# The pluperfect in native and non-native English: A comparative corpus study

Devyani Sharma

**ABSTRACT**

One of the challenges in characterizing non-native varieties of English is accounting for variant uses of ostensibly standard English forms. The present corpus study examines both quantitative and qualitative aspects of pluperfect use in Indian English (IndE), British (BrE), and American English (AmE). IndE is found to differ from native usage by additionally associating had+V-ed with present perfect and preterite meanings. Licensing of pluperfect contexts by time adverbials is also found to be significantly lower in IndE. AmE shows the lowest overall use of the pluperfect and the highest use of disambiguating adverbials. Thus, AmE and IndE show distinct paths of divergence from BrE. Variation within IndE exhibits a tendency for greater non-nativeness in regional (vs. national) press and in the bureaucratic (vs. press) registers, suggesting a multidimensional distribution of IndE nonstandardness in India. In a qualitative analysis, these nonstandard uses are shown to convey new pragmatic meanings deriving from ambiguity in the native system. Finally, these changes are evaluated in relation to the broader TMA system of IndE as well as TMA systems of other non-native Englishes which exhibit similar characteristics.

## Introduction

This paper examines a case of dialectal variation in a subpart of the tense-modality-aspect (TMA) system of Indian English (IndE). Rather than examining nonstandard forms in the dialect, which are relatively few, I examine new functions associated with an existing form, namely the use of the pluperfect had+V-ed construction, and contrast this usage with that found in native British (BrE) and American English (AmE). Research on non-native Englishes and creoles has recently begun to focus on the phenomenon of variation in the meanings associated with standard English forms. Shastri's (1992) survey of existing studies on features of Indian English observes that "the identification of variant functions of the same forms in Indian English is the immediate task before us" (Shastri, 1992:274). Bickerton (1980:113), in his discussion of decreolization, similarly notes that the development of new functions for existing forms must be treated as a distinct mode of change from the acquisition of new forms for existing meanings or functions.

Due to the complexity and variability of English use in multilingual settings, as well as a delayed recognition of systematicity in post-colonial, indigenized varieties of English (Williams, 1987; Kachru, 1992), quantitative studies on IndE and other non-native varieties are much less common than descriptive overviews of dialect features (some exceptions include Leitner, 1991; Sridhar, 1991; Shastri, 1996). In the present study, I explore techniques for quantifying functional and pragmatic variants of the pluperfect form in IndE. Perhaps even more than structural innovations, it is this class of variants that constitutes the basis of divergence from native usage in such dialects. I examine a corpus of present-day Indian English print texts for such a shift in usage, and employ both a quantitative and qualitative comparison with corpora of two native varieties of English — British and American — to establish the nature and degree of this change.

The central finding here is that the pluperfect in IndE may additionally be used with either a preterite or a present perfect meaning (Cf. Rickford and Rafal, 1996 for a similar phenomenon in AAVE). In addition, the comparison to native varieties reveals cross-dialectal variation in the use of this construction and the explicit disambiguation of its meaning (via time adverbials, for instance) across different varieties of English. In particular, AmE and IndE diverge from BrE usage of pluperfect in distinct ways; AmE exhibits greater restrictiveness along certain

dimensions, while IndE has extended the use of the construction beyond the BrE discourse constraints.  The new usage in IndE appears to derive from a nonstandard pragmatic viewpoint signaling and tense reorientation.  While language transfer from native language(s) is commonly cited as a source of non-native variation, I suggest that ambiguity inherent in <u>native</u> usage can play as important a role.  The inter-dialectal comparison is followed by an examination of variation within the IndE data as well, to provide a more detailed picture of the distribution of nonstandardness across distinct segments of IndE speech communities.

I also relate pluperfect usage to other features of the IndE TMA system, including the present perfect and the modals <u>would</u> and <u>could</u>. Innovations in their use resemble the shift in use of the pluperfect in certain ways, and thus may constitute part of a larger network of reorganization within the TMA system, rather than being isolated sets of alternations (Harris, 1984; Winford, 1993).  Finally, I relate these findings to TMA patterns in other nonstandard varieties of English, such as Trinidadian English, Singapore English and AAVE.

In the first section, I provide a brief overview of the introduction, spread, and currently conflicted status of English in India, along with certain structural characteristics of IndE identified previously in the literature.  A description of the corpus data and methodological approach is followed by the quantitative data on dialectal variation among the three dialects, as well as intra-dialectal variation within IndE, based on register and regionality.  The qualitative discussion examines the pragmatic meanings that accompany non-standard uses.  Finally, I locate these findings within the broader TMA system of IndE and developments in other non-native varieties.

# 1  English in India: history and structure

Kachru (1986:19) and Williams (1987:162) use the phrase "non-native institutionalized variety" for English that is rarely the mother tongue of speakers but which is acquired locally as a second language through educational institutions.  In these settings, non-native speakers may ultimately become most proficient in English or may use English in the widest range of functional domains, leading to a continuum that approaches nativeness at one extreme. Thus, although I follow the conventional use of "non-native" here, it is necessary to bear in mind that the native/non-native distinction in linguistic theory tends to reify continuous phenomena into discrete categories.

Many of the significant structural features of post-colonial pidgins, creoles, and non-native varieties of European languages can be systematically traced to the particular social and historical conditions of their emergence (Bickerton, 1984; Rickford, 1987; Bokamba, 1991; Mufwene, 1996). This section includes a brief overview of the historical conditions out of which IndE developed, along with a summary of previously noted TMA features.

Mehrotra (1998) identifies three broad periods of English in India: 1578-1834; 1835-1947;1947-present.  During the initial 250-year period, when English was first introduced into the subcontinent, the main domains of its use were trade, the military, and missionary work. The second phase of English in India began with an ordinance issued in 1835, which decreed that English be the medium of all schools and universities in India. The ordinance was based on a recommendation by T. B. Macaulay entitled <u>Minute on Indian Education</u>, which proposed the formation of "a class of persons, Indian in blood and colour, but English in taste, in morals and in intellect" (Kachru, 1983:22).  As in many post-colonial regions, the long-term result of this policy was an over-arching, class-based politicization of the English language throughout the country, leading to divided linguistic ideologies.[1]

The third, current period of the establishment of English in India begins immediately after independence in 1947.  As many Indians ceased to aspire to a British acrolectal standard in their English, the ideological distance from the native variety and the range in bilingual competence expanded.  This is the crucial point at which indigenized, non-native varieties of English come to be distinct from regular second-language acquisition (Sridhar et al., 1992).  Indian languages and vernaculars were initially favored as mediums of instruction; however, today English has not only remained in India, it has also spread further across various domains of society against the predictions of many at the time of independence. English dominates mass media and the publishing world and it is also a central, functional language in domains such as higher education, administration, court proceedings, and politics. Nevertheless, English has not displaced many indigenous languages, but rather plays a functional role in a complex, multilingual arena.[2]

This outlined history gives some indication as to why the language has persisted post-independence.  The roots of English were set early and deep; many current institutions including universities, administrative organizations, and the media were first established in English. The linguistic politics of the region offers no uncontroversial option as a national substitute language and the national infrastructure as well as the global economic bias toward English increases the demand for the medium.   At the same time, English is a unique linguistic marker of class divisions and power asymmetries in the nation, as is the case in many former British colonies (Dasgupta, 1993; see also Spitulnik (1998) for a discussion of this phenomenon in Zambia).

Few speakers of IndE still aspire to a native standard as their acrolect, as demonstrated in Sahgal (1991).  The range of semi-autonomous domains of English use sketched here suggests that multiple standards might exist for IndE, rather than a single scale of greater or lesser nonstandardness.[3]  In addition, regional standards with distinct sets of nonstandard features also seem to exist, partly influenced by widespread multilingualism.  These two issues — domains of use and regional variation — form the basis of the intra-dialectal section of this study.

Early studies of IndE tended to be restricted to prescriptive or pedagogical descriptions based on deficiency- or error-oriented approaches (see Kachru (1992) for a detailed discussion). In response to these relatively limited investigations, functional approaches developed more recently attempt to contextualize forms and patterns found in IndE within (a) their sociolinguistic context, (b)  their functional role within a multilingual setting, and (c) their possibly innovative use, as opposed to their failure to match a target native standard, which may have earlier been assumed to be a result of incomplete L2 acquisition (Sridhar and Sridhar, 1992).

Some of the more commonly noted features of IndE include null or extraneous articles, null arguments, uninverted questions, reduplication, and nonstandard verb-particle constructions, preposition use, and verb selectional restrictions.  Many observed structural features vary considerably from speaker to speaker based on their relative proficiency in English as well as the influence of their native language(s).  The TMA usage discussed in this study are more systematic and quantifiable than some IndE features, as they apply to a small set of forms that show regular variation in both spoken and written registers.

Leitner (1991) discusses several attested and speculative features of the IndE TMA system. Comparing modal use in the Kolhapur corpus with the Brown (American) and LOB (British) corpora, he found that shall and should occur with greater frequency in the Indian corpus, suggesting that this may be related to the common observation that IndE retains a high number of formality markers;  would, will, might, and ought appear less frequently in IndE.  In suggesting further research, Leitner observes that "past perfect might signal the notion of 'remote past' in Indian English" (Leitner, 1991:228).  A text corpus example is given in (1a).  The spoken IndE example in (1b), from personal interview data, has the meaning: My parents just visited; we saw them off last week. The present study aims to capture the attested range of meanings associated with this construction, but is not limited a priori to remote past usage, despite Leitner's speculation.

(1)    a.    This is the second time that such an object **had been** sighted here.        (KC A13 142)
       b.    My parents just visited. Just last week we **had seen** them off.  All of them have their
             visa, so they all are, like, coming and going.

With regard to the present perfect, Williams (1987) provides the IndE sentence in (2) to contrast native and non-native uses of the English present perfect construction.

(2)         I **have read** this book last month.                                    (Williams, 1987:183)

In general, the present perfect in Standard English cannot co-occur with punctual time adverbials (Comrie 1976:54).  Williams attributes this non-standard use to a 'lexicalization of completion', whereby the perfect form indicates completion of the activity rather than indicating the continuing impact of a past event on the current situation.

Finally, the use of progressive aspect with stative predicates to signal a current, ongoing state is one of the more salient and frequently noted features of IndE (Williams, 1987:172; Leitner, 1991; Kachru, 1992; Richter, 2000). An example from the Kolhapur Corpus is given in (3).

