable questions before the election. Who's responsible for not holding the Federation Cup for two years and also the junior national? Also, why the Nehru Gold Cup was not held after the one in Cochin in 1996 and the Rajiv Gandhi Cup since 1998? More important, why some major tournaments, which had been permanent fixtures in the soccer calendar, suffered for want of dates and top players?

[IND2000 Press Rep 17]

Since there is no obvious sign in the examples above that would suggest that a pragmatic or stylistic effect was intended, the highlighted patterns appear to have slipped into each of the texts rather unselfconsciously. While in examples (159) and (160) the use of question marks qualifies the underlined constructions formally as direct questions, the distinction between direct and embedded questions is blurred in example (161). Here, only the first of the two interrogative clauses shows inversion,

whereas the second lacks inversion. Moreover, there is no question mark that would qualify the sentence unambiguously as a direct question:

- (161) RELIEF at the return of Roop Lal after 26 years in a Pakistani jail should quickly lead to two questions — one, how many Pakistanis are we holding in our jails without trial and why such things should happen in a country governed, as it claims to be, by the rule of law. Roop Lal appears to have said to an anchorman who repeated it on television that he did not regret what he did as it was for his country. [IND2000 Press Edit 10]

A similar blurring of the distinction between direct and embedded interrogatives is found in the IndE student essays, cf.

- (162) A boy dressed in Western style is called as gentlemen [sic] and of advanced society. Then why we should not adopt western culture if western culture takes Indian society towards advancement. [IND2000 Ess 07]
- (163) A man has never protected the virtue of a woman. Rama did not protect the virtue of Sita; nor the five Pandavas of Draupadi. Both these women have protected their own virtue by the sheer force of their purity. Then how or on what basis we can say that man is the stronger sex. [IND2000 Ess 08]
- (164) Most of the Indians of this era ... part with their ancient culture, which they feel is the greatest hurdle in their development. But how wrong they are? [IND2000 Ess 07]

In examples (162) and (163), question marks are missing, which might reflect the fact that the underlined questions were meant to have strong assertive force qualifying them as rhetorical questions proper. While the author of example (162) wants to drive home the point that Indians should, in fact “adopt western culture,” the writer in example (163) strongly disagrees with the claim that “man is the stronger sex.” In example (164), a rhetorical question has been used for a similar purpose.

Generally speaking, rhetorical questions might be a pragmatic environment making the suspension of subject-auxiliary inversion particularly likely to happen on account of the proximity of their communicative load to that of declarative sentences, which usually have no inversion. The majority of the noninverted direct *wh*-questions in the spoken data are, in fact, rhetorical questions rather than sincere questions eliciting information. See in this regard examples (165) and (166) from a broadcast discussion on the volatile political situation in Sri Lanka around the turn of the millennium. In both cases, speakers use rhetorical questions with strong assertive force to express their disagreement with what has previously been stated:

- (165) <\$B><#>the Tamil community has suffered<,> after that uh just one incident<,> <#>L-T-T-E could not<,> accept peace<,> as a<,> way forward<,>

<#>do you say<,> that seventeen people's lives<,> is more important than the overall interest of the community

<\$C><#>I am not I am not telling like that<,> <#>why the hostilities started<,> the reason<,> that is the reason<,> <#>and what was assured in the accord Indo-Sri Lankan accord<,> that was not honoured by the Sri Lankan government at first<,> <{><[>that is the basic reason</[>

<\$E><#><[>yeah but Mr Vaiko</[></[> why it failed<,> <#>why it failed the Rajiv Gandhi Jayawardene accord why it failed

[IND2000 Broad Disc 06]

- (166) <\$B><#>but is the killing of seventeen L-T-T-E leaders<,> is it<,> more important than the overall interest of the Tamil community

<\$C><#><{1><[1>no</[1>

<\$B><#><[1>you know</[1></[1> what happened after that

<\$C><#>how it <{><[>started</[>

<\$B><#><[>the Tamil</[></[> community has suffered<,> after that uh just one incident

[IND2000 Broad Disc 06]

In the next set of examples, again from spontaneous speech, non-inverted rhetorical questions occur in turn-end position, functioning pragmatically to keep the conversation running smoothly by allocating speaking rights in an orderly fashion (cf. Green 1996: 159–61), cf.

- (167) <\$G><#>N-L-F-T is the outlaw organisation<,> <#>we believe the Indian democracy<,> the democratic process<,> and we believe the Indian constitution<,> <#>then how and why we should be<,> with them

[IND2000 Broad News 15]

- (168) <\$G><#>while that Gurudas Dasgupta the Left Front candidate in Panskura<,>accused Rajiv Gandhi and Sonjaji<,> so many times that and blamed them that they took money from buffers<,> the same time Mamata Banerjee defend and protested<,> protested uh Gurudas Dasgupta that and and defend<,> Rajivji and Sonjaji<,> <#>so in this condition<,> least support Mamata Banerjee <#>why you support that Gurudas Dasgupta

[IND2000 Broad News 16]

- (169) <\$E><#>but uh what uh what the judges expressed<,> was<,> a deep sense of worry<,> <#>it was indicated that<,> what they have heard was<,> was<,> a matter for<,> extreme worry for not only the judges but as citizens<,> that<,> some kind of<,> independent area is getting created<,> where the two governments for the last more than ten years have not been able to penetrate<,> and not rule it<,> <#>this is a this is a very worrisome subject<,> worrisome

development<,> <#>and not only as judges as citizens also they are worried<,> <#>and what the state governments have been doing for last ten years<,> when this situation was happening [IND2000 Broad News 18]

Example (170) below is the only one from the speech sections of the Primary Corpus in which subject-auxiliary inversion has been suspended consistently in a comparatively long turn. It is from a broadcast discussion on the conflicts in Kashmir, in which a member of the studio audience (speaker H; male) addresses several questions to the discussion panel. The pauses preceding each of the interrogative constructions suggest that direct questions (rather than embedded questions) were intended:

- (170) <\$H><#>I just want to ask Mrs uh Sayeed and Ambika Soni<,> <#>what what uh the father of Mrs Sayeed has done<,> for this <,>sol<./> solution of the problem of Kashmir <#>she says that there there is a mindset and that's not been solved but he was a<,> Home Minister of uh of India<,> <#>what efforts he made for the solution of the this problem<,> <#>and I also just want to ask Mrs Ambika Soni<,> <#>she has uh asked that in the period of first one years government has not done this thing and that thing<,> <#>what Congress has done<,> for the last forty fifty years I can say<,> for the solution of this uh problem [IND2000 Broad Disc 02]

To conclude, while there are various plausible explanations for such constructions to occur in writing and speech, be it learner difficulties, pragmatic motivations, influence from languages like Hindi or an inherent ambiguity in English concerning the distinction between direct and indirect questions, no overarching explanation suggests itself for the entire spectrum of variation observable in the data. From a quantitative view, the suspension of subject-auxiliary inversion is an overall rare phenomenon in the corpus, which points to a rather low degree of acceptability of this structure in professional writing and speech. In rapid speech and less heavily edited writing, the proportion of such constructions appears to be higher, as the distribution in the Primary Corpus suggests.

5.7.2 Embedded *wh*-questions

Let us turn to embedded *wh*-interrogatives, for which the distribution across text categories and registers returned a similar patterning as for direct *wh*-questions without inversion. Embedded inversion, too, remains an overall rare syntactic phenomenon in the material under investigation, but appears to be slightly more common in spontaneous speech than in edited writing. Interestingly, embedded inversion is rare in the student essays, which may reflect a high degree of exonormatively oriented scholastic

Table 5.68 Inversion in embedded *wh*-questions.

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad	IndE IND2000 Essays
# embedded questions	22 (100%)	20 (100%)	75 (100%)	40 (100%)
no inversion	21 (95.5%)	18 (90%)	65 (86.7%)	39 (97.5%)
inversion	1 (4.5%)	2 (10%)	10 (13.3%)	1 (2.5%)

mediation in this particular instance. See Table 5.68 for the distribution of the two grammatical variants in the Primary Corpus.⁷²

Example (171), the only one of its kind found in the 1978 press texts, is similar to example (162) in that the distinction between direct and indirect questions has been blurred. Note the question mark at the end of the sentence, which could mean that the underlined construction was intended as a direct question, cf.

- (171) The all-important question is why is she able to attract such big crowds or mobilize so many votes even while the memories of the emergency are so fresh?
[IND1978 Press Rep 06]

However, even in standard English, as Green 1996 points out, the inversion of subject and auxiliary verb in such a context would be a pragmatic possibility to communicate that “the individual to whom the answer is implied or assumed to be relevant is in fact ignorant of the answer” (141), which is why example (171) does not represent any IndE-specific usage pattern proper.⁷³ The same is true for the next two examples of embedded inversion found in the press material from 2000. The fact that both were found in the context of quoted speech harks back to the affinity of embedded inversion with spoken language. While in example (172) the pattern occurs within a direct quotation, it appears in indirect speech in example (173), cf.

- (172) Malleswari, here to receive a gift cheque worth Rs 5 lakh from Satyam Infoway, said she would have appreciated the recognition more had it been given before going to Sydney. “People ask me what am I going to do with all the money that I have been showered with. I wonder why the same people did not come to me earlier to ask how I was preparing for the games with so little money,” the iron lady quipped.
[IND2000 Press Rep 08]

72. Cf. Appendices 5.15–5.18 for the token frequencies of the individual embedded *wh*-questions.

73. According to Green 1996, the utterance *she wants to know who did I appoint* would be more acceptable than *she already knows who did I appoint*, because the second utterance contradicts the implication that the answer is not known to the addressee.

- (173) He said Leander Paes and he were practising “pretty hard” as they had not played for a long time. “It takes time to settle down. We are sure to come good in front of our home crowd and are sure contenders for the medal.” When asked who does he rate as the toughest pair in the world doubles, Bhupathi said, “we haven’t beaten any top players this year. We will take one match at a time but it will not be easy for any one to go all out in front of the home crowd.” [IND2000 Press Rep 08]

Two similar examples of embedded inversion were also found in spontaneous speech:

- (174) <\$H><#>but it’s not a question of succeeding<,> <#>it’s the only thing is why should it be done<,> <#>why should that if everybody <#>see it is that no religion no minority or majority nobody should target the other side<,> <#>why should it be done [IND2000 Broad Disc 05]
- (175) <\$D><#>no but as far as the product is concerned it’s definitely exploiting the anxiety<,> <#>and even the kind of fairness that comes on<,> it’s terrible <#>and there’s a lot of chemicals you know problems <#>and I think we should they should be educating the consumer on what the hell are they putting in<,> <#>you can’t change melanin [IND2000 Broad Disc 09]

One such example was encountered in the student essays:

- (176) This helps us in knowing the various kinds of dresses prevalent in the world and what are their importances. [IND2000 Ess 08]

Note that standard English generally allows for subject-operator inversion in embedded questions “when the *wh*-element is the subject complement or an obligatory adverbial, particularly if the subject is lengthy” (cf. Quirk et al. 1985: 1052). According to my reading of the data, the following two instances of embedded inversion exemplify the application of the lengthy-subject rule and fall within what is common practice in standard English:

- (177) <\$A><#>yeah it’s extremely sad that this has happened<,> <#>Kannan thanks so much for joining us on News Hour and<,> elaborating on what you think<,> are the reasons behind the split and the effects of it <#>thanks so much [IND2000 Broad Int 06]
- (178) <\$C><#>well that’s a very big question <#>we’ll handle it because we have uh similar uh ammunition dumps all over the country<,> <#>this is one of those things that has happened which should not have happened<,> <#>we find out what are those uhm measures needed to see that such things are not done<,> because<,> measures are <.>always</.> always there <#>they have been there <#>but somehow we found that they were not adequate in this particular case [IND2000 Broad News 06]

As embedded inversion, thus, is an inherent possibility in standard English question formation, IndE should be considered innovative only insofar as it extends the pragmatic possibilities of this device to apply to a wider range of usage contexts, for example if the subject is not lengthy:

- (179) <\$D><#>uh it was uh happening since ninety-two<,> here and there<,> and it's a family problem<,> <#>we try to solve uh find out what is what is the cause for that<,> [IND2000 Broad Int 07]
- (180) <\$C><#>right <#>but in terms really of the civilian population there<,> I mean give us a sense of what was the mood<,> <#>is there any sort of anticipation that now the L-T-T-E is coming and we'll<,> sort of be liberated once again [IND2000 Broad Int 10]
- (181) <\$C><#>you know I I think he has we have a fundamental difference in our interpretation of what is a politician<,> <#>a politician is not just a minister<,> <#>the person in the zilla parishads and the mandals<,> at the gram panchayati levels<,> are politicians <#>they're opinion-makers<,> [IND2000 Broad Disc 07]

However, more contrastive research on less heavily edited forms of writing and less formal speech on the basis of larger corpora will be needed to define those contexts more precisely. In the usage range studied here, embedded auxiliary inversion, like the suspension of inversion in direct *wh*-questions, is too rare a phenomenon to allow further insights into IndE-specific usage patterns.

5.8 More variation at the sentence level?

Only and *itself* as focus elements

In standard English, *only* may be used as a so-called focusing subjunct for information-management purposes to “draw attention to a part of a sentence as wide as the predication or as narrow as a single constituent of an element such as a premodifying adjective in a noun phrase as subject, or an auxiliary within a verb phrase” (Quirk et al. 1985:604). It is a ‘restrictive’ subjunct in that it indicates that “the utterance concerned is true in respect to the focused part” (604). As a restrictive subjunct, *only* must be distinguished from additive subjuncts like *also* which are used to signal that the focused utterance is “additionally true in respect of the part focused” (604). See the following illustrative examples of *only* in IndE contexts (the focused elements are given in bold print):

- (182) <\$A><#>Israel's Defense Minister says soldiers would no longer **only respond to Palestinian fire** but take the initiative [IND2000 Broad News 18]

- (183) <\$B><#>the Prime Minister has clarified that India's help to Sri Lanka will only take the form<,> of **humanitarian assistance**<,> but the allies are asking for a clearer definition<,> of what humanitarian assistance would entail
[IND2000 Broad News 16]

Examples (182) and (183) from two IndE broadcast news reports conform to the rules outlined by Quirk et al. 1985 in that the use of *only* restricts the application of the utterances with respect to the focused elements, in each case implying a sharp contrast that is made explicit in the respective context. In example (182), the Israeli Defense Minister is reported to have announced a change in military tactics against Palestinians. What used to be an exclusively defensive strategy (hence the choice of *only*) has been turned into a strategy including offensive operations. In example (183), India's Prime Minister is reported to grant Sri Lanka humanitarian assistance, the use of *only* in this context implying the existence of unspecified alternative forms of assistance that are apparently deemed irrelevant. While the focus in example (182) is on the entire predication, only a part of the predication has been focused in example (183). Both examples have in common that *only* appears in medial sentence position, focusing elements to its right.⁷⁴ Alternatively, *only* may also appear in initial or end-position of a clause, and it can be used to focus elements to its right or left, cf.

- (184) Only high export growth, especially in the IT sector, and rapid expansion in services could lend support to the optimism expressed by the Finance Minister.
[IND2000 Press Edit 03]
- (185) <\$A><#>so far the garment industry enjoyed only small-scale industry status<,> which meant<,> that there was a limit of one crore on plant and machinery investment
[IND2000 Broad News 14]
- (186) Her focus was **the script and her actors** only, she said.
[IND2000 Press Rev 07]

IndE has been claimed to behave differently in regard to using *only* in two ways. For one, left focus with *only* has been said to be particularly frequent. As Nihalani et al. (1979, 2004) put it, "[t]he very frequent use of 'only' for the purpose of emphasizing the word or phrase (usually indicating time) which precedes it is characteristic of IVE [Indian variants of English]" (1979: 131, 2004: 131). Moreover, if *only* occurs with left focus in IndE, so the argument goes, it does not necessarily imply a sense of contrast as it would in standard English. Bhatt (2000, 2004), for example, claims that combinations of the type NP+*only* no longer allow a contrastive reading in what he calls vernacular IndE but merely a 'presentational' reading, which he sees as a variety-specific independent innovation that extends the syntactic focus-marking pos-

74. According to Quirk et al. (1985: 490–500), the terms initial, medial and end position denote the positions preceding any other clause element (initial), between subject and verb (medial) and following all obligatory elements in a clause (final).

sibilities in IndE to compensate for “the absence of a nuclear stress rule that marks sentential focus” (Bhatt 2004: 1024) in IndE speech. Interference from Indian substrate languages has been mentioned as an additional influential factor in creating variability in this particular instance.⁷⁵

To test whether usages of the types described by Nihalani et al. (1979, 2004) and Bhatt (2000, 2004) could be encountered in the Primary Corpus and whether any meaningful variety-specific distributions would emerge, the Primary Corpus was consulted first to identify all NP+*only* constructions in the written and spoken data. Particular attention was paid to instances of left-focus *only* that could be given a non-contrastive, presentational reading along the lines of Bhatt (2000, 2004). Identifying such uses, however, proved difficult, since it was usually impossible to reconstruct the discourse-pragmatic load of *only* in the given contexts and decide whether a NP+*only* constructions had been intended to be understood contrastively (a possible option in standard English) or presentationally (the claimed IndE-specific constellation). In fact, all but one example of left-focus *only* in the corpus data could be given both readings without any major distortion of meaning. Table 5.69 gives the number of uses of *only* with left focus as compared with the total number of subjunct uses of *only* in the press sections of the Primary Corpus.⁷⁶

Even though the absolute frequencies in the Primary Corpus are small, it is fair to say that, as far as sentence position is concerned, quantitative differences between IndE, BrE and AmE are marginal. This puts Nihalani et al.’s (1979: 131–2, 2004: 131) first claim about left-focus *only* being a common feature in IndE into perspective. In the domain of quality newspaper prose, uses of *only* with left focus are rare in IndE – as elsewhere among journalists in the English-speaking world. See the two examples from written BrE first:

- (187) Now, his pupil and successor, Marcel Dupre, himself in his seventies and a pioneer of organ records, has re-recorded it there in a coupling with Widor’s fifth and Gothic symphonies which shows how well his master scored for his beloved Cavaille-Coll instrument. (Westminster **mono only**.)

[UK1961 Press Rev 10]

- (188) The minister of agriculture is proposing freedom of contract (on **new tenancies only**).

[UK1991 Press Edit 08]

In both examples, left-focus *only* constructions appear in sentence-final position in parentheses, which might be a textual clue hinting at the rather restricted use to which the pattern is put in BrE newspapers today. Indian press English appears to employ

75. Cf. Lange (2007: 113–4) for references to the use of focus particles in Hindi and Marathi and their potential influence on IndE.

76. Included in the count were uses of *only* as subjunct. Excluded were adjectival uses, conjunctions and the fixed expression *not only*.

Table 5.69 Subjunct *only* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
only [subjunct] of those:	99	79	90	105	97	81
NP+only	2	5	1	1	–	–

the structure somewhat more freely. Consider examples (189) and (190), found in the 1978 press material:

- (189) The two figures need not be the same, for, there are tourists who visit a location **during the day only** and do not utilise existing hotel facilities.
[IND1978 Press Rep 15]
- (190) She told pressmen that it was not a matter of acceptance of the party programme **on paper only**.
[IND1978 Press Rep 17]

In terms of semantics, both instances of NP+*only* in the examples above can be given contrastive readings, which renders examples (189) and (190) fully in line with standard English usage conventions. Similarly, the five tokens of left-focus *only* found in the press material from 2000 follow the norms outlined in Quirk et al. 1985. Each occurs in the context of quoted speech, and three of the five instances are found in the same news report. While it is, again, impossible to reconstruct whether or not a contrastive or presentational focus was intended, each instance can be given a meaningful contrastive reading when taking the wider context into account:

- (191) ‘Package to cost **Rs. 68 cr. only**’
Mr. Paswan maintained that the cost of this package for the employees would be **Rs. 68 crores a year only** and not Rs. 1,200 crores as put out by various sources ... Talking to presspersons, Mr. Paswan said that today’s meeting had more or less approved the package cleared by the Telecom Commission. The employees’ federation had wanted rent-free telephones with 250 free calls, bonus equivalent to 73 days wages and a debate and discussion on the corporatisation process. In return, the Telecom Commission proposed rent-free telephones with 250 free calls, **bonus for 70 days wages only** and co-operation in the corporatisation process. This package has more or less been approved today, but for the reduction in the number of free calls, the Minister said.
[IND2000 Press Rep 19]
- (192) Sets were minimal, almost inconspicuous. A chair for Tipu and half a dozen stools, where one expected colour and the flamboyance with which the Sultan dreamed. The elaborately worked tiger motifs were missing. As Yamuna saw

it, “Elaborate sets would have fought with everything else and would have become very fussy.” Her focus was **the script and her actors** only, she said.

