



# *Register Variation in Indian English*

Chandrika Balasubramanian

**Studies in Corpus Linguistics**  
**37**

**JOHN BENJAMINS PUBLISHING COMPANY**



30809112

# Register Variation in Indian English

Chandrika Balasubramanian  
Western Carolina University



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.

Cover design: Françoise Berserik  
Cover illustration from original painting *Random Order*  
by Lorenzo Pezzatini, Florence, 1996.

#### Library of Congress Cataloging-in-Publication Data

Balasubramanian, Chandrika.

Register variation in Indian English / by Chandrika Balasubramanian.

p. cm. (Studies in Corpus Linguistics, ISSN 1388-0373 ; v. 37)

Includes bibliographical references and index.

1. English language--Variation--India. 2. English language--India--Usage. 3. Languages in contact--India. 4. Bilingualism--India. 5. India--Languages. I. Title.

PE3502.I6B25 2009

427'.954--dc22

2009023005

ISBN 978 90 272 2311 1 (HB; alk. paper)

ISBN 978 90 272 8903 2 (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

List of figures	XV
List of tables	XVII
CHAPTER 1	
Introduction	1
1.0 English as a global language: Emergence of new varieties	1
1.1 Understanding New Englishes	2
1.2 International Corpus of English: Role in Dialect Research	5
1.3 Development of Corpus of Contemporary Indian English (CCIE): Outcomes of Project	5
1.3.1 Combining CCIE with ICE-India	6
1.4 English in India: Context for the book	7
1.5 Outline of the book	11
CHAPTER 2	
Literature review	13
2.0 Introduction	13
2.1 Methodologies employed in dialect studies	14
2.1.1 Regional and Social Dialectology	14
2.1.1.1 Regional Dialectology	14
2.1.1.2 Strengths and weaknesses of regional dialectology	15
2.1.2 Sociolinguistic studies	16
2.1.2.1 Strengths of sociolinguistic dialect studies	17
2.1.3 Similarities and differences between regional dialectology and social dialectology.	17
2.1.4 Corpus linguistics and dialectology: the study of register	19
2.1.4.1 Corpus linguistics and representativeness	21
2.1.4.2 Strengths and weaknesses of corpus linguistics as a methodology to study dialect	22
2.2 Variety studies today	23
2.2.1 Illustrative/Impressionistic studies	24
2.2.2 Empirical/Data-based studies	24

- 2.2.3 Corpus-based studies 25
- 2.3 Treatment of varieties in variety studies:
  - Representation of variety being examined and Implications 29
- 2.4 Situating the study of Indian English 30
  - 2.4.1 Previous literature on linguistic features examined in current study 31
- 2.5 Conclusion 38

## CHAPTER 3

**Corpus design and methodology**

39

- 3.0 Introduction and overview of methodology 39
- 3.1 Construction of the Corpus of Contemporary Indian English 41
  - 3.1.1 Written corpus 45
    - 3.1.1.1 Major sources of the written corpus 45
    - 3.1.1.2 Registers and sub-registers of the written corpus 48
      - 3.1.1.2.1 Register: Written news 48
      - 3.1.1.2.2 Register: Fiction 49
      - 3.1.1.2.3 Register: Written entertainment 50
      - 3.1.1.2.4 Register: Business correspondence 50
      - 3.1.1.2.5 Register: Personal correspondence 51
      - 3.1.1.2.6 Register: Written sports 51
      - 3.1.1.2.7 Register: Written travel news 51
    - 3.1.1.3 Summary of Written registers and their sources 51
  - 3.1.2 Spoken corpus 52
    - 3.1.2.1 Register: Spoken news 53
    - 3.1.2.2 Register: Spoken Academic English 53
    - 3.1.2.3 Register: Conversational English 55
    - 3.1.2.4 Register: Spoken sports 56
- 3.2 Methodology 57
  - 3.2.1 Initial processing of written texts 57
    - 3.2.1.1 Saving the texts 57
    - 3.2.1.2 Naming the texts 57
  - 3.2.2 Initial processing of spoken texts 59
    - 3.2.2.1 Recording spoken data 59
    - 3.2.2.2 Transcription 60
    - 3.2.2.3 Naming spoken files 60
  - 3.2.3 Combining CCIE with ICE-India 61
  - 3.2.4 Computer programs used for the initial analysis of both spoken and written files 64
  - 3.2.5 Initial processing of all corpus files 64