(3)         [The ministers] are fully corrupted. They **are having** vindictive motives.   (KC H31 1630)

Williams further notes that the use of progressives with stative verbs also occurs in West African varieties of English, attributing it to a regularization of the verbal system, and Harris (1984:56) cites certain Celtic dialects as also exhibiting this usage.  This feature is noteworthy due its typological rarity cross-lingusitically.  Though it is not particularly common in the Kolhapur Corpus, as noted in Richter (2000), it can be found in spoken usage (e.g. 'I can't ask them about it; they may be having some bad feelings.' from personal interview data).

Changes in the use of the perfect, the progressive, and modals in non-native Englishes may either be explained in terms of substrate transfer or in terms of a system-internal regularization tendency, or a combination.  The latter of the two processes may help to account for the occurrence of these features across unrelated dialects.  I return to these issues after presenting and analyzing the corpus data for the pluperfect.

## 2   Data and Methodology

The debate over models for the study of non-phonological variation has taken various forms over the past few decades (Cf. Sankoff, 1973; Lavandera, 1978; Romaine, 1984; Cheshire, 1987 amongst others).  Central to the debate is the question of referential meaning, which arises out of any attempt to straightforwardly extend methodologies for the study of phonological variation to syntactic variation.  The fundamental problem lies in identifying structural variants with exact semantic equivalence. In this study, rather than isolating two forms — one standard and one variant — I examine the non-equivalence in meanings associated with one particular form.

For the purposes of gaining an overview of pluperfect use in IndE, I divide the quantitative section of this study into two parts.  In the first part, I compare the distribution of uses of had+V-ed in IndE to its distribution in two native varieties: British and American English.  This is a necessary initial step, as some earlier studies of syntactic variation in IndE assume an overly strict interpretation of how the variable in question is used in the native standard variety.  Three assumptions are inherent in such studies:  (a) that the proposed variable does in fact exhibit systematic variation in contrast to native varieties; (b) that the standard is not variable; and (c) that there is a fixed standard equivalent for a given form in the non-native variety.  All three assumptions are problematic.  An idealized reification of native and non-native varieties can misrepresent both varieties and, most importantly, mask the possible relationships between variation in forms and functions in both (Cf. Mufwene, 1996).

With the recent advent of sufficiently large databases, some of these issues in the study of dialectal syntax are now partially resolvable.  Shastri (1996:79), in his corpus study of infinitival to-complementation showed that variation of the type he examined in IndE also existed to a lesser degree in native English.  If non-native uptake and expansion of existing variation in native usage is an important mechanism of change, capturing such native variability is crucial to any analysis of non-native usage.

The second part of the quantitative study examines intra-dialectal variation in IndE usage.  The qualitative discussion that follows turns to the pragmatic basis of the new usage, and its relation to other IndE TMA properties as well as other nonstandard TMA systems.  Before introducing the results, I provide some background information on the English past perfect, the electronic corpora, and the coding criteria used in the study.

### 2.1  The Past Perfect construction

The perfect construction indicates relational anteriority and relevance to a deictic zero point (Comrie, 1976:53; Binnick, 1991:161; Bybee, Perkins & Pagliuca, 1994:55).  The English past perfect construction is most generally used to refer to an event which occurred before a particular point in the past.  In one sense, alluded to by Comrie (1976:53), perfect aspect — whether past, present, or future — relates to two time points while tenses such as past or present simply refer to one, with no implication that it bears a necessary relation to others.  An intuitive visual representation of such generalizations is provided in Reichenbach (1947:290).  In his discussion of systems of time reference, Reichenbach notes that in order to distinguish among types of past and present tenses one must make reference to three, not two, time points.  In the diagram below, the pluperfect, preterite, and present perfect are contrasted in terms of the time points R (reference point), E (event point), and S (speech point).

a.    *Past Perfect*

I had seen John

b.    *Simple Past*

I saw John

c.    *Present Perfect*

I have seen John

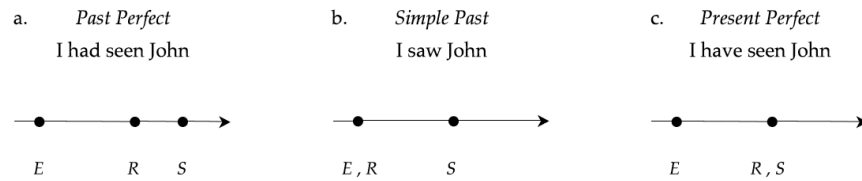$E$    $R$    $S$          $E, R$    $S$          $E$    $R, S$

FIGURE 1. Tense and aspect time relations in English (Reichenbach, 1947).

The distinguishing feature of the past perfect is that the Reference Point is distinct from both the Event Point and the Speech Point. This R point may be signaled to varying degrees of explicitness in the clause or the immediate context, which is a crucial issue in the interpretation of the past perfect.

In the aspectual system of English, a distinction is made between past and present perfect; as the term "past perfect" indicates, the construction could be seen as a straightforward compound of past tense and perfect aspect. Dahl (1985) argues against this view of the pluperfect as a compound. He observes that the pluperfect "may develop secondary or extended uses which are not characteristic of the perfect by itself" (e.g. counterfactuals) and that there seem to be "less strict constraints as concerns use with definite time adverbs in the pluperfect than in the Present Perfect" (Dahl, 1985:144). He also notes that the pluperfect occurs in some TMA systems that lack a regular perfect, as in French:

(4)    a.    Elle **a lu** ce livre.
            SHE HAS READ THIS BOOK = 'She read this book'          (SIMPLE PAST)
       b.    Elle **avait lu** ce livre.
            SHE HAD READ THIS BOOK = 'She had read this book'      (PLUPERFECT)

In such languages, the present perfect does not always convey the compound meaning of present tense and perfect aspect. Rather, the present perfect has partly or completely replaced the function of simple past. Comrie (1976) suggests that such shifts in correlation between form and function of TMA marking result from the "gradual relaxation" of such requirements as degree of recentness and presence of adverbials. Precisely these requirements are directly identifiable as differing in strictness between the dialects examined here, suggesting that these are some of the mechanisms of change in tense-aspect meaning.

Thus, the innovative use of the pluperfect for marking a more general remote past is a functional extension that has some cross-linguistic support. Given this typological pattern, it is important to consider the possibility of substrate language transfer influencing the direction of this development, in this case from Indo-Aryan languages. One characterization of the Hindi past perfect suggests that it expresses "the completion of an action before a definite moment [or an action] in the past, and also emphasizes the <u>remoteness</u> of the action from the present moment, its belonging to a finished segment of time — last year, yesterday, yesterday morning etc." (Katenina, 1960, my emphasis). Some Dravidian languages also show an extension of the perfect construction to past meaning. For instance, Sridhar (1990:229) observes with regard to Kannada that "it is worth noting that the present perfect form is used in Kannada in some of those contexts where the simple past would be appropriate in some languages such as English. This leads to the very common use by Kannada speakers of English of sentences such as 'I have bought the book yesterday' for 'I bought the book yesterday'."

The typological space of possible form-meaning pairings across tense-aspect systems, and particularly in substrate languages, constitutes an important backdrop for studying the emergence of non-standard systems. However, this does not necessarily mean that the shift evident in the data here is exclusively a result of language transfer. The shift of pluperfect use for preterite meaning, discussed in detail below, can be observed in unrelated dialects of English which have little or no mutual contact and different substrate languages. Furthermore, the qualitative analysis presented in this paper identifies some potential sources of IndE usage which derive from native English ambiguity.

## 2.2  The corpora

In order to establish whether the pluperfect in IndE varies systematically from native varieties at all, the first part of this corpus survey compares the use of had+V-ed in the press register in IndE to its use in comparable BrE and AmE press registers. Rather than assuming a priori that the use of had+V-ed deviates from native standards, this comparison can establish whether there is a difference in usage and additionally, whether any comparable variability in usage occurs within and among the native varieties themselves. The study is then extended to bureaucratic texts in the IndE corpus, in order to examine register variation within IndE.

The use of written rather than spoken data faces the problem of a generally lower frequency of many interesting nonstandard forms.  In particular, some of the structural variants listed in (1) show a high frequency in vernacular speech but are often too infrequent or unsystematic in print media to be quantifiable.  However, the most important motivation for selecting a text corpus is that the collected registers represent established domains of English use in India.  The analysis of speech data would be rendered problematic by the range of bilingual proficiency levels.  This problem of distinguishing between L2 proficiency limitations and stable nativizations of the variety is, to a large degree, avoided by the restriction to print media here.

The machine-readable corpora that are examined here are the Kolhapur Corpus of Indian English (texts from 1978), the Brown Corpus of present-day American English (texts from 1961), and the Lancaster-Oslo/Bergen (LOB) Corpus of present-day British English (texts from 1961). All three corpora were obtained from the International Computer Archive of Medieval and Modern English (ICAME) and were specifically selected because they are comparable in terms of design, text selection from various registers, and size.[4]  The LOB Corpus was created as a counterpart to the Brown Corpus and the Kolhapur Corpus is similarly designed for comparative studies of the three varieties.  All three are million-word corpora, containing 500 texts of approximately 2000 words each, distributed across 15 text categories.

The press register forms the basis for comparison in the first part of the study, mainly in order to maximize the range of contexts for the use of had with various time references, but within a comparable register range.  Also, Kachru refers to the national newspaper register as a "nativized" register (1996:22), implying that English is one of the languages that has been associated with these domains of society for long enough to genuinely reflect a nativized version of the language.

For the purposes of the inter-dialectal press register comparison, only data from the press reportage subsection ("A") of each corpus were extracted.  This comprises a set of 44 texts per corpus, with approximately 2000 words per text, amounting to a 88,000-word subcorpus.   Within the IndE corpus analysis, the IndE data from the press reportage section is contrasted with the bureaucratic register, comprised of data taken from the H ("Miscellaneous: Government Documents, Foundation Reports, Industry Reports") category of the Kolhapur Corpus. This is a sub-corpus consisting of 37 texts containing a total of 74,000 words.  While the restriction to these two registers limits the total corpus word-count to much less than 1 million, the number of actual tokens extracted and examined from each variety for the analysis is substantial enough to draw robust generalizations.

Aside from the cumulative effect or density of various types of co-occurring nonstandard features, much of the recognizable 'Indianness' in many of the registers in the Kolhapur Corpus lies in the wordy literary style and formality features that have often been cited as additional characteristics of IndE.  These features include lengthy or wordy prose style, Latinization (for instance, demise rather than death), excessive use of archaic forms and idioms, and a high frequency of formal style markers.

