

## How Indian is Indian English?: Indian Words in Registers of Indian English

Chandrika Balasubramanian<sup>1</sup>  
Sultan Qaboos University, Oman

Article Received: 14 July 2016; Article Accepted: 4 October 2016

### Abstract

The rising status of English as a world language has led to the emergence of several non-native or new varieties of English, with Indian English being a major new variety. Much work on Indian English has focused on establishing the *Indianness* of the language. As Kachru points out, studies of Indian English consider “linguistic interference and the Indian cultural context as essential for the understanding and description of the Indianness in this variety of English” (*The Indianness of Indian English* 391). Other early research on the Indianness of Indian English includes Verma, Dubey, Hosali, and Bhatt, to name just a few. The work represented by most previous studies on Indian English, which focus largely on the nature and degree of difference of Indian English from external norms, does not, however, provide a clear picture of the system underlying Indian English. The current study focuses on the Indianness of Indian English, but tests Kachru’s claims regarding the progressive Indianisation of Indian English. The first part of the study is an empirical investigation of Indian words in three spoken and three written registers of the language from a corpus compiled in the year 2000 and earlier, and identifies distinct semantic categories of words in the registers. Based on the results of this first analysis, the second part of the study analyses the occurrence of Indian words in a smaller corpus of a register of written Indian English compiled in 2016. The first part of the study shows that there are marked differences in the degree of Indianisation of Indian English among different registers, and the second part of the study shows that with respect to the register studied, written Indian English seems to be undergoing a process of un-Indianisation.

### Keywords

Indian English, Indian words, register, Entertainment News, Indianisation, un-Indianisation

---

<sup>1</sup> Chandrika Balasubramanian (PhD) is an assistant professor at Sultan Qaboos University, Oman, where she has been teaching since 2013. Her research focuses on variation in international Englishes, with a specific focus on the English used in the Indian subcontinent. Her 2009 publication, *Register Variation in Indian English*, was the first large-scale empirical investigation of variation within this international variety. Email: chandribala@squ.edu.om.

## Introduction

The global spread of English is a historically unprecedented phenomenon; no other language in modern history has achieved the global presence that English has now. Accompanying this unprecedented spread is the emergence of many new varieties of English in different parts of the world. Several names for such new varieties have also emerged: New Englishes, Institutionalised Second Language Varieties and International Englishes are just a few. As Mesthrie and Bhatt put it, “It has become customary to use the plural form ‘Englishes’ to stress the diversity to be found in the language today, and to stress that English no longer has one single base of authority, prestige and normativity” (3). Kachru, in *The Alchemy of English*, explains that these New Englishes “have a long history of acculturation in new cultural and geographical contexts; they have a large range of functions in the local educational, administrative, and legal systems” (19), with the result that they have developed nativised discourse and style types different from traditional native varieties like British or American English. Today, it is a widely accepted fact that second language varieties, or Kachru’s Outer Circle variety speakers, far outnumber traditional native speakers of the language. By Crystal’s estimate, Outer Circle speakers number between 300-500 million, while Inner Circle speakers number 320-380 million. A natural extension of this spread of English has been the establishment of a strong research tradition dedicated to the study of the new varieties of English. While initial studies including those such as Bansal, Bakshi, Baumgardner and D’Souza concerned themselves with how the new varieties differed from native varieties, later studies such as Balasubramanian, Lange, Hundt and Gut, and Sedlatschek have concerned themselves more with how the new varieties function as systems unto themselves, as opposed to erroneous forms of native varieties. Studies now also focus on how new varieties vary internally, just as traditional native varieties like British and American English do.

Indian English is one new variety that has been extensively studied. Much recent scholarship on Indian English has focused on how English has become nativised in India, and is now no longer regarded as a foreign language. As Y. Kachru and Nelson explain, “nativization brings forth a new variety of English... and is triggered by language contact and the natural inclination in users to mould the medium to express relevant contextual realities” (31).

## Literature Review

For several decades now, researchers have described how English has spread and become nativised in diverse parts of the world. One way in which nativisation of the language occurs, it is claimed, is by the adoption of lexical items from the L1 into English. Y. Kachru explains that the mixing of L1 with English results from “the process of nativization and acculturation of the English language in Asia” (223). Y. Kachru goes on to explain that “Nativization of English in the Outer

and Expanding Circles manifests itself in incorporation of the language in multiple styles and genres” (223).

Much previous research on Indian English holds that a characteristic feature of the nativisation of English in India is the presence of Indian words in it. Kachru explains that South Asian English is the only variety of non-native English with a strong research tradition of studies focusing on the lexicon. As early as 1886, the Hobson-Jobson Dictionary, for example, “provided linguistic entertainment by its lexical explanations” (*The Alchemy of English* 41). More recently, Schneider explains, “there is the usual share of fauna and flora words... and there is a wide range of ‘Indianisms,’ words which denote elements of the indigenous culture and lifestyle” (165). In describing the natural and invaluable function of Indian words in Indian English, as early as 1968, the linguist P.E. Dustoor said:

Our mental climate will always foster plants that do not flourish in England or America; and such plants, just because they are somewhat exotic, add to the charm of a garden. All lovers of English will, therefore, encourage them to grow in the world-wide garden of English. (126)

On the occurrence of Indian words in Indian English, in *Asian Englishes Beyond the Canon*, Kachru explains that there exist two groups of these words. The first is a group of words that have been “assimilated across varieties of language, specifically in British and American English,” while the second group of words consists of those that “occur frequently in various registers of the South Asian varieties of English” (523). Kachru further divides the second group into three classes: single lexical items, hybridised lexical items, and “English lexical items used with extended or restricted semantic connotations” (525). Hosali mentions the same categories and calls them “loan words,” “loan compounds” and “collocatives.” As an example of the third class, she uses the sentence “I had my ears bored so that I could wear my diamond ear-rings” (qtd. in Balasubramanian 32). She explains how the meaning of “bore” is extended to mean “pierce.” While the first two types of loan words will be examined in this study, the third type cannot be studied because of the limitations of the methodology employed in the current research; as described below, for the most part, Indian words for the current study were identified by Microsoft Word’s identification of the words as errors. An analysis of collocatives would involve a much closer reading of all the texts in the corpus, and is beyond the scope of the current paper. It would, however, make an interesting study in the future.