3.2.6	Analysis of the files: Computer programs for different kinds of analyses	65
3.2.7	Kinds of analyses	66
3.2.7.1	Type I analysis: Identifying Indian and non-Indian variants	66
3.2.7.2	Type II analysis: General descriptions of patterns of occurrences	66
3.2.8	Linguistic features and methodology employed in their initial analysis	67
3.2.8.1	Features examined in study	67
3.2.8.2	Rationale for choice of and description of linguistic features	68
3.2.8.3	Lexical features	69
3.2.8.3.1	Keep used meaning Put	69
3.2.8.3.2	Stay used meaning live	70
3.2.8.3.3	Indian words	71
3.2.8.4	Grammatical features	74
3.2.8.4.1	Stative verbs in the progressive	74
3.2.8.4.2	Patterns of occurrence of the present and past perfect	75
3.2.8.4.3	Prepositional verbs and preposition use	76
3.2.8.4.4	Articles	76
3.2.8.4.5	Use of future instead of simple present	78
3.2.8.4.6	Use of the progressive instead of simple or perfect aspect verbs	78
3.2.8.4.7	Absence of subject-auxiliary inversion in WH-question formation	79
3.2.8.4.8	Transitive verbs used intransitively	80
3.2.8.4.9	Relative clauses	81
3.2.8.4.10	Patterns of occurrence of modals	81
3.2.8.4.11	Initial and non-initial existential There	81
3.2.8.4.12	Invariant tag Isn't it? and No?	82
3.2.8.4.13	Also and Too	83
3.3	Conclusion	84

#### CHAPTER 4

### Lexical and grammatical features in spoken and written Indian English 85

4.0	Introduction	85
4.1	Analysis of lexical features	86
4.1.1	Keep meaning Put: Type I analysis	86

- 4.1.2 Stay meaning Live: Type I analysis 87
- 4.1.3 Indian words: Type II Analysis 89
- 4.2 Grammatical features 89
  - 4.2.1. Stative verbs in the progressive: Type I Analysis 90
  - 4.2.2 Past and present perfect: Type I Analysis 91
    - 4.2.2.1 Past Perfect 91
    - 4.2.2.2 Present Perfect 92
  - 4.2.3 Prepositional verbs and prepositions: Type II Analysis 93
    - 4.2.3.1 Productivity of particular verbs and prepositions 93
    - 4.2.3.2 Prepositional verbs in British and American English versus Indian English 98
    - 4.2.3.3 Most common prepositional verbs 99
  - 4.2.4 Article use: Type I Analysis 100
    - 4.2.4.1 Articles before ordinal numbers 100
    - 4.2.4.2 Articles before noun phrases (spoken corpus only) 101
    - 4.2.4.3 Articles before lot of 101
    - 4.2.4.4 Articles before little 102
    - 4.2.4.5 Articles before few 102
    - 4.2.4.6 Articles before number of 103
    - 4.2.4.7 Summary of article analysis 103
  - 4.2.5 Future instead of simple present: Type I Analysis 104
  - 4.2.6 Progressives of non stative verbs: Type II Analysis 105
  - 4.2.7 Absence of subject-auxiliary inversion in WH-questions: Type I Analysis 106
  - 4.2.8 Transitive verbs used intransitively: Type I Analysis 106
  - 4.2.9 Relative constructions: Type II Analysis 108
    - 4.2.9.1 Common relativizers 108
      - 4.2.9.1.1 *Who* versus *which* and *that* 108
      - 4.2.9.1.2 *Which* versus *that* 109
    - 4.2.9.2 Frequency of relative clauses 109
    - 4.2.9.3 Non-standard relativizers 109
    - 4.2.9.4 Summary of Relative Clause Analysis 109
  - 4.2.10 Modals: Type II Analysis 110
    - 4.2.10.1 Observations about the distribution of modals and semi-modals in Indian English 111
  - 4.2.11 Initial and non-initial Existential *There*: Type I Analysis 111
  - 4.2.12 Invariant tag *isn't it?* and *No?*: Type I Analysis 112
    - 4.2.12.1 *Isn't it* 112
    - 4.2.12.2 *No* 113

