1-1-1997

Optimality and the syntax of lectal variation

Rajesh M. Bhatt

Optimality and the syntax of lectal variation

Optimality and the Syntax of Lectal Variation*

Rakesh M. Bhatt

1. Introduction

This paper presents an account of English language variation in India. It has been a relatively daunting task to demonstrate that Indian vernacular English is just as systematic and logical as any other variety of English, say "standard" Indian, British or American English. This paper focuses on the syntactic differences between two varieties of Indian English — the standard and the vernacular — restricting the discussion to the syntax of null arguments *(pro drop)* and *wh*-question formation. These data, I will argue, are confounding for the mainstream syntactic models (Chomsky 1981, 1986, 1995), but when the same data are viewed from the optimality vantagepoint (Prince and Smolensky 1993), the differences between the two varieties follow as a natural consequence of the architecture of the theory.

Despite the advances in sociolinguistic theory over the past several decades, we still find in the literature numerous instances of syntactic variation presented as grammatical anomalies. Quirk (1990), echoing Prator (1968), claims, for example, that "the English of the teachers (in India and Nigeria) themselves inevitably bears the stamp of locally acquired *deviations from the standard* (British English) language ..." (ibid:8, emphasis added).[1] This discourse, we now know, is not new; it is a

reproduction of the early 60s discourse on Black English (African American Vernacular English), as evidenced in the works of Bereiter, Engelmann, and Jensen. What has changed in this ideological discourse is the geopolitical setting — from the inner city schools of the United States to English language education in India. With respect to AAVE, Labov (1970), Wolfram & Fasold (1974), among many others, were able to demonstrate successfully the empirical, methodological and theoretical-conceptual problems with the deficit model of Bereiter et al. This paper replicates the spirit of Labov and Wolfram & Fasold in dealing with English language variation in India. The following two principles guide the rationale of the study presented in this paper:

(1)   (a) Principle of error correction (Labov 1982:173)
A scientist who becomes aware of a widespread idea or social practice with important consequences. that is invalidated by his (or her) data is obliged to bring this error to the attention of the widest possible audience.
(b) Principle of linguistic gratuity (Wolfram 1993:227)
Investigators who have obtained linguistic data from members of a speech community should actively pursue positive ways in which they can return linguistic favors to the community.

The goal of this paper is to present a tidy demonstration of the syntactic differences between two varieties of English in India: the standard Indian English (SIE) and the Indian vernacular English (IVE). Thereafter, an optimality-theoretic (Prince and Smolensky 1993) account is presented which is able to yield the empirical generalization (and the intuition) that the grammar of IVE is just as systematic and logical as that of SIE.[2] This paper also presents, even if only tangentially, a strong argument for using Optimality Theory (OT) as a framework of research on language variation and change.[3]

[1] It is instructive to see how certain ideological startegies and rhetorical methods are continulaly manipulated to legitimize and rationalize the power of theoretical constructs like "standard", "deviant", etc. For a discussion, see Kachru (1996) and Bhatt (1995a).

[2] The "real" question here is whether this intuition actually plays a role in the grammatical process or whether it dissolves into taxonomic artifacts.

[3] Although I recognize that eventually a restrictive theory of language use is obligated to declare the precise nature of the "context of situation", which presumably yields observed realization of linguistic

Before discussing the methodology, the data, the generalizations, and the analysis, a brief socio-historical description of English in India will shed some light on the regional cultural identity of its users — one that is unrelated to the Judeo-Christian and Western ethos and its canons — and on the process of acculturation of the English language in local (Indian) contexts of use.

2.   English in India: Socio-historical Context[4]

English came to India around 1600 via the establishment of the East India Company. Although initially severely limited in the numbers of its speakers, English bilingualism increased with various strategies of trade and proselytizing, especially during 1614-18th century (cf. Duff 1837, Richter 1908, Law 1915). The proselytizing strategy was chiefly instrumental in introducing English bilingualism to the Indian subcontinent. Proselytization was rationalized in several ways; Grant (1831-31:60-61) had the following to say:

> The true curse of darkness is the introduction of light. The Hindoos err, because they are ignorant and their errors have never been laid before them. The communication of our light and knowledge to them, would prove the best remedy for their disorders.