**Table 1:** Rate of occurrence of selected contracted forms in the press registers (sections A,B, and C) of three corpora

| CONTRACTED FORM | AMERICAN | BRITISH | INDIAN |
|---|---|---|---|
| will  —>  'll | 5.4% | 2.5% | 0% |
| have  —>  've | 4.0% | 1.9% | 0.1% |
| not  —>  n't | 28.7% | 18.2% | 0% |

An example of this relative stylistic formality is in Table 1, which shows the degree to which certain forms are contracted in press reporting in the three varieties of English. IndE shows a consistently lower use of contracted forms than native varieties; Biber (1987) and Hundt (1996) report a similar difference between BrE and AmE contraction in written registers. Many of these contracted forms in native British and American press reporting occur within direct quotes. Interestingly, IndE appears to paraphrase direct quotes much more frequently than either AmE or BrE, which may account for the lack (or avoidance) of contracted forms. This difference in rates of indirect speech reporting also plays a role in increased pluperfect use, as will be shown in the quantitative section.

**Table 2:** Average sentence length (words per sentence) in three corpora

| REGISTER | AMERICAN | BRITISH | INDIAN |
|---|---|---|---|
| Press (reportage) | 7.9 | 8.7 | **10.6** |
| Press (editorial) | 8.1 | 9.2 | **10.8** |
| Skills, Trades, Hobbies | 8.3 | 9.9 | **11.1** |
| Scholarly/scientific writing | 8.3 | 9.8 | **10.3** |
| Detective Fiction | 8.8 | 10.1 | **11.5** |
| Humor | 8.7 | 10.3 | **11.2** |

Another illustration of IndE stylistic formality is in Table 2, which shows that IndE uses longer sentences in all registers, without exception. Furthermore, genre differences in average sentence length are similar within each of the three dialects. Interestingly, BrE occupies an intermediate position and AmE appears to use shorter sentences, on average, in every register. This pattern, in which AmE and IndE represent the two distinct patterns with respect to BrE, is reflected in several of the results of the present study, as will be seen in the analysis that follows.

Although these examples indicate a formality and perhaps standardness in IndE deriving from an earlier high prestige variety of BrE, variation across lexicon, morphology, syntax and semantics evident in the corpus also shows an unambiguous indigenization of the variety. A simple example is code-switching and borrowed forms in the IndE corpus. For the same amount of text, the national press data in fact contains more non-English lexical forms (170) in comparison to regional press (159); although the range of lexical items is greater in the regional press. Thus, even this relatively standard and formal register of press is firmly indigenized in certain respects. In summary, then, the three corpora are very similar in terms of size and types of texts. While the Kolhapur corpus resembles the native corpora in structure and content, it exhibits a number of distinguishing structural characteristics as well, in terms of both formality and indigenized traits.

## 2.3 Coding criteria and variables

In looking at the use of had+V-ed across the three corpora, my main interest is in extracting the range of variability in meanings associated with use of the standard pluperfect construction. In other words, I am interested in the association of nonstandard meanings with a standard form. To this end, I coded for the semantic function of had+V-ed and also for various internal and external independent variables. In Shastri's (1988:40) terms, this identifies "opaque", rather than "transparent" featuures. Transparent features, such as borrowings or code-mixing in IndE, clearly signal the presence of non-nativeness. By contrast, "opaque" features are not readily observable; he suggests that for opaque syntactic, semantic and pragmatic features, "it is perhaps not the form that is at variance but the function" (Shastri, 1992:274).

In order to delimit the appropriate dataset, all tokens of had+V-ed were extracted, leaving out all non-pluperfect uses of the form had. This includes the following uses of had: obligative (I had to be at the airport on time), possessive/stative (They had two seconds remaining on the clock), causative (She had the offending item removed), and counterfactual/subjunctive (If I had

known, I would have left sooner) uses.  Of these, only counterfactuals may potentially relate to the use of the pluperfect. For instance, in AmE, the general reduction in the use of the pluperfect construction, as observed in Elsness (1997:229), may be accompanied by an increase in use of the simple past form in the protasis of counterfactuals.  However, an exploratory comparison of counterfactuals in the three corpora did not reveal any noteworthy patterns.  The rates of occurrence of counterfactuals for the press subsections A, B, and C in the three corpora were low and not notably different.  An additional complication in including them is that they do not share the temporal frame of the pluperfect; their discourse context, tense environment and time adverbial requirements — features which are coded for each pluperfect example in the data here — are distinct. While it may be fruitful to conduct a close analysis of counterfactual usage in these three varieties, the discussion of its role in the changing uses of had+V-ed is restricted here to the qualitative discussion of IndE TMA features.

Each remaining instance of had+V-ed was examined in its textual context to ascertain its intended meaning in that context. This judgement was more subjective than the usual coding of structural variables, which can be objectively classified. However, since the variation in question is in meaning, not form, this approach was necessary (See Harris 1984:46 for a discussion of methodological issues such as the need for subjective coding and the infrequency of tokens in the study of syntactic variation).[5]

Most importantly, judging the context of each token, regardless of whether it occurred in a native or non-native text, allowed for variability in use within native varieties as well.  As it turned out, very few tokens of the native variety occurred in contexts where a pluperfect meaning was not inferable.  The fact that even those few do occur in native usage, however, is important to the study and is discussed in the analysis. The three standard English meaning equivalents found with the had+V-ed construction are listed in (5).

(5)    a.    PLUPERFECT (standard meaning)
       b.    PRESENT PERFECT
       c.    PRETERITE

(5a) refers to the standard native meaning. The other two variants were other meanings that were found to be associated with the had+V-ed construction in the data.[6]  The method of establishing the meaning of the had+V-ed construction in each example involved a close examination of the discourse context in which it occurred for a clear indication of a distinct reference point (R) and event point (E) to license the pluperfect meaning. For example, in the standard use in (6), the past event (E) is the taking up of reins by General Rafael.  The announcement of the arrests represents an intervening reference point (R) between that event and the current speech event (S), namely the report itself.  Finally, the past (R) event is embedded within the adverbial clause "a few hours after …", which unambiguously distinguishes the two points in past time.

(6)    had+V-ed WITH STANDARD PLUPERFECT MEANING:
       The arrests were announced a few hours after Trujillo's son, Gen. Rafael, **had taken**
       up the reins of power his father held for more than 30 years.            (LOB A31 17)

Examples of each of the two types of nonstandard uses of the pluperfect construction found in the Kolhapur Corpus are given in (7).

(7)    had+V-ed WITH PRESENT PERFECT MEANING:
    a.    Meanwhile there was rain in Chikmagalur since morning today and reports
          about rain **had been received** from some other areas also.  Chikmagalur goes to
          the polls tomorrow to decided the political future of Mrs. Indira Gandhi.
                                                                            (KC A43 157)

    b.    This afternoon fresh warnings were flashed to West Delhi areas where the New
          Friends Colony and Maharani Bagh area lie right in the way of the coming
          waters.  Never before in the Capital's history these colonies **had faced** such a
          flood threat.  The Okhla industrial complex in this sector is deserted.   (KC A1 10)

<u>had</u>+V-<u>ed</u> WITH PAST MEANING:

c.   "India's struggle for freedom was led since 1920 by Mahatma Gandhi who believed in non-violence," said the document.  Apart from this single instance Gandhiji did not find a place elsewhere.  The capsule **had been silent** about the role of the early pioneers in the freedom struggle… The document only mentions the names of…                                              (KC A34 106)

The main feature of the excerpts in (7) is that there is no evidence of a reference point (R) in the past, which is distinct from the event (E) being alluded to as well as the present speech point (S).  For example, in (7b), the clauses surrounding the pluperfect construction describe the approaching floods in present tense, and the past event point (of never before having faced such a threat) only lies anterior to the present reference point, not an intervening past one.  Similarly, in (7c), no reference point intervenes between the present speech or reporting point and the writing of the time capsule (in fact, the nature of this particular use is such that the pluperfect use can even be replaced with present tense).

In the presentation of the data here, nonstandard uses of <u>had</u>+V-<u>ed</u> with either present perfect or preterite meanings are generally grouped together as "other" in opposition to "standard" meaning, as it was unclear in a few cases of nonstandard use whether present perfect or preterite would be more appropriate in the given context.

Four contextual cues were also coded for each example, in order to examine the degree to which each variety explicitly marks the relative anteriority of pluperfect contexts. The few instances of present perfect were classified under 'present' in (8b) and (8c). 'Preceding' and 'following' are interpreted linearly; thus an embedded clause in the preceding sentence counts as the closest preceding clause. These internal constraints may collectively convey the pluperfect meaning but are not obligatory and therefore reflect degrees of variability in style and inferable discourse meaning.

(8)   a.   DISAMBIGUATING ADVERBIAL  (present/absent)
       b.   PRECEDING TENSE  (pluperfect, past, present, future)
       c.   FOLLOWING TENSE (pluperfect, past, present, future)
       d.   REPORTED SPEECH VERB  (yes/no)

The category of explicit 'disambiguating adverbials' includes time-marking adverbs (<u>earlier, afterwards, already, later</u>), as in (9a), and time-marking adverbial clauses, as in (9b).

(9)   a.   **Five minutes earlier** the London Scottish defence **had** been split asunder.  (LOB A23 198)
       a'.  Compare to:   * **Five minutes earlier** the London Scottish defence **has** been split asunder.

       b.   There were also brief discussions on Laos, Berlin and other foreign questions, **after** Mr. Kennedy **had** informed Mr. Macmillan of his discussions with Mr. Krushschov.
                                                                                    (LOB A04 20)
       b'.  Compare to:   * …**after** Mr. Kennedy **has** informed…

Adverbial phrases which could signal any type of past event, such as <u>over the past ten years</u> in the example in (10), were not counted as disambiguating adverbials.

(10) a.   He complained that the gap between earnings of manual workers in local government, compared with the average in all industries for men over 21, **had** widened **over the past ten years.**                              (LOB A04 170)
       b.   Compare to:   …**has** widened **over the past ten years**…

Since time adverbials only optionally accompany pluperfects, they can act as a measure of explicit signaling of the pluperfect meaning as opposed to inferable discourse meaning, which may be subject to variable interpretation, allowing changes in usage.

The tense of both preceding and following clauses also provide contextual but again, not deterministic, cues to reference and event points (Comrie, 1985:65; Rickford et al., 1996:228); these were also included as independent, internal variables.