[IND2000 Press Rev 07]

- (193) She has no plans to include any other game in her academy, saying, “there won’t be any other game in my academy. Being a weightlifter myself, I’ll impart coaching to **the young weightlifters** only.” [IND2000 Press Rep 08]

It is only in the spoken data and in the student essays that left-focus *only* occurs in environments approximating more obviously the innovative usages described by Nihalani et al. (1979, 2004) and Bhatt (2000, 2004). See the distribution of NP+*only* constructions in those text categories, cf. Table 5.70.

According to Table 5.70, the category of Indian student writing comes closest to what has been claimed to be characteristic of IndE. Here, 7 of 45 subjunct uses of *only* display left focus. In the IndE broadcast sections, three out of 68 subjunct uses of *only* show left focus, with the three example of NP+*only* found in passages of spontaneous speech. On the other hand, no single occurrence of left-focus *only* was found in the BrE speech data. Consider example (194) from conversational IndE:

- (194) <\$D><#>we we are in all the ways we are helping the poor<,> and the farmers in the form of<,> always we are helping<,> by giving<,> input subsidies to the farmers<,> <#>and then we are giving fodder to the to the to the cattle that they’re having<,> <#>and then we are giving<,> most of them they’re getting getting uh the oil extensions<,> <#>and we are giving uh uh employment generation the area where there is no employment<,> <#>we are trying all level <#>Andhra Pradesh is the first state<,> who want to help the people<,> who are in trouble <#>we are giving most of our<,> attention to **these farmers** only <,> this year

[IND2000 Broad Int 07]

In example (194), speaker D (a state minister; female) is explaining her ministry’s agenda to support farmers in Andhra Pradesh. In the last utterance of the turn, a NP+*only* focusing construction has been used. The pause preceding the utterance-final adverbial *this year* provides an audible clue that left focus has been intended. By the norms of standard English, the use of *only* to focus an object to its left in clause mid-position would be unusual (cf. Quirk et al. 1985:606–7). According to Bhatt (2000, 2004), the focus of *only* might have been intended to be presentational in that attention was meant to be drawn to the noun phrase *these farmers* and, hence, to the main topic of the conversation without implying any sense of contrast. A contrastive reading along the lines of *we are giving most of our attention to these farmers only* (i.e. *not to any other group*), possible in standard English, would be highly unlikely in vernacular IndE.

The next two examples were found in two IndE student essays and are similar to the examples provided by Bhatt (2004:1025) to illustrate the possibility of

Table 5.70 Subjunct *only* [Primary Corpus; broadcast and essays].

	IndE IND2000 Broad	IndE IND2000 Ess	BrE UK1991 Broad
only [subjunct] of those:	68	45	61
NP+only	3	7	–

presentationally-focused noun phrases in subject position. In each case, the context suggests that *only* focuses the subject left of it and not the verb to its right:

- (195) Indian Culture had also great influence in an innocent girl child life. She had to work the whole day in her house and **the male child** only was allowed to go to study in a school. [IND2000 Ess 06]
- (196) When Swami Vivekanand proclaimed the supremacy of Indian culture and tradition and was able to move the people a lot. [sic] Several scholars like Max Muller, Charles Wilkens, William Jones etc. have also added compliments to the glory of India. **These few examples** only reveal that how much our traditions have helped us to establish a dignified identity ... [IND2000 Ess 07]

Again, if we follow Bhatt (2000, 2004), such uses of *only* in vernacular IndE would most likely not imply a contrast but be interpreted as functioning presentationally to highlight the subject. However, while in example (195) a standard English contrastive reading is also possible without distorting the message of the statement (*the male child only was allowed to go to study in a school (i.e. not the girl)*), only a presentational NP+*only* interpretation is plausible for example (196), as a contrastive reading would suggest nonsensically that no other Western scholars but those named had ever spoken favorably of “Indian culture and tradition”. Example (196), hence, comes closest to what Bhatt (2000, 2004) describes as an innovation in vernacular IndE. It is the only such example in the Primary Corpus.

That the overall rarity of presentational *only* in the Primary Corpus is not coincidental – and left-focus *only* with ‘presentational’ rather than ‘contrastive’ meaning indeed not as common in contemporary IndE as might be expected when looking at previous descriptions of the variety – is further borne out by the results produced by Lange 2007 on the basis of ICE-India. In the written section of ICE-India, Lange (2007: 107) identified 665 instances of standard English *only* with contrastive focus alongside a single unambiguous example of *only* employed as a presentational focus-marking device. Incidentally, this example was also found in a student essay. In the speech sections of ICE-India, on the other hand, innovative uses of *only* were more recurrent. Here, 114 tokens of presentational *only* were found alongside 1,250 tokens of contrastive *only*. Those research findings corroborate what the smaller Primary

Corpus hints at: innovative *only* has a foothold in spoken and less heavily planned IndE but appears to be unacceptable to many users in more carefully planned forms of language use. According to Lange 2007, this could be the case because users of IndE might feel that “*only* is stigmatized precisely because of its high frequency in spoken IE [IndE]” (116). That presentational *only* is, indeed, among the more well-known features of IndE is shown by the fact that the item is occasionally used symbolically to represent the ‘Indianness’ of IndE, as in the title of a more recent publication on IndE that seeks to “capture the new grammar and vocabulary, the essence, the unintended humour, and the way Indian-English has sallied forth in the country” (John 2007: 16): *Entry from backside only*.⁷⁷

Further interesting uses of left-focus *only* are displayed in examples (197) and (198). Here, the choice of *only* can be interpreted to reinforce the highlighting function of the chosen cleft constructions. Contrastive readings are, again, possible:

(197) It is **the period of childhood** only when we are too naughty.
[IND2000 Ess 01]

(198) It is **during childhood** only that the foundation of one’s character is laid.
[IND2000 Ess 02]

The last set of examples from the Primary Corpus illustrates that *only* may not only focus noun phrases to its left, but also the temporal adverb *then*:

(199) There has to be cooperation from both sides **then** only will the team be able to get make its countrymen feel proud of it. [IND2000 Ess 04]

(200) <\$C><#>this outcry is not enough<,> <#>you will have to come forward<,>
<#>you will have to understand<,> the real nuances of the bill<,> <#>and
then only<,> you you should fight for it [IND2000 Broad Disc 08]

77. Also see in this context the following humorous anecdote found in the travelogue *Shiva Moon* by the German writer Helge Timmerberg (2006: 10), which speaks of the salience of left-focus presentational *only* in IndE in cross-cultural communication:

Die zweite Möglichkeit, in New Delhi anzukommen, ist das “Imperial”. Das schönste Hotel der Welt, eine Mischung aus Mogul und Kolonial, Maharadscha und Offizier, Turban und Krone, Schönheit und Macht. Mein letztes Gespräch am Mahagonitresen der Rezeption verlief so:

“How much is the room?”

“Single or double, Sir?”

“Single.”

“Two hundred and eighty US only, Sir.”

Das “only” war nicht ironisch gemeint, das sagen sie immer, nach jedem Preis. Alles in Indien kostet irgendwas “only”, außerdem sind zweihundertachtzig Dollar für ein Hotel wie das “Imperial” im internationalen Vergleich tatsächlich “only”, aber für mich war das die Hälfte meiner Miete zu Haus.

- (201) <\$C><#>we have said it<,> that it is not the appropriate time<,> to see that the government is losing its authority again within the gap of<,> only five to six months<,> but certainly it is an<,> alarm and it is a notice to the government<,> that<,> any decisions to be taken in future<,> the decisions<,> which involve the people's interest<,> should be discussed in the allies parties' meetings in details<,> and then only the decisions could be announced

[IND2000 Broad News 05]

A look at the collocation *then only will* + N + V (vis-à-vis *only then will* + N + V) and the expression *then and then only* (vis-à-vis *then and only then*) on the World Wide Web (March 30, 2004) allows further glimpses of the availability of *only* in IndE functioning to focus the adverb *then* to its left (and right). Notably rare in written BrE and AmE under the '.uk' and '.us' top-level country domains, the combination *then* plus *only* is particularly frequent under the '.in' domain, but is also recurrent in SgE and partly also in SAfrE (at least as far as the fixed expression *then and then only* is concerned), cf. Table 5.71.

But note that register-specific constraints seem to be in operation limiting the availability of left-focus *only* after *then* in IndE. The left-focus constellation is clearly less common in India's online quality newspapers (Google search dated March 30, 2004); cf. Table 5.72.

While it would be interesting to know to what degree users of contemporary IndE find those patterns acceptable today, answering this question will have to be referred to future attitude surveys. In more carefully planned language use, it can be concluded, left-focus *only* is still quite rare around the turn of the millennium.

Let me add in this context that innovative uses of the reflexive pronoun *itself* as a presentational focus element were just as scarce as those of presentational *only* in the usage range under investigation, even though it has been claimed that "[the] use of the reflexive pronoun 'itself' for the purpose of emphasizing the word or phrase (usually indicating time, but sometimes place) which precedes it is characteristic of the IVE [Indian] variety of English" (Nihalani et al. 1979: 105, 2004: 105). See the distribution of the form in the Primary Corpus (cf. Table 5.73).

Examples (202) and (203) are the only two examples of the claimed Indianism in the Primary Corpus. Both were encountered in spontaneous speech:

- (202) <\$C><#>no let's let's just go let's look at some of the other areas the other more contentious areas which are the laws itself<,>

[IND2000 Broad Disc 10]

- (203) <\$B><#>they claimed<,> that in the last three months<,> more than six shivsainiks have been killed in the state<,> three of them in the last three days itself

[IND2000 Broad News 05]

Table 5.71 Subjunct *only* in selected collocations [Internet: country domains].

	.in	.sg	.uk	.za	.au	.us
then only will (+ N +V)	104	126	265	245	147	48
only then will (+ N +V)	317	384	103000	24900	29700	11700
	1 : 3	1 : 3	1 : 389	1 : 102	1 : 202	1 : 244
then and then only	37	3	94	30	100	52
then and only then	63	35	28900	215	14200	11900
	1 : 2	1 : 12	1 : 307	1 : 7	1 : 142	1 : 229

Table 5.72 Subjunct *only* in selected collocations [Internet: newspapers].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU
then only will (+ N +V)	1	7	2	–	10	1	–
only then will (+ N +V)	8	380	95	9	122	87	197
then and then only	–	3	–	–	2	–	–
then and only then	2	12	2	2	13	3	27

Table 5.73 *Itself* as focus element.

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad News	IndE IND2000 Broad I/D	IndE IND2000 Ess
itself [all uses]	20	23	8	8	5
of those: focus element	–	–	1	1	–

Lange 2007, too, found innovative uses of *itself* to be more frequent in the speech sections of ICE-India than in writing. In spoken IndE, there were 51 instances of innovative *itself* vis-à-vis 119 instances of uses of *itself* as a reflexive and intensifying pronoun. In written IndE, on the other hand, only 16 innovative uses alongside 99 standard English uses were identified. Note that differences between speech and writing are not as pronounced for innovative *itself* as they are for innovative *only*, which leads Lange 2007 to the conclusion that “[t]he pattern NP + *itself* in locative, temporal and other expressions is well represented in both spoken and written IE [IndE]” (106). Since innovative uses of *itself* can also be encountered recurrently in the Kolhapur Corpus from 1978, we seem to be dealing with a usage that “is also stable over time” (Lange 2007: 106).⁷⁸ Its overall token frequency in contemporary IndE is low,

⁷⁸ Lange (2007: 100) identified 20 innovative and 312 standard uses of *itself* in the Kolhapur Corpus.

though – lower, in fact, than that of presentational *only*, which, on direct comparison with presentational *itself*, must be considered grammatically less complex and, hence, “more flexible in its use, less dependent on specific contexts, and therefore probably easier to appropriate for new functions” (Lange 2007: 114).

5.9 Chapter summary

Contemporary IndE syntax, on the whole, differs only moderately from standard English in quantitative terms. The overwhelming majority of tokens of definite and indefinite articles, nouns, tense forms, interrogative constructions and focus elements in the Primary Corpus have been shown to participate in the behavior outlined in Quirk et al. 1985, which, after all, makes IndE a variety of *English*. That said, numerous stable usage patterns were uncovered in the qualitative analysis of the corpora that are variety-specific in terms of exploiting and extending the semantic-pragmatic possibilities of English. Examples of such IndE innovative modifications include the frequent use of the definite article with proper names and with nouns denoting institutions of human life (e.g., *police*, *society*), the use of the past perfect with preterite meaning in text-, turn- and narrative-initial positions to situate discourse firmly in the past, and the use of *only* as presentational focusing device. Also, contemporary IndE is currently seeing the phrases *majority of* and, less forcefully, *lot of* developing into phrasal quantifiers. IndE has also added a variety-specific sense to *tuitions*, ‘private lessons’, and prefers the expression *last but not the least* to the standard English variant *last but not least* across text categories and registers.

Several usage differences between contemporary IndE and the other varieties considered in this study appear to reflect divergent diachronic developments among Englishes worldwide. Examples of features that have dropped out of favor in BrE and AmE but continue life in the Indian subcontinent to variant degrees of integration are the use of *legislation*, *equipment*, *agitation*, and *machinery* as count nouns, the use of the plural *aircrafts* and the use of *police* with singular verbal and pronominal concord. Contemporary IndE also shows a liking for singular verbal and plural pronominal concord with collective nouns, whereas BrE employs plural and variable verbal concord more frequently. In this regard, IndE approximates what has been said to be typical of AmE usage – a likeness that may have emerged due to divergent diachronic developments in the British Isles. On the other hand, IndE has been slightly more reluctant than BrE to reintroduce the mandative subjunctive into writing under the influence from AmE, although the mandative subjunctive has made inroads into IndE within the domain of quality newspaper prose and is nowadays particularly common in Chennai’s *The Hindu*. Other examples of IndE journalism opening up to ongoing change in international newspaper English include the suspension of backshift in reported speech and the use of direct quotes in news reportage, which are both more common phenomena in the IndE press texts from 2000 than in the material from 1978. In those instances, IndE newspapers are slightly more advanced than BrE news-

papers, which might be indicative of a relatively strong orientation toward AmE usage norms within the domain of IndE journalism.

Furthermore, the case studies have revealed that IndE varies considerably in its proximity to the codified norms of standard English when moving across text types and registers. Several phenomena appear to be restricted to less heavily edited forms of speech and writing, for example variation along the count and number divides, uses of *since* elements with the present tense, uses of the present perfect with adverbials of the past, variation in the formation of direct *wh*-questions and uses of *only* as presentational focus adverb. On the whole, India's quality newspapers adhere most closely to the international usage standard, whereas the student essays are furthest removed. However, the comparison of the IndE press material from 1978 and 2000 has also brought to light that India's quality newspapers were opening up to forms not conforming to standard English in the two decades between 1978 and 2000, albeit moderately. IndE press reportage was, for example, slightly more variable around the turn of the millennium than twenty-two years earlier along the count-number divides, in the use of the past perfect with preterite meaning, and in question formation.

Several other syntactic features which have been pointed out as characteristic features of IndE were rare in the corpora, for example plural uses of *information*, *employment* or *furniture*. Their near absence in the present data suggests a rather low degree of integration of those features in present-day IndE, although they might well be more common in text types and registers not represented in the corpus material used for the present study.

Conclusion

This study started out by criticizing the feature-list approach to IndE, which continues to date as the predominant way of describing and characterizing contemporary IndE. The case studies presented in the previous chapters show that feature-list descriptions are much too general and static to account for what IndE essentially is – not a carbon-copy of the Queen’s English or “‘normal English’ with more or less learner-specific deviations” (Schmied 1994: 217), but a viable brand of English that is variable, innovative at times, and always capable of changing according to the needs of its users. Many claimed Indianisms from the feature lists were too rare to be attested in the IndE databases. Those that did occur were often not frequent (although there were exceptions) and did not always qualify as independent innovations. Saying this means putting previous findings into perspective: IndE is a variety of *English* that has more in common with other varieties than a look at any feature list would suggest. However, the case studies also show that it is not enough to focus on this quantitative aspect alone. While largely in tune with the codified rules and norms of standard English, IndE users do take the freedom to modify *their* English to extend the range of their language and enlarge its possibilities whenever a communicative need to do so should arise. The combination of data-based empirical work and qualitative case studies in the previous chapters has, in my view, successfully captured the variant degrees to which IndE users operate on local and global usage norms in such situations and contexts.

The present study has produced plenty of evidence to show that IndE is capable of innovation. As for the lexicon, IndE users borrow words and expressions from Indian languages (e.g. *lakh*, *crore*, *bandh*, *gherao*, *hartal*, *rasta roko*), create new words (e.g. *upgradation*, *speed money*, *timepass*, *senti*, *funda*, *hydel*, *incharge*) and use existing words with new variety-specific meanings and functions (e.g. *ticket*, *meet*, *to take out* ‘to lead’, *hi-fi* ‘posh’, mediopassive uses of *release*). IndE users have also carved out their very own usage preferences for individual words and compounds that are less common elsewhere in the English-speaking world (e.g. *parliament session*, *sympathy wave*, *telephonic conversation*, *thrice*, *upliftment*) or have been dropped there altogether (e.g. *lecturership*, *to chalk out*, *to chart out*). But the case studies have achieved more than simply to collect words. By drawing on closed micro-corpora like the Primary Corpus, the Kolhapur Corpus and ICE-India and on several larger Internet corpora, it has been possible to locate some of the inroads of lexical innovation into IndE across text categories, registers and regions and, in doing so, describe the behavior of individual lexical items in greater detail. The use of the Primary Corpus, for example, has revealed that loanwords in IndE quality newspapers do not stratify evenly but tend

to cluster in press reviews, where they are put to use in formal and informal contexts, for instance in texts representing 'cine journalism'. In this regard, the comparison of the parallel press corpora from 1978 and 2000 has brought to light that India's quality newspapers were starting to use informal loanwords such as *filmi*, *masala*, *baap*, *chalu* and *zindabad* around the year 2000, whereas no such informal item could be traced in the 1978 data. This finding suggests that informal loanwords might slowly be gaining in acceptance among Indian journalists, a small but sure sign for the stabilization of the local IndE usage norms at the turn of the millennium. The online corpora, on the other hand, have been helpful in describing the status of rare lexical items in the Primary Corpus in much more detail than previous descriptions have achieved. For instance, the Hindi loanword *Vidhan Sabha* 'state assembly', representing the use of Hindi-based political terminology in contemporary IndE, has not gained pan-Indian currency yet, although its spread has been encouraged by its use in government publications. Similarly, the loan translation *timepass* 'pastime' appears to be a favored variant in Hindi-speaking areas but is rarer elsewhere in the subcontinent, where the standard English variant *pastime* is still favored. The opposite is true for the nominal compound *petrol bunk* 'petrol pump', which appears to be a Southern Indian favorite that is rarely found elsewhere in newspapers across the subcontinent. Although several Hindi-based loan items appear to be common across India today (e.g. *lathicharge*, *bandh*, *dharna*, *rasta roko*), there is evidence that new coinages based on English words (e.g. *eve-teasing*, *upgradation*, *to delink*) spread somewhat more easily across India and South Asia, hence functioning as supraregional 'lexical links'. While it would, of course, be futile to base any more general claims about the nature of the IndE lexicon on those few examples alone, the evidence documented in this study suggests that the IndE lexicon is regionally quite diversified, which calls for further in-depth lexicographical research across regions and linguistic areas to throw more light on this intriguing issue.