4.2.13	Circumstance Adverbials <i>Also</i> and <i>Too</i> : Type II Analysis	114
4.2.13.1	Results on general distribution	114
4.2.13.2	Results on position	115
4.2.13.3	Results on position of <i>also</i> and focus of sentence	116
4.3	Conclusion	117
CHAPTER 5		
	Register variation	119
	Lexical features	119
5.0	Introduction	119
5.1	<i>Keep</i> meaning <i>Put</i> : Type I analysis	120
5.2	<i>Stay</i> meaning <i>Live</i>	123
5.3	Indian words	126
5.3.1	Discussion of Indian words in Conversational English	130
5.3.2	Discussion of Indian words in Spoken Academic Language	133
5.3.3	Discussion of Indian words in Spoken News	134
5.3.4	Discussion of Indian words in Written News	135
5.3.5	Discussion of Indian words in Written Academic English	137
5.3.6	Discussion of Indian words in Business Correspondence	138
5.3.7	Discussion of Indian words in Personal Correspondence	139
5.3.8	Discussion of Indian words in Written Entertainment News	141
5.3.9	Discussion of Indian words in Fiction	143
5.3.10	Discussion of Indian words in Written Travel News	144
5.3.11	Discussion of Indian words across registers	146
5.4	Conclusion	147
CHAPTER 6		
	Register variation	149
	Grammatical features	149
6.0	Introduction	149
6.1	Stative verbs in the progressive: Type I analysis	149
6.2	Past and present perfect	156
6.2.1	Past perfect	156
6.2.2	Present perfect	157
6.3	Articles	165
6.3.1	Article use before ordinal numbers	165
6.3.2	Articles before phrases	171
6.3.3	Articles before quantifiers	171
6.3.4	Summary of article analysis	177
6.4	Use of the future instead of the simple present (for habitual present)	178



6.11 Invariant tag <i>isn't it</i>	204
6.11.1 <i>Isn't it</i>	204
6.11.2 <i>No</i>	205
6.12 <i>Also</i> and <i>Too</i>	209
6.12.1 General distribution results	209
6.12.2 Positions of <i>also</i> and <i>too</i>	212
6.12.2.1 <i>Also</i>	213
6.12.2.2 <i>Too</i>	213
6.12.3 <i>Also</i> and focus	214
6.13 Conclusion	225
CHAPTER 7	
Conclusion	227
7.0 Introduction	227
7.1 Summary of Type I analyses	227
7.2 Summary of Type II analyses	229
7.3 Conclusions to be drawn from Type I and Type II analyses	232
7.4 Co-occurrence of Indian variants	232
7.5 What, then is, Indian English?	233
7.6 Revisiting Kachru's Concentric Circles	234
7.7 Limitations of the study	235
7.8 Directions for future research	236
References	237
APPENDIX 1	
Methodologies employed in the study of language varieties	243
APPENDIX 2	
Review of variety studies	249
APPENDIX 3	
List of Indian Fiction	255
APPENDIX 4	
Tape insert	259
APPENDIX 5	
Corpus contributor questionnaire	261