In *The Alchemy of English*, Kachru explains that single lexical items “vary in their frequency of occurrence. They are essentially register-dependent and therefore are normally used when referring to contexts which are typically South Asian” (42). Kachru describes hybridised lexical items as those with at least one

word from a South Asian language and one word in English. Further, he claims that they are very common in the English used in India. In his category of hybridised items, he includes “open-set items without grammatical constraints on the selection... and closed-system items involving bound morphemes and showing certain grammatical constraints” (*The Alchemy of English* 42-43).

Indian words in previous research on Indian English have been examined mainly for their stylistic implication. Many researchers (Verma, Kachru, Dubey) have examined the occurrence of Indian words in Indian English based on the assumption that they occur because the English used in India “has proved to be ineffective in conveying aspects or messages from a culture alien to it” (Dubey 21). Similarly, Baumgardner regards the presence of non-English words in a variety of English (such as Pakistani English) as lexical innovations “through which local social and cultural phenomena can find expression” (175).

Kachru has provided a basis for the study of Indian words in Indian English, and often comments on the presence of Indian words as contributing to its Indianness. However, a detailed analysis of the types of Indian words, and the contexts in which they occur does not exist yet. Mesthrie and Bhatt explain that “vocabulary retentions in New Englishes tend to cluster in semantic fields pertaining to local customs and culture, including terms for food, clothing, music and dance” (110). While they go on to provide a selection of such terms from different varieties, they do not provide systematic descriptions of any varieties. Mesthrie, with respect to South African Indian English (SAIE), mentions that in written SAIE, “terms denoting religious and cultural festivals” and words pertaining to culture, religion, clothing, food and domestic utensils “regularly occur without glosses” (7). Mesthrie provides us with a one thousand word lexicon of SAIE, but does not, however, provide an indication of the use of these Indian words across registers.

While there is, then, a wealth of literature on Indianisms in Indian English, a detailed analysis of the types of Indian words, and the contexts in which they occur does not exist yet.

### **Aims of the Study**

The aims of the current study are two-fold. Firstly, the study aims to determine the degree to which six registers of Indian English are Indianised (based on the occurrence of Indian words in the registers). For this part of the study, three spoken and three written registers were chosen from a Corpus of Contemporary Indian English (CCIE) compiled in the year 2000 combined with certain sections of ICE-India (compiled earlier), as part of a larger study of register variation in Indian English. Secondly, based on the results obtained from the first analysis, the next part of the study analyses the occurrence of Indian words in 40,000 words of written Entertainment News, with a view to determining how much the

language in this particular context has changed over the past 16 years. Specifically, then, the study has the following aims:

- i) To determine the degree to which six registers of Indian English, including Written Entertainment News, have been Indianised, based on an occurrence of Indian words and hybridised constructions in them;
- ii) To determine the number of Indian words and hybridised constructions in a 40,000 word corpus of Written Entertainment News compiled in 2016; and
- iii) To determine the degree to which the language in Written Entertainment News, has changed between the year 2000 and the year 2016.

It is important to note that this is not a study on code-switching or language mixing, and does not attempt to propose a framework for where and how languages mix. How multilinguals take advantage of the different linguistic repertoires at their disposal to create certain effects creatively, whether in poetry, fiction, or in advertising has been the subject of many papers (Y. Kachru), and will not be dealt with in the current paper. The focus of this study, rather, is lexical.

### **Methodology**

For the first part of the study, a corpus of three spoken and three written registers with a word count of just over a million words was used. The corpus is a combination of CCIE, compiled in the year 2000, and certain sections of ICE-India, compiled in the 1990s. The methodology employed in the construction of CCIE is described in detail in Balasubramanian, while the construction of ICE-India was described in Greenbaum, so these details will not be provided here. The six registers analysed in the first part of the study include Conversational English, Spoken Academic English, Spoken News, Written News, Written Academic English and Written Entertainment News. The reasons for separating Written Entertainment News from Written News are several: the former contains many more features akin to Conversational English in its informality, and it also includes interviews which, although edited, are more conversational in nature. Given that the texts chosen for this register are neither strictly conversational nor strictly written news, they were included in a separate register. For the first part of the study, the Written Entertainment News register was compiled entirely in the year 2000.

For the second part of the study, a smaller corpus of 40,000 words of Written Entertainment News from the year 2016 was compiled. Sources for this corpus were various, and included the following:

- *Deccan Herald*, a newspaper published in the southern state of Karnataka, published exclusively in English
- *The Hindu*, a newspaper published in English in the southern state of Tamil Nadu
- *Deccan Chronicle*, a newspaper published in English in the southern state of Andhra Pradesh
- *The Indian Express*, a paper published in Calcutta in English
- *The Times of India*, a national newspaper published in English in New Delhi
- *Filmfare*, one of India's top magazines, published in Mumbai. The magazine focuses on Bollywood, cinemas, the lives, personal and professional, of anyone connected with India's large film industry
- *Femina*, a popular women's magazine published in Mumbai
- *Midday*, an online newspaper from Mumbai
- *Zee News* – an online newspaper
- *IndiaTV News* – an online newspaper

The publications were chosen to get as wide a representation of India as possible. Further, care was taken to ensure that the register compiled in 2016 closely resembled the one compiled in 2000; most of the sources are the same. For both corpora, all files were saved as Microsoft Word documents to enable identification of Indian words using the software's spell check function.

Table 1 below shows the registers and their word counts for both corpora. For the original corpus, information is provided in parentheses as to how many words came from CCIE, and how many from ICE-India. As is clear from the table, files in Spoken Academic English and Spoken News were partly from ICE-India.

Table 1: Registers Analysed

	<b>Register</b>	<b>Word Count</b>
Original Corpus (2000 and earlier)	Conversational English	233,912 (CCIE)
	Spoken Academic English	75,484 (35,484 CCIE; 40,000 ICE)
	Spoken News	98,074 (58,074 CCIE; 40,000 ICE)
	Written Academic English	120,000 (ICE)
	Written News	447,546 (CCIE)
	Written Entertainment News	86,378 (CCIE)
	<b>TOTAL</b>	<b>1,061,394</b>
New Corpus (2016)	Written Entertainment News	40,042

As mentioned earlier, the analysis of the Indian words in the files included creating a list of words identified by Microsoft Word as misspellings. All files

were also read to ensure that Indian words that were spelled the same as an English word were not missed in the analysis. This rather laborious method was used for the identification of both single lexical items and what Kachru terms hybridised items. The Indian words were then categorised into 10 different semantic groups:

- Food
- Clothing
- Arts (dance, music, etc.)
- Religion
- People
- Discourse Markers
- Greetings
- Politics
- Larger chunks of language (between 2 and 10 words in length)
- Other – words that clearly did not belong in any of the above categories, and could not be classified in any other way.