Tracing structural innovation in lexicosyntax and morphosyntax has been a more challenging task. While there can be no doubt about the innovative potential of contemporary IndE in syntax, the evidence presented in the previous chapters suggests that the impact of independent innovation in the standard usage range should be considered rather moderate. IndE has created new particle verbs (e.g. *to term as* 'to term') and expanded the functional load of adverbs (e.g. *only*), verbs (e.g. *release*, *wish*, *avail*, *enjoy*), pronouns (e.g. *itself*), nouns (e.g. *tuition*) and tense forms (e.g. the past perfect). But again: those innovative forms are in the minority when looking at the corpus data from a bird's-eye view. Of all verb-particle combinations with *up*, *down*, *off*, *away* and *out* in the Primary Corpus, for example, only three percent were found to differ in use from the codified norms of standard English. The majority of those 'different' uses appeared in student writing and, upon a closer look, qualified as unstable learner errors. Only few combinations filtered out from the corpus material emerged as stable IndE innovations, for instance *contest from* (*a constituency*) 'to run for office (in a particular constituency)' or *take out* (*a protest demonstration*) 'to lead (a protest demonstration)'. Also recurrent (although still minority variants) in

present-day IndE are the combinations *request for*, *stress on*, *emphasize on*, *discuss about*, *investigate into* and *urge for*, which have probably come into being on the analogy of their related noun-particle combinations available in standard English. Methodologically, the online corpora were a particularly useful tool to establish that those combinations are minor variants in IndE within the domain of published writing, where they function alongside the codified standard variants and are usually employed in unselfconscious ways. Moreover, the online corpora have revealed that those particle verbs are nowadays not used exclusively in IndE but throughout the English-speaking world. But even if differences in world Englishes are mainly quantitative in nature, IndE, on the whole, seems to make more frequent use of those forms today than other varieties (for example South African English) do. This, in turn, suggests a relatively high degree of integration of those items in IndE despite their rarity and despite the prescriptive pressure exerted on IndE users by exonormatively oriented forces such as usage handbooks or style guides.

Several other lexicosyntactic and morphosyntactic features that have been claimed to be IndE-typical features are nearly absent from the data, for example the patterns '*want* plus *that*-clause', '*contemplate* plus *to*-infinitive', '*look forward* plus *to*-infinitive' and plural uses of the noncount nouns *information*, *employment* and *furniture*. While this is not to say that those phenomena do not occur at all in IndE today, their low frequency in the standard usage range suggests that their acceptability among users of IndE is still rather limited. Similarly, none of the case studies designed to measure the impact of L1 interference on IndE supports the view that interference from Indian languages should be considered a major driving force for syntactic variation. Contemporary IndE does certainly not generally 'underuse' definite or indefinite articles nor 'overuse' the past perfect. The investigation of the corpus material has drawn a much more nuanced and diversified picture by illustrating that the impact of L1 interference affects IndE differently across different text categories and registers. Uses of the past perfect that are possibly motivated by substrate influence are, for example, more common in the Primary Corpus in less heavily edited forms of speech and writing, for instance in the broadcast interviews and discussions and, in particular, in the student essays. India's quality newspapers, on the other hand, adhere largely to the norms of standard English. Hence, while it is well possible to delineate local contexts in which the past perfect is used in IndE-specific ways, IndE at the same time remains embedded in an international context, which itself is currently undergoing processes of variation and change affecting the past perfect (for example in the context of reported speech). Therefore, even if interference from India's background languages is an ever-present possibility to create variation, the IndE past perfect 'at large' cannot be characterized monolithically in reference to L1 interference alone. Likewise, interference from Indian languages may be a possible influential factor reinforcing the omission of surface articles, the suspension of inversion in direct *wh*-questions or uses of the particle *as* as quotative marker after verbs of naming and quoting, for example *term* and *rename*. However, it has been difficult to gauge the exact impact of substrate influence in all

those cases as well because English itself is inherently variable in those areas of the language system.

One of the major findings presented in the previous chapters is that present-day IndE carries numerous traces of its origin in the brand of educated English evolving in India in the second half of the nineteenth century. Those findings go against Shastri et al.'s (1986) all too general claim that "the Indianness of Indian English is a post-Independence phenomenon" (3) and support the view that the second half of the nineteenth century should be considered as marking "the beginnings of the evolution of 'educated' Indian English" (Mukherjee 2007: 167). Lexical examples are *lecturership* 'lectureship', *detenu* 'detainee', *to demit* 'to console', *to chart out* 'to chart', *to chalk out* 'to chalk', *untoward* 'unexpected' and *thrice* 'three times'. None of those items is listed in the major dictionaries of contemporary English any longer, although all of them used to be available in English, as a look at the OED and its quotations base has shown. In IndE, those words are nowadays frequent across registers and text categories, as are the verbs *to bunk* (as in *to bunk school*) and *to avail* 'to take advantage of' (as in *to avail an opportunity*). But *avail* is a nice example also to illustrate that IndE does not simply 'retain' words that have been dropped in other parts of the English-speaking world; it is capable of modifying them according to the communicative needs of its users. *Avail* may nowadays be used in IndE in informal contexts in the sense 'to use, to take advantage of'. Divergent diachronic developments also account for contemporary usage differences between IndE and the major varieties in the use of the function words *amid/amidst* and *until/till*. IndE, unlike BrE and AmE, generally prefers *amidst* (to *amid*) and *till* (to *until*) today. In the first case, international English is seeing the slow but steady retirement of the variant *amidst*, a development that South Asian Englishes in general appear to be following only slowly. In the second case, the internationally established stylistic difference between *until* (the formal variant) and *till* (the informal variant) most probably goes back to a development in BrE and AmE in the second half of the nineteenth century, which has apparently not been followed as forcefully in IndE as elsewhere in the English-speaking world. Another such example is the combination *with a view to*, which is frequently used with the bare infinitive in IndE today (as occasionally happens also in AmE) whereas BrE has turned this phrase quite rigorously into a multi-item preposition. Likewise, the recurrent count uses of *legislation*, *equipment*, *agitation*, and *machinery* and the relatively frequent use of singular verbal and pronominal concord with *police* are best explained as retentions of features that used to be more common in BrE and AmE in the nineteenth and early twentieth centuries, although more diachronically sensitive research is called for to substantiate this claim further. The fact that those phenomena have frequently been pointed out as learner errors in IndE teaching materials and style guides illustrates the continuing exonormative pressure exerted on IndE today. However, treating such forms as errors is to deny IndE its independent development. While IndE remains embedded in an international context, it is not a copy of any other English and should not be treated as such.

As regards the postcolonial force of AmE on contemporary IndE, only moderate influence from AmE has been felt in the usage range under investigation. Examples of external influence from AmE include the occasional use of AmE-based spellings in IndE writing, especially in texts representing academic and scientific writing, and mild repercussions of informal AmE conversational language in IndE speech. The corpora also returned several examples of IndE journalism opening up to changes in international newspaper English that are led by AmE, for instance the use of contracted forms, the suspension of backshift in reported speech and the growing popularity of direct quotes in news reportage. In the latter two instances, IndE newspapers are slightly more advanced than BrE newspapers, which suggests a relatively strong orientation toward AmE usage norms within the category of IndE journalism as opposed to other domains of written use. On the other hand, IndE has been more reluctant than other varieties in reintroducing the mandative subjunctive under the external influence from AmE, although the mandative subjunctive has made noticeable inroads into IndE within the domain of quality newspaper prose and is nowadays particularly common in Chennai's *The Hindu*. Interestingly, IndE was found to approximate AmE most closely across text categories and registers in favoring singular over plural verbal and pronominal concord with collective nouns, whereas BrE users often employ plural and variable verbal concord. But rather than representing the effects of the contemporary global pressure of AmE, this linguistic tidbit must be given a historical reading: it appears that IndE, like AmE, was independent enough around the turn of the nineteenth and twentieth centuries not to follow a usage preference that was emerging in BrE around that time.

How different is IndE in sum? From a structural point of view, there is little evidence in the corpus data that would suggest a major structural divergence of IndE away from standard English. This puts a question-mark on the claim that IndE should be treated as a fully autonomous independent system, as S. K. Verma (1978, 1982) has suggested, and is more in line with Leitner's (1992b) and Mukherjee's (2007) proposals to treat IndE as a semi-autonomous variety of English shaped simultaneously by local as well as global forces. What the corpus material suggests for the standard usage range is that independent innovation happens – but it happens moderately at all levels, and when it can be pinpointed, as is the case for particle verbs such as *to term as* and *to rename as*, for the past perfect and for *only* and *itself*, structural nativization seems to be happening along the lines of what is a possible option for English in general. The functions of definite and indefinite articles or the functions of the past perfect are heavily context-dependent to begin with, so seeing that IndE exploits those dimensions of variability to its own ends should not come as a big surprise. There is plenty of evidence in the corpus data to suggest that IndE users do exactly that – they draw freely on the possibilities of English and make their own choices and this in a context-sensitive fashion. But they also take their decisions with a strong view to the codified usage standard, which is bound to put a limit to the impact of structural nativization in the

future. Methodologically, this has consequences for describing IndE as a variety in its own right. IndE cannot be understood by enumerating decontextualized abstract features. IndE can only be characterized adequately in reference to context-specific variational profiles which, as the case studies in the previous chapters have revealed, can be assessed through investigations based on carefully designed micro-corpora such as the Primary Corpus or ICE-India alongside significantly larger databases, for example online corpora. It is my hope that the present study will come to be seen as a valuable contribution toward describing contemporary IndE in exactly that way.

References

- Abbi, Anvita, Gupta, R. S. & Kidwai, Ayesha (eds). 2001. *Linguistic Structure and Language Dynamics in South Asia. Papers from the Proceedings of SALA XVIII Roundtable*. Delhi: Motilal.
- Advani, Lal Krishna. 2008. *My Country, My Life*. New Delhi: Rupa.
- Aggarwal, Kailash L. 1988. English and India's three-language formula: An empirical perspective. *World Englishes* 7(3): 289–98.
- Agnihotri, Rama Kant. 1992. Acquisition of articles in learning English as a second language: A cross-cultural study. In Koul (ed.), 179–92.
- Agnihotri, Rama Kant, Khanna, Amrit Lal & Mukherjee, Aditi. 1984. The use of articles in Indian English: Errors and pedagogical implications. *IRAL* 22(2): 115–29.
- Agnihotri, Rama Kant, Khanna, Amrit Lal & Mukherjee, Aditi. 1988. *Tense in Indian English. A Sociolinguistic Perspective*. New Delhi: Bahri.
- Agnihotri, Rama Kant & Khanna, Amrit Lal. 1994. *Second-Language Acquisition: Socio-Cultural and Linguistic Aspects of English in India*. New Delhi: Sage.
- Agnihotri, Rama Kant & Khanna, Amrit Lal. 1997. *Problematizing English in India*. New Delhi: Sage.
- Algeo, John. 1988. British and American grammatical differences. *International Journal of Lexicography* 1(1): 1–31.
- Algeo, John. 1992. British and American mandative constructions. In Blank (ed.), 599–617.
- Algeo, John. 2006. *British or American English? A Handbook of Word and Grammar Patterns*. Cambridge: CUP.
- Annamalai, E. 1994. English in India: Unplanned development. In Kandiah & Kwan-Terry (eds), 261–77.
- Annamalai, E. 2001. *Managing Multilingualism in India. Political and Linguistic Manifestations*. New Delhi: Sage.
- Annamalai, E. 2008. Contexts of multilingualism. In Kachru et al. (eds), 223–34.
- Aramavudan, Srinivas. 2007. *Guru English. South Asian Religion in a Cosmopolitan Language*. New Delhi: Penguin.
- Bailey, Richard W. & Görlach, Manfred (eds). 1982. *English as a World Language*. Ann Arbor MI: University of Michigan Press.
- Bakshi, Raj N. 1991. Indian English. *English Today* 7(3): 43–6.
- Bansal, R. K. 1983. *Studies in Phonetics and Spoken Language. National Lectures in English 1980–1981*. Hyderabad: CIEFL.
- Bartsch, Renate. 1985. The influence of language standardization on linguistic norms. *Studia Linguistica* 39: 23–50.
- Bauer, Laurie. 1994. *Watching English Change. An Introduction to the Study of Linguistic Change in Standard Englishes in the Twentieth Century*. Harlow: Longman.
- Bauer, Laurie. 2002. *An Introduction to International Varieties of English*. Edinburgh: EUP.

- Baumgardner, Robert J. 1992. 'To *shariat* or not to *shariat*?': Bilingual functional shifts in Pakistani English. *World Englishes* 11(2–3): 129–40.
- Baumgardner, Robert J. (ed.). 1993a. *The English Language in Pakistan*. Karachi: OUP.
- Baumgardner, Robert J. 1993b. The indigenization of English in Pakistan. In Baumgardner (ed.), 41–54.
- Baumgardner, Robert J. 1993c. The Urduization of English in Pakistan. In Baumgardner (ed.), 83–203.
- Baumgardner, Robert J. (ed.). 1996. *South Asian English. Structure, Use and Users*. Delhi: OUP.
- Bautista, Maria Lourdes S., Llamzon, Teodoro A. & Sibayan, Bonifacio P. (eds). 2000. *Parangal nang Brother Andrew. Festschrift for Andrew Gonzalez on His Sixtieth Birthday*. Manila: Linguistic Society of the Philippines.
- Beliya, G. D. (ed.). 2001. *Press in India 2001. 45th Annual Report of The Registrar of Newspapers for India*. New Delhi: Ministry of Information and Broadcasting, Government of India.
- Benson, Morton, Benson, Evelyn & Ilson, Robert. 1986. *Lexicographic Description of English*. Amsterdam: John Benjamins.
- Bhaskararao, Peri. 2002. English in contemporary India. *Asian/Pacific Book Development* 33(2): 5–7.
- Bhatia, Tej K. 2000. *Advertising in Rural India*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Bhatt, Rakesh M. 2000. Optimal expressions in Indian English. *English Language and Linguistics* 4(1): 69–95.
- Bhatt, Rakesh M. 2004. Indian English: Syntax. In Kortmann et al. (eds), 1016–30.
- Bhaya Nair, Rukmini. 2008. Language and youth culture. In Kachru et al. (eds), 466–94.
- Biber, Douglas et al. (eds). 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson.
- Biswas, Gopa. 2004. Language policy in southeast Asia: A case study of India. In Mansoor, Meraj & Tahir (eds), 106–11.
- Blaicher, Günther & Glaser, Brigitte (eds). 1994. *Anglistentag 1993 Eichstätt: Proceedings*. Tübingen: Niemeyer.
- Blair, David & Collins, Peter (eds). 2001. *English in Australia*. Amsterdam: John Benjamins.
- Blank, Claudia (ed.). 1992. *Language and Civilization: A Concerted Profusion of Essays and Studies; In Honour of Otto Hietsch*. Frankfurt: Peter Lang.
- Burchfield, Robert (ed.). 1994. *The Cambridge History of the English Language*, Vol. 5. Cambridge: CUP.
- Burchfield, Robert (ed.). 1996. *The New Fowler's Modern English Usage*, 3rd edn. Oxford: OUP.
- Cambridge International Dictionary of English*. 2000. CD-Rom. Cambridge: CUP.
- Carls, Uwe. 1979. Select bibliography of Indian English (up to 1978). *Zeitschrift für Anglistik und Amerikanistik* 27: 327–40.
- Carls, Uwe. 1987. Die Entwicklung des Englischen in Indien von einer Fremdsprache zu einer Zweitsprache. In Hansen (ed.), 115–24.
- Carls, Uwe. 1994. Types of lexical peculiarities in Indian English. In Blaicher & Glaser (eds), 205–16.
- Carls, Uwe. 1999. Compounding in Indian English. In Carls & Lucko (eds), 141–53.
- Carls, Uwe. 2000. *Anti- and pro-* in Indian English. In Dalton-Puffer & Ritt (eds), 35–44.
- Carls, Uwe & Lucko, Peter (eds). 1999. *Form, Function and Variation in English. Studies in Honour of Klaus Hansen*. Frankfurt: Peter Lang.
- Carney, Edward. 1994. *A Survey of English Spelling*. London: Routledge.

- Cheshire, Jenny (ed.). 1991. *English Around the World. Sociolinguistic Perspectives*. Cambridge: CUP.
- Chesterman, Andrew. 1993. Articles and no articles. In Jucker (ed.), 13–24.
- Clyne, Michael (ed.). 1992. *Pluricentric Languages: Differing Norms in Different Nations*. Berlin: Mouton de Gruyter.
- Coelho, Gail M. 1997. Anglo-Indian English: A nativized variety of Indian English. *Language in Society* 26(4): 561–89.
- Collins, Peter & Peters, Pam. 1988. The Australian corpus project. In Kytö, Ihalainen & Rissanen (eds), 103–20.
- Comrie, Bernard. 1985. *Tense*. Cambridge: CUP.
- Concise Oxford Dictionary*. 2000. CD-Rom. Oxford: OUP.
- Corbett, Greville. 1983. *Hierarchies, Targets and Controllers*. Beckenham: Croom Helm.
- Crowther, Jonathan (ed.). 1996. *Oxford Advanced Learner's Dictionary of Current English. Fifth Edition with a Special Supplement of Indian English*. Oxford: OUP.
- Crystal, David. 1997. *English as a Global Language*. Cambridge: CUP.
- Crystal, David. 2005. Indian English. Interview. By Jill Kitson. February 19, 2005 <<http://www.abc.net.au/rn/arts/ling/stories/s1298284.htm>> (27 July 2005).
- Dahl, Trine. 1993. Adjective compounding as a text level process in English technical texts. In Jucker (ed.), 45–56.
- Dalton-Puffer, Christiane & Ritt, Nikolaus (eds). 2000. *Words: Structure, Meaning, Function. A Festschrift for Dieter Kastovsky*. Berlin: Mouton de Gruyter.
- Das, Amita. 1994. *India: Impact of the West*. Bombay: Popular Prakashan.
- Das, Gurcharan. 2002. *The Elephant Paradigm. India Wrestles with Change*. New Delhi: Penguin.
- Das, Sisir Kumar. 1982. Indian English. In Pride (ed.), 141–9.
- Dasgupta, Jyotirinda. 1970. *Language Conflict and National Development: Group Politics and National Language Policy in India*. Bombay: OUP.
- Dasgupta, Probal. 1993. *The Otherness of English. India's Auntie Tongue Syndrome*. New Delhi: Sage.
- Daswani, C. J. 1978. Some theoretical implications for investigating Indian English. *Indian Writing in English*. In Mohan (ed.), 115–28.
- Dayal, Ravi. 2006. Wealth of Indian English writing. In Verghese (ed.), 297–308.
- Dhillon, Rajinder S. 1998. *Common Errors in English Language*. New Delhi: Dhillon Group of Publications.
- Dissanayake, Wimal. 2008. Language in Cinema. In Kachru, Kachru & Sridhar (eds), 395–406.
- D'souza, Jean. 1987. English in India's language modernization. *World Englishes* 6(1): 63–70.
- D'souza, Jean. 1996. Creativity and language planning: The case of Indian English and Singapore English. *Language Problems and Language Planning* 20(3): 244–62.
- D'souza, Jean. 1997. Indian English: Some myths, some realities. *English World-Wide* 18(1): 91–105.
- D'souza, Jean. 2001a. Contextualizing range and depth in Indian English. *World Englishes* 20(2): 145–59.
- D'souza, Jean. 2001b. Indian English: The reality behind the myths. In Abbi, Gupta & Kidwai (eds), 241–55.
- Dubey, Vinod S. 1989. *Newspaper English in India*. New Delhi: Bahri Publications.