## APPENDIX 6

Header for spoken file	263
------------------------	-----

## APPENDIX 7

Tables of Indian words in registers of Indian English	265
---	-----

Index	281
-------	-----

## List of figures

Figure 1.	The Corpus of Contemporary Indian English and its Registers	43
Figure 2.	Written Registers of the CCIE	44
Figure 3.	Spoken Registers of the CCIE	44
Figure 4.	Map of India	48
Figure 5.	Combined Corpus with Registers	63
Figure 6.	Written registers of Combined Corpus	63
Figure 7.	Spoken registers of Combined Corpus	64
Figure 8.	Indian words in Spoken and Written Indian English (Normalized to 1,000,000)	89
Figure 9.	Progressive and Perfect aspect verb phrases in Indian corpus	105
Figure 10.	Modals in Indian Corpus	110
Figure 11.	Semi-modals in Indian Corpus	111
Figure 12.	Also and too in entire, spoken, and written corpora (Normalized to 1,000,000 words)	114
Figure 13.	Distribution of Indian words across spoken and written registers (Normalized to 1,000,000)	127
Figure 14.	Distribution of Indian words across spoken registers (Normalized to 1,000,000)	127
Figure 15.	Distribution of Indian words across written registers (Normalized to 1,000,000)	128
Figure 16.	Indian words in Conversational English (Normalized to 1,000,000 words)	132
Figure 17.	Counts of Indian word categories in Spoken Academic Language	133
Figure 18.	Counts of Indian word categories in Spoken News	135
Figure 19.	Counts of Indian word categories in Written News	136
Figure 20.	Counts of Indian word categories in Written Academic English	137
Figure 21.	Counts of Indian word categories in Business Correspondence	138
Figure 22.	Counts of Indian word categories in Personal Correspondence	140

Figure 23.	Counts of Indian word categories in Written Entertainment News	141
Figure 24.	Counts of Indian word categories in Fiction	143
Figure 25.	Counts of Indian word categories in Written Travel News	145
Figure 26.	Progressive vs. perfect in written registers (frequency in 1000s, normalized to 1,000,000)	180
Figure 27.	Progressive vs. perfect in spoken registers (frequency in 1000s, normalized to 1,000,000)	181
Figure 28.	Relativizers in British and American Conversation vs. Indian Conversation (in thousands, normalized to 1,000,000)	190
Figure 29.	Relativizers in British and American Fiction vs. Indian Fiction (in thousands, normalized to 1,000,000)	191
Figure 30.	Relativizers in British and American Written News versus Indian Written News (in thousands, normalized to 1,000,000)	192
Figure 31.	Relativizers in British and American Written Academic English versus Indian Written Academic English (in thousands, normalized to 1,000,000)	193
Figure 32.	Modals and Semi modals in Indian spoken registers	195
Figure 33.	Modals and Semi modals in Indian written registers	196
Figure 34.	Indian vs. American and British: Also Conveng = Conversational English; wracad = Written Academic English; wrnews = Written News	209
Figure 35.	Indian vs. American and British: Too Conveng = Conversational English; wracad = Written Academic English; wrnews = Written News	211
Figure 36.	American and British (also and tooConveng) = Conversational English; wracad = Written Academic English; wrnews = Written News	211
Figure 37.	Indian also and tooConveng = Conversational English; wracad = Written Academic English; wrnews = Written News	212
Figure 38.	Continuum of Indian English	232
Figure 39.	India's New Concentric Circles	234

## List of tables

Table i.	Corpus Based Investigations of New Varieties of English	26
Table 1.	Overview of the project	39
Table 2.	The Corpus of Contemporary Indian English and its Registers	41
Table 3.	Written registers and their sources	52
Table 4.	Combined corpus	62
Table 5.	Features examined in the study	68
Table 6.	<i>Keep</i> meaning <i>Put</i> in the entire corpus	87
Table 7.	<i>Stay</i> meaning <i>Live</i> in the entire corpus	88
Table 8.	Stative progressives	91
Table 9.	Past perfect in entire corpus	91
Table 10.	Present perfect in entire corpus	92
Table 11.	Prepositional verbs in corpus: Verbs	94
Table 12.	Prepositions that occur in prepositional verbs in Indian Corpus	97
Table 13.	Comparing common prepositions in British and American English and Indian English	99
Table 14.	Articles before ordinal numbers	101
Table 15.	Articles before noun phrases	101
Table 16.	Articles before <i>lot of</i>	102
Table 17.	Articles before <i>little</i>	102
Table 18.	Articles before <i>few</i>	103
Table 19.	Articles before <i>number of</i>	103
Table 20.	Total articles examined	104
Table 21.	Use of the future instead of the simple present	105
Table 22.	Absence of subject-auxiliary inversion	106
Table 23.	Transitive verbs used intransitively	107
Table 24.	Non-initial existential <i>There</i>	112
Table 25.	<i>Isn't it</i> as invariant tag	112
Table 26.	<i>No</i> as invariant tag	113
Table 27.	<i>Also</i> in the entire corpus	115
Table 28.	<i>Too</i> in the entire corpus	115