After 1765 when East India Company established political control in India, and especially in early 19th century, the spread of English was aided and abetted by support from prominent Indians

---

expressions of a certain communicative act; I submit that such an attempt is beyond the scope of this paper. As a very brief, yet bold speculation I suggest that some articulated theory of diglossia, along the lines of Ferguson (1959) — where certain (H/L) forms are tagged with certain (H/L) functional domains — may account for the observed choices among the competing candidates of linguistic expressions (e.g., the alternation in the use between differentiated and undifferentiated tags; see Bhatt 1995b, for some discussion).

[4] Most of the discussion in this section is taken from Kachru (1983, 1996).

led by Raja Ram Mohan Roy, Dwarka Nath Tagore, and Rajunath Hari Navalkar, who preferred English to Indian languages for academic, scientific, and other intellectual inquiry. This local demand for English, coupled with Thomas B. Macaulay's Minute of 1835, led to the use of English in all official and educational domains. Macaulay's Minute, the first language policy in India, introduced English for the following purpose:

> To sum up what I have said, ... that we ought to employ them (Indians) in teaching what is best worth knowing; that English is better worth knowing than Sanscrit or Arabic; ... We must at present do our best to form a class who may be interpreters between us and the millions we govern; a class of persons, Indian in blood and colour, but English in taste, in opinions, in morals, and in intellect.

Although English instruction created bilinguals, it is worth pointing out that the models for pedagogy and acquisition were not native speakers. As Kachru (1996:907) notes: "Whatever the assumptions, in reality the teaching of English was primarily in the hands of the locals, and not with the native speakers of the language. ... It was, therefore, not unusual to find teachers with Irish, Welsh, or Scottish backgrounds overseeing the local teachers and educators involved in the teaching of English, who provided the models for the teachers, both in class and outside it." And, further, as the use and users of English increased, so did its acculturation to non-Western sociolinguistic contexts.

By the time India got its independence from Britain in 1947, English was firmly established as a medium of instruction and administration. With respect to the role of English in post-Colonial India, precious little changed: English still enjoys the status of associate official language; it continues to be the language of the legal system and the Parliament; it is one of the three mandatory languages introduced in schools; English newspapers are published in twenty seven of the twenty nine states and union territories, and they command the highest circulation in terms of the total reading public; the percentage of books published in English is higher than the percentage of books published in any other language; and, finally, in 1971, 74% of India's scientific journals and 83% of nonscientific journals were published in

English (Kachru 1990: 35-36). Presently, India is the third largest English-using nation (60 mil) after the USA and the UK.

As a result of over 200 years of contact with native Indian languages, English has become an Indian language, both in its structure and use. And like other natural languages, English in India displays a hierarchy of varieties — from standard (monitored) to vernacular (unmonitored). The standard and the vernacular are stable systems; the difference between them is a function of the formality of the context, in the sense of Labov (1972). Thus, the Standard Indian English is the variety used self-consciously by educated speakers in any formal domain of interaction, whereas the Indian Vernacular English is the variety used by the same speakers in routine social interactions, without exercising any conscious control of language use.

3.   Methodology

The proposal of language variation adopted in this paper is premised on two standard assumptions: (i) linguistic competence is the knowledge of what constitutes as optimal linguistic expression within a structured range of plausible alternatives, and (ii) the grammar of IVE is a product of the dynamics of language contact. The grammar is defined as a structured collection of behavioral tendencies; the job of the grammarian, then, is (a) to collate the observed tendencies into categorical paradigms of patterns (=descriptive adequacy), and (b) to explain why the patterns in fact obtain (explanatory adequacy). Given these assumptions, I propose the following hypotheses:

(2a)   IVE is just as systematic and logical as SIE;
(2b)   The grammars of IVE and SIE are constrained by the same set of grammatical constraints;
(2c)   The differences in the two varieties is a function of how each grammar prioritizes these constraints.