- Dubey, Vinod S. 1991. The lexical style of Indian newspaper English. *World Englishes* 10(1): 19–32.
- Dutt, Sandeep. 1999. *Good Schools of India. A Guide to the Best Residential Schools*. Dehra Dun: The English Book Depot.
- Edwardes, Michael. 1967. *British India 1772–1947. A Survey of the Nature and Effects of Alien Rule*. New Delhi: Rupa.
- Ferguson, Charles A. 1996. English in South Asia: Imperialist legacy and regional asset. In Baumgardner (ed.), 29–39.
- Foley, J. A. et al. (eds). 1998. *English in New Cultural Contexts. Reflections from Singapore*. Singapore: OUP.
- García, Ofelia & Otheguy, Ricardo (eds). 1989. *English across Cultures, Cultures across English. A Reader in Cross-Cultural Communication*. Berlin: Mouton de Gruyter.
- Gargesh, Ravinder. 2006. South Asian Englishes. In Kachru, Kachru & Nelson (eds), 90–113.
- George C., Jacob. 1998. Indian English: Certain lexical and grammatical variations. *Modern English Teacher* 7(4): 15–9.
- Goffin, R. C. 1934. *Some Notes on Indian English*. Oxford: Clarendon Press.
- Görlach, Manfred. 1985. Lexicographical problems of new Englishes and English-related pidgin and creole languages. *English World-Wide* 6(1): 1–36.
- Görlach, Manfred. 1995. *More Englishes. New Studies in Varieties of English 1988–1994*. Amsterdam: John Benjamins.
- Gramley, Stephan & Pätzold, Kurt-Michael. 1992. *A Survey of Modern English*. London: Routledge.
- Graustein, Gottfried & Leitner, Gerhard (eds). 1989. *Reference Grammars and Modern Linguistic Theory*. Tübingen: Max Niemeyer.
- Graustein, Gottfried & Thiele, Wolfgang (eds). 1989. *Englische Textlinguistik und Varietätenforschung. Linguistische Arbeitsberichte*. Leipzig: Karl-Marx-Universität.
- Green, Georgia M. 1996. *Pragmatics and Natural Language Understanding*, 2nd edn. Mahwah NJ: Lawrence Erlbaum.
- Greenbaum, Sidney (ed.). 1996. *Comparing Englishes Worldwide. The International Corpus of English*. Oxford: Clarendon Press.
- Gupta, Anjuli. 1991. The failure of English as a lingua franca in India. In Gupta & Kapoor (eds), 58–80.
- Gupta, R. S. 1991. English and Indian languages: Code mixing. In Gupta & Kapoor (eds), 207–15.
- Gupta, R. S. & Kapoor, Kapil (eds). 1991. *English in India. Issues and Problems*. Delhi: Academic Foundation.
- Hankin, Nigel. 2003. *Hanklyn-Janklin*. New Delhi: Tara Press.
- Hansen, Klaus (ed.). 1987. *Studien zur Sprachvariation (unter besonderer Berücksichtigung des Englischen)*. Berlin: Humboldt-Universität.
- Hansen, Klaus, Carls, Uwe & Peter Lucko. 1996. *Die Differenzierung des Englischen in nationale Varianten, Eine Einführung*. Berlin: Erich Schmidt.
- Harris, John. 1996. Syntactic variation and dialect divergence. In Singh (ed.), 31–57.
- Hawkins, R. E. 1984. *Common Indian Words in English*. Delhi: OUP.
- Hickey, Raymond (ed.). 2004. *Legacies of Colonial English*. Cambridge: CUP.
- Hirschmüller, Helmut. 1989. The use of complex prepositions in Indian English as compared with British and American English. *Linguistische Arbeitsberichte der Karl-Marx-Universität Leipzig* 69: 52–7.

- Hoffmann, Sebastian & Mukherjee, Joybrato 2007. Ditransitive verbs in Indian English and British English: A corpus-linguistic study. *Arbeiten aus Anglistik und Amerikanistik* 32(1): 5–24.
- Hosali, Priya. 1997. *Nuances of English in India. What the Butler Really Said*. Pune: Centre for Communication Studies.
- Hosali, Priya. 2004. Butler English: Morphology and syntax. In Kortmann et al. (eds), 1031–44.
- Hosali, Priya. 2005. Butler English. *English Today* 21(1): 34–9.
- Hosali, Priya & Tongue, Ray. 1989. *A Dictionary of Collocations for Indian Users of English*. Hyderabad: CIEFL, Department of Correspondence Courses.
- Hundt, Marianne. 1996. Has BrE been catching up with AmE over the past thirty years? In Ljung (ed.), 135–51.
- Hundt, Marianne. 1998a. *New Zealand English Grammar. Fact or Fiction?* Amsterdam: John Benjamins.
- Hundt, Marianne. 1998b. On the use of the mandative subjunctive in four major varieties of English. In Lindquist et al. (eds), 159–75.
- Hundt, Marianne. 2007. *English Mediopassive Constructions. A Cognitive, Corpus-based Study of their Origin, Spread, and Current Status*. Amsterdam: Rodopi.
- Hundt, Marianne, Nesselhauf, Nadja & Biewer, Carolin (eds). 2007. *Corpus Linguistics and the Web*. Amsterdam: Rodopi.
- Jenkins, Jennifer. 2003. *World Englishes. A Resource Book for Students*. London: Routledge.
- Jespersen, Otto. 1949. *A Modern English Grammar on Historical Principles*, Vols. I – VII. Copenhagen: Ejnar Munksgaard.
- Johansson, Stig. 1979. American and British English grammar: An elicitation experiment. *English Studies* 60: 195–215.
- Johansson, Stig & Stenström, Anna-Brita (eds). 1991. *English Computer Corpora: Selected Papers and Research Guide*. Berlin: Mouton de Gruyter.
- Johansson, Stig & Norheim, Else Helene. 1988. The subjunctive in British and American English. *ICAME Journal* 12: 27–36.
- John, Binoo K. 2007. *Entry from Backside Only. Hazaar Fundas of Indian-English*. New Delhi: Penguin.
- Joshi, Svati (ed.). 1994. *Rethinking English. Essays in Literature, Language, History*. Delhi: OUP.
- Jucker, Andreas H. 1992. *Social Stylistics: Syntactic Variation in British Newspapers*. Berlin: Mouton de Gruyter.
- Jucker, Andreas H. (ed.). 1993. *The Noun Phrase in English, Its Structure and Variability*. Heidelberg: Universitätsverlag C. Winter.
- Juul, Arne. 1975. *On Concord of Number in Modern English*. Copenhagen: Nova.
- Kachru, Braj B. 1965. The Indianness in Indian English. *Word* 21: 391–410.
- Kachru, Braj B. 1978. Lexical innovations in South Asian English. In Mohan (ed.), 80–100.
- Kachru, Braj B. 1982a. English in South Asia. In Bailey & Görlach (eds), 353–83.
- Kachru, Braj B. (ed.). 1982b. *The Other Tongue. English across Cultures*. Urbana IL: University of Illinois Press.
- Kachru, Braj B. 1983. *The Indianization of English. The English Language in India*. Delhi: OUP.
- Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English in the outer circle. In Quirk & Widdowson (ed.), 11–30.
- Kachru, Braj B. 1986. *The Alchemy of English: The Spread, Functions and Models of Non-native Englishes*. Oxford: Pergamon.
- Kachru, Braj B. 1994. English in South Asia. In Burchfield (ed.), 497–553.

- Kachru, Braj B. 1996. Series editor's preface. In Baumgardner (ed.), vi–xv.
- Kachru, Braj B. 2005. *Asian Englishes: Beyond the Canon*. Hong Kong: Hong Kong University Press.
- Kachru, Braj B., Kachru, Yamuna & Nelson, Cecil L. (eds). 2006. *The Handbook of World Englishes*. Oxford: Blackwell.
- Kachru, Braj B., Yamuna Kachru, Yamuna & Sridhar, Shikaripur N. (eds). 2008. *Language in South Asia*. Cambridge: CUP.
- Kachru, Yamuna. 1987. Cross-cultural texts, discourse strategies and discourse interpretation. In Smith (ed.), 87–100.
- Kachru, Yamuna. 2003. On definite reference in World Englishes. *World Englishes* 22(4): 497–510.
- Kachru, Yamuna. 2008. Hindi – Urdu – Hindustani. In Kachru, Kachru & Sridhar (eds), 81–102.
- Kachru, Yamuna & Nelson, Cecil L. 2006. *World Englishes in Asian Contexts*. Hong Kong: Hong Kong University Press.
- Kamdar, Mira. 2007. *Planet India. How the Fastest-growing Democracy is Transforming America and the World*. New York NY: Scribner.
- Kandiah, Thiru. 1998. Why new Englishes? In Foley et al. (eds), 1–40.
- Kandiah, Thiru & Kwan-Terry, John (eds). 1994. *English and Language Planning: A Southeast Asian Contribution*. Singapore: Times Academic Press.
- Kapoor, Kapil. 1993. *Language, Linguistics and Literature. The Indian Perspective*. Delhi: Academic Foundation.
- Kapoor, Kapil & Gupta, R. S. 1991. Introduction. In Gupta & Kapoor (eds), 13–24.
- Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. London: Longman.
- Khan, Z. 1994. English-Hindi code-mixing as a communicative strategy in Indian cine-journalism in English and its influence on bilingual EFL undergraduate students in India.” *JEFL* 13–14: 150–78.
- Kirkpatrick, Andy. 2007. *World Englishes. Implications for International Communication and English Language Teaching*. Cambridge: CUP.
- Kohli, Devindra. 1989. Indian English: A *khichri* of words. *Times Literary Supplement* December 1–7: 1332; 1340.
- Kortmann, Bernd et al. (eds). 2004. *A Handbook of Varieties of English*, Vol. 2: *Morphology and Syntax*. Berlin: Mouton de Gruyter.
- Kortmann, Bernd & Szmrecsanyi, Benedikt. 2004. “Global synopsis: Morphological and syntactic variation in English.” In Kortmann et al. (eds), 1142–1202.
- Kortmann, Bernd & Schneider, Edgar W. 2008. General introduction. In Meshtrie (ed.), 1–7.
- Koul, Omkar N. (ed.). 1992. *English in India, Theory and Applied Issues*. New Delhi: Creative.
- Krishnaswamy, N. & Burde, Archana S. 1998. *The Politics of Indians' English. Linguistic Colonialism and the Expanding English Empire*. Delhi: OUP.
- Kytö, Merja, Ihalainen, Ossi & Riissanen, Matti (eds). 1988. *Corpus Linguistics Hard and Soft. Proceedings of the Eighth International Conference on English Language Research on Computerized Corpora*. Amsterdam: Rodopi.
- Labru, G. L. 1984. *Indian Newspaper English*. Delhi: B. R. Publishing Corporation.
- Lal, Vinay & Nandy, Ashis (eds). 2006. *Fingerprinting Popular Culture. The Mythic and the Iconic in Indian Cinema*. New Delhi: OUP.
- Lange, Claudia. 2007. Focus marking in Indian English. *English World-Wide* 28(1): 89–118.

- Leech, Geoffrey. 2007. New resources, or just better ones? The holy grail of representativeness. In Hundt, Nesselhauf & Biewer (eds), 133–49.
- Lehmann, Thomas. 1989. *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Cultures.
- Leitner, Gerhard. 1985. Das Englische in Indien – ein Überblick über die Literatur seit ca. 1975. *ACOLIT* 17: 18–38.
- Leitner, Gerhard. 1987. Zur Grammatik des Englischen in Indien. *Zeitschrift für Anglistik und Amerikanistik* 35: 33–50.
- Leitner, Gerhard. 1988. Wie anders ist das Englische in Indien? In Ludwig (ed.), 386–96.
- Leitner, Gerhard. 1989a. Core grammar versus variety grammar – The case of English. In Graustein & Leitner (eds), 163–83.
- Leitner, Gerhard. 1989b. Der Tempusgebrauch im indischen Englisch als Problem für das ‘international English’. *Linguistische Arbeitsberichte (Sektion Theoretische und angewandte Sprachwissenschaft Karl-Marx-Universität Leipzig)* 69: 43–52.
- Leitner, Gerhard. 1991a. ‘Have I your permission to leave, sir?’ – Die Herausforderung der neuen Formen des Englischen für die Anglistik. *Die Neueren Sprachen* 90(1): 23–38.
- Leitner, Gerhard. 1991b. The Kolhapur Corpus of Indian English – Intravarietal description and/or intervarietal comparison. In Johansson & Stenström (eds), 215–31.
- Leitner, Gerhard. 1992a. English as a pluricentric language. In Clyne (ed.), 179–237.
- Leitner, Gerhard. (ed.). 1992b. *New Directions in English Language Corpora. Methodology, Results, Software Developments*. Berlin: Mouton de Gruyter.
- Leitner, Gerhard. 1994. *Begin and start* in British, American and Indian English. *Hermes* 13: 99–122.
- Levin, Magnus. 1998a. Concord with collective nouns in British and American English. In Lindquist et al. (eds), 193–204.
- Levin, Magnus. 1998b. On concord with collective nouns in English. In Renouf (ed.), 99–111.
- Lewis, Ivor. 1991. *Sahibs, Nabobs and Boxwallahs. A Dictionary of the Words of Anglo-India*. Bombay: OUP.
- Lindquist, Hans et al. (eds). 1998. *The Major Varieties of English. Papers from MAVEN 97, Växjö 20–22 November 1997*. Växjö: Växjö University.
- Ljung, Magnus (ed.). 1996. *Corpus-based Studies in English. Papers from the 17th International Conference on English Language Research on Computerized Corpora (ICAME 17)*. Amsterdam: Rodopi.
- Lüdeling, Anke, Evert, Stefan & Baroni, Marco 2007. Using web data for linguistic purposes. *Corpus Linguistics and the Web*. In Hundt, Nesselhauf & Biewer (eds), 7–24.
- Ludwig, Hans-Werner (ed.). 1988. *Anglistentag 1987 Tübingen. Vorträge*. Essen: Hoffmann.
- Lukmani, Yasmeen. 1992. Indian English: The written record. In Blank (ed.), 155–65.
- Mair, Christian. 1992a. A methodological framework for research on the use of nonstandard language in fiction. *Arbeiten aus Anglistik und Amerikanistik* 17(1): 103–23.
- Mair, Christian. 1992b. Problems in the compilation of a corpus of standard Caribbean English: A pilot study. In Leitner (ed.), 75–96.
- Mair, Christian. 1996. Machine-readable text corpora and the linguistic description of languages. In Zuell, Harkness & Hoffmeyer-Zlotnik (ed.), 65–76.
- Mair, Christian. 1998. Corpora and the study of the major varieties of English: Issues and results. In Lindquist et al. (eds), 139–57.

- Mair, Christian. 2002. Three changing patterns of verb complementation in late modern English: A real-time study based on matching text corpora. *English Language and Linguistics* 6(1): 105–31.
- Mair, Christian. 2003a. Kreolismen und verbales Identitätsmanagement im geschriebenen jamaikanischen Englisch. In Vogel, Napp & Lutterer (eds), 79–96.
- Mair, Christian. (ed.). 2003b. *The Politics of English as a World Language*. Amsterdam: Rodopi.
- Mair, Christian. 2006a. *Twentieth-Century English. History, Variation and Standardization*. Cambridge: CUP.
- Mair, Christian. 2006b. Tracking ongoing grammatical change and recent diversification in present-day standard English: The complementary role of small and large corpora. In Renouf (ed.), 355–76.
- Mair, Christian. 2007. Change and variation in present-day English: Integrating the analysis of closed corpora and web-based monitoring. In Hundt, Nesselhauf & Biewer (eds), 233–47.
- Mansoor, Sabiha, Meraj, Shaheen & Tahir, Aliya (eds). 2004. *Language Policy, Planning and Practice: A South Asian Perspective*. Karachi: Aga Khan University and OUP.
- Marshall, P. J., ed. 1996. *The Cambridge Illustrated History of the British Empire*. Cambridge: CUP.
- Masani, Zareer. 1987. *Indian Tales of the Raj*. Berkeley CA: University of California Press.
- Mason, John. 1978. *ISC English Practice Revised. A Senior Secondary Course*. New Delhi: OUP.
- Mathew, Rama. 1997. English in India: A response to Mark Tully. *ELT Journal* 51(2): 165–8.
- Mathur, Ashna Rani (ed.). 2006. *The Indian Media. Illusion, Delusion and Reality. Essays in Honour of Prem Bhatia*. New Delhi: Rupa.
- McArthur, Tom. 2002. *The Oxford Guide to World English*. Oxford: OUP.
- McArthur, Tom. 2006. English world-wide in the twentieth century. In Mugglestone (ed.), 360–93.
- McCrum, Robert, Cran, William & MacNeil, Robert. 1986. *The Story of English. New and Revised Edition*. Boston MA: Faber and Faber.
- McEnery, Tony, Xiao, Richard & Tono, Yukio. 2006. *Corpus-based Language Studies. An Advanced Resource Book*. London: Routledge.
- Mehrotra, Raja Ram. 1982. Indian English: A sociolinguistic profile. In Pride (ed.), 150–73.
- Mehrotra, Raja Ram. 1989. Indian literature in English. In García & Otheguy (eds), 421–39.
- Mehrotra, Raja Ram. 1998. *Indian English. Texts and Interpretations*. Amsterdam: John Benjamins.
- Mehrotra, Raja Ram. 2000. Indian Pidgin English: Myth and reality. *English Today* 16(3): 49–52.
- Melchers, Gunnel & Shaw, Philip. 2003. *World Englishes*. London: Arnold.
- Merriam-Webster's Collegiate Dictionary. 2000. CD-Rom. Springfield IL: Merriam-Webster, Incorporated.
- Merriam-Webster's Dictionary of English Usage. 1994. Springfield IL: Merriam-Webster, Incorporated.
- Mesthrie, Rajend (ed.). 2008. *Varieties of English 4. Africa, South and Southeast Asia*. Berlin: Mouton de Gruyter.
- Mobärg, Mats. 1998. Media exposure vs. educational prescription: The case of British and American English in Sweden. In Lindquist et al. (eds), 249–62.
- Modiano, Marko. 1996. The Americanization of Euro-English. *World Englishes* 15(2): 207–15.

- Mohan, Ramesh (ed.). 1978. *Indian Writing in English. Papers Read at the Seminar on Indian English Held at the Central Institute of English and Foreign Languages, Hyderabad, July 1972*. Bombay: Orient Longman.
- Mugglestone, Lynda (ed.). 2006. *The Oxford History of English*. Oxford: OUP.
- Mukherjee, Joybrato. 2007. Steady states in the evolution of new Englishes. Present-day Indian English as an equilibrium. *Journal of English Linguistics* 35(2): 157–87.
- Mukherjee, Joybrato & Hoffmann, Sebastian. 2006. Describing verb-complementational profiles of new Englishes. *English World-Wide* 27(2): 147–73.
- Muthiah, S. 1991. *Words in Indian English. A Reader's Guide*. New Delhi: Indus.
- Nelson, Gerald. 1996. The design of the corpus. In Greenbaum 1996 (ed.), 27–35.
- Nichols, Ann Eljenholm. 1987. The suasive subjunctive: Alive and well in the upper Midwest. *American Speech* 62: 140–52.
- Nihalani, Paroo, Tongue, Ray K. & Hosali, Priya. 1979. *Indian and British English. A Handbook of Usage and Pronunciation*. New Delhi: OUP.
- Nihalani, Paroo, Tongue, Ray K., Hosali, Priya & Crowther, Jonathan. 2004. *Indian and British English. A Handbook of Usage and Pronunciation, Second Edition*. New Delhi: OUP.
- Olavarria de Ersson, Eugenia & Shaw, Philip. 2003. Verb complementation patterns in Indian Standard English. *English World-Wide* 24: 137–61.
- Övergaard, Gerd. 1995. *The Mandative Subjunctive in American and British English in the 20th Century*. Uppsala: Acta Univ. Ups.
- Oxford English Dictionary*. 1999. Second edn. CD-Rom. Oxford: OUP.
- Pandharipande, Rajeshwari. 1987. On nativization of English. *World Englishes* 6(2): 149–58.
- Parasher, S. V. 1980. Mother tongue English diglossia: A case study of educated Indian bilinguals' language use. *Anthropological Linguistics* 22(4): 151–62.
- Parasher, S. V. 1981. English in India: A sociolinguistic reappraisal. *Zeitschrift für Anglistik und Amerikanistik* 29(4): 330–42.
- Parasher, S. V. 1991. *Indian English. Functions and Form*. New Delhi: Bahri.
- Parasher, S. V. 1994 [1983]. Indian English: Certain grammatical, lexical and stylistic features. In Agnihotri & Khanna (eds), 145–64.
- Pattanayak, D. P. 1978. Some observations on English in India: Its form and function. *Indian Writing in English*. In Mohan (ed.), 184–91.
- Paul, Premila. 2003. The master's language and its Indian users. In Mair (ed.), 359–65.
- Peters, Pam. 1995. *The Cambridge Australian English Style Guide*. Cambridge: CUP.
- Peters, Pam. 1998. The survival of the subjunctive: Evidence of its use in Australia and elsewhere. *English World-Wide* 19(1): 87–103.
- Platt, John et al. (eds). 1984. *The New Englishes*. London: Routledge and Kegan Paul.
- Pride, John (ed.). 1982. *New Englishes*. Rowley MA: Newbury House Publishers.
- Quirk, Randolph et al. (eds). 1985. *A Comprehensive Grammar of the English Language*. Harlow: Longman.
- Quirk, Randolph & Widdowson, Henry G. (eds). 1985. *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: CUP.
- Rahman, Tariq. 2002. *Language Ideology and Power. Language-learning among the Muslims of Pakistan and North India*. Karachi: OUP.
- Ramaiah, L. S. 1988. *Indian English. A Bibliographical Guide to Resources*. Delhi: Gian Publishing House.
- Ramamurthi, V. Lalitha. 1991. An error analysis of the wrong usage of English noun phrases. *International Journal of Dravidian Linguistics* 20(1): 117–29.