Table 29.	<i>Also</i> and focus	116
Table 30.	<i>Keep</i> meaning <i>Put</i>	121
Table 31.	<i>Stay</i> versus <i>Live</i>	124
Table 32.	Frequency of Indian Words in Registers of Indian English (Normalized to 1,000,000)	129
Table 33.	Stative verbs in the Progressive	152
Table 34.	Past Perfect	158
Table 35.	Present Perfect	161
Table 36.	Articles before ordinal numbers – both spoken and written corpus	166
Table 37.	Articles before noun phrases – spoken corpus only	172
Table 38.	Articles before quantifiers in spoken and written registers	174
Table 39.	Article use in general	177
Table 40.	Future for simple present	179
Table 41.	Absence of subject-auxiliary inversion in WH-question formation	183
Table 42.	Transitive verbs used intransitively across all registers	186
Table 42a.	Frequency of transitive verbs occurring intransitively across registers of Indian English	189
Table 43.	Distribution of modal verbs across registers of British and American English versus Indian English	197
Table 44.	Non-initial Existential <i>there</i>	199
Table 45.	Invariant tag <i>isn't it</i>	206
Table 46.	<i>No</i> as invariant tag	207
Table 47.	<i>Also</i> : Distribution	216
Table 48.	<i>Too</i> : Distribution	217
Table 49.	<i>Also</i> : Focus	218
Table 50.	General conclusions about Type 1 analyses	230
Table 51.	Conclusions about Individual Registers for Type 1 analyses	230

## CHAPTER 1

# Introduction

### 1.0 English as a global language: Emergence of new varieties

The rapid spread of English and its rising status as a world language has been a subject of discussion in the field of Applied Linguistics for several years, and accompanying this spread has been the emergence of several new varieties of English (called NVEs, New Englishes, non-native Englishes, etc.). As Schneider (2003) describes, “present-day English as a global language is more than the world’s predominant lingua franca – it is also a language which is currently growing roots in a great many countries and communities around the world, being appropriated by local speakers, and in that process it is diversifying and developing new dialects...” (p. 233). The way regional and social factors have influenced the growth of New Varieties of English and fostered change has formed the subject matter of sociolinguistics and dialectology from both theoretical and practical standpoints, and today, nobody would deny the fact that “World English exists as a political and cultural reality” (Crystal, 2003, p. xii). Irrespective of where it developed, according to Platt et al. (1984, p. 2–3), a New English has the following characteristics:

1. It has developed through an education system. This means that it has been taught as a subject and, in many cases, also used as a medium of instruction in regions where languages other than English were the main languages.
2. It has developed in an area where a native variety of English was not the language spoken by most of the population.
3. It is used for a range of functions among those who speak or write it in the region where it is used.
4. It has become ‘localized’ or ‘nativized’ by adopting some language features of its own, such as sounds, intonation patterns, sentence structures, words, and expressions.

Today, fundamental to our understanding of World Englishes, is an acceptance of the concept that English is not monolithic. How could it be, if it is used by millions of people in areas of the world as different and apart as India, on the one hand, and Kenya, on the other?

### 1.1 Understanding New Englishes

Braj Kachru (1988) first explained the spread of English in the world and the different roles it plays in different countries by suggesting we think of three concentric circles: the inner circle, the outer circle, and the expanding circle. Kachru suggested that the inner circle included countries where English is spoken as the primary language – countries like the UK and the USA, countries that are the “traditional cultural and linguistic bases of English” (Kachru, 1992, p. 356). The outer circle includes countries like India and Singapore, where English functions as an important second language; these outer circle countries are typically multilingual, and English also plays the role of common language of communication. Platt’s New Englishes, then, fit into Kachru’s outer circle. The last circle, the expanding circle, includes countries like China and Russia, countries “which recognize the importance of English as an international language, though they do not have a history of colonization by members of the inner circle, nor have they given English any special administrative status” (Crystal, 2003, p. 60). Today, by many estimates, outer circle English speakers (300–500 million) far exceed inner circle speakers (320–380 million, Crystal, 2003). Crystal also says that there are 75 territories which are either members of the inner or outer circles in which English holds a special place.