Three kinds of data were collected: (a) recordings of spontaneous speech (à la Labov 1972a); (b) data from published sources, like Kachru (1983) and Trudgill and Hannah (1985); (c) introspective judgments (Labov 1972b). The recorded data were collected using a portable DAT recorder to ensure the highest quality recordings. Altogether nine speakers (five men and four women) participated in the conversations. They all belonged to

educated middle-class families, and spoke, in addition to English, fluent Hindi. Their permission to use the recorded material in an anonymous fashion was obtained. The main topics discussed, although not restricted to, were: neighborhood disputes, wedding in the family, trip to a summer resort, and pollution in New Delhi, India. The conversations vary in length from approximately 10-35 minutes, representing approximately 7 hours of collected material. Furthermore, where recordings were not possible, notes were taken of what was said, and in what context. Finally, the data were collated, and a catalogue of the following syntactic properties was drawn:

(3a)   inversion/adjunction in wh-questions,
(3b)   referential null (topic) subjects (pro-drop),
(3c)   null expletives subjects ('silent' *it*)

The second kind of data comes from published sources like Kachru (1983) and Trudgill and Hannah (1985). Both of these sources were consulted, where possible, for comparisons with the spontaneous speech data (cf. also, Sells et al. 1994). Finally, judgments on crucial data (inversion in indirect questions, and subject and object pro-drop), unavailable in the published sources, were elicited from 27 native speakers of Indian English, which included high school English teachers, professionals (three doctors, two engineers) and two linguists.[5] In a small test instrument containing 4 items, subjects were given sentences with uninverted direct and inverted indirect questions (e.g., 'Nobody knows what is Indian government doing these days.'), and subject and/or object pro-drop sentences (e.g., Q: 'Do you have some tickets?'; A: 'Sorry, sold already.'). They were then asked to report whether the sentences were spoken by a speaker of Indian English or a speaker of British English. 4 out of 7 English teachers did not accept any instance of Indian vernacular English. The results are given in (4) below:

---

[5] Following Wolfram (1986), Sells et al. (1994, 1996), I have drawn comparisons of introspective data with spontaneous speech data to minimize the risk of hypo-and hyper- correction.

(4)

No. = 27	Direct Questions w/o inversion	Indirect Question w/ inversion	Subject/ Object pro-drop	Dummy subject pro-drop
Indian English	23	20	22	19
British English	1	1	1	0
Neither	4	7	5	8

The data in the table in (4) above demonstrates a surprisingly high awareness of endocentric (=Indian) norm of English. The result of this pilot study does replicate Kachru's (1976) study in which more than 55% of Indian graduate students reported using the variety of English they speak as "Indian English", compared to 29% labeling their's "British English". What the data in (4) suggest is that most speakers of Indian English are aware of the vernacular use of English (reported in (4) as Indian English) as well as the educated use of English, which I, for expository purposes, refer to as Standard Indian English (reported in (4) as British English). In the next section, I present syntactic differences between these two varieties, SIE and IVE, and show how the "standard" GB accounts (à la Chomsky 1981, 1986, 1993) fail to capture these differences in a systematic way.

4. The Data, the Generalizations, and the Standard Accounts

4.1. Extraction Facts

In Standard Indian English (henceforth, SIE), root questions are formed by moving the wh-phrase to the left-periphery (Spec-CP) of the clause followed by, in non-subject extractions, the auxiliary (in Comp). Some examples are given in (5) below:

(5a)   What$_i$ has$_j$ he t$_j$ eaten t$_i$?

(5b)   What$_i$ *do* you want t$_i$?

(5c)   [How much interest]$_i$ *did* they charge you t$_i$?