- Renouf, Antoinette (ed.). 1998. *Explorations in Corpus Linguistics*. Amsterdam: Rodopi.
- Renouf, Antoinette (ed.). 2006. *The Changing Face of Corpus Linguistics*. Amsterdam: Rodopi.
- Reppen, Randi, Susan M. Fitzmaurice and Douglas Biber, eds. 2002. *Using Corpora to Explore Linguistic Variation*. Amsterdam: John Benjamins.
- Rogers, Chandrika K. 2002. Syntactic features of Indian English. An examination of written Indian English. In Reppen, Fitzmaurice & Biber (eds), 187–202.
- Sahgal, Anju. 1991. Patterns of language use in a bilingual setting in India. In Cheshire (ed.), 299–306.
- Sahgal, Anju & Agnihotri, Rama Kant. 1985. Syntax – the common bond. Acceptability of syntactic deviances in Indian English. *English World-Wide* 6(1): 117–29.
- Sand, Andrea. 1998. *Linguistic Variation in Jamaica: A Corpus-based Study of Radio and Newspaper Usage*. Tübingen: Gunter Narr.
- Sand, Andrea. 2003. The definite article in Irish English and other contact varieties of English. In Tristram (ed.), 413–30.
- Sand, Andrea. 2004. Shared morpho-syntactic features in contact varieties of English: Article use. *World Englishes* 23(2): 281–98.
- Sand, Andrea & Siemund, Rainer. 1992. LOB – 30 Years On ... *ICAME Journal* 16: 119–22.
- Sanjay, B. P. 2006. Press history in India. In Mathur (ed.), 13–26.
- Sayder, Stefan. 1989. The subjunctive in Indian, British and American English: A corpus-based study. In Graustein & Thiele (eds), 58–66.
- Schilk, Marco. 2006. Collocations in Indian English: A corpus-based sample analysis. *Anglia* 124(2): 276–316.
- Schmied, Josef (ed.). 1989a. *English in East and Central Africa I*. Bayreuth: Eckhard Breitinger.
- Schmied, Josef. 1989b. Second-language varieties across the Indian Ocean. In Schmied (ed.), 85–97.
- Schmied, Josef. 1994. Syntactic style variation in Indian English. In Blaicher & Glaser (eds), 217–32.
- Schneider, Edgar W. 2000. Corpus linguistics in the Asian context: Exemplary analyses of the Kolhapur corpus of Indian English. In Bautista, Llamzon & Sibayan (eds), 115–37.
- Schneider, Edgar W. 2003. The dynamics of New Englishes: From identity construction to dialect birth. *Language* 79(2): 233–81.
- Schneider, Edgar W. 2004. How to trace structural nativization: Particle verbs in world Englishes. *World Englishes* 23(2): 227–49.
- Schneider, Edgar W. 2007. *Postcolonial English. Varieties Around the World*. Cambridge: CUP.
- Sharma, Devyani. 2001. The pluperfect in native and non-native English: A comparative corpus study. *Language Variation and Change* 13: 343–73.
- Sharma, Devyani. 2005. Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition* 27: 535–566.
- Shastri, S. V. 1983. Towards a definition of Indian English. *CIEFL Newsletter* 18(3–4): 1–8.
- Shastri, S. V. 1985. Research in progress towards a description of Indian English: A standard corpus in machine-readable form. *English World-Wide* 6(2): 275–8.
- Shastri, S. V. 1988a. Code-mixing in the process of Indianization of English: A corpus-based study. *Indian Linguistics* 1–4: 34–53.
- Shastri, S. V. 1988b. The Kolhapur corpus of Indian English and work done on its basis so far. *ICAME Journal* 12: 15–26.
- Shastri, S. V. 1988c. Preface. In Ramaiah, vii–ix.
- Shastri, S. V. 1992. Opaque and transparent features of Indian English. In Leitner (ed.), 263–75.

- Shastri, S. V. 1996. Using computer corpora in the description of language with special reference to complementation in Indian English. In Baumgardner (ed.), 70–81.
- Shastri, S. V., Patilkulkarni, C. T. & Shastri, Geeta S. 1986. *Manual of Information to Accompany the Kolhapur Corpus of Indian English, for Use with Digital Computers*. Kolhapur: Shivaji University.
- Shekar, Chandra & Hedge, M. N. 1996. Cultural and linguistic diversity among Asian Indians: A case of Indian English. *Topics in Language Disorders* 16(4): 54–64.
- Sheorey, Ravi. 2006. *Learning and Teaching English in India*. New Delhi: Sage.
- Siemund, Rainer. 1995. 'For who the bell tolls' – Or why corpus linguistics should carry the bell in the study of language change in present-day English. *Arbeiten aus Anglistik und Amerikanistik* 20(2): 351–77.
- Sinclair, John (ed.) 2001. *Collins Cobuild English Dictionary for Advanced Learners*. Glasgow: HarperCollins.
- Singh, Rajendra. 1994. Indian English: Some conceptual issues. In Agnihotri & Khanna (eds), 369–81.
- Singh, Rajendra. 1995. A note on code-switching and stratification in North India. In Singh, Dasgupta & Lele (eds), 131–4.
- Singh, Rajendra (ed.) 1996. *Towards A Critical Sociolinguistics*, 31–57. Amsterdam: John Benjamins.
- Singh, Rajendra, Dasgupta, Probal & Lele, Jayant K. (eds). 1995. *Explorations in Indian Sociolinguistics*. New Delhi: Sage.
- Singh, Ravendra (ed.). 2000. *Statistical Abstract India 1999*. Delhi: Central Statistical Organisation, Ministry of Statistics and Programme Implementation, Government of India.
- Singh, Udaya Narayan. 2006. *India Writes. A Story of Linguistic and Literary Plurality*. New Delhi: National Book Trust India.
- Sinha, Surendra Prasad. 1978. *English in India. A Historical Study with Particular Reference to English Education in India*. Patna: Janaki Prakashan.
- Smith, Larry (ed.). 1987. *Discourse across Cultures, Strategies in World Englishes*. New York NY: Prentice Hall.
- Smith-Pearse, T. L. H. 1968 [1934]. *The English Errors of Indian Students*. Madras: OUP.
- Spear, Percival. 1965. *A History of India Volume 2. From the Sixteenth to the Twentieth Century*. New Delhi: Penguin.
- Spitzbardt, Harry. 1976. *English in India*. Halle: VEB Max Niemeyer Verlag.
- Spitzbardt, Harry. 1987. Standardisierte Besonderheiten des Indischen Englisch. In Hansen (ed.), 125–35.
- Sridhar, Kamal K. 1982. English in a South Indian urban context. In Kachru (ed.), 141–53.
- Sridhar, Shikaripur N. 1990. *Kannada: Descriptive Grammar*. London: Routledge.
- Sridhar, Shikaripur N. 1992. The ecology of bilingual competence: Language interaction in the syntax of indigenized varieties of English. *World Englishes* 11(2–3): 141–50.
- Srivastava, Sanjay. 1998. *Constructing Post-Colonial India. National Character and the Doon School*. London: Routledge.
- Subba Rao, G. 1950. *Indian Words in English: A Study in Indo-British Cultural and Linguistic Relations*. Oxford: Clarendon.
- Subramanian, K. 1978. Penchant for the florid. In Mohan (ed.), 203–6.
- Svartvik, Jan. 1998. Varieties of English: Major and minor. In Lindquist et al. (ed.), 15–27.
- Svartvik, Jan & Leech, Geoffrey. 2006. *English. One Tongue, Many Voices*. Basingstoke: Palgrave Macmillan.

- Talaat, Mubina. 1993. Lexical variation in Pakistani English. In Baumgardner (ed.), 55–62.
- Taylor, Brian. 2001. Australian English in interaction with other Englishes. In Blair & Collins (eds), 317–40.
- Taylor, Harvey M. & Krishna Rao, A.V. 1981. Indian English 'errors' and American English usage. *IRAL* 19: 137–47.
- Tharoor, Shashi. 2007. *The Elephant, the Tiger, and the Cell Phone. Reflections on India, the Emerging 21st-Century Power*. New York NY: Arcade Publishing.
- Tharu, Susie. 1994. The arrangement of an alliance: English and the making of Indian literatures. In Joshi (ed.), 160–80.
- Thelwall, Mike. 2005. Creating and using web corpora. *International Journal of Corpus Linguistics* 10(4): 517–41.
- Thomason, Sarah G. 2001. *Language Contact. An Introduction*. Washington DC: Georgetown University Press.
- Tickoo, Makhan L. (ed.). 1991a. *Languages and Standards: Issues, Attitudes, Case Studies*. Singapore: SEAMEO Regional Language Centre.
- Tickoo, Makhan L. (ed.). 1991b. Stakeholders and standards: Englishes for tomorrow's India. In Tickoo (ed.), 131–45.
- Timmerberg, Helge. 2006. *Shiva Moon. Eine Reise durch Indien*. Berlin: Rowohlt.
- Tongue, Ray & Hosali, Priya. 1981. A dictionary of collocations (continued). *CIEFL Newsletter* 17(1–2): 24–7.
- Tristram, Hildegard L. C. (ed.). 2003. *The Celtic Englishes* III. Heidelberg: Winter.
- Trivedi, Harish. 2006. All kinds of Hindi: The evolving language of Hindi cinema. In Lal & Nandy (eds), 51–86.
- Trudgill, Peter & Hannah, Jean. 2002. *International English. A Guide to the Varieties of Standard English*, 4th edn. London: Arnold.
- Tully, Mark. 1997. English: An advantage to India? *ELT Journal* 51(2): 157–64.
- Turner, John F. 1980. The marked subjunctive in contemporary English. *Studia Neophilologica* 52: 271–7.
- Vaish, Viniti. 2008. *Biliteracy and Globalization. English Language Education in India*. Cleveland: Multilingual Matters.
- Valentine, Tamara M. 1991. Getting the message across: Discourse markers in Indian English. *World Englishes* 10(3): 325–34.
- Varma, Pavan K. 2004. *Being Indian. Inside Real India*. London: William Heinemann.
- Venezky, Richard L. 1999. *The American Way of Spelling: The Structure and Origins of American English Orthography*. New York NY: The Guildford Press.
- Vergheese, B. G. (ed.). 2006. *Tomorrow's India. Another Tryst with Destiny*. New Delhi: Viking.
- Verma, Dharendra. 1933. *Hindi bhasa ka itihas*. Allahabad: Hindustani Academy.
- Verma, Shivendra Kishore. 1978. Syntactic irregularities in Indian English. In Mohan (ed.), 207–20.
- Verma, Shivendra Kishore. 1982. Swadeshi English: Form and function. In Pride (ed.), 174–87.
- Viereck, Wolfgang, Viereck, Karin & Ramisch, Heinrich. 2002. *Dtv-Atlas Englische Sprache*. München: Deutscher Taschenbuch Verlag.
- Vijayanunni, M. (ed.). 1997. *Census of India 1991, Series 1 – India, Paper 1 of 1997, Language. India and States (Table C–7)*. New Delhi: Registrar General & Census Commissioner, India.
- Virtanen, Tuija & Lindgrén, Signe-Anita. 1998. British or American English? Investigating what EFL students say and what they do. In Lindquist et al. (eds), 273–81.

- Visser, F. T. 1963. *An Historical Syntax of the English Language*. Leiden: Brill.
- Vogel, Elisabeth, Napp, Antonia & Lutterer, Wolfram (eds). 2003. *Zwischen Ausgrenzung und Hybridisierung*. Würzburg: Ergon.
- Wagenaar, H. W. et al. (eds). 1993. *Allied Chambers Transliterated Hindi – Hindi – English Dictionary*. New Delhi: Allied Chambers (India).
- Whitworth, George Clifford. 1982 [1907]. *Indian English: An Examination of the Errors of Idiom Made by Indians in Writing English*. New Delhi: Bahri Publications.
- Wilson, Kenneth G. 1993. *The Columbia Guide to Standard American English*. New York NY: Columbia University Press.
- Wiltshire, Caroline R. 2005. The 'Indian English' of Tibeto-Burman language speakers. *English World-Wide* 26(3): 275–300.
- Wren, P. C., Martin, H. & Prasada Rao, N. D. V. 2001. *High School English Grammar and Composition. New Edition*. New Delhi: S. Chand and Company.
- Yadurajan, K. S. 2001. *Current English. A Guide for the User of English in India*. New Delhi: OUP.
- Yule, Henry & Burnell, A. C. 1994 [1886]. *Hobson-Jobson. A Glossary of Colloquial Anglo-Indian Words and Phrases, and of Kindred Terms, Etymological, Historical, Geographical and Discursive*. New Delhi: Rupa.
- Zuell, Cornelia, Harkness, Janet & Hoffmeyer-Zlotnik, Jürgen H. P. (eds). 1996. *Text Analysis and Computers*. Mannheim: ZUMA.

Appendices

Appendix 2

Appendix 2.1 Primary Corpus: Press Categories.

IND2000	IND1978	Lob	Flob	Brown	Frown
Reports:					
01: Hindu	A01: Hindu	A01	A01	A01	A01
02: Statesman	A02: Statesman	A02	A02	A02	A02
03: Hindustan Times	A03: Hindustan Times	A03	A03	A03	A03
04: Indian Express	A04: Indian Express	A04	A04	A04	A04
05: Times of India	A05: Times of India	A05	A05	A05	A05
06: Times of India	A06: Times of India	A06	A06	A06	A06
07: Hindu	A07: Hindu	A07	A07	A11	A15
08: Times of India	A08: Times of India	A08	A08	A12	A16
09: Times of India	A09: Times of India	A17	A17	A07	A07
10: Indian Express	A10: Indian Express	A18	A18	A08	A08
11: Hindu	A12: Hindu	A11	A11	A16	A22
12: Statesman	A13: Statesman	A12	A12	A17	A23
13: Hindustan Times	A14: Hindustan Times	A13	A13	A18	A24
14: Times of India	A15: Times of India	A14	A14	A19	A25
15: Economic Times	A16: Economic Times	A15	A15	A26	A34
16: Times of India	A17: Times of India	A16	A16	A27	A35
17: Times of India	A20: Times of India	A22	A22	A13	A17
18: Indian Express	A21: Times of India	A23	A23	A14	A18
19: Hindu	A22: Hindu	A20	A20	A20	A11
20: Economic Times	A24: Economic Times	A21	A21	A21	A12
Editorials:					
01: Hindu	B01: Hindu	B01	B01	B01	B01
02: Statesman	B02: Statesman	B02	B02	B02	B02
03: Hindustan Times	B03: Hindustan Times	B03	B03	B03	B03
04: Times of India	B05: Times of India	B05	B04	B04	B04
05: Hindustan Times	B07: Hindustan Times	B07	B07	B11	B11
06: Hindustan Times	B08: Hindustan Times	B08	B08	B13	B13
07: Hindu	B09: Hindu	B09	B09	B15	B21
08: Statesman	B10: Statesman	B10	B10	B16	B23
09: Times of India	B11: Times of India	B11	B11	B17	B24
10: Sunday S'man	B12: Sunday S'man	B12	B12	B20	B20
Reviews:					
01: Hindu	C01: Hindu	C01	C01	C01	C01
02: Hindu	C02: Hindu	C02	C02	C02	C02
03: Statesman	C03: Statesman	C03	C03	C03	C03
04: Times of India	C04: Times of India	C04	C04	C04	C04
05: Times of India	C05: Times of India	C05	C05	C05	C05
06: Hindustan Times	C06: Hindustan Times	C06	C06	C06	C06
07: Hindu	C07: Hindu	C07	C07	C07	C07
08: Hindustan Times	C08: Hindustan Times	C08	C08	C08	C08
09: Times of India	C09: Times of India	C09	C09	C09	C09
10: Indian Express	C10: Times of India	C10	C10	C10	C10

Appendix 3

Appendix 3.1 The loanword *thonga* [Internet: newspapers; January 16, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE
thonga	–	–	–	–	–	2

Appendix 3.2 Spelling variants of *tehsil/taluk* [Internet: newspapers; January 24, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE
<tehsil>	250	84	21	193	1930	7
<tahsil>	6	2	4	2	1	–
<taluk>	65	920	1050	10	35	–
<talug>	–	1	–	–	1	1
<taluka>	907	77	17	97	147	13
<taluga>	1	–	–	–	1	–

Appendix 3.3 Loanwords for political institutions [Internet: newspapers; February 7, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE
panchayat samiti	9	24	2	1	490	16
Lok Sabha	623	4740	553	646	6900	557
Rajya Sabha	278	2730	279	343	3710	285
Vidhan Sabha	75	175	11	19	3480	17

Appendix 3.4 *Vidhan Sabha* vs. *assembly* [World Wide Web; 14 February 2001].

	ASSEMBLY	VIDHAN SABHA
Tamil Nadu	1310	–
Orissa	1220	–
Assam	460	–
Andhra Pradesh	406	–
Karnataka	244	–
Pondicherry	241	–
Meghalaya	65	–
Sikkim	59	–
Arunachal Pradesh	39	–
Lakshadweep	2	–
Andaman/Nicobar Is.	1	–
Dadra and Nagar H.	1	–
Daman and Diu	1	–

Appendix 3.5 *Bandh* and *dharna* [Internet: newspapers; March 20, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN
bandh	598	1040	28	3080	1600	272	12	18	8
general strike	74	174	24	558	239	16	160	57	22
dharna (total)	765	833	143	1040	3870	77	3	–	2
a dharna	347	335	84	233	2640	47	–	–	–
a sit-in	22	36	3	21	70	20	355	41	9
the dharna	115	121	21	69	826	10	2	–	1
the sit-in	3	6	1	9	15	3	36	2	4

Appendix 3.6 *Hartal* and *gherao* [Internet: newspapers; March 20, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN
a gherao	3	1	–	1	23	1	4	–	1
a picketing	2	14	1	1	1	–	–	36	–
the gherao	3	3	2	6	81	11	1	–	1
the picketing	1	9	1	5	6	–	4	16	–
gheraoed	133	69	13	132	434	75	5	1	23
picketed	8	31	1	94	27	1	11	22	–
hartal (total)	32	177	7	32	87	5	13	127	223
hartal call	4	14	1	1	7	–	–	–	22
strike call	91	111	16	347	323	19	456	10	20

Appendix 3.7 *Rasta roko*, *road roko* and *road blockade*

[Internet: newspapers; March 20, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN
rasta roko	38	55	41	10	58	1	–	–	–
rasta rokos	5	8	–	–	4	1	–	–	–
road roko	1	92	1	11	5	–	–	–	–
road rokos	2	6	–	1	1	–	–	–	–
road blockade	27	46	18	16	78	19	22	–	2
road blockades	17	13	8	12	45	9	6	–	–

Appendix 3.8 Integration of *speed money* [Internet: newspapers; August 4, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN
Total	1	5	1	2	22	1	1	–	–
<speedmoney>	–	–	–	–	–	–	–	–	–
<speed-money>	–	–	–	–	1	–	–	–	–
<speed Money>	–	–	–	1	–	–	–	–	–
<speed money>	1	5	1	1	21	1	1	–	–
inverted commas	1	1	–	1	12	–	–	–	–
italics	1	–	–	–	–	–	–	–	–

Appendix 3.9 *Speed money* [Internet: country domains; August 4, 2001].

	.in	.sg	.uk	.za	.au	.us
speed money	60	2	6	–	6	–

Appendix 3.10 Integration of *timepass* [Internet: newspapers; March 19, 2003].*

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN
Total	10	20	5	23	15	1	1	–	–
<time pass>	3	9	1	3	7	–	–	–	–
<time-pass>	3	4	3	10	5	–	–	–	–
<timepass>	4	7	1	10	3	1	1	–	–
inverted commas	1	4	–	1	5	–	–	–	–
italics	2	–	–	–	–	–	–	–	–
pastime	6	278	39	–	10	97			

* Excluded from the count were all occurrences of *timepass* denoting a column of the same name in the *Indian Express* as well as free collocations of time and pass such as *first-time pass*.

Appendix 3.11 *Timepass* [Internet: country domains; 19 March 2003].