Finally, reported or quoted speech seemed to favor choice of past perfect over present perfect. The notion of reported speech was generously interpreted, in order to fully compare the similar pattern, observable in all three corpora, of shifting from past to pluperfect in any type of reported contexts. The examples from the three corpora in (11) show clauses in which the reporting of an event is the main trigger for shifting the reference point (R) to the (already past) moment of reporting by the original speaker.

(11) a.   The committee **noted** that the probe into elementary watershed basins **had not so far been** given the importance it deserves.                      (KC A38 119)

   b.   The conservation leader **said** other mistakes in spraying **had caused** serious damage in Ohio and Wyoming.                                          (Brown A25 11)

   c.   Coun. Collinghan <u>said</u> there **had been** an improvement on the square, where the buses were now parking at the rear instead of in front of the library. (LOB A43 116)

In all three examples in (11), the state of affairs described in the pluperfect clause is still in effect at the time of the article. The only intervening point is the moment of quotation, which reorients the deictic zero point to the speaker's reporting of the event. If reported speech had been found to categorically require <u>had</u>+V-<u>ed</u> across the corpora, it would have been omitted from the data as a categorical internal constraint. However, reported speech does also occur with present perfect, as in (12), suggesting that it is a constraining but not strictly determining factor.

(12)   Mr. Gajendra Prasad Himansu, State Minister for Irrigation, while replying to Mr. Pasupati Singh **said** that the Chief Engineer **has been directed** to examine and to take necessary action on the report.                              (KC A35 48)

Thus, reported speech was included with other internal constraints so as to examine whether it triggers pluperfect use to varying degrees.

In summary, the coding of the meaning of each pluperfect form is based on the textual context and temporal frame of the <u>had</u>+V-<u>ed</u> construction, while the internal variables contribute additional information regarding how this context is signaled.[7]

(13)   a.   CORPUS (American, British, Indian)
       b.   REGIONALITY (national, regional)
       c.   REGISTER (press, bureaucracy)

The external variables in (13) were included in order to examine variation within the Kolhapur Corpus; they constitute the second part of the quantitative study, which examines variation within IndE.[8]


# 3  Quantitative Results

### 3.1  Inter-varietal comparison

Figure 2 shows the raw totals of <u>had</u>+V-<u>ed</u> identified in the three sets of press reportage data. These include both standard (pluperfect) and nonstandard uses of the <u>had</u>+V-<u>ed</u> construction. As these figures hold for equivalent amounts of text in each corpus (88,000 words per corpus), we can conclude from Figure 2 that the Indian press corpus shows a higher total use of <u>had</u>+V-<u>ed</u> than the native press corpora, and the American corpus shows the lowest of the three. A lower use of both present and past perfect in AmE in contrast to BrE has been observed in other studies as well (Elsness, 1997; Harris, 1984:38; Dusková, 1976:59). This data shows that IndE has a higher rate of use of the pluperfect construction than both of these varieties.

The graphs in Figures 2-4 are in the form of percentages. The actual N (number of tokens) is included at the top of each percentage bar. For the analysis of internal variables, 100 tokens from the two native varieties were coded and almost all tokens in the IndE press corpus were coded (230). More data was coded for IndE in order to allow for more detailed analyses within different registers of the Indian corpus for the second part of the quantitative section. In the results that

follow, statistical significance in each case was measured non-parametrically, by performing a chi-square test on the data in question. The results of the significance tests are included below each figure or table.
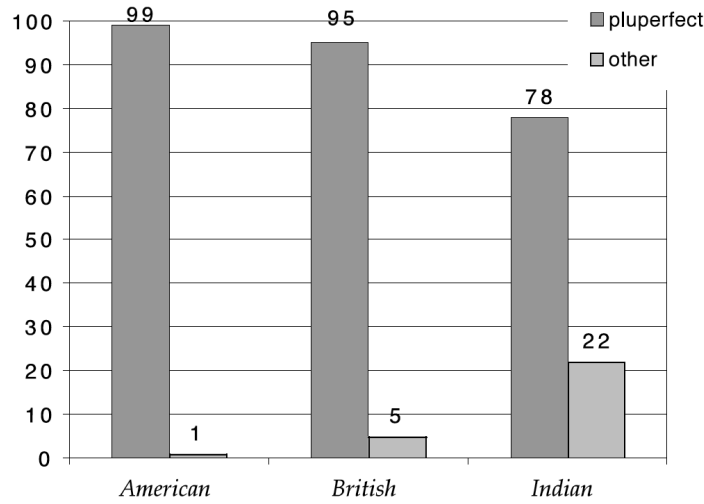
FIGURE 3. Meanings associated with *had* + V-*ed* in percent (IndE vs. AmE/BrE, $p \le$ .001; BrE vs. AmE, not significant).
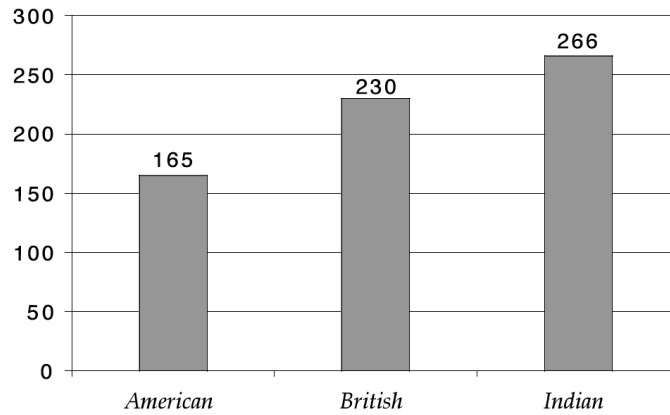
FIGURE 2. Total number of tokens of *had* + V-*ed* in the three corpora.

Figure 3 reflects the meanings associated with each token — coded, as mentioned, based on the E, R, and S points evident in the discourse context. This distribution indicates that the higher proportion of had+V-ed in the Indian corpus in Figure 2 is distributed over a wider range of meanings. The "Other" category for IndE is comprised of 28 (12%) tokens used with present perfect meaning and 23 (10%) with preterite meaning. The difference between the IndE data and each of the native varieties is significant, while the slight difference between native varieties is not. Only a total of 6 tokens in the other corpora appeared in contexts where it was impossible to establish a distinct event point (E) and reference point (R) for the context of use. However, even this small number indicates crucially that some ambiguity in pluperfect meaning can arise in native varieties. Two of these examples from the British corpus are given in (14).

(14)   a.   (Report on the newly elected Mayor of Huddersfield)    Clr. Brook began his education in two of the local schools. As a boy and a man his interests **had covered** a very wide field… His work on the council over the many years of his service **had been** outstanding in more than one respect.                    (LOB A30 31)

b.  Only a few hours after Mr. Lloyd and his 24-strong delegation landed at Accra this morning, hundreds of shop assistants demonstrated outside the British-owned Kingsway Stores, the largest in town.  The stores **had been hit** by the same strike wave that has paralysed the port of Takoradi for the past week.  Roots of the discontent: The Austerity Budget, including a compulsory savings scheme…

(LOB A2 61)

The underlying factors motivating these instances of ambiguous usage in BrE are examined in the qualitative part of the analysis, as a possible source of the non-native usage patterns.
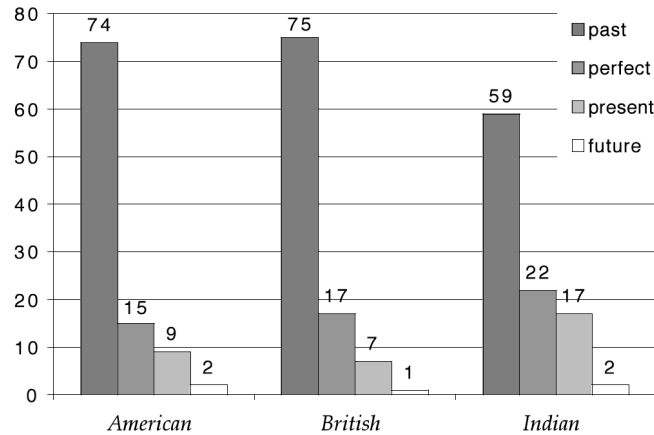


FIGURE 4.  Preceding tense with *had* + V-*ed* in percent (IndE vs. AmE/BrE, $p \le .01$; BrE vs. AmE, not significant).
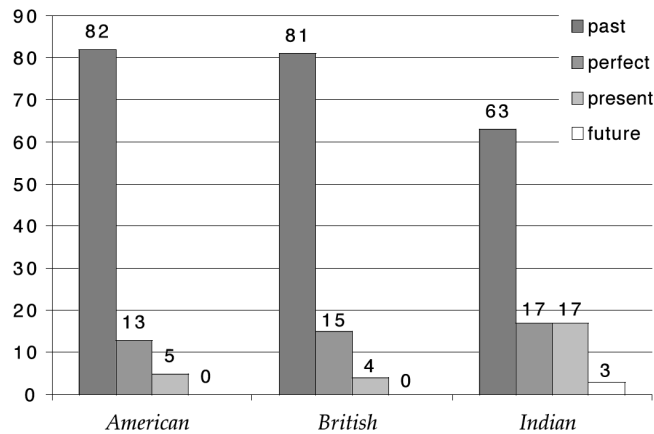


FIGURE 5.  Following tense with *had* + V-*ed* in percent (IndE vs. AmE/BrE, $p \le .01$; BrE vs. AmE, not significant).

In order to examine possible factors in the inter-dialectal differences in Figures 2 and 3, we can examine the contextual factors that were listed in (8).  In Figures 4 and 5, the frequency with which different tenses precede and follow all coded instances of had+V-ed are summarized. In native varieties, the past tense context is the most felicitous for a further (past perfect) reference into more specific past tense events. In fact, the nearly identical distribution of tense frequencies across the BrE and AmE corpora in both Figure 4 and Figure 5 is remarkable.  Although past tense is the most common tense environment for had+V-ed in IndE as well, other tenses are more frequently found adjacent to these clauses.  In particular, present tense shows a greater presence in both preceding and following environments.  These tense differences relate to the use of the had+V-ed construction with alternative meanings, as a preterite or present perfect meaning for

had+V-*ed* would not require a pre-set past tense environment as stringently, if at all.  Support for this comes from the additional finding that 71% (10/14) of the tokens in IndE which occur with present or future tense in <u>both</u> their preceding and following environments are used with a non-standard meaning.  Some of these tense combinations, in fact, never occur when the pluperfect is used in BrE or AmE, for instance, future-pluperfect-future. Again, the contrast between the IndE distribution and native English distributions of tense environments is statistically significant in both Figure 4 and Figure 5, while the difference among native varieties is not.  In terms of discourse organization and sequence-of-tense (SOT) phenomena, preceding tense environment (Figure 4) is the more relevant of the two for conditioning pluperfect use.