(5d)   Why do you look worried?

Embedded indirect questions in SIE also involves movement of the wh-phrase to the left-periphery (Spec-CP) of the embedded clause, without, however, any auxiliary following it (in Comp). Some examples are given in (6) below:

(6a)   They know who$_i$ Vijay has invited t$_i$ tonight.

(6b)   I wonder where$_i$ he works t$_i$.

(6c)   I asked him what$_i$ he ate t$_i$ for breakfast.

(6d)   Do you know where$_i$ he is going?

The well-known empirical generalization about data such as (5) and (6) is that Inversion is restricted to matrix sentences; it does not apply in embedded contexts. This generalization is expressed in the standard GB accounts (May 1985, Chomsky 1986, Rizzi 1990) in terms of Wh-Criterion, given in (7) and the relevant structural configuration shown in (8) below.

(7a)   Each +wh X^0 must be in a Specifier-head relation with a wh-phrase.

(7b)   Each wh-phrase must be in a Specifier-head relation with a +wh X^0.

(8)
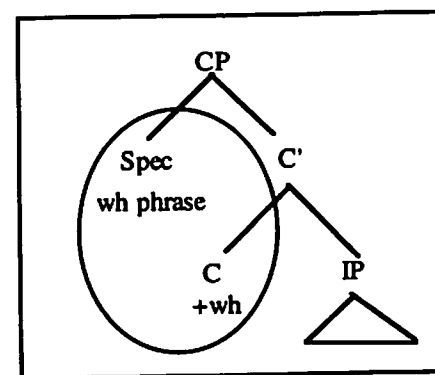
According to the Wh-Criterion, the data in (5) and (6) are explained by assuming that INFL is specified [+wh] (see Rizzi 1990), and the role of inversion in matrix context is to carry the [+wh] feature to a position where it can satisfy (7a). In (5b), the [+wh] INFL is moved to C and the empty verb *do* is inserted to

support the stranded INFL. In embedded questions the [+wh] feature is specified on the embedded C by the selection properties of the matrix verb. Wh-movement to the embedded Spec-CP satisfies (7b). Inversion is excluded since C is content-full (i.e., has [+wh], and therefore movement of INFL to C would violate the Projection Principle. In (9a) and (9b) below, I show the wh-movement operations in direct and embedded contexts.

(9a)
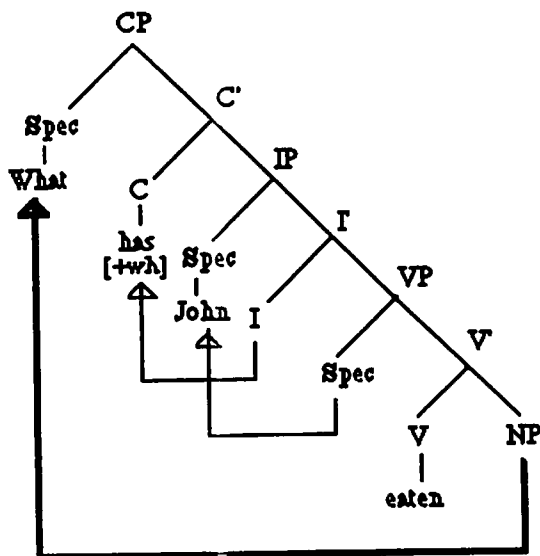
In Indian Vernacular English (henceforth, IVE), on the other hand, root questions are formed also by moving the wh-phrase to the left-periphery (Spec-CP) of the clause. However, there is no auxiliary (in Comp) following the left-moved wh-phrase. Some examples are given in (10) below:

(10a)   What$_i$ he has eaten t$_i$?
(10b)   What$_i$ you want t$_i$?
(10c)   [How much interest]$_i$ they charged you t$_i$?
(10d)   Why you look worried?