	.in	.sg	.uk	.za	.au	.us
timepass	56	1	13	–	3	3

Appendix 3.12 *Unproper, debtful and dissentment* [Internet: newspapers; April 19, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN	BrE GU	AmE WP
improper	176	294	81	437	556	43	600	–	–	396	662
unproper	–	–	–	–	–	–	1	–	8	–	–
the dissent	11	13	1	6	18	5	16	–	–	12	27
dissentment	–	–	–	–	–	–	–	–	–	–	–
debtful	–	–	–	–	–	–	–	–	–	–	–
indebted	75	125	14	93	252	31	294	–	6	598	60

Appendix 3.13 *Unproper, debtful and dissenment* [Internet: country domains; April 19, 2001].

	.in	.sg	.uk	.za	.au	.us
improper	92100	17900	669000	73000	618000	1820000
unproper	14	23	1060	4	371	87
the dissent	164	61	15000	149	9760	86600
dissentment	2	4	4	4	1	–
debtful	1	–	4	2	5	5
indebted	37900	896	474000	52800	232000	214000

Appendix 3.14 *Govtt, deptt. and asstt.* [Internet: newspapers; November 25, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN	BrE GU	AmE WP
govt.	6430	12900	1060	39600	7020	258	–	–	–	24	476
govtt.	1	–	–	–	–	1	–	–	–	–	–
dept.	404	987	101	3950	1130	4	978	3	6	156	793
deptt.	10	1	–	118	129	1	106	–	3	–	–
asst.	57	1440	16	295	114	2	13	–	1	23	24
asstt.	5	8	–	64	100	–	3	–	2	–	–

Appendix 3.15 *Govtt, deptt. and asstt.* [Internet: country domains; November 25, 2002].

	.in	.sg	.uk	.za	.au	.us
govt.	1870000	58600	899000	418000	3840000	1400000
govtt.	38	–	1	–	4	7
dept.	814000	269000	6680000	318000	318000	8840000
deptt.	602000	19	545	9	151	383
asst.	221000	69000	320000	12100	127000	1290000
asstt.	157000	12	160	3	322	16

Appendix 3.16 *Senti, funda and enthu* [Internet: newspapers; September 15, 2003].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN	BrE GU	AmE WP
senti	2	2	–	1	3	–	–	–	–	–	–
funda	1	54	6	123	13	–	14	–	–	–	–
enthu	–	2	–	3	2	–	–	–	–	–	–

Appendix 3.17 *Age-bar and by-heart* [Internet: newspapers; September 26, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	PakE DA	SrLaE ST	BaE IN	BrE GU	AmE WP
age-barred	–	–	–	–	–	–	–	–	–	–	–
age-barring	–	–	–	–	–	–	–	–	–	–	–
to by-heart	–	–	–	–	–	–	–	–	–	–	–

Appendix 3.18 *Age-bar and by-heart* [Internet: newspapers; September 26, 2001].

	.in	.sg	.uk	.za	.au	.us
age-barred	–	–	–	–	–	–
age-barring	–	–	–	–	–	–
to by-heart	–	–	–	–	–	–

Appendix 3.19 *Bottle of beer vs. beer bottle* [Internet: newspapers; April 29, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
bottle of beer								
drink	9	6	–	6	9	–	15	11
container	–	–	–	–	1	–	–	–
ambiguous	1	–	–	–	1	–	–	–
beer bottle								
drink	–	–	–	–	2	–	–	–
container	–	4	1	3	5	–	29	8
ambiguous	–	–	–	–	2	–	–	–

Appendix 3.20 Nominal compounds: IndE, BrE, AmE [Internet: newspapers; June 16, 2001].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
schedule flight	–	–	–	1	–	–	–	–
scheduled flight	13	14	1	3	16	1	42	3
charter bus	–	–	–	–	1	–	–	2
chartered bus	5	2	–	63	6	2	–	5
parliament elections	5	14	8	5	44	2	65	3
parliamentary elections	139	145	43	286	520	73	619	101
milch cow	3	18	–	6	11	3	12	1
milched cow	–	–	–	–	–	–	–	–
ice water	1	1	–	1	4	–	6	35
iced water	1	–	–	4	2	–	11	–
archive material	–	1	–	1	4	–	20	–
archived material	–	–	–	–	–	–	8	2
archival material	4	8	1	12	14	3	3	6
department store	30	23	5	79	60	5	232	194
departmental store	14	15	3	88	71	8	–	–

Appendix 3.21 Nominal compounds: PakE, BaE, SrLE [Internet: newspapers; June 16, 2001].

	PakE DA	BaE IN	SrLE ST
ice water	–	–	2
iced water	2	–	2
archive material	–	–	1
archived material	–	–	–
archival material	1	–	2
department store	17	3	23
departmental store	12	1	4

Appendix 3.22 Nominal compounds [Internet: country domains; June 16, 2001].

	.in	.sg	.uk	.za	.au	.us
ice water	18	95	2050	141	875	1570
iced water	3	46	818	107	909	42
archive material	9	5	6000	126	344	34
the archive material	–	1	247	4	15	1
archived material	6	2	787	74	302	120
the archived material	1	1	93	–	23	13
archival material	28	30	2510	192	2230	475
the archival material	8	3	155	11	134	14
department store	39	720	16600	519	6600	5940
departmental store	36	246	188	29	37	4

Appendix 3.23 Mediopassive *release* [Internet: newspapers; November 29, 2003].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
the film will release	4	1	–	5	1	2	–	–
the film will be released	6	43	14	12	61	9	12	2
the movie will release	–	–	–	–	1	–	–	–
the movie will be released	–	5	3	1	11	–	–	1

Appendix 3.24 Mediopassive *release* [Internet: country domains; November 29, 2003].

	.in	.sg	.uk	.za	.au	.us
the film will release	15	–	3	3	3	4
the film will be released	25	17	622	92	18300	448
the movie will release	19	–	3	–	–	–
the movie will be released	5	9	628	38	157	70

Appendix 3.25 Common formal lexical items [Internet: newspapers; December 17, 2000].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
detenu	13	43	–	11	7	–	–	–
detainee	33	13	4	33	105	1	280	113
miscreants	40	509	231	29	1740	170	81	23
culprits	84	1350	167	61	2750	178	443	76
criminals	110	1520	206	192	3080	361	3840	714
fructify	41	49	6	9	62	5	3	–
materialise	264	288	25	66	496	41	512	185
demit	23	11	3	–	20	2	–	–
condole	39	124	4	70	126	8	–	1
untoward	276	348	71	152	946	59	116	50
unexpected	701	767	103	292	1190	145	2520	1060
cessation	146	198	30	53	249	23	326	88
stopping	588	718	122	254	1280	128	2060	1060
thrice	665	353	68	120	979	115	73	33
three times	741	457	102	202	1120	63	2610	1630

Appendix 3.26 Rare formal lexical items [Internet: newspapers; December 17, 2000].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
ancient	371	9380	2790	695	5220	1660	12100	11100
old	2800	110000	29000	6790	47800	31300	123000	313000
demise	485	356	67	70	902	51	1460	408
death	7270	6490	2300	3840	18400	2040	45500	7780

Appendix 3.26 (continued)

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
bosom	7	193	303	64	147	44	443	443
chest	161	3050	984	338	3090	1210	6730	8580
comely	–	70	8	4	99	15	93	295
pretty [graphic]	457	8620	2230	805	4050	2650	24500	62300
resplendent	11	619	155	31	210	107	283	406
dazzling	36	1840	579	120	970	447	1150	2050

Appendix 3.27 Contractions in IndE, BrE, AmE [Primary Corpus: Press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
'd (= had)	–	–	–	6	–	5
'd (= would)	1	4	2	5	3	9
'll (= will)	–	4	8	13	27	25
'm (= am)	–	–	7	9	8	28
're (= are)	1	3	5	8	18	51
's (= has)	–	–	2	4	5	25
's (= is)	4	51	34	61	54	175
's (= us)	2	–	2	–	1	7
've (= have)	–	3	8	12	12	29
ain't	–	–	–	1	1	4
aren't	2	4	2	3	2	7
can't	4	7	5	12	16	27
couldn't	–	4	3	1	1	11
didn't	–	6	6	10	13	36
doesn't	2	9	10	9	14	2
don't	6	12	20	23	17	51
hadn't	1	2	–	–	1	4
hasn't	–	2	1	1	1	10
haven't	–	2	2	6	1	5
isn't	1	4	3	9	4	18
mustn't	–	–	–	–	–	–
needn't	–	–	–	–	–	3
shouldn't	–	–	–	1	1	3
wasn't	1	1	3	4	5	10
weren't	–	1	1	–	2	3
won't	–	4	4	4	9	28
wouldn't	–	2	3	2	4	6
	11 types 25 tok.	19 types 125 tok.	21 types 131 tok.	22 types 204 tok.	24 types 220 tok.	26 types 602 tok.

Appendix 3.28 Contractions in IndE and BrE [Primary Corpus: Broadcast].

	IndE IND2000 Broad News	IndE IND2000 Broad I/D	BrE UK1991 Broad News	BrE UK1961 Broad I/D
'd (= had)	–	5	28	33
'd (= would)	–	9	18	29
'll (= will)	11	51	46	32
'm (= am)	8	90	18	72
're (= are)	20	146	67	150
's (= has)	11	20	64	50
's (= is)	84	386	247	439
's (= us)	17	48	2	13
've (= have)	11	51	63	148
ain't	–	–	–	–
aren't	1	5	2	11
can't	–	22	10	31
couldn't	–	3	4	11
didn't	–	19	13	49
doesn't	3	21	10	23
don't	5	113	39	109
hadn't	1	–	1	2
hasn't	–	10	4	5
haven't	3	3	7	15
isn't	1	7	8	30
mustn't	–	–	2	3
needn't	–	1	–	–
shouldn't	–	4	2	6
wasn't	–	2	5	25
weren't	–	1	–	5
won't	1	8	5	12
wouldn't	1	9	10	17
	15 types 178 tok.	24 types 1034 tok.	24 types 675 tok.	25 types 1320 tok.

Appendix 3.29 BrE and AmE spellings [KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
colour	263	272	221	2	5
color	4	–	1	286	281
favour	156	176	217	2	3
favor	4	2	2	214	234
centre	319	204	246	2	7
center	17	1	11	305	239

Appendix 3.29 (continued)

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
theatre	121	93	186	34	40
theatre	–	–	–	67	77
defence	113	124	103	1	1
defense	–	–	–	184	188
offence	70	42	55	1	5
offense	1	–	–	14	41

Appendix 3.30 BrE and AmE spellings [Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
colour	20	23	13	8	–	–
color	–	–	–	–	16	7
favour	17	14	16	38	–	–
favor	–	–	–	1	22	20
centre	49	59	24	26	–	–
center	–	1	–	–	31	21
theatre	20	20	22	35	3	7
theatre	–	–	–	2	18	11
defence	25	15	26	18	–	–
defense	–	1	–	–	20	23
offence	5	5	2	3	–	–
offense	–	–	–	–	7	11

Appendix 3.31 Consonant doubling in IndE, BrE and AmE [KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
travell*	44	46	72	8	3
travel*	–	–	–	42	54
cancell*	13	12	10	2	–
cancel*	–	–	–	7	16
modell*	5	8	13	–	–
model*	–	–	–	4	17
channell*	3	3	1	–	2
channel*	2	–	1	3	7

Appendix 3.32 Spelling variants: <practice>/<practise>, <licence>/<license>
[KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
practice (n)	137	133	217	82	168
practise (n)	1	4	1	–	–
practice (v)	3	8	1	39	42
practise (v)	51	35	30	3	1
licence (n)	21	23	66	–	–
license (n)	1	1	4	41	39
licence (v)	–	–	2	–	–
license (v)	13	10	46	17	22

Appendix 3.33 Spelling variants: <practice>/<practise>, <licence>/<license>
[Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
practice (n)	9	16	10	10	11	13
practise (n)	–	–	–	–	–	–
practice (v)	–	–	–	–	6	2
practise (v)	4	2	1	2	–	–
licence (n)	7	2	1	3	–	–
license (n)	–	1	–	–	3	2
licence (v)	–	–	–	–	–	–
license (v)	1	–	–	2	2	1

Appendix 3.34 Spelling variants <iz>/ <is> in IndE [Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND200 Ess
<is>	179	183	12
<iz>	22	7	7

Appendix 3.35 Spelling variants <iz>/ <is> in BrE and AmE [Primary Corpus].

	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
<is>	78	115	–	–
<iz>	13	–	121	125

Appendix 3.36 <organis*>/<organiz*>, <recognis*>/<recogniz*>

[KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
organis*	367	141	160	1	1
organiz*	120	72	56	281	315
recognis*	96	104	135	1	2
recogniz*	41	70	51	169	190

Appendix 3.37 AmE spellings in academic and scientific writing [KOL, LOB and FLOB].

CATEGORY J	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991
organis*	29	7	22
organiz*	75	16	30
recognis*	13	11	18
recogniz*	14	29	18

Appendix 3.38 Lexical Americanisms and Briticisms

[KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
railway	129	68	61	13	5
railroad	3	1	2	79	46
working day	4	2	5	1	–
work(-)day, workday	1	–	–	1	3
public transport	–	2	6	1	–
public transportation	–	–	–	1	2
pavement	16	18	19	13	10
sidewalk	2	2	2	26	19
footpath	1	4	5	1	1

Appendix 3.39 Lexical Americanisms and Briticism

[Internet: newspapers; December 28, 2003].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
railway	3380	3660	428	1270	7040	724	2950	242
railroad	21	12	6	15	40	2	164	677
working day	90	48	17	106	189	14	253	37
work(-)day	11	10	3	15	20	1	26	49
workday	5	9	2	6	3	–	6	118
public transport	102	172	46	328	284	32	1060	16
public transportation	7	13	5	15	19	2	7	152
pavement	96	155	35	297	284	57	642	191
sidewalk	19	23	5	31	23	4	111	318
footpath	47	39	35	64	242	14	417	14

Appendix 3.40 *Film* vs. *movie* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
film	333	252	182	135	207
movie	34	7	35	62	120

Appendix 3.41 *Film* vs. *movie* [Internet: newspapers; December 30, 2003].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
film star	165	179	46	145	399	17	258	23
movie star	32	32	8	45	54	1	148	152

Appendix 3.42 *Garbage* vs. *rubbish* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
rubbish	–	17	13	4	4
garbage	12	2	4	6	29

Appendix 3.43 *Lorry/truck, flat/apartment and holidays/vacation*

[KOL, BROWN, FROWN, LOB, FLOB].

	IndE Kolhapur 1978	BrE LOB 1961	BrE FLOB 1991	AmE Brown 1961	AmE Frown 1992
lorry	8	12	22	–	–
truck	19	13	20	90	25
flats	5	30	12	–	–
apartments	3	4	5	17	10
summer holiday(s)	2	5	4	–	–
summer vacation(s)	1	–	–	2	1
holiday	31	103	106	30	21
vacation	7	3	3	55	23

Appendix 3.44 *Lorry/truck, flat/apartment and holidays/vacation*

[Internet: newspapers; December 29, 2003].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
lorry	86	241	216	60	133	42	1400	5
truck	1300	453	214	1810	3970	161	1610	1380
lorryload	–	24	–	–	–	–	36	–
truckload	5	142	30	10	225	43	271	365
lorry-load		102	15	–	1	1	33	1
truck-load	1	32	21	–	77	7	18	73
two-bedroom flat	3	41	539	3	22	9	78	11
two-bedroom apartment	5	35	67	–	4	10	35	647
summer holiday(s)	42	27	14	40	66	7	361	11
summer vacation(s)	87	38	22	83	271	13	18	76
on holidays	37	24	21	29	75	13	48	11
on vacation	24	30	10	33	54	6	48	227

Appendix 4**Appendix 4.1** Verbs combining with *up*.

bear (Nihalani et al. 1979:31), build (Nihalani et al. 1979:39), check (Nihalani et al. 1979:47), close (Nihalani et al. 1979:50), come (Nihalani et al. 1979:51), connect (Nihalani et al. 1979:55), cope with (Nihalani et al. 1979:57), cover (Nihalani et al. 1979:59), cram (Nihalani et al. 1979:60), end (Nihalani et al. 1979:73), fill (Nihalani et al. 1979:80), follow (Whitworth 1982 [1907]:40), give (Whitworth 1982 [1907]:40; Nihalani et al. 1979:89), heal (Nihalani et al. 1979:95), hold (Whitworth 1982 [1907]:40), jam (Nihalani et al. 1979:105), light (Whitworth 1982 [1907]:40; Nihalani et al. 1979:114), look for (Nihalani et al. 1979:116), make (Whitworth 1982 [1907]:40), pay (Nihalani et al. 1979:137), pick (Whitworth 1982 [1907]:40; Nihalani et al. 1979:138–9), put (Whitworth 1982 [1907]:40; Nihalani et al. 1979:145), raise

(Nihalani et al. 1979:146), read (Nihalani et al. 1979:147), ring (Nihalani et al. 1979:152), rise (Nihalani et al. 1979:153), rouse (Nihalani et al. 1979:153), stand (Whitworth 1982 [1907]:59), stick (Whitworth 1982 [1907]:40), take (Whitworth 1982 [1907]:40; Nihalani et al. 1979:173–4), tear (Whitworth 1982 [1907]:40), use (Whitworth 1982 [1907]:40; Nihalani et al. 1979:187).

Appendix 4.2 Verbs combining with *down*.

bow (Whitworth 1982 [1907]:40; Nihalani et al. 1979:37), break (Whitworth 1982 [1907]:40), bring (Whitworth 1982 [1907]:40; Nihalani et al. 1979:38), (Nihalani et al. 1979:71), fall (Whitworth 1982 [1907]:40, 59; Nihalani et al. 1979:78), get (Nihalani et al. 1979:88; Das 1982:144), list (Nihalani et al. 1979:115), note (Nihalani et al. 1979:128), settle (Nihalani et al. 1979:159), shoot (Nihalani et al. 1979:160), slash (Nihalani et al. 1979:163), trod (Whitworth 1982 [1907]:40).

Appendix 4.3 Verbs combining with *out*.

air (Nihalani et al. 1979:19), borrow (Nihalani et al. 1979:37), carve (Whitworth 1982 [1907]:59), eat (Nihalani et al. 1979:135), find (Whitworth 1982 [1907]:40; Nihalani et al. 1979:80), go (Whitworth 1982 [1907]:40; Nihalani et al. 1979:90, 130, 143) , lengthen (Nihalani et al. 1979:113), mark (Nihalani et al. 1979:119), moot (Nihalani et al. 1979:124–5), opt for (Nihalani et al. 1979:133), point (Whitworth 1982 [1907]:40), pull (Whitworth 1982 [1907]:40; Nihalani et al. 1979:144), solve (Nihalani et al. 1979:164), spot (Whitworth 1982 [1907]:40; Nihalani et al. 1979:166), trace (Nihalani et al. 1979:183), transfer (Nihalani et al. 1979:183), wear (Whitworth 1982 [1907]:40).

Appendix 4.4 Verbs combining with *away*.

explain (Nihalani et al. 1979:77), fly (Nihalani et al. 1979:82) , knock (Nihalani et al. 1979:108), marry (Whitworth 1982 [1907]:40; Nihalani et al. 1979:120), sell (Nihalani et al. 1979:158), shirk (Nihalani et al. 1979:160), spend (Nihalani et al. 1979:27; 165).