**Table 3**: Presence of disambiguating adverbials in *had*+V-*ed* clauses in three corpora (N in parentheses) χ²: AmE vs. BrE/IndE **significant** (p ≤ .001);   IndE vs. BrE not  significant

| DISAMBIGUATING ADVERBIAL | AMERICAN | BRITISH | INDIAN |
|---|---|---|---|
| Present | 59%  (59) | 36%  (36) | 30%  (70) |
| Absent | 41%  (41) | 64%  (64) | 70%  (160) |

The presence or absence of explicitly disambiguating adverbs or adverbial clauses, as explained in the examples in (10)-(11), was coded for every token to measure the degree to which pluperfect meaning is explicitly signaled. Disambiguating adverbials were found to be most frequently present in the American corpus (59%), significantly less common in the British corpus (36%), and least common in the Indian corpus (30%).

The importance of time adverbials as well as adjacent tense in the pragmatic interpretation of aspect meanings has been commonly noted (Comrie, 1985:65; Winford, 1993:151), and Crystal (1966) suggests that "time relations in English are handled more by the careful use of adverbials… than by any other means" (1966:7).  Consequently, the general absence of adverbial marking may facilitate the reinterpretation of pragmatic meanings of aspectual constructions. Given the data here, this absence may permit an expansion of the use of the pluperfect in both BrE and IndE.  BrE does show a higher rate of pluperfect use, but with only slightly more nonstandard meanings.  The significantly greater degree of change in function in IndE may be partly explained by the compounding effect of its status as a non-native language in contact with the more past-like pluperfect constructions found in Indo-Aryan and Dravidian-type systems, described earlier.  Thus, stylistic variation such as the use of time adverbials can create environments that allow a substrate language-based functional reanalysis.

Finally, the occurrence of <u>had</u>+V-<u>ed</u> within reported speech was also coded, as this seemed to be an important constraint on use of pluperfect rather than present perfect or past. The pattern seen in other internal variables so far, such that AmE diverges systematically from IndE, is again present here, but not to a significant degree.

**Table 4:** Frequency of occurrence of *had*+V-*ed* in reported speech in three corpora (N in parentheses) χ²: AmE/BrE/IndE not significant

| REPORTED SPEECH | AMERICAN | BRITISH | INDIAN |
|---|---|---|---|
| Reported | 39%  (39) | 41%  (41) | 44%  (102) |
| Not reported | 61%  (61) | 59%  (59) | 56%  (128) |

In Table 4, we can see that IndE, which does not use tense or adverbial signaling as strictly as the other two varieties, has the highest proportion of pluperfect usage in reported speech contexts, possibly accounting for some of the overall higher rate of <u>had</u>+V-<u>ed</u>.  This observation is only a slight tendency in statistical terms, and therefore does not represent a strong influence in the usage patterns. However, as mentioned in relation to Table 1, a feature of IndE press reporting reporting is its avoidance of direct quotations, as compared to AmE and BrE.  This tendency towards formality leads to more reported speech contexts, allowing for a greater rate of

pluperfect use.  Also, IndE shows a distinct pattern of maintaining the 'reported speech pluperfect' over several sentences beyond the original reporting verb, as a stylistic device.  This pattern of use is most commonly initiated by a reported speech predicate, after which that 'voice', rather than the point of view of the reporter, is maintained by the pluperfect.

(15)     J. R. Jayawardene told the Sri Lanka Parliament the day the new Constitution came into force earlier this month that it **had enlarged** the people's democratic freedom.  It **had extended** their sovereignty to include the process of government, Fundamental Rights and the franchise. It **had ensured** that the President and Parliament did not extend their terms without public approval in a referendum. Prolonged rule through Emergency regulations, as in the past, was out.  Parliament must now debate and vote on every Public Security Ordinance and, after three months, it must be passed by a two-thirds majority.  Fundamental Rights **had been made** justiciable, there was a provision for an Ombudsman (Parliamentary Commissioner for Administration) and the method of voting by Proportional Representation **had been** introduced.  (KC A2 2-13)

In the IndE example in (15), the quotative voice of the initial reported speaker is maintained over several clauses, even with a break in tense type in between.  This usage differs from native usage only in degree, however; as (16) shows, the pluperfect may occur in a sentence subsequent to the quotative sentence in BrE as well.

(16)     Mr. Healey denounced the Government for using double standards. Ministers **had rebuked** the unaligned nations for not condemning the new Russian tests, but they themselves had no condemnation for events in Angola or Algeria.                                                    (LOB A5 196)

     In keeping with the restricted use of perfect in AmE, this type of extended quotative pluperfect beyond a single clause appears to be relatively rare in AmE.  Thus, while the results for reported speech are not statistically significant, the development of stylistic devices can create or inhibit new contexts for use of the pluperfect form.
     In sum, the pattern that emerges across the internal variables is that AmE and BrE are almost identical in their tense contexts for the pluperfect, but diverge more in the use of time adverbials.  In the case of time adverbials, and reported speech to a lesser degree, BrE lies between AmE and IndE, while tense ordering is almost identical for BrE and AmE.  More specifically, IndE tends to extend the pluperfect construction to both the shared <u>aspectual</u> meaning of present perfect and the shared <u>tense</u> meaning of simple past.  This 'leakage' in functions correlates with a less stringent set of restrictions in IndE on contextual cueing of pluperfect meaning via time adverbials.  By contrast, AmE seems to have restricted the use of the pluperfect further than BrE, exhibiting stricter requirements for explicit marking and a generally lower overall use of the construction.[9]
     The next section presents the second set of quantitative results from the corpus survey, namely variation in pluperfect use within IndE, followed by a qualitative discussion of the discourse basis of this change as well as its relation to other TMA phenomena in IndE and other nonstandard varieties.

### 3.2  Intravarietal variation: pluperfect use within the IndE corpus

Two external variables were included in the coding of the corpus data in order to allow a quantitative examination of variation within IndE.  The distribution of a variable can provide some indication of the relevant or meaningful dimensions of a speech community, an issue that is of particular interest in recently emerging varieties of English.  The two intra-variety factors discussed here are regionality and register.

**Table 5**: Usage of *had*+V-*ed* in Indian English according to regionality. $\chi^2$: **significant** (p ≤ .05)

| GEOGRAPHIC RANGE | STANDARD MEANING | OTHER MEANING |
|---|---|---|
| National texts | 83%   (100) | 17%   (20) |
| Regional texts | 72%   (79) | 26%   (31) |

All the IndE texts, in both press and bureaucratic registers, are classified in the Kolhapur Corpus as either national or regional.  Examples of national texts are central government documents or major city newspapers with a national distribution.  Regional texts are, for example, local newspapers or state government documents.   The data in Table 5 shows a statistically significant difference between national and regional usage in IndE.   The nonstandard use of had+V-ed is one of many variables that occur more frequently in regional texts.  This difference between regional and national usage is noteworthy, as it reflects a lectal range that seems to be sensitive to geographic location relative to centralized, urban standards.  Given the multilingual landscape and high register nature of English in this spectrum, these results support the view that IndE standardness is higher in urban areas.  All the same, the occurrence of nonstandard had+V-ed in the 'national', i.e. centralized, urban, text sources (press and government) indicates that the feature is not entirely absent in mainstream usage.

The second external variable that was contrasted was register.  I compared the press register to the register named "Miscellaneous" in the corpus but which I classify here as bureaucratic documents.  The 'bureaucratic' register includes texts from the H category of the Kolhapur Corpus, which includes government documents, foundation reports, and industry reports, both national and regional.  The predominant text type within the bureaucracy register is central and state government documents.

The selection of these particular registers to explore register-based variation derives from a consideration of nativization of registers and the earlier discussion of the historical establishment of domains of English use in India.  Kachru (1983:26) cites the civil services as possibly playing a significant role in the establishment of a native standard in India, as administration is one of the oldest and most widespread institutions of English use both pre- and, even more expansively, post-independence.  Furthermore, 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 on a scale of register, the Indian bureaucratic register in fact reflects a potentially more indigenized standard than, say, that of the national English language media.  Meanwhile, the national English press in India tends to have closer professional ties with international English media and may also derive its standard from native varieties of journalistic prose.

According to these differences in the function of English in the two registers, we would expect to find that print media show fewer nonstandard features than government and other bureaucratic registers. All the same, English plays a central role in both government and press in India, and these can be considered two of the commonest domains of regular IndE usage.  By contrast, the national/regional distinction in Table 5 corresponds to greater and lesser degrees of overall English use, and therefore the discrepancy there might be expected to be higher.

**Table 6:** Usage of *had*+V-*ed* in Indian English according to register. $\chi^2$:  not significant

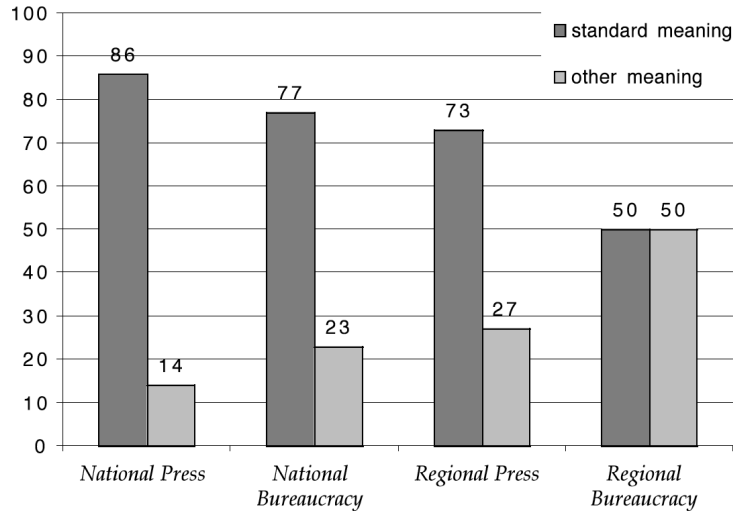| REGISTER | STANDARD MEANING | OTHER MEANING |
|---|---|---|
| Press texts | 79%  (154) | 19%  (42) |
| Bureaucratic texts | 73%  (25) | 27%  (9) |

FIGURE 6.  Cross-classification of IndE external variables in percent.