(9b)
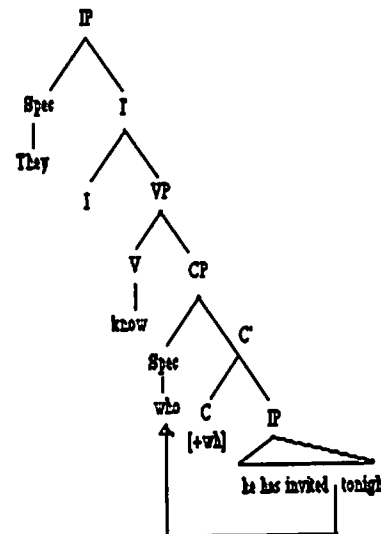
The embedded (indirect) questions in IVE involves wh-movement to the left-periphery (Spec-CP) of the embedded clause. The wh-phrase, surprisingly, is followed by the auxiliary, i.e., wh-movement in embedded contexts is accompanied by auxiliary movement (inversion) to, presumably, Comp. The relevant data is given in (11) below:

(11a)   They know who$_i$ has$_j$ Vijay t$_j$ invited t$_i$ tonight.
(11b)   I wonder where$_i$ does he work t$_i$.
(11c)   I asked John what$_i$ did he eat t$_i$ for breakfast.
(11d)   Do you know where$_i$ is$_i$ he t$_j$ going t$_i$?

The simple empirical generalization that emerges from data in (10) and (11) is that in IVE, inversion is restricted to embedded questions; it does not apply in matrix questions. The interesting empirical fact is that the question formation strategy in IVE is just the mirror image of that in SIE. Following the work of Déprez (1991) and Bakovíc (1995), I will assume that XP-movements that require a following head are movements to Specifier, while all other movements are adjunctions. The relevant

structural configurations for questions in IVE are given in (12a) and (12b) below:

(12a) Adjunction

he has eaten t
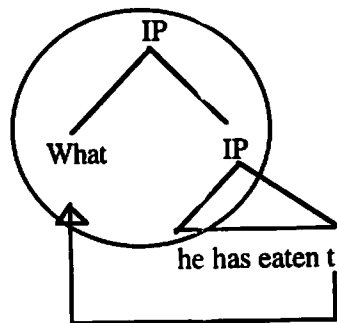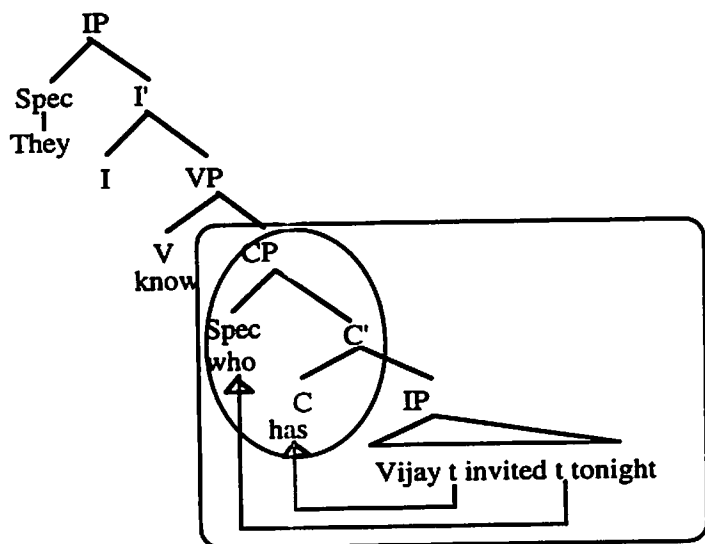
(12b) Inversion

Given (10), (11) and (12a,b), it is not possible to maintain the Wh-Criterion (cf. 7a,b, above) for the following reasons:

(13a) Adjunction data (10a-d) do not follow from the Wh-Criterion—violates (7b).

(13b) Inversion data (11a-d) violates the Projection Principle.

Given the standard GB account, the grand prediction is that the data such as given in (10) and (11) above are simply not English. In the next section, I present more data from IVE that is problematic for the standard GB-theoretic accounts.