The categories were chosen because of the seeming frequency with which certain kinds of words occurred. These are words which, as previous research has pointed out, are “a vehicle of Indian culture to express culturally determined networks of activities that are typically Indian” (Verma 73). While other categories are self-explanatory, a note of explanation is necessary about Discourse Markers. The words in this category included those that had no semantic function within the sentence they occurred in other than the role of fillers, often used by the speakers to continue holding the floor.

The Larger Chunks of Language category typically included sentences or portions of sentences, as long as they were longer than 2 words. For the Written Entertainment News register, the analysis did NOT include the names of films or songs, which, given the nature of the register, occur extremely frequently.

In this analysis, words like *avatar*, *karma* and *mantra*, words of Indian origin, but that have been assimilated into the English language, have been excluded. Further, it is important to note that all the registers had words that did not fit into any of the above-mentioned categories. This included words that were less substantial than those in Larger Chunks of Language. However, since they were few in number, they are not included in a discussion of the results. The Other category had almost no words in any register other than Written Entertainment News.

## Results

The results from the first analysis are presented first, followed by those obtained for the second analysis. Following a general discussion of the overall results, detailed results are presented in terms of graphs depicting the distribution of semantic categories of words in each register, with counts provided normalised to 1,000,000 words. The graphs are followed by a discussion of the results and example sentences with Indian words in each semantic category in each register.

Figure 1 below shows the distribution of Indian words in the different semantic categories in all the registers studied. As mentioned earlier, frequencies of Indian words are normalised to 1,000,000. What is immediately obvious from the figure is that in general, there are more Indian words in the spoken registers than in the written registers, a finding that is not surprising. Further, it is clear that Conversational English has far more Indian words than do the other two spoken registers, Spoken Academic English and Spoken News. That Spoken Academic English and Spoken News closely resemble their written counterparts is not surprising, given the relative formality of these registers. Researchers of world Englishes have often claimed that innovations, both grammatical and lexical, begin in the spoken registers and then spread to the written modes. This analysis confirms this claim. Of the written registers, what is, perhaps, most interesting is the high proportion of Indian words in Written Entertainment News (2000). This, and the fact that the proportion of Indian words in Written Entertainment News (2016) has by far fewer Indian words than the register with files compiled in 2000, is discussed further below. First, however, is a discussion of the results obtained for the individual registers.

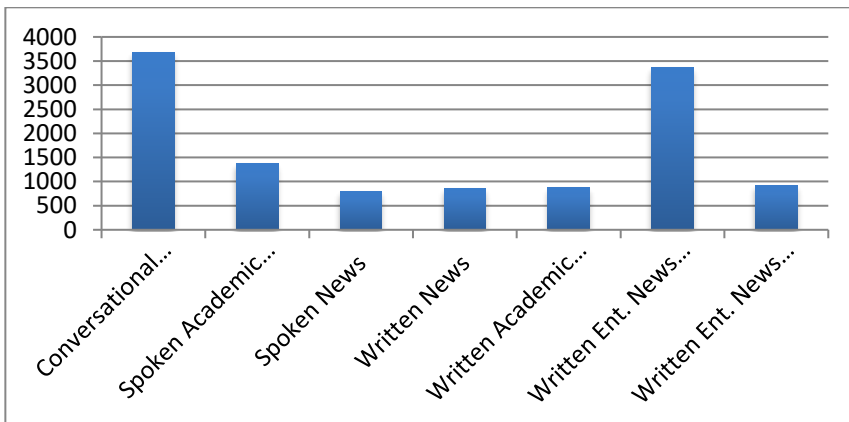


Figure 1: Distribution of Indian words across registers



*Conversational English*

Figure 2 below shows the distribution of Indian words across the 9 semantic categories studied.

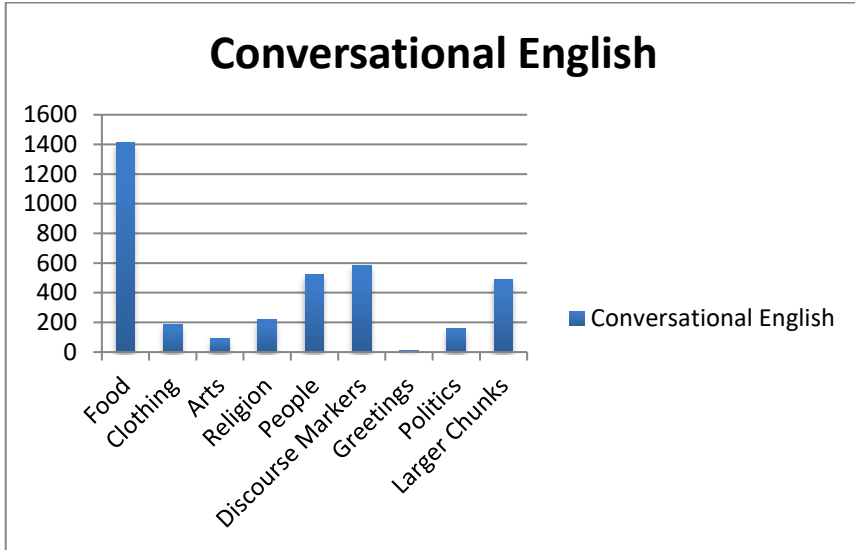


Figure 2: Distribution of Indian words in Conversational English

That food is the semantic category that has the highest number of Indian words will not come as a surprise to anyone who knows Indians; food is widely considered as central to their lives. The words in this category come mostly from Hindi, although there are words in the South Indian languages, too. The high frequency of Hindi words as opposed to words in other languages, despite the fact that there were more corpus contributors who came from language backgrounds other than Hindi, is worthy of further research but beyond the scope of the current paper.

Other semantic categories of Indian words which feature in Conversational English include Discourse Markers, People and Larger Chunks of Language. Discourse Markers such as different forms of “yes” and “no” (when not performing their normal discourse functions), and terms used to refer to people casually (such as “hey” in English) were also found in many texts. Words in the category Religion also featured in this register and they usually occurred in Sanskrit.