Table 6 shows the distribution of had usage across the press and bureaucratic registers; Figure 6 shows a more detailed cross-classification of both the external variables.[10]  The difference in percentage of nonstandard aspectual uses of <u>had</u> in Table 6 supports the prediction that administrative bureaucracy might show greater nonstandardness due to its extension across geographic and socioeconomic divisions. In native English registers, on the other hand, bureaucratic data falls high on a scale of standardness and/or formality, perhaps even more so than press.  This difference suggests further that the relative formality of text registers cannot be assumed to be universal, but may instead be subject to considerable sociocultural determination. In other words, degrees of non-standardness may not always correlate directly with a universally predictable stylistic cline of formality, as is common in many studies of social and stylistic variation (Cf. Finegan and Biber, 1994:317).  However, it is important to note that the greater nonstandardness in bureaucracy relative to the press is only a tendency and is not statistically significant.  This result may in fact lend credence to the idea that the press and civil services involve semi-autonomous standards, rather than occupying points on a unified continuum.

## 4   Qualitative discussion

### 4.1  Pragmatics of *had* usage in IndE

The actual tense and aspect contexts in which these nonstandard uses of the pluperfect construction occur deserve closer examination.  Is a new, specialized semantic meaning associated with these uses, as was found for Trinidadian English in Winford (1993), or is the pluperfect being licensed more freely in the IndE pragmatic or discourse framework?  In this section, a qualitative discussion of the semantics and pragmatics of the nonstandard uses supports the latter hypothesis.

At first glance, the 51 nonstandard uses of the <u>had</u>+V-<u>ed</u> construction that were found in this study seemed to fall into a number of groupings, listed in (17).  The examples below are given without context and are intended only to indicate the types of clauses that were classified into different groups.  For all groups, an intervening R point is not distinct in the discourse context.

(17)   a.   SUSTAINED STATE OF AFFAIRS.  The instances in this category often include explicit extended time adverbials that are more commonly associated with present perfect. In all cases, the state of affairs is still in effect.

   i.   Similar concessions **had been in force** for years in the southern States…
   ii.  Politics in Bihar, for decades, **had been caste-ridden** …
   iii. The students **had been opposing** the State Government decision…

b.    COVERTLY REPORTED SPEECH.    These tokens occur without a quoted speaker but with a remotely inferable reported speech context.

    i.    The agitators… **had also disrupted** road traffic.
    ii.    The government **had spent** Rs 300 on the development of each site…
    iii.  Good care **had been taken** to use the leverage of canalisation to promote growth.

c.    REMOTE PAST.    The event point (E) in these pluperfect uses is located in the remote past.  Many of these examples refer explicitly to a remote time point.

    i.    In the past, there **had been** criticism in the J.P.P. Executive Committee over the issue…
    ii.   The National Cooperative Development Corporation **had contributed** Rs.35.45 lakhs [100,000] during 1976-7.

d.    RELATIVE TO IMMEDIATELY PRIOR DESCRIPTION.    In these examples, an event is described in detail, which then becomes the R point for a subsequent pluperfect clause.  Many examples include a reference to the prior description (e.g. "such an object").

    i.    This is the second time that such an object **had been sighted** here.
    ii.   Rarely **had** a Ranji final **taken** such a course on the opening day itself.
    iii.  Never before in the Capital's history these colonies **had faced** such a flood threat.

Unlike Winford's (1993) study of *have* in Trinidadian English, the IndE data are not explicable according to the four commonly noted meanings of the perfect: result/stative, experiential/existential, persistent situation, and recent past/hot news (Comrie, 1976:57).  No single, <u>semantic</u> meaning can be attributed to all the uses in (17) without distorting the actual range of contexts in which the construction appears.  Note from (17c), furthermore, that Leitner's (1991:228) suggestion that pluperfect is used for remote past in IndE turns out to be partially true, but only comprises a subset of all the uses found in the data.

Rather than comprising a heterogeneous set of unclear uses or failures to match a native standard, I argue these nonstandard uses found in IndE can be unified into a regular pattern of pragmatic viewpoint marking.  Furthermore, no nonstandard use is entirely novel; rather, the new uses derive from the existing ambiguities present in the native variety.

The intuitive nonstandardness of many of the examples is the lack of an intervening R(eference) point to license the use of the past perfect, thereby imbuing it with an apparently present perfect or preterite meaning.  However, most of the IndE examples treat some focus in the discourse as a pseudo-R point, which reorients the viewpoint and distances other events from that focus.  Two examples are discussed in (18) and (19).

(18)    One of the scientists, Dr P. S. Sehra, said it was a strange luminous object moving from north-west to south-east at 9.12 p.m. The object, sighted from the P.R.L campus, had a glowing double head and a long, narrow tail.  It was seen for about 10 to 15 seconds;
    a.    the object **was** "bluish yellow" in colour and the elevation was about 70,
    b.    he **added**.
    c.    This **is** the second time
    d.    that such an object **had been** sighted here.
    e.    The first **was** sighted on April 3. Professor D. Lal, director of the Physical Research Laboratory, and his colleagues, Dr J. N. Desai and Dr J. N. Goswami.  (KC A13 138-144)

The two past tense clauses in (18a) and (18b) refer to two different event points, one at which a peculiar object was sighted and a second at which this sighting was reported by Dr. Sehra to the current speaker/reporter.  In (18c), present tense relates the sighting of the object to the present time (S).  However, in (18d), rather than continuing to relate the first ever sighting to the present time (S) knowledge, it is shifted relative to Dr. Sehra's second sighting.  This pseudo-R point is the main focus of the report and disrupts the maintenance of present relevance in the

reporting. This shift to an apparent R point, in lieu of maintaining a fixed relevance to S, is comparable to the pattern in (19).

(19) a.  This Bill **is** very timely.
    b.  I hope the cultivators **will be benefitted** after passing this Bill to a greater extent.
    c.  Other Bills like the Manipur Town and Country Planning Bill **had also been passed** for raising the standard of living of the people in the State.
    d.  I **hope** the Bills which are under consideration in the House today **will be passed** for the benefit of the people.

This example shows an even more radical shift, namely from future to pluperfect and back to future (with a present quotative verb). Not a single instance of the tense ordering [future-pluperfect-present/future] appears in the BrE or AmE corpus data. In (19a), the debate over a bill is the E point or state of affairs being described. In (19b), a future tense reference is made to projected benefits. When a remote past reference is introduced in (19c), the original event (E) point and the focus of the narrative — debating the bill — becomes a pseudo-R point. Note that the passing of earlier bills is still in effect at the present (S) point and would usually be referred to with present relevance, i.e. with present perfect. Thus, as in (18d), the anterior E in (19c) is introduced as being relevant to the previous, focal E, rather than the current S as might be expected.

The generalization that unifies an otherwise disparate group of uses is that the reference point (R) shifts back in a context in which it would, in standard English press reportage, be maintained as coterminous with S. In other words, the point of view shifts relative to the <u>new focus</u> instead of being maintained relative to the speech point or moment of reporting.

On this view, a primary effect of the <u>had</u>+V-<u>ed</u> construction is to distance a completed, past event from the narrative focus. This permits a wider range of discourse uses than the standard, native English pluperfect. Standard native English conforms to a general principle of <u>permanence of the reference point</u>, which eliminates such tense combinations as (20).

(20)  * I had mailed the letter when John has come

This principle very broadly states that in a coherent segment of discourse, R is maintained and E and S may shift relative to it (Reichenbach, 1947:293). However, shifting of the R point does standardly occur under certain circumstances in native varieties as well, as in the case of indirect speech in subordinate clauses. Thus, (21b) can be used to express the meaning of (21a).

(21)  a.  Sue thought, "Max has left."
      b.  Sue thought that Max had left.

For this reason, a simple principle that R must be maintained is too strong to allow for the range of observable sequence-of-tense (SOT) phenomena. In fact, a range of contextual factors — including temporal adverbs, telicity of the verb, stativity, and result entailments —determines SOT phenomena and other characteristics of the perfect (Cf. XN 'extended now' theory; McCoard, 1978:151). The quantitative data in this study have already shown that sequence of tense considerations and adverbs play some part in the choice of pluperfect form. Thus, R-shifting is an inherent component of native English but may be subject to variable usage. Now, rather than seeing the IndE examples as entirely novel, we can consider their salient R-shifted characteristic as extending out of ambiguities already present in R-shifting in native, standard English.

The BrE example from (14b), repeated in (22), demonstrates this variability in native usage. Here, the pluperfect form may be associated with present perfect meaning.

(22)  Only a few hours after Mr. Lloyd and his 24-strong delegation landed at Accra this morning, hundreds of shop assistants demonstrated outside the British-owned Kingsway Stores, the largest in town. The stores **had been hit** by the same strike wave that has paralysed the port of Takoradi for the past week. Roots of the discontent: The Austerity Budget, including a compulsory savings scheme…                                    (LOB A2 61)

The use of the pluperfect in (22) is ambiguous between present perfect and past perfect meanings.  Either the demonstration of shop assistants *or* the arrival of Mr. Lloyd introduces a new reference (R) point, licensing the pluperfect shift in tense viewpoint such that the "strike wave" refers to the time period prior to that.  However, the strike wave is not clearly restricted to that time period and continues to have present relevance, permitting a present perfect reading. This ambiguity in the meaning of the pluperfect has been previously noted in Tichy (1980:361); it can either be equivalent to a past-in-the-past or a present-perfect-in-the-past.

It is this ambiguity in narrative or discourse reorientation, observable to a small degree in native usage in the BrE and AmE data, that is prevalent in many of the IndE examples, resulting in a local, anteriority-marking function of had. This demonstrates the importance of taking into account variability and ambiguity in the lexifier or native variety itself when considering possible sources of new form-function relationships in a new variety (Mufwene, 1996).

## 4.2   Related changes within the TMA system of IndE

The functional reallocation of the had+V-ed construction must be considered within the context of other forms and meanings in the IndE TMA system.  Harris (1984) makes this argument in his study of the Hiberno English TMA system:

> Of necessity the [sociolinguistic variable] model encourages an atomistic view of variation, whereby pairs of apparently alternating standard and nonstandard forms are studied in isolation from other forms in the relevant grammatical subsystem.  This 'worm's eye' approach encourages the impression that differences between the standard and a particular vernacular are merely superficial and tends to obscure whatever deep-seated divergences there might exist between the two varieties.  In contrast, the likelihood of radical structural differences coming to light is increased as the angle of observation is increased to encompass the wider grammatical subsystems in which the apparently alternating forms are embedded.
>
> (Harris, 1984:32)

Winford (1993) also endorses a systematic approach to TMA analysis along these lines, adding that functional equivalences and pragmatic categories across two dialects are crucial for a clear understanding of variation in TMA marking.