4.2. Pro-Drop

With respect to argument pro-drop, SIE works like other regional standard British and American varieties: Finite clauses without subject are disallowed, as shown in (14a) and (14b) below:[6]

(14a) *pro* dances well.

(14b) *He said that *pro* would come tomorrow.

There are numerous reports in the literature linking pro-drop to rich agreement. In other words, there seem to be a general association between subject (pro) drop and rich (person, number, gender) agreement (=Licensing). The agreement affixes can recover the phi-feature (person / number / gender) content of the dropped subject (=Identification). Languages which have rich subject-agreement morphology, like Spanish, Italian, and Hindi allow pro-drop, whereas SIE, as well as other standard English varieties, has impoverished agreement morphology, and, therefore, does not allow pro-drop.[7]

Under the standard GB account (cf., Rizzi 1986, Jaeggli and Safir 1989), there are two requirements for pro-drop. The first is the "Licensing" requirement; that Pro-drop is allowed if that position is Case-governed by a "licensing" head, which can vary from language to language. Thus, INFL is a licensing head in

---

6 Mufwene (1988) discusses several instances in casual speech style of English where subject pro-drop is possible, e.g., 'Just stopped by to say hello!'

7 Chinese is an exception to this generalization: it has no agreement morphology, and yet is allows empty categories. According to Huang (1984), Chinese uses a different mechanism to license pro: it is variable bound to a zero topic.

Spanish, but not in English, and therefore pro-drop is not "licensed" (possible) in English. The second requirement is the "Identification" requirement; that the content of the pro must be fully recoverable. One way to achieve identification is when pro is coindexed with features of person and number on its Case-governing head. Again, the impoverished English Agr is unable to identify/recover the content of pro. These two requirements, licensing and identification, predict the ungrammaticality of (14a) and (14b).

The pro-dropping facts of IVE are interesting. IVE, like Spanish and Italian, allows pro-drop, as shown in (15a), (15b), and (15c).[8]

(15a) He played cricket all day today — and now *pro*   does not want to work on his homework!

(15b) Subject and Object pro-drop
  A: You got tickets?
  B: No, *pro* sold *pro* already.

(15c) A: Is he in his office?
  B: Sorry, *pro* left just now only.

The data in (15) pose two empirical problems to standard GB account. The first problem is that like SIE, IVE is morphologically impoverished, and therefore should not license pro-drop; but it does. It is possible to stipulate that INFL in IVE is a licensing head, just as in Spanish. This stipulation, however, is fraught with empirical problems. Unlike Spanish and Italian, IVE does not have Subject-Verb inversion, e.g., '*Speaks he.' Further, unlike Spanish and Italian, IVE does not show any trace of *that-t* effects. Thus, sentences such as 'who did you say that came' in IVE are ungrammatical, although similar sentences in Spanish and Italian are not. The second problem is that IVE does behave

---

[8] Platt at al. (1984) discuss similar data for Malaysian and Hong Kong English, as shown below:
  (a) Dis Australians, you see dem hold hand hold hand, honey here, honey there, darling here, darling dere, next moment *pro* separated already.
  (b) If you don't like *pro*, *yaya* (nursemaid) will give you water.
  (c) In Australia, people never carry umbrella — so if you carry *pro* they will laugh at you.

like Spanish and Italian in that it does not require semantically empty subjects (discussed in next subsection).

Although the pro-dropping facts in IVE do not follow standard explanations of Licensing and Identification, on closer examination we notice that the absence of an overt subject in IVE is not free — it is required when the subject is coreferential with the discourse topic (cf. Huang 1984, Grimshaw & Samek-Lodovici 1995). Briefly, discourse topic here is defined as what the sentence is all about. The distribution of pro-drop in IVE is similar to Italian as argued in Samek-Lodovici (in preparation) and Grimshaw & Samek-Lodovici (1995). Compare the Italian data in (16) and the IVE data in (17):

(16) Q: E' partita [la madre di Gianni]?
        Did John's mother leave?
   A: Si, **pro**   /*lei   e' partita
        Yes, (she) left.