Indian words in the People category were also common. In this register, people were often relatives; Indian languages have numerous kinship terms that show exact relationships better expressed in an Indian language than in English. For example, in most Indian languages, the terms for maternal and paternal relatives are different. In Hindi, maternal uncle and aunt are *mama* and *mausi*.

Kinship terms for uncle and aunt on the paternal side are further divided into father's older brother or sister and younger brother and sister. A father's younger brother and sister are *kaka* and *kaki*, while the father's older brother and sister are *chacha* and *chachi*. Using Indian words in English to denote kinship terms, therefore, definitely shows how Indian words in Indian English act as "a vehicle of Indian culture to express culturally determined networks... that are typically Indian" (Verna 73).

Larger Chunks of Language were also very common in this register. In future research, a study on code-switching with a focus on determining where and why code-switching occurs is definitely needed. This study simply shows that this is a common phenomenon in this particular register, and often, there are situations where speakers switch back and forth between English and an Indian language within the same sentence.

Other semantic categories in this register which had Indian words in them include Clothing, Greetings, Politics and Music. The words occurred in different languages, often, but not always, depending on the first language of the speaker. As mentioned earlier, Hindi occurred most frequently; this finding needs further research, given that (as described in Balasubramaniam), the contributors of the Conversational English section of the corpus were mainly speakers from the four South Indian states.<sup>2</sup> Below are some sentences from Conversational English with Indian words in them. Following each sentence in parentheses is the semantic category the word is from, a translation of the Indian words as well as the language in which the Indian words occurred. For ease of reading, the Indian words in the sentences as well as their translation in parentheses have been italicised.

- You get two *parathas* a plate, no *alu parathas*. (Food; *kind of bread*; Hindi; Conversation2 from CCIE)
- You'll stitch the *salwaar* if I bought the material? (Clothing; *kind of pants*; Hindi; Conversation8 from CCIE)
- Do you have any more film-based *ghazals*? (Arts; *genre of music*; Hindi; Service Encounters20 from CCIE)
- After Khamoshi, I kept hearing *shlokas* in my head. (Religion; *verses*; Sanskrit; FInt7 from CCIE)
- She's my *didi*, my inspiration, my blessing. (People; *older sister*; Hindi; FInt8 from CCIE)

<sup>2</sup> Due to their shared ancestry, some Hindi words and phrases may have structural and semantic similarities with those in other Indo-Aryan languages such as Urdu, Punjabi, Bengali, etc. However, unless their use is particularly associated with any of those languages, the origin of such words and phrases is considered as solely Hindi in this paper.

- You can't eat much; you can't imagine, *yaar!* (Discourse Markers; *fellow*; Hindi; Service Encounters17; CCIE)
- *Namaskara*, how are you? (Greetings; *bello*; Kannada; Conversation 8 from CCIE)
- He was a *tehsildar*, a pretty respectable individual. (Politics; *governmental official*; Hindi; Oral Interviews23 from CCIE)
- Methodology *ille, enna panninderke!* (Larger chunks; *not methodology, what are you doing?* Tamil; Conversation1 from CCIE)
- That was all just *tamasha*. (Other; *spectacle*; Hindi; Conversation6 from CCIE)

### *Spoken Academic English*

Figure 3 below shows the distribution of different semantic classes of Indian words in Spoken Academic English.

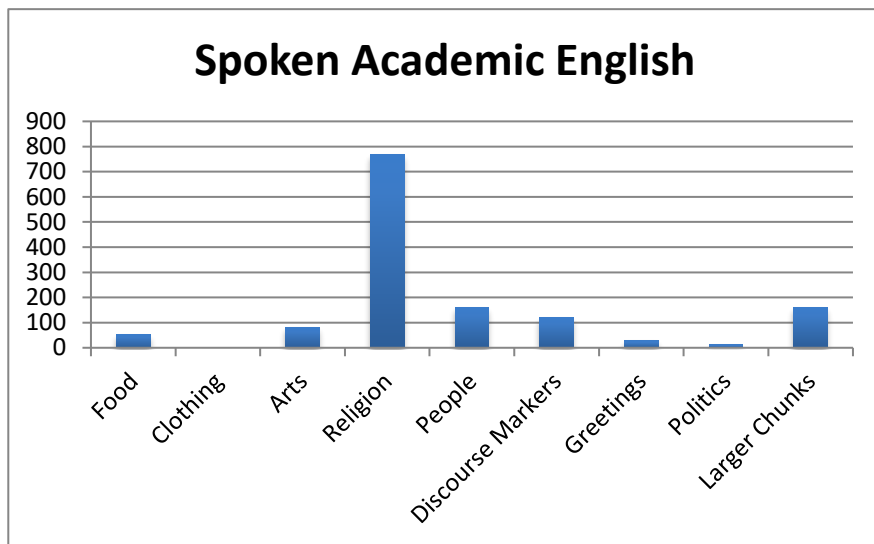


Figure 3: Distribution of Indian words in Spoken Academic English

As is clear from the figure, Indian words are relatively uncommon in this register except for the semantic category Religion. The reason for this is that the study of history, politics, or even literature in India often focuses on India's religious past.

Other categories that do have some Indian words include People, Greetings, and Larger Chunks. As with Conversational English, most words occurred in Hindi. Below are examples of sentences from this register.

- Feed me with coconut *payasam*. (Food; *type of dessert*; Tamil and Telugu; S1B-018 from ICE-India)
- You probably might say *shehnai*. (Arts; *musical instrument*; Hindi; S1b-009 from ICE-India)
- ... the weapon for freedom was *ahimsa*. (Religion; *non-violence*; Hindi; S1B-011 from ICE-India)
- It is not the land of *rishis* anymore. (People; *holy man*; Hindi; Lectures3 from CCIE)
- *Achha*, one who doesn't pay the... (Discourse Markers; *yes, ok*; Hindi; S1B-005 from ICE-India)
- *Namaste*. Welcome to Yoga Time. (Greetings; *Good day*; Hindi; Lecture5 from CCIE)
- It comes to the district, it comes to the village *panchayat*. (Politics; *governing body*; Hindi; Lecture3 from CCIE)
- Not director *sa'ab kab aaye ga?* (Larger Chunks; *when will the director come?* Hindi; Lecture4 from CCIE)

### Spoken News

Figure 4 below shows the distribution of different semantic categories of Indian words in Spoken News.