The sheer number of speakers who speak English as a second (and in countries like India, many speak it as a first) language alone necessitates moving away from the traditional notions such as “native” and “non-native.” Graddol, Leith, & Swan (1996) state that “In practice it is difficult to draw hard and fast boundaries between... the ‘native’/‘non-native’ distinction” because in contexts like India and Singapore, “some (notionally) non-native speakers become familiar with English from an early age and use the language routinely” (p. 13). Other terms synonymous with “native” and “non-native” are “L1 English” and “L2 English.” This distinction, too, is no longer entirely valid, given that in a number of traditionally non-English speaking countries like India and Singapore, there are increasing numbers of people who speak English as a *first* language, i.e., an L1. As D’Souza (1997) explains,

“If non-native speakers are those to whom the language is not ‘native’, then we should accept that speakers from America and Australia should be considered ‘non-native’, too. English now belongs to the world, and multilinguals who speak English do not need to apologize for not being monolinguals. Terms like ‘non-native’, ‘second-language’ etc. can no longer be used unless one defines them anew in every instance of use. Their meanings cannot be taken for granted.”

(p. 102). Terms like New Englishes, Other Englishes, and New Varieties of English then, have gained prominence.



D'Souza, who uses the term New Varieties of English, or NVEs, explains that "the NVEs are a phenomenon peculiar to the ex-British and American dominions where English stayed on after the colonial masters left, stayed on to become more than a second language, though not really a first" (1997, p. 241). D'Souza further comments that in these contexts, English "took on certain characteristics that seem to justify the study of these varieties as phenomena different from those in all other countries in which English is spoken" (p. 241). Schneider (2003) proposes a model for understanding the shared processes that he explains all New Englishes go through in their development. He explains that the development of any New English entails five processes: "Foundation, Exonormative Stabilization, Nativization, Endonormative Stabilization, and Differentiation" (p. 243). He explains that the first process, Foundation, is the initial phase where "English begins to be used on a regular basis in a country that was not English-speaking before" (p. 244) and that characterizing this phase is a "complex contact situation" (p. 244). In this phase, contact between the two language groups remains restricted, with cross-cultural communication being achieved by just a few people. Further, during this phase, indigenous languages do not influence the English spoken by the settlers. During Phase 2, Exonormative Stabilization, the "external norm, usually written and spoken British English as used by educated speakers, is accepted as a linguistic standard of reference" (p. 245). Also, this phase is characterised by Structural Nativization, where "as soon as a population group starts to shift to a new language, some transfer phenomena at the level of phonology and structure are bound to occur" (p. 246). Phase 3 is Nativization, which, according to Schneider, is "the most important, the most vibrant one, the central phase of both cultural and linguistic transformation in which both parties realize that something fundamental has been changing..." (p. 247). It is during this phase of Nativization that the New English starts to construct its identity independent of the "native" English. It is during this phase, then, that characteristic "features" of the new English emerge. Phase 4, or Endonormative Stabilization, is "marked by the graduate adoption and acceptance of an indigenous linguistic norm, supported by a new locally rooted self-confidence..." (Schneider, 2003, p. 249).

During the fifth phase, Differentiation, "the focus of an individual's identity construction narrows down, from the national to the immediate community scale... Consequently, new varieties of the formerly new variety emerge as carriers of new group identities within the overall community" (p. 253).

Despite this view of New Englishes as evolving, growing entities, the study of New Englishes over the past several decades has not gone beyond studying the features that supposedly characterize these varieties, i.e., determining and describing what features emerged during Schneider's phase 3 (Nativization). Khubchandani and Hosali (1999) state that "in the past few decades, various aspects of