(17) A: Is he in his office?
   B: Sorry, *pro* left just now only.

In Italian (16), as argued by Grimshaw & Samek-Lodovici (1995), pro-dropping is restricted to those arguments which are topic-connected. Topic-connected arguments must obligatorily drop. In IVE (17), however, where in B's response the subject is (referring to) the topic, it can optionally be dropped. The generalization, then, for pro-dropping in IVE is that pro-drop is restricted to those arguments (subject/objects) that are topic-connected. The difference between Italian and IVE is that in the former, topic-connected arguments drop obligatorily, whereas in IVE the dropping of topic-connected arguments is only optional.

4.3.   Null Expletive (*it*) Subjects

Turning now to null expletive subjects, SIE requires dummy subjects in finite clauses, as shown in (18a) whereas IVE does not require dummy subjects in finite clauses, as show in (18b).

(18a) *pro* is clear that he will not come.
(18b) Here *pro* is not safe to wait.

Under the standard account (18a) is ungrammatical due to the violation of formal "licensing" and "identification" requirements. The grammaticality of (18b) has no account under "licensing" and "identification" requirements (cf., Sells, Rickford & Wasow 1994).

## 5. Optimality Theory — A Description

The starting point in our discussion of the framework is the overall rationalist approach in which Optimality Theory (henceforth, OT) is grounded. The rationalist approach is premised on the assumption that grammatical intuitions provide privileged access to the system underlying language performance. Under this approach, the collection of the grammatical judgments of the "idealized native speaker/hearer" represents linguistic competence — the knowledge which underlies the use of language. The idealization, of course, and unfortunately, leaves no room for language variation, or its account thereof. And further, when linguist's introspection conflicts with actually observed utterances, the former prevails in the construction of grammars.

Although OT is rationalist in spirit, it departs from the traditional frameworks in its ability to accommodate linguistic variation, as will become clear momentarily. OT differs from orthodox rule/principle-based approach (à *la* Chomsky 1965, 1981, 1995) in the manner discussed below.

Optimality theory (Prince and Smolensky 1993) is about how grammars are defined by constraint hierarchies (McCarthy 1995). Universal Grammar in OT is expected to provide a finite set of potentially conflicting (violable) constraints on structural well-formedness. Languages differ from each other in terms of how each ranks the set of violable constraints. Thus, in essence, different configurations of constraint ranking yield, in principle, different grammars, as shown schematically in (19). If so, it follows that minimally different constraint rankings will give rise to dialect variation, theoretically. Adopting OT thus provides a mechanism to faithfully account for the subtle grammatical differences between SIE and IVE, without risking empirical coverage.

(19)
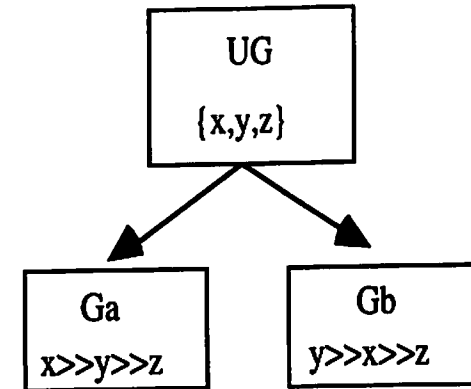
OT, instead of using categorical constraints to express empirical generalizations, uses "violable" (soft) constraints. These soft constraints are violable in just those contexts in which they conflict with a higher ranked constraint. The core ideas of OT can be summed up in the following way: constraints can be violated; constraints are ranked; and the optimal form is grammatical. Generative grammar consists of ranked constraints which examine (via *Eval*) all possible candidate structural descriptions freely generated by input-output function (*Gen*). This is illustrated in (22) below.[9] The output that has the least violations (=0, in the best case scenario) is optimal, i.e., grammatical.