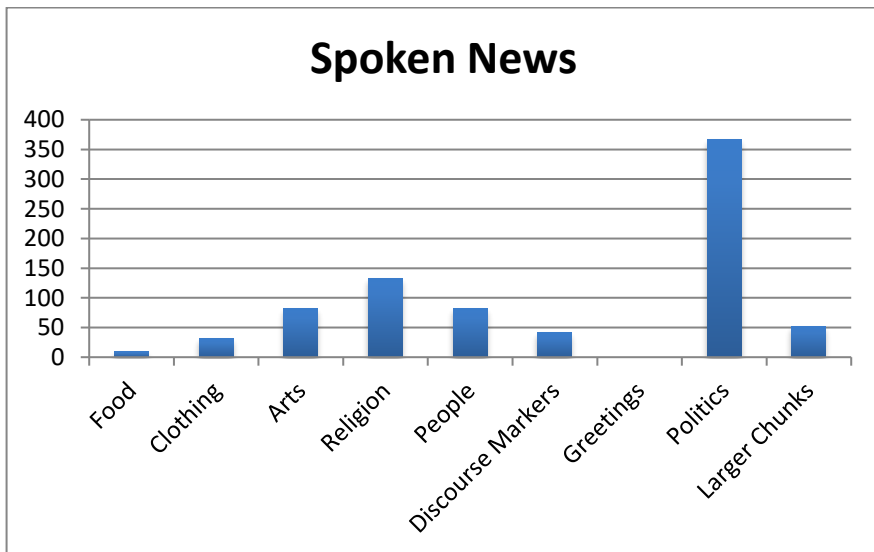


Figure 4: Distribution of Indian words in Spoken News

As one might expect, the largest number of words in this register belonged to the category Politics, and most words occurred in Hindi. There were also a fair

number of words in Religion, once again emphasizing the importance religion plays in India. Following are some sentences from this register:

- ... small forest produces like *tendu* leaves and *sal* seeds. (Food; *kinds of crops*; Hindi; S2B-009 from ICE-India)
- Jamid Akhtar, his face covered with a *burka*, was produced in court. (Clothing; *covering cloth*; Hindi, Urdu; News2 from CCIE )
- The noted violinist of the Delhi *gharana*... (Arts; *house of music*; Hindi, Urdu; S2B-005 from ICE-India)
- The nation is profoundly shocked at this terrible tragedy, *Shri* Rajiv Gandhi's assassination... (People, *Mr.*, Hindi; S2B-020 from ICE-India)
- *Achha*, other thing, Akilesh, what is the profile of Dandi? (Discourse Marker; *ok, yes*; Hindi; News7 from CCIE)
- Because they seem to be more *kabab me haddi* business. (Larger Chunks; *third wheel*; Hindi; News5 from CCIE)

*Written Academic English*

Figure 5 below shows the distribution of different semantic categories of words in Written Academic English.

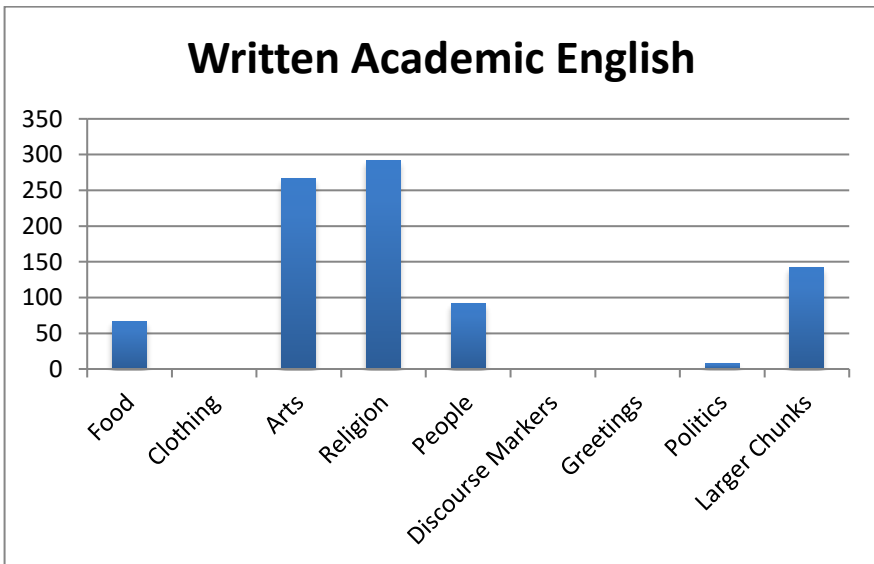


Figure 5: Distribution of Indian words in Written Academic English

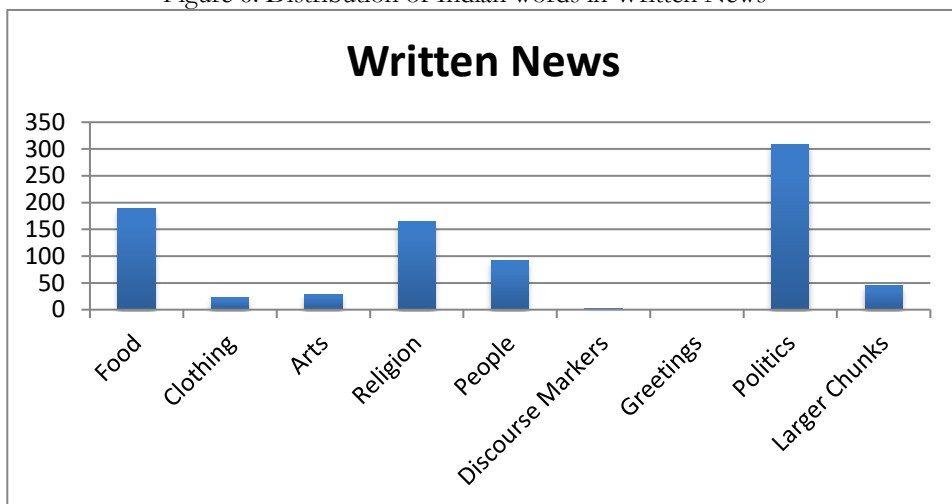
As the figure shows, Indian words feature more prominently in this register than one might have imagined. When one examines the semantic categories in which they occur, however, reasons for this become clear. Religion, as with the other registers, is the category with the most Indian words. Arts is another category with a large number of words, and this is because many of the texts in this register, as described in Balasubramaniam came from the Humanities section of ICE-India which focused on music and dance, which are central to Indian culture. A logical conclusion, then, is that the presence of this category of words in this register is not register-specific, but topic-specific. An absence of words in Greetings, Clothing and Discourse Markers indicates their obvious irrelevance to this register. Below are example sentences from this register.