Appendix 4.5 Unrecorded particle verbs [Internet: newspapers; June 10, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
depict out	–	–	–	–	–	–	–	–
throw out the baby with	2	1	–	–	2	–	9	3
throw away the baby with	–	–	–	–	–	–	–	–
speak his mind out	1	–	–	–	2	–	–	–
take a decision	728	633	193	131	3910	334	75	11
take up a decision	1	–	–	–	–	–	–	–
corruption grows	1	–	–	–	–	–	1	–
corruption grows up	–	–	–	–	–	–	–	–
lifted the ban	27	27	12	7	110	12	47	7
lifted up the ban	–	–	–	–	–	–	–	–
fulfill up	–	–	–	–	–	–	–	–
heads will roll	7	3	1	3	11	1	15	8
heads will roll down	–	–	–	–	–	–	–	–

Appendix 4.6 Claimed Indianisms: particle verbs.

to admit (in) (Nihalani et al. 1979: 100), to answer (to) (Nihalani et al. 1979: 22), to apply (to/for) (Nihalani et al. 1979: 22), to appoint (as) (Nihalani et al. 1979: 22), to approach (to) (Nihalani et al. 1979: 23, Yadurajan 2001: 44), to aspire (for) (Whitworth 1982 [1907]: 60), to associate (to) (Nihalani et al. 1979: 25), to assume (on) (Nihalani et al. 1979: 25), to attend (on) (Nihalani et al. 1979: 26, Yadurajan 2001: 18), to avail (with) (Yadurajan 2001: 18), to award (with) (Nihalani et al. 1979: 27; Yadurajan 2001: 44), to be baffled (with) (Nihalani et al. 1979: 29), to be charged (for) (Nihalani et al. 1979: 46), to be concerned (with; about) (Nihalani et al. 1979: 54), to be gratifying (for) (Whitworth 1982 [1907]: 64), to be provided (to) (Whitworth 1982 [1907]: 62), to be remanded (to) (Nihalani et al. 1979: 150), to be vested (with) (Whitworth 1982 [1907]: 62), to begin (from) (Whitworth 1982 [1907]: 59; Nihalani et al. 1979: 33), to benefit (to) (Nihalani et al. 1979: 33, Yadurajan 2001: 44), to bestow (to) (Whitworth 1982 [1907]: 76), to blame (with) (Whitworth 1982 [1907]: 78), to call (as) (Nihalani et al. 1979: 24), to combat (against) (Whitworth 1982 [1907]: 61; Nihalani et al. 1979: 19, 51; Yadurajan 2001: 44), to commence (from) (Nihalani et al. 1979: 52), to complain (against) (Whitworth 1982 [1907]: 60), to comprise (of) (Nihalani et al. 1979: 53; Yadurajan 2001: 40), to concede (to) (Nihalani et al. 1979: 54), to conceive (about) (Nihalani et al. 1979: 54), to confess (about) (Nihalani et al. 1979: 54), to confide (in) (Whitworth 1982 [1907]: 61), to confirm (of) (Nihalani et al. 1979: 54), to congratulate (for) (Whitworth 1982 [1907]: 60; Nihalani et al. 1979: 54; Yadurajan 2001: 42), to consent (with) (Whitworth 1982 [1907]: 62), to consist (of) (Yadurajan 2001: 44), to contain (of) (Nihalani et al. 1979: 55), to contest (for) (Nihalani et al. 1979: 56; Yadurajan 2001: 43–4), to continue (on) (Nihalani et al. 1979: 56), to contrast (to/with) (Whitworth 1982 [1907]: 62), to contribute (in) (Whitworth 1982 [1907]: 68), to correspond (with/to) (Nihalani et al. 1979: 57), to criticize (on) (Nihalani et al. 1979: 60), to cure (from) (Nihalani et al. 1979: 61), to cut (off) (Whitworth 1982 [1907]: 40), to demand (about/for) (Nihalani et al. 1979: 65; Yadurajan 2001: 53), to deter (in) (Whitworth 1982 [1907]: 61), to devote (on/in/for/to) (Whitworth 1982 [1907]: 59), to disagree (from) (Nihalani et al. 1979: 67), to discontinue (with) (Nihalani et al. 1979: 67), to discuss (about/on) (Nihalani et al. 1979: 67; Yadurajan 2001: 43–4), to dispose (off) (Yadurajan 2001: 83), to elaborate (on) (Nihalani et al. 1979: 71–2), to emphasize (on) (Nihalani et al. 1979: 72, Yadurajan 2001: 44), to enter (into) (Nihalani et al. 1979: 74), to expect (of) (Whitworth 1982 [1907]: 70), to file (in) (Nihalani et al. 1979: 80), to furnish (about) (Nihalani et al. 1979: 87), to guess (of) (Whitworth 1982 [1907]: 63), to hesitate (from) (Nihalani et al. 1979: 96), to investigate (into) (Nihalani et al. 1979: 104), to invite (for) (Nihalani et al. 1979: 104; Yadurajan 2001: 44), to lack (of) (Nihalani et al. 1979: 110), to leave (behind) (Nihalani et al. 1979: 112), to mention (about/as to) (Nihalani et al. 1979: 122), to object (on) (Nihalani et al. 1979: 129), to order (for) (Nihalani et al. 1979: 133; Yadurajan 2001: 43–4), to originate (from) (Whitworth 1982 [1907]: 59), to pass (by/over) (Whitworth 1982 [1907]: 40), to pity (on) (Nihalani et al. 1979: 140), to play in/into the hands of people (Whitworth 1982 [1907]: 66), to proceed (on) (Nihalani et al. 1979: 143), to protest (Yadurajan 2001: 19), to provide (to) (Whitworth 1982 [1907]: 59), to pursue (for) (Nihalani et al. 1979: 144), to reach (to) (Yadurajan 2001: 44), to regret (for) (Nihalani et al. 1979: 149), to rename (as) (Nihalani et al. 1979: 150), to repent (for) (Nihalani et al. 1979: 150; Yadurajan 2001: 19), to reply (Yadurajan 2001: 18), to request (for) (Nihalani et al. 1979: 151; Yadurajan 2001: 44), to resemble (to) (Nihalani et al. 1979: 151), to resolve (on) (Whitworth 1982 [1907]: 40), to scorn (at) (Nihalani et al. 1979: 26, 156), to seek admission (in) (Whitworth 1982 [1907]: 68), to set in/into motion (Whitworth 1982 [1907]: 67), to stipulate (on) (Nihalani et al. 1979: 168), to stress (on) (Nihalani et al. 1979: 169; Yadurajan 2001: 44), to study (about) (Nihalani et al. 1979: 169), to suit (to) (Nihalani et al. 1979: 170), to swear (in) (Whitworth 1982 [1907]: 40), to take a fancy (for) (Whitworth 1982 [1907]: 64), to take law in/into your hands (Whitworth 1982 [1907]: 66, 73), to term (as) (Nihalani et al. 1979: 24), to translate (in) (Whitworth 1982 [1907]: 67), to trespass (into) (Nihalani et al. 1979: 183), to urge (for) (Nihalani et al. 1979: 187), to warn (from/to) (Whitworth 1982 [1907]: 59; Nihalani et al. 1979: 191)

Appendix 4.7 Variation in verb particles [Internet: newspapers; October 16, 2002].

	PakE DA	BaE IN	SrLE ST
admitted to hospital	204	35	13
admitted in hospital	4	2	–
admitted into hospital	–	1	–
admission to unversity	3	–	1
admission in university	2	–	–
admission into university	–	–	–
was translated into	15	3	3
was translated in	4	–	–
took the law into	9	–	2
took the law in	8	–	–
aspired to the	1	–	–
aspired for the	3	–	–

Appendix 4.8 Variation in verb particles [Internet: country domains; October 16, 2002].

	.in	.sg	.uk	.za	.au	.us
admitted to hospital	131	505	84300	4230	32100	572
admitted in hospital	175	–	37	37	18	7
admitted into hospital	10	36	599	34	287	7
admission to unversity	27	54	500	89	441	135
admission in university	8	1	18	–	–	–
admission into university	4	13	25	1	16	–
was translated into	171	108	12100	497	4580	1550
was translated in	27	22	483	20	160	78
took the law into	7	20	857	126	214	108
took the law in	3	1	8	15	7	15
aspired to the	77	11	741	47	365	123
aspired for the	1	4	16	3	4	4

Appendix 4.9 Verb particles with *urge* in IndE, BrE, AmE
[Internet: newspapers; October 18, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
urging for the	2	2	5	1	16	–	–	–
urging the	483	594	164	74	3500	161	596	482
for : (for)	1:242	1:297	1:33	1:74	1:219	–	–	–

Appendix 4.10 Verb particles with *urge* in PakE, BaE, SrLE

[Internet: newspapers; October 18, 2002].

	PakE DA	BaE IN	SrLE ST
urging for the	2	–	–
urging the	738	113	35
for : (for)	1:369	–	–

Appendix 4.11 Verb particles with *urge* worldwide

[Internet: country domains; October 18, 2002].

	.in	.sg	.uk	.za	.au	.us
urging for the	4	3	83	10	41	81
urging the	453	438	53400	871	21500	50500
for : (for)	1:113	1:146	1:643	1:87	1:525	1:624

Appendix 4.12 Verb particles with *discuss* in IndE, BrE, AmE

[Internet: newspapers; October 18, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
discussing about the	1	5	–	–	13	–	–	–
discussing the	371	394	96	112	873	183	817	787
about : (about)	1:371	1:79	–	–	1:67	–	–	–

Appendix 4.13 Verb particles with *discuss* in PakE, BaE, SrLE

[Internet: newspapers; October 18, 2002].

	PakE DA	BaE IN	SrLE ST
discussing about the	2	2	–
discussing the	176	76	36
about : (about)	1:88	1:38	–

Appendix 4.14 Verb particles with *discuss* worldwide

[Internet: country domains; October 18, 2002].

	.in	.sg	.uk	.za	.au	.us
discussing about the	103	56	515	39	99	94
discussing the	6550	973	443000	18200	131000	112000
about : (about)	1:64	1:17	1:860	1:467	1:1323	1 : 1191

Appendix 4.15 Verb particles with *investigate* in IndE, BrE, AmE
[Internet: newspapers; October 18, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
investigating into the	38	15	11	11	251	3	–	–
investigating the	574	285	159	152	5330	305	3580	771
into : (into)	1:15	1:19	1:14	1:14	1:21	1:102	–	–

Appendix 4.16 Verb particles with *investigate* in PakE, BaE, SrLE
[Internet: newspapers; October 18, 2002].

	PakE DA	BaE IN	SrLE ST
investigating into the	47	7	6
investigating the	693	103	42
into : (into)	1:15	1:15	1:7

Appendix 4.17 Verb particles with *investigate* worldwide
[Internet: country domains; October 18, 2002].

	.in	.sg	.uk	.za	.au	.us
investigating into the	44	80	158	7	86	33
investigating the	3610	3620	517000	27100	176000	68800
into : (into)	1:82	1:45	1:3272	1:3871	1:2047	1:2085

Appendix 4.18 Verb particles with *emphasis/ze* in IndE, BrE, AmE
[Internet: newspapers; October 18, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
emphasising on the	22	19	9	6	120	8	–	–
emphasising the	110	271	41	24	461	47	256	2
emphasizing on the	–	–	1	1	3	–	–	–
emphasizing the	5	–	2	12	14	11	4	193
on : (on)	1:5	1:14	1:4	1:5	1:4	1:7	–	–

Appendix 4.19 Verb particles with *emphasis/ze* in PakE, BaE, SrLE

[Internet: newspapers; October 18, 2002].

	PakE DA	BaE IN	SrLE ST
emphasising on the	4	7	–
emphasising the	92	27	21
emphasizing on the	11	1	1
emphasizing the	230	14	7
on : (on)	1:21	1:5	1:28

Appendix 4.20 Verb particles with *emphasis/ze* worldwide

[Internet: country domains; October 18, 2002].

	.in	.sg	.uk	.za	.au	.us
emphasising on the	44	33	249	26	31	–
emphasising the	501	422	99400	6620	39500	834
emphasizing on the	75	34	99	10	33	42
emphasizing the	739	468	26100	846	10100	30300
on : (on)	1:10	1:13	1:361	1:207	1:775	1:741

Appendix 4.21 Verb particles with *request* in IndE, BrE, AmE

[Internet: newspapers; October 18, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
requesting for the	8	5	6	–	17	2	–	–
requesting the	115	122	38	22	394	69	38	55
for : (for)	1:14	1:24	1:6	–	1:23	1:35	–	–

Appendix 4.22 Verb particles with *request* in PakE, BaE, SrLE

[Internet: newspapers; October 18, 2002].

	PakE DA	BaE IN	SrLE ST
requesting for the	16	–	1
requesting the	236	24	33
for : (for)	1:15	–	1:33

Appendix 4.23 Verb particles with *request* worldwide

[Internet: country domains; October 18, 2002].

	.in	.sg	.uk	.za	.au	.us
requesting for the	123	113	184	27	70	533
requesting the	743	415	59900	6020	41700	226000
for : (for)	1:6	1:4	1:326	1:223	1:596	1:424

Appendix 4.24 Verb particles with *stress* in IndE, BrE, AmE

[Internet: newspapers; October 18, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
stressing on the	41	37	23	12	167	24	–	–
stressing the	118	212	46	42	590	60	193	116
on : (on)	1:3	1:6	1:2	1:4	1:4	1:3	–	–

Appendix 4.25 Verb particles with *stress* in PakE, BaE, SrLE

[Internet: newspapers; October 18, 2002].

	PakE DA	BaE IN	SrLE ST
stressing on the	46	7	6
stressing the	474	101	27
on : (on)	1:10	1:14	1:5

Appendix 4.26 Verb particles with *stress* worldwide

[Internet: country domains; October 18, 2002].

	.in	.sg	.uk	.za	.au	.us
stressing on the	109	57	122	4	45	15
stressing the	622	372	52900	996	15300	9800
on : (on)	1:6	1:7	1:434	1:249	1:340	1:653

Appendix 4.27 Collocations of *contest* [Internet: country domains; February 2, 2003].

	.in	.sg	.uk	.za	.au	.us
contesting elections	210	7	471	125	254	155
contesting elections from PLACE	7	–	1	–	–	–
he contested from PLACE	12	2	–	–	–	–
is contesting from PLACE	9	–	4	–	1	1

Appendix 4.28 Collocations of *cater* [Internet: newspapers; March 9, 2002].

	IndE IE	IndE HI	IndE DH	IndE HT	IndE TR	IndE TE	BrE GU	AmE WP
caters the needs	–	–	–	–	1	1	–	–
caters for the needs	–	7	–	–	–	–	2	–
caters to the needs	3	125	36	–	116	9	1	3

Appendix 4.29 Verb particles with *protest* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
protest N	–	–	–	–	–	1
protest against	3	6	2	2	–	–
protest at	–	–	–	–	–	–
protest about	–	–	–	2	–	–
protest over	–	1	–	–	1	–

Appendix 4.30 Verb particles with *protest* [Primary Corpus; broadcast].

	IndE IND2000 Broad	BrE UK1991 Broad
protest N	5	–
protest against	6	–
protest at	–	–
protest about	–	2
protest over	–	–

Appendix 4.31 Verb particles with *protest* [KOL, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
protest N	3	–	1	3	9
protest against	12	10	6	3	2
protest at	1	1	3	–	–
protest about	–	–	5	–	–
protest over	–	–	–	–	–

Appendix 4.32 Complementation of *want* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
want (verb)	24	53	35	53	32	73
want (NP) that	–	1	–	–	–	–
want (standard)	24	52	35	53	32	73
want NP (N/adj)	4	13	12	18	11	23
want NP to-inf	5	7	3	8	3	5
want NP –ed	1	1	–	4	–	–
want NP –ing	–	–	–	–	–	1
want to-inf	14	31	20	23	18	44
wanna	–	–	–	–	–	–

Appendix 4.33 Complementation of *want* [Primary Corpus; broadcast and essays].

	IndE IND2000 Broad	IndE IND2000 Ess	BrE UK1991 Broad
want (verb)	118	58	138
want (NP) that	2	1	–
want (standard)	116	57	139
want NP (N/adj)	33	4	26
want NP to-inf	15	1	24
want NP –ed	–	1	1
want NP –ing	2	–	1
want to-inf	66	51	86
wanna	–	–	–

Appendix 4.34 Complementation of *want* [KOL].

	IndE KOL 1978
want (verb)	592
want (NP) that	7
want (standard)	585

Appendix 4.35 Variation in complementation: infinitive vs. *V-ing*

failure in (Shastri 1996:75–6), no hope doing s.th. (Whitworth 1982 [1907]:140f.), pleasure in doing s.th. (Whitworth 1982 [1907]:140f.), satisfaction of doing s.th. (Whitworth 1982 [1907]:140f.), sentiment of doing s.th. (Whitworth 1982 [1907]:140f.); admit s.o. to do s.th. (Whitworth 1982 [1907]:140f.), aim at doing s.th. (Whitworth 1982 [1907]:140f.), avoid doing s.th. (Whitworth 1982 [1907]:140f.), be averse to doing s.th. (Whitworth 1982 [1907]:140f.), be described as s.th. (Whitworth 1982 [1907]:140f.), be eligible (Shastri 1996:75–6), be found to do s.th. (Parasher 1991:242–3), be indifferent about doing s.th.

(Whitworth 1982 [1907]:140f.), be interested (Shastri 1996:75–6), be offended at s.th. (Whitworth 1982 [1907]:140f.), be prohibited from doing s.th. (Whitworth 1982 [1907]:140f.), be ready to do s.th. (Whitworth 1982 [1907]:140f.), be unanimous in s.th. (Whitworth 1982 [1907]:140f.), contemplate (Nihalani et al. 1979:55), fail (Whitworth 1982 [1907]:140f, Shastri 1996:75–6), forbid (Shastri 1996:75–6), help s.o. in doing s.th. (Parasher 1991:242–3), hesitate (Shastri 1996:75–6), know (swimming) (Nihalani et al. 1979:109), learn (driving) (Nihalani et al. 1979:112), like (Shastri 1996:75–6, Parasher (1991:242–3), look forward to (Nihalani et al. 1979:116), make s.o. do s.th. (Whitworth 1982 [1907]:140f.), miss (Nihalani et al. 1979:124), persist (Nihalani et al. 1979:138), persuade (Shastri 1996:75–6), pose as doing s.th. (Whitworth 1982 [1907]:140f.), prevent (Whitworth 1982 [1907]:140f, Shastri 1996:75–6), refuse to do s.th. (Whitworth 1982 [1907]:140f.), regard s.th. as s.th. (Whitworth 1982 [1907]:140f.), spend time –ing (Nihalani et al. 1979:165), strike s.o. as s.th. (Whitworth 1982 [1907]:140f.), succeed in doing s.th. (Whitworth 1982 [1907]:140f.), suspect (Nihalani et al. 1979:171), try to do s.th. (Whitworth 1982 [1907]:140f.), with a view to (Parasher 1991:150–1)

Appendix 4.36 Complementation of *contemplate*

[KOL, ICE-IND, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	IndE ICE-IND 1990s	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
contemplate –ing	–	1	1	1	–	–
contemplate NP/wh	9	7	16	19	13	22
contemplate to inf	–	1	–	–	–	–
contemplate that	–	–	–	1	–	–
contemplate (intr.)	2	2	–	–	1	1

Appendix 4.37 Complementation of *with a view*

[KOL, ICE-IND, BROWN, FROWN, LOB, FLOB].

	IndE KOL 1978	IndE ICE-IND 1990s	BrE LOB 1961	BrE FLOB 1991	AmE BROWN 1961	AmE FROWN 1992
with a view to + V	5	4	–	1	–	–
with a view to + V-ing	24	2	14	4	1	1

Appendix 5

Appendix 5.1 Definite and indefinite articles [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
a/an	2000	2161	2282	2187	2307	2261
the	6427	6319	5402	5123	5708	4848

Appendix 5.2 Definite and indefinite articles [Primary Corpus; broadcast, essays].

	IndE IND2000 Broad News	IndE IND2000 Broad I/D	IndE IND2000 Ess	BrE UK1991 Broad News	BrE UK1991 Broad I/D
a/an	961	1003	587	1097	1119
the	3337	2206	956	3035	2307

Appendix 5.3 Variation in article usage: *to (the) outside world*

[Internet: country domains; February 10, 2003].

	.in	.sg	.uk	.za	.au	.us
to outside world	88	38	359	45	231	113
to the outside world	584	491	177000	11200	51500	15500
(the) : the	1:7	1:13	1:493	1:249	1:223	1:137

Appendix 5.4 Variation in article usage: *since (the) early 1990s*

[Internet: country domains; February 10, 2003].

	.in	.sg	.uk	.za	.au	.us
since early 1990s	6	3	218	7	32	12
since the early 1990s	30	61	13000	153	604	269
(the) : the	1:5	1:20	1:60	1:22	1:19	1:22

Appendix 5.5 Variation in article usage: *at (the) political level*

[Internet: country domains; February 10, 2003].

	.in	.sg	.uk	.za	.au	.us
at political level	49	6	583	117	99	2
at the political level	255	121	14600	436	12300	143
(the) : the	1:5	1:20	1:25	1:4	1:124	1:72

Appendix 5.6 Variation in article usage: *live in (the) United States*

[Internet: country domains; February 10, 2003].