In the introductory overview of the IndE TMA system, I noted that modals and certain other tenses have been investigated to some degree in the Kolhapur Corpus and in other studies as well (Katikar, 1984; Williams, 1987; Leitner, 1991). The three TMA features of IndE that deserve reconsideration in light of the pluperfect usage discussed here are the use of the modals could and would, the present perfect construction, and the progressive.

In his study of IndE modals, Leitner (1991) primarily restricts his discussion to frequencies of forms and does not identify a noteworthy development in the functions of would and could in IndE, namely that these modals can be used with the standard meanings of will and can respectively. The shared extension in both cases is the use of these two modals with non-conditional meanings, and may derive partly from an overextension of the polite use of the subjunctive into other modal contexts to index politeness or formality.

The two examples in (23) are personal communications from different IndE speakers, and show instances of would being used with a simple future meaning.  These modal uses are another instance of a new function being associated with a standard form.

(23) a.   He **would be coming** to receive you on 15th and he **would be bringing** the ticket too.  This is because M. would have gone to Delhi by then and we just want to ensure that the ticket is not misplaced.  He **would be seeing** you off too.

  b.   I **would be visiting** your place tomorrow…  I **would be reaching** around 9:30am. I completed my work just now and **would be free** tomorrow.

The modals could and would in Standard English parallel uses of the pluperfect construction in certain contexts.  They are used as the past equivalents of can and will, as in the reported speech examples in (24a); the equivalent alternation between simple past and pluperfect is shown in (24b).

(24)    REPORTED SPEECH:
      a.  I can/will bring the book.
      a'.  I told her I **could**/**would** bring the book.

      b.  I brought the book.
      b'.  I told her I **had brought** bring the book.

    Furthermore, both modals <u>could</u> and <u>would</u> are used generally in counterfactual or irrealis contexts, as is <u>had</u>.  This is shown in the examples in (25).

(25)    COUNTERFACTUALS:
      a.  I can/will bring the book.
      a'.  If I **could**/**would** bring the book …
      a''.  I **could**/**would** have brought the book.

      b.  I brought the book.
      b.  If I **had brought** the book …

The IndE uses of <u>could</u> and <u>would</u> in non-irrealis contexts parallel the IndE uses of <u>had</u>, which occur similarly in non-pluperfect contexts. Together, the use of <u>could</u>, <u>would</u>, and <u>had</u> in new environments forms an overlapping set of extensions in IndE that may be mutually reinforcing.
    Another potentially related shift in the TMA system, illustrated earlier in example (2), is what Williams (1987:183) has described as a lexicalization of completion.  The examples in (26) show that, in IndE, the present perfect can indicate a completed past action rather than a past action with ongoing relevance in the present.  In both examples, the punctual adverbial phrases (<u>four years ago</u> and <u>six months back</u>) standardly require past tense.  (26a) is from an online discussion and (26b) is from personal interview data.

(26)  a.  We **have known**, four years ago, that weightlifting was going to be an Olympic
        sport. We **have known,** four years ago, that Malleshwari was a medal prospect
        — heck, four years ago, she was the world champion.[11]

    b.  Actually, I **have come** here about six months back.

    If <u>have</u> has come to represent lexicalized or grammaticalized completion and has lost the connotation of present relevance, it essentially constitutes perfective (rather than perfect) marking. Dahl (1985:78) describes perfective marking as the denotation of  "a single event with a well-defined result or end-state, located in the past."  Perfective represents an action as a whole — sometimes considered a type of Aktionsart — while the perfect indicates relational anteriority and relevance to a deictic zero point. (Binnick 1991:161; Bybee et al. 1994:55).  This depiction of IndE present perfect use mirrors the innovative uses of the past perfect discussed here, as these new uses have been shown to contribute information such as remoteness and prior completion. Together, past and present perfect use in IndE begin to resemble a generalized perfective viewpoint aspect.
    If this is the case, the progressive form in IndE may additionally fit into this picture, as the progressive in IndE appears often to signal nonperfectivity.  Michaelis (1998) argues that although the progressive is often associated with imperfectivity cross-linguistically, the standard English progressive is not a marker of imperfective aspect but rather phasal aspect, which accounts for the ungrammaticality of stative progressives.  In IndE, however, progressives do occur in stative or habitual contexts; e.g., sentences such as: <u>they are having vindictive motives</u>, and <u>they may be having bad feelings</u>, from the discussion of example (3).  The consistent use of the progressive form with the future (a nonperfective category) in (23) is also noteworthy in this respect.
    Pending more detailed analyses, this preliminary contrast suggests that it may be possible to conceive of the variation seen in the pluperfect case as a part of a more systematic tendency, whereby, amongst other factors, a perfective/nonperfective distinction is privileged in the new system.  Michaelis claims that "the viewpoint aspects perfective and imperfective are covert

categories of English; that is, they are not directly encoded morphologically" (Michaelis, 1998:59); in some non-native Englishes, there may in fact be an emergent viewpoint aspect distinction. This section has attempted to contextualize IndE pluperfect use within its TMA system; naturally, these speculations regarding connections to other TMA features are preliminary at this point and await a more complete investigation.

### 4.3   Similarities in other nonstandard TMA systems

A final consideration in the data in this study of IndE TMA forms is the existence of comparable phenomena in sometimes unrelated non-native or nonstandard varieties.  In their study of had+V-ed in narratives of African-American adolescents, Rickford et al. (1996) identify a use of the construction that marks preterite rather than pluperfect.  They cite Cukor-Avila & Bailey (1995) as referring to this construction as "innovative had + past", a description they extend in order to explain the usage in AAVE narratives.  Rickford and Rafal depict the innovative use as conveying narrative reorientation or signaling discourse peaks by foreshadowing complicating action or evaluative points.  In comparison to the data here, which indicates a local, anteriority-marking function and a possible lexicalization of completion, the innovative functions of preterite had in AAVE narratives are slightly different. However, the absence of a canonical, intermediate reference point (R) and extension to preterite meaning is comparable.

   Another interesting parallel occurs in Trinidadian English.[12]  Consider the similarities between the preterite use of had in IndE, with the possibly related shifts found in would and could, and the summary of features presented by Solomon (1972) in the following excerpt, which I quote in full to include the theoretical issue he raises.

> What makes [an analysis of dialect mixture] difficult is the tacit assumption that the changes occasioned by the functional shift are changes, quantitative or qualitative, in the incidence of forms or "sets" of forms, without consideration of the cognitive value of the forms in the different dialects. To put it another way, although there is interference there is no linguistic variable that can describe the shift from English comes to French vient. The shift is from a system where a certain conceptual distinction is absent to one where it is present, and the formal correlates of the concept are not amenable to meaningful manipulation…
>
>    Examples are in fact abundant: in Trinidadian "English", "could" is frequently used with the meaning, approximately, of SE "can" or French Creole pe; could have is used equivalently to French Creole te pe, roughly translatable as "could" or "was able".  "Had gone" is used with the meaning of French Creole "te ale", which is not directly translatable into English but means that the event happened in the past but without any implication of recentness or relevance to the moment of speaking.
>
> <div align="right">(Solomon 1972:4)</div>

   Winford (1993:172) also points out precisely these alternate meanings of could and would in Trinidadian English.  The descriptions of could and would and of pluperfect use in Trinidadian English both match the usage found in the IndE data here. Furthermore, Solomon's description addresses contact with French Creole, which shows evidence of influence from the French type of system, a pattern towards which IndE appears to have shifted as well.  Foley et. al (1998:142) note an identical extension of the use of could and would in Singapore English, providing the example:  I am sorry to tell you that I would have to turn down your application.

   Finally, the anterior, completive aspect marking of perfect constructions discussed for present and past perfect in IndE has also been observed in Philippine English and Singapore English (Platt, Weber and Ho, 1984).

   Although some L1 influence from Indian languages is certainly present in various structural deviations of IndE, the cross-dialectal similarities presented in this section suggest it may not be the exclusive explanation for the TMA changes discussed here.  First, TMA systems are rather diverse, particularly between Dravidian and Indo-Aryan languages, and thus may at best only reinforce or inhibit certain ongoing shifts through structural similarities.  More importantly, an equivalence appears to exist between several TMA changes in IndE and those in unrelated nonstandard varieties worldwide.  Based on these observations, a more accurate representation of the semantic extension in pluperfect use is that universal principles of reanalysis and generalization begin to operate on inherent ambiguities in certain constructions in the

native/lexifier language, accompanied by social mechanisms of spread and possibly secondary L1 reinforcement.

## 5  Conclusions

This study has examined several aspects of pluperfect usage in IndE.  The non-native corpus data additionally associate present perfect and preterite meanings with the had+V-ed construction, showing significant differences from native usage.  Furthermore, the frequency of cases in which pluperfect contexts are explicitly licensed by time adverbials is noticeably lower in IndE than in native varieties.  Differences between AmE and BrE also emerged from the analysis of the three corpora, namely that AmE shows a lower overall use of had+V-ed and also a stricter set of constraints on signaling such contexts.

Variation within IndE was found to show a tendency for greater non-nativeness in regional press and in the bureaucratic register.  However, both registers examined and both types of press exhibited some degree of non-standardness in pluperfect use, indicating that while variation may exist within the variety, this usage is fairly widespread.  The unifying characteristic amongst the nonstandard uses was found to lie in distancing from a given narrative focus, resulting in a generalized remoteness and completion marking.  A comparison of the use of the pluperfect construction with other non-native characteristics of the TMA system of IndE revealed the possibility of a systematic semantic shift which may cover other variants in meaning associated with would and could and present perfect have as well.  This shift may be attributed to an emergent anterior, completive (perfective, rather than perfect) marking and a concomitant loss of the necessity for a distinct reference point in the context, allowing both present perfect and preterite meanings to be associated with all three constructions.   Finally, the brief comparison of these patterns to the TMA systems of other non-native or nonstandard varieties of English brought out a number of shared characteristics, suggesting further directions for research on possible causes of these regular shifts in varieties of English.

## References

Biber, Douglas. (1987). A Textual comparison of British and American writing. *American Speech 62:* 99-119.

Bickerton, Derek. (1980). Decreolisation and the Creole Continuum. In A. Valdman and A. Highfield (eds.) *Theoretical Orientations in Creole Studies.*  New York: Academic Press

Binnick, Robert.  (1991).  *Time and the Verb.*  Oxford: Oxford University Press.