- These include *tur*, gram, pea, bean, etc. (Food; *a kind of food crop*; Hindi; W1A-019 from ICE-India)
- *Nada* is the basic material of music, and *varna* is that of literature, and both are inseparable in a way. (Arts; *musical terms*; Sanskrit; W2A-010 from ICE-India)
- So Britishers passed the abolition of *sati*. (Religion; *practice of sati*; Hindi; W1A-011 from ICE-India)
- *Shri* Shahu Chhatrapati, as a grandfather... (People; *Mr.*; Hindi; W2A-006 from ICE-India)
- Press gives important current events of world, country, *Taluk*, district, etc. (Politics; *political area*; all languages; W1A-004 from ICE-India)
- Sometimes the above form may be followed by another conversational opener as follows: *aur sunaiya, kyaa baal hai?* Please tell me, what is the news? (Larger Chunks; translation provided in text; Hindi; W2A-003 from ICE-India)

### *Written News*

Figure 6 below shows the distribution of different semantic categories of words in Written News.

Figure 6: Distribution of Indian words in Written News



As the figure shows, Indian words in this register frequently belong to the Food, Religion and Politics categories. The presence of political words is because many words in this category are both topic- and register-specific. Food has a presence because several articles included in the corpus focused on various agricultural projects, several of which have, as far as I am aware, no English equivalents. This, then, does indeed support claims made by researchers such as Dubey who said Indian words occur in Indian English because without them, Indian English has “proved to be ineffective in conveying aspects or messages from a culture alien to it” (23). The fact that Religion features as prominently as it does, reiterates the point made earlier about religion playing a major role in life in India in general. Following are example sentences from this register.

- *Jowar, bhajra, ragi* and maize at Rs. 415 against.... (Food; *types of grains*; Hindi; ETBus4 from CCIE)
- She is surprisingly well-clad in a *lehenga*. (Clothing; *type of skirt*; Hindi; FEEd2 from CCIE)
- Banani Mitra gave a good start with her *bhajans* (Arts; *type of song*; all languages; STf2 from CCIE)
- The Muslims offered *namaḥ* at the Babri mosque. (Religion; *prayer*; Hindi, Urdu; DCBus1 from CCIE)
- The principal reality today is that Mian *Sabeb* is praying for your victory. (People; *sir*; Hindi and Urdu; HED3 from CCIE)
- The underlying principle of the *panchayati raj* is the use of local knowledge... (Politics; *form of government*; Hindi; HED8 from CCIE)

*Written Entertainment News 2000*

The distribution of Indian words across different categories in Written Entertainment News year 2000 is shown in Figure 7 below.

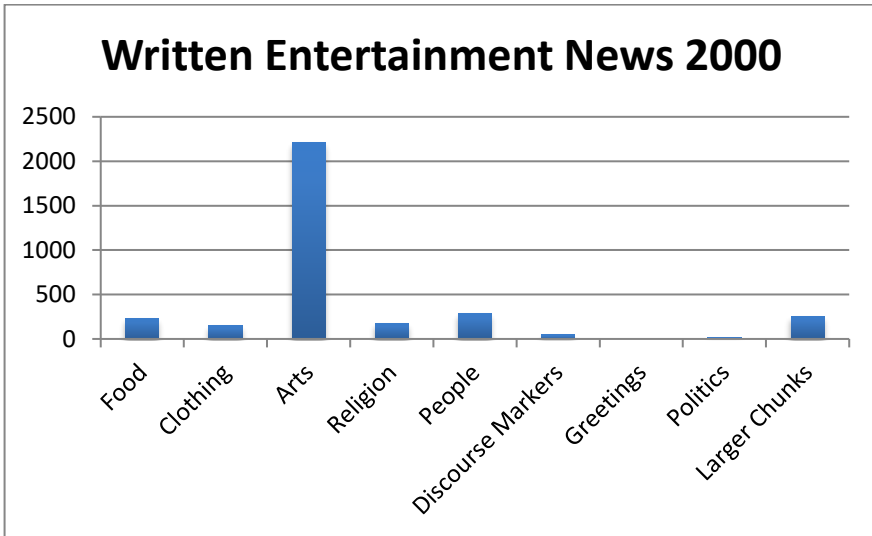


Figure 7: Distribution of Indian words in Written Entertainment News 2000

As mentioned earlier, this is the written register with the most Indian words in it. What is immediately obvious is that the register has its largest category of words in Arts. This is explained when one considers the nature of the texts in this register. Most of the texts discussed issues related to the Indian film industry. Since music plays a huge role in the film industry, the preponderance of Indian words in this register is reasonable. The main Indian language used in this category is Sanskrit because of the influence this language still has on Indian music. This is followed by Hindi. The other semantic categories which have large numbers of Indian words in them include Food, Clothing, Religion, People and Larger Chunks of Language. With all these semantic categories, the Indian words occurred in Hindi more often than in other Indian languages irrespective of the source. In other words, even sources published in the southern Indian states used Hindi words more often than they did words from any of the Dravidian languages. While this finding is an extremely interesting one and warrants further research, this is beyond the scope of the current paper.

What is interesting about this register is the relative lack of Discourse Markers. Given the informality of the language in this register, I expected it to more closely resemble a spoken register like Conversational English in this regard. Four instances of Discourse Markers were found in this register of the corpus, and all were in Hindi.



The interesting class of words in this register is Other. On examining this category, two things become obvious. The first is that Hindi is a popular language to use. In fact, in this register more than 90% of the words in this category, regardless of their source, were in Hindi. The second fact, however, is the more interesting. In this register speakers seemed to make a stylistic choice to use Hindi instead of English in a lot of contexts. Also, this is the only register examined in the first analysis with an example of Kachru's hybridisation; although Kachru claimed that hybridisation is a "very productive" process, the current analysis proves otherwise. Examples of sentences with Hindi words used in this manner include the following:

- He succeeds in *patoing* Rohit's girlfriend (Hindi; *get around her by buttering her up*; FENT2 from CCIE).
- At a *tehzeeb-bhari* soiree, Sudha and Suri took us to the residence of *pukka* Hyderabadites. (Hindi; *culture-filled*; *perfect*; DHent2 from CCIE)
- That's surely worth a *dekho*. (Hindi; *a look*; FEent4 from CCIE).
- For more *masti* and *mazaa*, on to Kamal Hasaan. (Hindi; literally, these two words mean *spice*. Used here, they refer to gossip about a certain film star; BMEnt3 from CCIE).