	.in	.sg	.uk	.za	.au	.us
live in United States	2	–	333	21	34	111
live in the United States	88	85	12200	190	474	9270
(the) : the	1:44	–	1:37	1:9	1:14	1:84

Appendix 5.7 Countability and number: list of variables

The following nouns show variation along the count/noncount and singular/plural divides according to Nihalani et al. 1979: abuse (Nihalani et al. 1979: 16), advice (Nihalani et al. 1979: 17–8), agitation (Nihalani

et al. 1979:19), aircraft (Nihalani et al. 1979:20), brutality (Nihalani et al. 1979:39), chaos (Nihalani et al. 1979:46), circumstance (Nihalani et al. 1979:49), clothes (Nihalani et al. 1979:50), condition (Nihalani et al. 1979:54), consideration (Nihalani et al. 1979:55), content (Nihalani et al. 1979:55), coordination (Nihalani et al. 1979:56), currency (Nihalani et al. 1979:61), damage (Nihalani et al. 1979:62–3), delay (Nihalani et al. 1979:64), demand (Nihalani et al. 1979:65), effect (Nihalani et al. 1979:71), employment (Nihalani et al. 1979:72), entertainment (Nihalani et al. 1979:74–5), equipment (Nihalani et al. 1979:75), evidence (Nihalani et al. 1979:75–6), experience (Nihalani et al. 1979:76), fruit (Nihalani et al. 1979:86), furniture (Nihalani et al. 1979:87), gain (Nihalani et al. 1979:88), hair (Nihalani et al. 1979:93), information (Nihalani et al. 1979:101–2), instruction (Nihalani et al. 1979:102–3), intention (Nihalani et al. 1979:103), jewellery (Nihalani et al. 1979:106), land (Nihalani et al. 1979:110–1), legislation (Nihalani et al. 1979:113), loot (Nihalani et al. 1979:117), machinery (Nihalani et al. 1979:118), mail (Nihalani et al. 1979:118), mood (Nihalani et al. 1979:124), news (Nihalani et al. 1979:126–7), noise (Nihalani et al. 1979:127), opinion (Nihalani et al. 1979:132–3), opposition (Nihalani et al. 1979:133), pants (Nihalani et al. 1979:136), personnel (Nihalani et al. 1979:138), progress (Nihalani et al. 1979:143), property (Nihalani et al. 1979:143), quality (Nihalani et al. 1979:145), quarter (Nihalani et al. 1979:146), reason (Nihalani et al. 1979:148), secrecy (Nihalani et al. 1979:158), selection (Nihalani et al. 1979:158), spectacles/specs (Nihalani et al. 1979:165), staff (Nihalani et al. 1979:166), talent (Nihalani et al. 1979:174), television (Nihalani et al. 1979:176), training (Nihalani et al. 1979:183), trouble (Nihalani et al. 1979:184), trousers (Nihalani et al. 1979:184), tuition (Nihalani et al. 1979:184–5), wood (Nihalani et al. 1979:195–6), work (Nihalani et al. 1979:196). The following additional nouns and phrases are claimed to display variation along the count/noncount and singular/plural divides by Yadurajan (2001:109–11): footwear, luggage, at one's own expense, to find fault with, to send word to, to have no option. According to Parasher (1991:167), 'bread' is commonly treated as a count noun in IndE. Nouns mentioned by Quirk et al. (1985:252) include: anger, behaviour, business, courage, harm, homework, photography, publicity, resistance, safety, traffic, violence.

Appendix 5.8 Direct and indirect speech after *say* [Primary Corpus; press].

	IndE IND1978 Press	IndE IND2000 Press	BrE UK1961 Press	BrE UK1991 Press	AmE US1961 Press	AmE US1992 Press
said (rep. verb)	91	264	146	213	234	333
said (+ dir. sp.)	18	109	24	112	85	165
said (+ ind. sp.)	73	155	122	101	149	168

Appendix 5.9 Tense with *after* [Primary Corpus; broadcast].

	IndE IND2000 Broad	IndE UK1991 Broad
after + past tense	21	16
after + past perfect	2	5

Appendix 5.10 Functions of the past perfect in IndE [Primary Corpus].

	IndE IND1978 Press	IndE IND2000 Press	IndE IND2000 Broad News	IndE IND2000 Broad I/D	IndE IND2000 Ess
total	206	176	73	35	16
past before past (standard)	181	150	63	14	–
cond./hypo. (standard)	13	8	–	3	2
nonstandard uses	12	18	7	18	14

Appendix 5.11 Inversion in direct questions [Primary Corpus; press (1978)].

	#questions	#inversion	#no inversion
how	9	9	–
why	14	14	–
what	16	16	–
where	2	2	–
when	1	1	–
who(m)/whose	2	2	–
which	–	–	–
	44 (100%)	44 (100%)	–

Appendix 5.12 Inversion in direct questions [Primary Corpus; press (2000)].

	#questions	#inversion	#no inversion
how	19	18	1
why	8	5	3
what	9	9	–
where	2	2	–
when	–	–	–
who(m)/whose	–	–	–
which	1	1	–
	39 (100%)	35 (88.6%)	4 (11.4%)

Appendix 5.13 Inversion in direct questions [Primary Corpus; broadcast].

	#questions	#inversion	#no inversion
how	41	39	2
why	34	31	3
what	26	23	3
where	11	11	–
when	1	1	–
who(m)/whose	1	1	–
which	–	–	–
	114	106	8
	(100%)	(92.4%)	(7.8%)

Appendix 5.14 Inversion in direct questions [Primary Corpus; essays].

	#questions	#inversion	#no inversion
	5	2	3
	(100%)	(40%)	(60%)

Appendix 5.15 Inversion in embedded questions [Primary Corpus; press (1978)].

	#questions	#no inversion	# inversion
how	13	13	–
why	2	2	–
what	7	6	1
where	–	–	–
when	–	–	–
who(m)/whose	–	–	–
which	–	–	–
	22	21	1
	(100%)	(95.5%)	(4.5%)

Appendix 5.16 Inversion in embedded questions [Primary Corpus; press (2000)].

	#questions	#no inversion	# inversion
how	9	9	–
why	4	3	1
what	2	2	–
where	3	3	–
when	–	–	–
who(m)/whose	1	–	1
which	1	1	–
	20	18	2
	(100%)	(90.0%)	(10.0%)

Appendix 5.17 Inversion in embedded questions [Primary Corpus; broadcast].

	#questions	#no inversion	# inversion
how	15	15	–
why	35	26	9
what	20	19	1
where	3	3	–
when	1	1	–
who(m)/whose	1	1	–
which	–	–	–
	75	65	10
	(100%)	(86.7%)	(13.3%)

Appendix 5.18 Inversion in embedded questions [Primary Corpus; essays].

	#questions	#no inversion	# inversion
	40	39	1
	(100%)	(97.5%)	(2.5%)

Index

A

- accommodation 9
- Agnihotri, Rama Kant 23, 29
- American English (AmE) 29,
 - 36–41, 86, 87, 96–102,
 - 109–112, 116–147, 155–160, 164,
 - 165, 169, 170, 174, 178, 180,
 - 181, 184, 186–189, 194–201,
 - 215–227, 233–237, 240–246,
 - 249–257, 260–264, 275, 276,
 - 279, 283–289, 301, 306–309,
 - 314–315
- Anand, Mulk Raj 18
- Angloversals 49
- Arabic 51, 62, 113
- archaism 80, 116, 235
- Australian English (AusE) 118,
 - 123, 134, 186, 187, 189, 223, 235,
 - 236, 240, 256, 284–289

B

- backshift 257, 261–263, 274,
 - 308, 315
- Bangladeshi English (BaE) 64–
 - 68, 70, 85, 88, 89, 93–95, 104,
 - 111, 121, 145, 156, 158, 160, 178
- Bengali 22, 23, 64
- Bhatt, Rakesh M. 291, 300–304
- bilingual/-ism 13, 22
- borrowing 10, 36, 37, 51–71, 76,
 - 78, 113, 114, 145, 311, 312
- British English (BrE) 29,
 - 32–41, 86, 87, 96–112, 116–147,
 - 155–160, 164, 165, 170,
 - 174–189, 194–196, 199–201,
 - 215–228, 233, 235, 236,
 - 240–257, 260–264, 274–276,
 - 279, 283–289, 301, 306–308,
 - 314–315

C

- Carls, Uwe 77, 96–103

- Channel V 58
- cline of proficiency 26–27,
 - 33, 43
- Clive, Robert 9–10
- cluster simplification
 - lots of/lot of* 213, 243
 - at all costs/at all cost* 242–
 - 243
- colonial lag 140, 289
- creative writing 12, 16, 18, 24

D

- differentiation → *Evolutionary Model of English*
- discourse marker 144
- divergent diachronic
 - development 50, 121, 131,
 - 145–147, 160, 196, 234, 243,
 - 257, 308, 314
- Diwani Act 10

E

- Early Modern English 158
- East India Company 8–11, 14
- endonormative stabilization → *Evolutionary Model of English*
- English-medium 1, 9, 15, 20,
 - 24, 43, 201, 287
- error → *learner difficulty*
- Evolutionary Model of English (Schneider) 9, 11–15, 19, 25
- exonormative stabilization → *Evolutionary Model of English*
- Expanding Circle → *Three Circles Model of English*

F

- formal English 16, 43, 55, 56,
 - 60, 71, 89, 96, 97, 114–118, 121,
 - 124, 145, 172, 176, 196, 287,
 - 288, 312

- forms of address
 - Sri* 60
 - sahib* 60–61
 - ji* 60
- Foundation Phase → *Evolutionary Model of English*
- fronting 291
- Gandhi, M. K. 17–18, 65
- grammaticalization 243, 248,
 - 254
- Great Rebellion 14–15

H

- Hastings, Warren 10
- Hindi 1, 17–23, 47, 48, 51,
 - 57–59, 62–65, 68, 76, 82, 145,
 - 197, 203, 209, 227, 258, 259,
 - 267, 312
- Hindustani 18
- Hinglish 16, 57–59, 81–82
- Hobson-Jobson* (Yule/Burnell)
 - 16, 62, 65
- Hundt, Marianne 109–110,
 - 245–255, 284–288

I

- Independence 1, 3, 8, 17–18,
 - 20–22, 26, 38
- independent innovation 3, 24,
 - 70, 75, 77, 84, 121, 149, 152, 157,
 - 159, 176, 178, 195, 216, 224,
 - 235–238, 242, 243, 274, 300,
 - 311, 312, 315
- India Act 10
- Indian Constitution 18
- informal English 16, 44, 56–58,
 - 61, 78, 89, 112, 113, 115, 116, 121,
 - 122, 124, 145, 147, 176, 178, 181,
 - 190, 275–277, 288, 291, 300,
 - 303, 304, 312, 315

Inner Circle → *Three Circles Model of English*

J

Jamaican English 204, 208, 215, 218, 222, 223

K

Kachru, Braj B. 7, 26–30, 49, 61, 67, 198, 209, 259, 268

Kannada 19, 22, 23

Kashmiri 58

Kortmann, Bernd 27, 290

Krishnaswamy/Burde 2, 10–13, 16–17, 22

L

L1 influence → *substrate influence*

language mixing

code-mixing 24, 36, 57–59, 82

code-switching 59

learner difficulty 49, 126, 152, 153, 162, 182, 191, 195, 198, 202, 213, 225–227, 259, 260, 266, 272, 280, 296, 312, 314

Leitner, Gerhard 24, 39–40, 132, 204, 259, 264, 270, 315

lingua franca 1, 64

loanword → *borrowing*

M

Macaulay, Thomas Babington 13–14

Mair, Christian 40–41, 49

Marathi 301

Mehrotra, Raja Ram 21, 55

Minute on Indian Education 13–14

MTV 58

Mukherjee, Joybrato 14, 19, 24–25, 37–38, 315

multilingual/-ism 7, 21, 22, 25

N

Narayan, R. K. 18

nativization → *Evolutionary Model of English*; 35–37, 40,

46, 47, 49, 108, 166, 315

Nehru, Jawaharlal 18

Netspeak 145

New Zealand English (NZE)

39, 110, 137, 245, 248–253

Nihalani et al. 32–34

O

official language 1, 8, 13, 18–20

Official Languages Act 19

orthography 58, 62, 78, 79,

81, 88, 94, 113, 114, 132–135, 146, 315

Outer Circle → *Three Circles Model of English*

P

Pakistani English (PakE) 64–

68, 70, 85, 88, 89, 91, 93–95, 104, 111, 121, 145, 155, 158,

160, 171

parallel corpora 35, 40–42, 44

Parasher, S. V. 22–23

Persian 51, 113

politeness 60

prescriptive 33, 34, 73, 108,

171–172, 186, 224, 253, 256, 274, 287, 289, 313

Punjabi 58

Q

Queen Elizabeth I 8

R

Rao, Raja 18

Regulating Act 10

reported speech → *backshift*

Roy, Raja Rammohan 13

S

Sahgal, Anju 22–23, 29

Sand, Andrea 39, 199–201, 218–219

Sanskrit 19, 43, 51, 115

Schmied, Josef 28–29, 38, 311

Schneider, Edgar W. 9–15, 19, 25, 37, 39

Sharma, Devyani 39, 198, 205–206, 258

Shastri, S. V. 8, 35–41

Singaporean English (SgE) 87, 104, 123, 156, 167, 186, 187, 200, 218, 223, 235, 236, 240, 242, 256, 285, 287, 306

South African English (SAfrE)

87, 123, 140, 146, 176, 186, 189, 196, 223, 235, 236, 240, 256, 286, 287, 289, 306, 313

spelling → *orthography*

Sri Lankan English (SrLE) 64,

67, 68, 70, 85, 86, 89, 94, 95, 104, 111, 119, 121, 145, 156, 160, 178

standard English 3, 4, 27,

30–34, 81–87, 93, 96, 104, 107–110, 116, 121, 124, 127–133, 147, 163–176, 180, 193, 195, 197, 203–219, 222–227, 229–233, 238–243, 257, 258, 263, 266–268, 270–278, 289–293, 297–304, 308, 309, 311–313, 315

Stephens, Thomas 8

substrate influence 28, 47, 115, 169, 171, 172, 197, 199, 202, 213, 227, 264, 274, 290, 301, 313

T

tag questions 28–30, 35

Tamil 22, 23, 48, 59, 64, 197, 227

Telugu 19, 171

Three Circles Model of English

(Kachru) 7, 29, 39, 49

Three Language Formula 1, 20

topicalization → *fronting*

Trudgill/Hannah 30–31

U

Urdu 18, 19, 48, 51, 61, 64, 67, 258

V

vernacular English → *informal English*

W

word formation

Latinate prefixes 95,
99–100, 147

nominal compounding 67,
68, 70, 77, 78, 80–84, 95,
96, 100–104, 139–142,
145–147, 311, 312

adjectival compounding 93,
95, 97–99, 147

affixation 60

shortening 87, 93

functional shift 93–94

world Englishes 3, 7, 9, 24, 25,
29, 31, 39, 44, 46, 49, 84, 86,
109, 110, 124, 131, 132, 185, 188,

194, 197, 198, 216, 224, 234,
243, 248, 249, 274, 283, 284,
308, 313

Y

Yadurajan, K. S. 32–34

youth slang 145

In the series *Varieties of English Around the World* the following titles have been published thus far or are scheduled for publication:

- T9 **MÜHLHÄUSLER, Peter, Thomas E. DUTTON and Suzanne ROMAINE:** Tok Pisin Texts. From the beginning to the present. 2003. x, 286 pp.
- T8 **McCLURE, J. Derrick:** Doric. The dialect of North-East Scotland. 2002. vi, 222 pp.
- T7 **MEHROTRA, Raja Ram:** Indian English. Texts and Interpretation. 1998. x, 148 pp.
- T6 **WINER, Lise:** Trinidad and Tobago. 1993. xii, 368 pp.
- T5 **WAKELIN, Martyn F.:** The Southwest of England. 1986. xii, 231 pp.
- T4 **PLATT, John, Heidi WEBER and Mian Lian HO:** Singapore and Malaysia. 1983. iv, 138 pp.
- T3 **MACAFEE, Caroline:** Glasgow. 1983. v, 167 pp.
- T2 **HOLM, John:** Central American English. (Julius Groos) Heidelberg, 1982. iv, 184 pp., + tape.
- T1 **TODD, Loreto:** Cameroon. (Julius Groos) Heidelberg, 1982. 180 pp., 1 map.
- G39 **PETERS, Pam, Peter C. COLLINS and Adam SMITH (eds.):** Comparative Studies in Australian and New Zealand English. Grammar and beyond. x, 397 pp. + index. *Expected June 2009*
- G38 **SEDLATSCHEK, Andreas:** Contemporary Indian English. Variation and change. 2009. xix, 363 pp.
- G37 **SCHREIER, Daniel:** St Helenian English. Origins, evolution and variation. 2008. xv, 312 pp.
- G36 **MURRAY, Thomas E. and Beth Lee SIMON (eds.):** Language Variation and Change in the American Midland. A New Look at 'Heartland' English. 2006. xii, 320 pp.
- G35 **HICKEY, Raymond:** Dublin English. Evolution and change. 2005. x, 270 pp. (incl. CD-Rom).
- G34 **MÜHLEISEN, Susanne and Bettina MIGGE (eds.):** Politeness and Face in Caribbean Creoles. 2005. viii, 293 pp.
- G33 **LIM, Lisa (ed.):** Singapore English. A grammatical description. 2004. xiv, 174 pp.
- G32 **HACKERT, Stephanie:** Urban Bahamian Creole. System and variation. 2004. xiv, 256 pp.
- G31 **THOMPSON, Roger M.:** Filipino English and Taglish. Language switching from multiple perspectives. 2003. xiv, 288 pp.
- G30 **ACETO, Michael and Jeffrey P. WILLIAMS (eds.):** Contact Englishes of the Eastern Caribbean. 2003. xx, 322 pp.
- G29 **NELSON, Gerald, Sean WALLIS and Bas AARTS:** Exploring Natural Language. Working with the British Component of the International Corpus of English. 2002. xviii, 344 pp.
- G28 **GÖRLACH, Manfred:** Still More Englishes. 2002. xiv, 240 pp.
- G27 **LANEHART, Sonja L. (ed.):** Sociocultural and Historical Contexts of African American English. 2001. xviii, 373 pp.
- G26 **BLAIR, David and Peter C. COLLINS (eds.):** English in Australia. 2001. vi, 368 pp.
- G25 **BELL, Allan and Koenraad KUIPER (eds.):** New Zealand English. JB/Victoria UP, 2000. 368 pp.
- G24 **HUBER, Magnus:** Ghanaian Pidgin English in its West African Context. A sociohistorical and structural analysis. 1999. xviii, 322 pp. (incl. CD-rom).
- G23 **HUNDT, Marianne:** New Zealand English Grammar – Fact or Fiction? A corpus-based study in morphosyntactic variation. 1998. xvi, 212 pp.
- G22 **GÖRLACH, Manfred:** Even More Englishes. Studies 1996–1997. With a foreword by John Spencer. 1998. x, 260 pp.
- G21 **KALLEN, Jeffrey L. (ed.):** Focus on Ireland. 1997. xviii, 260 pp.
- G20 **MACAULAY, Ronald K.S.:** Standards and Variation in Urban Speech. Examples from Lowland Scots. 1997. x, 201.
- G19 **SCHNEIDER, Edgar W. (ed.):** Englishes around the World. Studies in honour of Manfred Görlach. Volume 2: Caribbean, Africa, Asia, Australasia. 1997. viii, 358 pp.
- G18 **SCHNEIDER, Edgar W. (ed.):** Englishes around the World. Studies in honour of Manfred Görlach. Volume 1: General studies, British Isles, North America. 1997. vi, 329 pp.
- G17 **PATRICK, Peter L.:** Urban Jamaican Creole. Variation in the Mesolect. 1999. xx, 329 pp.
- G16 **SCHNEIDER, Edgar W. (ed.):** Focus on the USA. 1996. vi, 368 pp.
- G15 **DE KLERK, Vivian (ed.):** Focus on South Africa. 1996. iv, 328 pp.
- G14 **McCLURE, J. Derrick:** Scots and its Literature. 1996. vi, 218 pp.
- G13 **GÖRLACH, Manfred:** More Englishes. New studies in varieties of English 1988–1994. 1995. 276 pp.
- G12 **GLAUSER, Beat, Edgar W. SCHNEIDER and Manfred GÖRLACH:** A New Bibliography of Writings on Varieties of English, 1984–1992/93. 1993. 208 pp.
- G11 **CLARKE, Sandra (ed.):** Focus on Canada. 1993. xii, 302 pp.