Bybee, Joan, R. Perkins and W. Pagliuca (eds.).  (1994).  *The evolution of grammar: tense, aspect, and modality in the languages of the world.*  Chicago: University of Chicago Press.

Cheshire, Jenny.  (1987).  Syntactic variation, the linguistic variable, and sociolinguistic theory. *Linguistics 25,* 257-282.

Comrie, Bernard.  (1976).  *Aspect.*  Cambridge: Cambridge University Press.

Comrie, Bernard.  (1985).  *Tense.*  Cambridge: Cambridge University Press.

Cukor-Avila, P. and G. Bailey.  (1995).  Grammaticalization in AAVE.  In Ahlers et al. (eds.) *Proceedings of the 21ª Annual Meeting of the Berkeley Linguistics Society (BLS 21).*  Berkeley: Department of Linguistics, University of California, 401-13.

Crystal, David.  (1966).  Specification of English Tenses. *Journal of Linguistics* 2:1-34.

Dahl, Östen.  (1985).  *Tense and aspect systems.*  Oxford:  Blackwell Press.

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

Dusková, L. (1976).  On some differences in the use of the perfect and preterite between British and American English. *Prague Studies in Mathematical Linguistics* 5: 53-68.

Elsness, Johan. (1997). *The Perfect and the Preterite in Contemporary and Earlier English.* Berlin: Mouton de Guyter.

Finegan, Edward and Douglas Biber. (1994). Register and Social Dialect Variation: An Integrated Approach. In D. Biber and E. Finegan (eds.). *Sociolinguistic perspectives on register.* Oxford: Oxford University Press.

Foley, J.A. et al. (1998). *English in New Cultural Contexts: Reflections from Singapore.* Oxford: Oxford University Press

Harris, John. (1996[1984]). Syntactic Variation and Dialect Divergence. In . R. Singh (ed.). *Towards a Critical Sociolinguistics.* John Benjamins Publishing Co., Amsterdam, 31-57.

Hundt, Marianne. (1996). Has BrE Been Catching Up with AmE Over the Past Thirty Years? In M. Ljung (ed.). *Corpus-based Studies in English.* Papers from the 17th International Conference on English Language Research on Computerized Corpora (ICAME 17), Stockholm. Amsterdam: Rodolpi.

Kachru, Braj. (1992 [1982]). *The Other Tongue: English across Cultures.* Urbana: University of Illinois Press.

Kachru, Braj. (1983). *The Indianization of English: The English language in India.* Oxford: Oxford University Press.

Kachru, Braj. (1996). South Asian English: Toward an Identity in Diaspora. In R. J. Baumgardner (ed.) *South Asian English: Structure, uses, and users.* Urbana: University of Illinois Press.

Katenina, T. E. (1960). *Jazyk Chindi.* Moscow: Nauka.

Katikar, P. (1984). The meanings of modals in Indian English. PhD Thesis. Shivaji University, Kolhapur.

Labov, William. (1982). Building on empirical foundations. In W. P. Lehmann & Y. Malkiel (eds.), *Perspectives on historical linguistics.* Amsterdam: John Benjamins. 79-92.

Lavandera, Beatriz R. (1978). Where does the Sociolinguistic Variable Stop? *Language in Society 7,* 171-182

Leitner, Gerhard. (1991). The Kolhapur Corpus of Indian English — intra-varietal description and/or intervarietal comparison. In S. Johansson and A-B. Stenström (eds.) *English Computer Corpora: Selected Papers and Research Guide.* New York: Mouton de Gruyter.

Mehrotra, Raja Ram. (1998). *Indian English.* Amsterdam: John Benjamins Publishing Company

McCoard, Robert. (1978). *The English Perfect: Tense-Choice and Pragmatic Inferences.* North-Holland Publishing Co.

Michaelis, Laura. (1998). *Aspectual Grammar and Past-Time Reference.* New York: Routledge.

Mufwene, Salikoko. (1996). Creolization and grammaticization: what creolistics could contribute to research on grammaticization. In P. Baker and A. Syea (eds.) *Changing Meanings, Changing Functions.* London: University of Westminster Press.

Platt, J., Weber, & Ho. (1984). *The New Englishes.* London: Routledge.

Reichenbach, Hans. (1947). *Elements of Symbolic Logic.* New York: The Macmillan Company.

Richter, Claudia. (2000). The Progressive Form in Indian, British, and American English. Unpublished M.A. Thesis, Zurich University.

Rickford, John. (1987). *Dimensions of a creole continuum.* Stanford: Stanford University Press.

Rickford, John, and Christine Théberge Rafal. (1996). Preterite *had* + V-ed in the Narratives of African-American Adolescents. *American Speech* 71.3.

Romaine, Suzanne. (1984). On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Linguistica 18,* The Hague:Mouton, 409-437.

Sahgal, Anju. (1991). Patterns of language use in a bilingual setting in India. In Cheshire, J. (ed.). *English Around the World.* Cambridge: Cambridge University Press.

Sankoff, Gillian. (1973). Above and beyond phonology in variable rules. In C. J. Bailey & R. Shuy (eds.). *NWAVE* 1972, 44-61.

Shastri, S. V. (1988a). The Kolhapur Corpus of Indian English and work done on its basis so far. *ICAME Journal* 12: 15-16.

Shastri, S. V. (1992). Opaque and Transparent features of Indian English. In G. Leitner (ed.) *New Directions in English Language Corpora: Methodology, Results, Software Developments.* New York: Mouton de Gruyter.

Shastri, S. V. (1996). Using Computer Corpora in the Description of Language with Special Reference to Complementation in Indian English. In Baumgardner, R. J. (ed.) *South Asian English: Structure, uses, and users.* Urbana: University of Illinois Press

Solomon, Dennis. (1972). Form, content, and the post-creole continuum. Ms. Conference on Creole Languages and Educational Development, U.W.I. Trinidad.

Spitulnik, Debra. (1998). Mediating Unity and Diversity: The Production of Language Ideologies in Zambian Broadcasting. In B. Schiefflin, K. Woolard, P. Kroskrity (eds.) *Language Ideologies: Practice and Theory.* Oxford: Oxford University Press.

Sridhar, S. N. (1990). *Kannada.* New York: Routledge.

Sridhar, Kamal. (1991). Speech acts in an indigenised variety: sociocultural values and language variation. In J. Cheshire (ed). *English Around the World.* Cambridge: Cambridge University Press.

Sridhar, Kamal, and S. N. Sridhar. (1992). Bridging the Paradigm Gap: Second-Language Acquisition Theory and Indigenized Varieties of English. In B. Kachru (ed.) *The Other Tongue: English Across Cultures.* Urbana: University of Illinois Press

Tichy, P. (1979). The logic of temporal discourse. *Linguistics and Philosophy* Vol. 3. 343-69.

Trudgill, Peter, and Jean Hannah. (1994). *International English : a guide to varieties of standard English.* London: E. Arnold Press.

Williams, Jessica. (1987). Non-native Varieties of English: A Special Case of Language Acquisition. *English World-Wide* 8:2 161-199.

Winford, Donald. (1993). Variability in the use of perfect *have* in Trinidadian English: A problem of categorial and semantic mismatch. *Language Variation and Change 5*, 141-187.

---

[1] The debate out of which Macaulay's <u>Minute</u> emerged has been reinterpreted over the decades but still persists today in various aspects of language planning and education. The view that English is an inevitably indigenized part of the Indian linguistic landscape stands in contrast to one that emphasizes the need for linguistic self-determination distinct from the 'colonial' language. An article posted as recently as July 1998 at http://www.indolink.com/Analysis/politics24.html refers to the conflict between "hybrid Macaulyite" (implying a sort of betrayal of national interests) and "swadeshi" (approximately "selfmade"; a term that was associated with the Indian independence movement) persuasions in the country. The use of the term "Macaulyite" reflects the persistence of this event in present-day Indian consciousness.

[2] The 3-language formula for education has been interpreted variably due to regional differences as well as a North-South divide in attitudes towards Hindi (a Northern, Indo-Aryan language), which predates English as a lingua franca and an alternate national language. An example is Tamil Nadu, a state which witnessed language riots in the 1960s in which 70 people were killed and Hindi and English were both temporarily banished from the state (Kachru, 1983:90). The three language formula is reduced to a two- or even one-language formula in such regions, for political and practical reasons.

[3] See also Rickford (1987:27) on multidimensionality in Guyanese Creole.

[4] For more information on the corpora, check the following website:
http://www.hit.uib.no/corpora.html

[5] In order to control for subjective bias, a subset of the data was independently recoded following the same criteria by a second coder; the two sets of results matched with 94% accuracy.

[6] According to Labov's original formulation of the Principle of Accountability, variants such as (14b) and (14c) should ideally be contrasted with all occurrences of regular present perfect and preterite forms as well; i.e. alternants of these forms ought to be included. However, Labov later notes that this stipulation is too strict to apply to all studies of variation, as the set of possible variants for certain syntactic features cannot be clearly restricted. He cites the study of invariant <u>be</u> in AAVE as one such case (Labov, 1982:30). In this study as well, inclusion of all other past uses is both theoretically and methodologically intractable. I restrict the study to comparisons of sets of occurrences, listed in (14), across dialects.

[7] The potential for circularity arises here, because unlike structural variables, which are more commonly studied, semantic variables may derive part of their actual meaning from some of these internal constraints. However, I include the internal variables in (21) as none of them constitutes an explicit coding criterion and there appears to be important variation in the use of (optional) contextual cues in different varieties. The basic semantic coding relies primarily on the greater discourse context, not on the internal variables.

[8] As this study exclusively uses corpus material, the external variables are regrettably restricted to these broad categories, as detailed social correlates for tokens are irretrievable from the corpus texts.

[9] In spite of the interesting patterns of deviation among varieties, it must be noted that IndE shows fairly high levels of standard usage in general.  Other researchers on IndE and other indigenized Englishes have, in fact, observed that the analysis of such varieties can be quantitatively intractable due to inconsistency in levels of nonstandardness and the high degree of standardness in text registers (Solomon, 1972; Leitner, 1991). However, this tendency to standardness must also be seen as a real indication that differences between non-native varieties and native varieties often occur at the same low levels as differences among native varieties.

[10] The corpus subsection H contained only a few state government documents, which is why the numbers for regional bureaucracy are rather low.

[11] From http://www.rediff.com/sports/olyday.htm, September 19, 2000.

[12] I follow Solomon (1972) in using the term Trinidadian English. Winford (1993) treats the entire continuum as Trinidadian English and the creole vernacular as Trinidadian Creole.