The first sentence illustrates the process of hybridisation where the verb used is a Hindi one – *patao* – loosely translated to *getting around her by buttering her up*. This word is used in the progressive form, with the English morpheme *-ing* attached to it. The other sentences illustrate what Kachru terms "lexical innovations... with at least one item from a South Asian language and the one from English" (*The Alchemy of English* 42). The phrase *tehzeeb-bhari* soiree translates to *culture-filled* soiree, while the phrase *a dekho* translates to *a look*. In all these sentences, there are simple and obvious English words for the words in Hindi. Reading these sentences in this context, however, it becomes clear that the writer uses these words for effect. This is particularly clear with the last sentence, where the word *dekho* in Hindi is the verb *to see*. In the sentence, however, it is used in place of the English noun *look*.

Written Entertainment News is different from other registers in one final way. Many of the Indian words used in this register are very register-specific and would not have an equivalent translation in English. The occurrence of certain words in this register shows that the film industry in India follows certain protocols on completion of films. All the people involved in the making of a movie go through a certain procedure (often with religious components) before the movie is first shown. Words such as these include the following:

- The *mahurat* of the film will be early next month. (*opening ceremony*; BMEnt1 from CCIE)
- It starts with the *nariyal* breaking ceremony. (*coconut*; FEEnt2 from CCIE)

What is also interesting about this register is that unlike the other written registers discussed thus far, Written Entertainment News does not provide translations for the words, or separate them in any way from the rest of the discourse with the use of quotation marks or italics. This is true of both Larger Chunks of Language and other Indian words. This is also true of film and song titles, though they are not discussed in this analysis. When Indian words occur, they simply form a part of the discourse. Below are sentences from all semantic categories (except for Other, which have been discussed above) found in the register.

- She demands *nimbu pani* right away to announce her pregnancy. (Food; *lime juice*; Hindi; FEEnt4 from CCIE)
- Pick out a handloom *kurta* to beat the sun-n-sweat. (Clothing; *shirt*; Hindi; FEEnt7 from CCIE)
- The song and dance in the *antakshari*, the *nok-jbok* between the lead pair... (Arts; *musical game; ribaldry*; Hindi; FEEnt14 from CCIE)
- ... for the *mangalsurta* or those shiny bowls of *sindoor*... (Religion; *necklace worn by married women, sacred vermilion*; Hindi; FEEnt4 from CCIE)
- All those *desis* out there in *Amrika*, get ready! (People; *Indians*; Hindi; an Indian version of *America*; FEEnt18 from CCIE)
- Awards, *naah?* For starters, we had to get... (Discourse markers; *Isn't that right?* Hindi; FEEnt7 from CCIE)
- *Kya biddu, itna star ka beta hain, magar dialog aisa bolta hain!* (Larger Chunks; *what fellow, you are the son of such a big star, but you speak the dialog like this!* Hindi; FEEnt11 from CCIE)

Given all these features, Written Entertainment News is the one that shows the greatest degree of Indianisation, and is by far the most interesting register to analyse for Indian words. It is these facts that prompted the second analysis – the analysis of a 40,000 word corpus of Written Entertainment News compiled in 2016. The objective was to determine if there were any differences between the language used in this register with the passage of time.

#### *Written Entertainment News in 2000 and 2016: A Comparison*

Table 2 below shows the distribution of different semantic categories of Indian words in the Written Entertainment News register compiled in 2000 and that compiled in 2016. The last column of the table shows a startling difference in the frequency of use of Indian words in the two registers. A closer examination of

the table reveals that with the exception of Politics, all semantic categories of words have fewer Indian words in 2016 than they did in 2000. While no claim is being made here about the significance or lack of significance (statistically) of these differences, one does wonder why, with the passage of time, this register of Indian English might be less Indian.

Table 2: Indian words in Written Entertainment News in 2000 and 2016; all counts normalised to 1,000,000

	F	C	A	R	P	DM	G	Pol	LC	Total
2000	232	151	2211	174	289	46	0	12	255	3370
2016	150	125	100	150	200	25	0	25	150	925

F - Food; C - Clothing; A - Arts; R - Religion; P - People; DM - Discourse Markers; G - Greetings; Pol - Politics; LC - Larger Chunks of Language

A close examination of the texts of the register compiled in 2016 reveals other differences. As mentioned earlier, in the texts of Written Entertainment News compiled in 2000, Indian words were not separated from the rest of the text in any way. They were not translated or italicised and there were no quotation marks used to separate them from the rest of the text. In the texts from 2016, however, this is no longer the case. Movie and song titles were invariably italicised and Indian words in the different semantic categories studied occurred within inverted commas. Further, in many circumstances, translations of the words were provided. It is also noteworthy that the 40,000 words came from 110 different files, and out of this total, only 42, or 38% had Indian words in them at all. For the corpus compiled in 2000, there were 38 files (much longer in length), and all of them had Indian words in them. A similarity between the two corpora is that both had many words in the Other category, and with both decades, most of these words were in Hindi.

Another noteworthy difference, and one that warrants further study, is the subject matter of the texts studied in the different time periods. While the corpus compiled in 2000 spoke almost exclusively about Bollywood and the other smaller film industries in India, the corpus compiled in 2016, with files from the same sources as those in the earlier corpus, discussed various topics such as the “hottest” restaurants in town, what fashion choices various film personalities were making, or matters as seemingly trivial as what the contents of a film personality’s purse were. Further, while certain subjects such as extra-marital relationships and sex were not discussed in the earlier publications, several articles in the 2016 corpus dealt with these issues. What this trend seems to point to is a movement of entertainment publications in India towards similarly oriented publications in the West. As topics in the Indian entertainment publications become more similar to those discussed in the West, the language used in them is also found to be more similar to the English used in the western entertainment

register. In other words, the less Indian the topic, the less Indian the language seems to have become. The endonormative stability of Indian English lexis suggested by Lambert that existed between 1930 and 2000 seems, at least based on this very preliminary analysis, to have been destabilised somewhat in certain contexts.

As we have done for the analyses discussed previously, following are sentences illustrating what has been described above.

- Avoid including dishes that have leafy veggies, like *palak paneer*, green chicken, etc. on the menu. (Food; *spinach and cheese*; Hindi; Femina6)
- I'd buy a Raw Mango *sari* from their shop-in-shop at Cinnamon on Ulsoor (Clothing; *feminine garment*; all languages; Femina5)
- This one is a heart touching *qawwali* about unrequited love. (Arts; *genre of music*; Hindi, Urdu; Deccan Herald7).
- "Sultan" will hit the theaters this *Eid* (Religion; *Muslim festival*; Hindi, Urdu; IndiaTV News9)
- Those days Amitabh, Jaya Bhaduri (he was dating her then) and Gulzar *saab* formed a lovely trio. (People; *sir*; Hindi; Filmfare3)
- People are saying that *arey* he came back. (Discourse Markers; *oh wow*; Hindi; Filmfare12)
- They know that when there is a shot, they will have to be silent. *Wo art ko samajhte hain*. They have tremendous respect of actors. (Larger Chunks; *they think that is art*; Hindi; Filmfare7)

### Discussion and Conclusion

That English on the Indian subcontinent takes on local colour to accommodate local needs is undeniable. That these local needs are register specific is also undeniable. As Kachru (1986) points out, matrimonial advertisements in India are necessarily culture-bound, and the English used in the construction of such an advertisement is necessarily Indian in its use of certain lexical items. Matrimonial advertisements, then, naturally have a preponderance of words that relate to religion and caste, and this is also illustrated in Mesthrie and Bhatt. While previous research provides no systematic account of the kinds of Indian words in various registers, the current study attempted to do just that. The first analysis of three written and three spoken registers exemplifies this idea that certain registers of Indian English are necessarily more Indian than are others. Further, the semantic categories of words in the six registers that were analysed do correspond with what previous research on Indian English lexis says: that the lexical items present in the English in these registers pertain to local customs and culture.

The second analysis, however, seems to indicate that with the passage of time, the English used in certain written contexts in India is undergoing what appears to be a process of un-Indianisation. The analysis also raises the question of whether the English used in certain registers is becoming less culture-bound. Y. Kachru (2006) cautioned that “the influence of global culture is all-pervasive in the audio-visual media” (226). The seeming reduction in Indianisms between 2000 and 2016 could serve as an example of just how pervasive the global culture, with the USA “as the leading economic and military superpower and the main agent of today’s economic and cultural globalization” (Schneider 1), is, and how, at least in the written medium, India recognises this influence and works at using a language that is accessible to as large a global population as possible. Schneider (2007) sums up this idea succinctly and explains that “as long as the *raison d’être* of Indian English is still essentially utilitarian and it is not a medium for community solidarity, the language is not likely to change its character and status in the near future” (173). What Kachru (2005) said of South Asia – “The blending and hybridization of two or more linguistic systems is accomplished with dexterity and immense effect” (2) – might be true of spoken Indian English, but not of written registers of Indian English, and the preliminary diachronic investigation conducted in the present paper has posed more questions than it has answered.

### **Limitations and Directions for Future Research**

The conclusions reached in this paper need to be vigorously tested with larger corpora and more registers. While this tentative conclusion can be reached from the current analysis, future research needs to examine spoken registers of Indian English from a diachronic perspective. Further, an examination of the use of Hindi over other Indian languages also needs to be thoroughly examined.

### **Works Cited**

- Balasubramanian, Chandrika. *Register Variation in Indian English*. Amsterdam: John Benjamins, 2009.
- Bakshi, Raj Nath. “Indian English.” *English Today* 7.3 (1991): 43-46.
- Bansal, Ram Krishna. “*The Intelligibility of Indian English*.” Eric Document # 108284 (1976).
- Baumgardner, Robert. “Pakistani English: Acceptability and the Norm.” *World Englishes* 14 (1996): 261-71.
- Bhatt, Rakesh. “Optimal Expressions in Indian English.” *English Language and Linguistics* 4 (2000): 69-95.
- Crystal, David. *English as a Global Language*. Cambridge: Cambridge UP, 2003.

- D'Souza, Jean. "Indian English: Some Myths, some Realities." *English World Wide* 18 (1997): 91-105.
- Dubey, V. "The Lexical Style of Indian English Newspapers." *World Englishes* 10 (1991): 19-32.
- Dustoor, Phiroz E. *The World of Words*. New York: Asia Publishing House, 1968.
- Hosali, Priya. "Some Syntactic and Lexico-semantic Features of an Indian Variant of English." *Central Institute of English and Foreign Language Bulletin* 3 (1991): 65-83.
- Hundt, Marianne and Ulrike Gut. *Mapping Unity and Diversity World-Wide: Corpus-Based Studies of New Englishes*. Amsterdam: John Benjamins, 2012.
- Kachru, Braj. "The Indianness of Indian English." *Journal of the International Linguistic Association* 21.3 (1969): 391-423.
- . *The Alchemy of English: The Spread, Functions, and Models of Non-Native Englishes*. Urbana: University of Illinois Press, 1986.
- . "The Sacred Cows of English." *English Today* 4 (1988): 3-8
- . *Asian Englishes Beyond the Canon*. Oxford: Oxford UP, 2005.
- Kachru, Yamuna. "Mixers Lyricizing in Hinglish: Blending and Fusion in Indian Pop Culture." *World Englishes* 25.2 (2006): 223-33.
- Kachru, Yamuna and Cecil Nelson. *World Englishes in Asian Contexts*. Hong Kong: Hong Kong UP, 2006.
- Lambert, James. "Diachronic Stability in Indian English Lexis." *World Englishes* 00.0 (2014): 112-127.
- Lange, Claudia. "Focus Marking in Indian English." *English World Wide* 28.1 (2007): 89-118.
- Mesthrie, Rajend. "South African Indian English." *English Today* 34 (1993): 12-16.
- Mesthrie, Rajend and Rakesh Bhatt. *World Englishes: The Study of New Linguistic Varieties*. Cambridge: Cambridge UP, 2008.
- Schneider, Edgar. *Postcolonial English: Varieties Around the World*. Cambridge: Cambridge UP, 2007.
- Sedlatschedk, Andreas. *Contemporary Indian English: Variation and Change*. Amsterdam: John Benjamins, 2007.
- Verma, Shivendra Kumar. "Swadeshi English: Form and Functions." *Indian English* 41 (1980): 73-84.