(20) OT Grammar:

INPUT ➤ [GEN] ➤ [EVAL] ➤ Interpretation
{11 .....ln}   candidate   optimal
            structures   structures

Before I close the discussion in this section, let me give an illustration of how OT accounts for language variation (cf., also Anttila (1995)): Consider two grammars, Grammar *A* and Grammar *B*, both of which have three constraints {x, y, z}.

---

[9] This diagram is taken from one of the handouts of the talk given at the MIT-OT workshop in May 1995. Regrettably, I have lost the handout, and the reference of the speaker.

Assume further, that in Grammar *A* these constraints are ranked in such a way that {x} dominates {y} dominates {z} [= x >> y, y >> z, x >> z]. In other words, Grammar *A* imposes a total order on the constraints: x >> y >> z. Now, assume that for a certain input we get two competing output candidates: *cand 1* and *cand 2*. Tableau 1 shows the competition between the two candidates. *Cand 1*, violates the highest ranking constraint {x}, which is lethal.   Grammar *A*, therefore, chooses *cand 2* straightforwardly as the optimal, grammatical, option.

Tableau 1: output = *cand2* :

Candidates		x	y	z
a.	*cand1*	*!		
b. ⇒	*cand2*		*	

Now consider the other grammar, Grammar *B*. Assume that it, too, has the same three (universal) constraints {x, y, z}, however, this grammar imposes slightly different ordering, viz., the constraint {y} dominates {x} dominates {z}. Now for the same input, as in Grammar *A* , we get the same two candidate competing outputs: *cand 1* and *cand 2*. The optimal output, as shown below in Tableau 2, is *cand 1*, because in this grammar *cand 2* violates a higher ranked constraint {y}, leading to its rejection as optimal.

Tableau 2: output = *cand1*:

Candidates		y	x	z
a. ⇒	*cand1*		*	
b.	*cand2*	*!		

In the next section I will propose a set of potentially conflicting meta-linguistic constraints and show how their interactions yield well-formed utterances in both SIE and IVE.

6.   The OT Account

I now present the analysis of the data following Labov (1972a) who has argued that (syntactic) variation is usually not free or indeterminate; it can often be shown to be systematic. In that spirit, I propose, following Sells, Rickford and Wasow (1994, 1996), the (universal) constraints listed in (21) to account for the syntactic variation discussed in section 4.0.

(21a)	**OP-SCOPE**:	Operators (e.g., wh-phrase) must take scope over the entire proposition (=c-command VP/IP at S-Structure).
(21b)	**OP-SPEC**:	Operators must be in Specifier position.
(21c)	**STAY**:	No movement (=trace) is allowed.
(21d)	**OB-HD**:	Heads of selected projections must be filled (either by trace or overt material)
(21e)	**SUBJECT**:	The canonical subject position (=highest A-Specifier in an "extended projection" (Grimshaw 1991) must be filled. (=EPP, *à la* Chomsky 1981, but violable).
(21f)	**DROP TOP**:	Leave arguments coreferent with the topic structurally unrealized.
(21g)	**PARSE**:	Parse input constituents
(21h)	**FULL INT**:	Parse lexical conceptual structure. (Failed by expletives and auxiliary do)

Before I show how these constraints interact to yield the syntactic differences between SIE and IVE, a couple of theoretical assumptions need to be explicitly stated. First, the constraints listed in (21) are not necessarily 'surface-true'; this is expected since the constraints which are always surface-true are going to be those which either do not conflict with any other constraint, or are always victorious in any conflict by virtue of the fact that they are always ranked higher than those with which they conflict (Grimshaw 1994).

Second, and importantly, the constraints listed above in (21) are universal; the grammar of every language has them and that all rankings of them are possible. This assumption follows a central meta-principle of OT, called UNIV, which says that constraints are universal. The constraints listed under (21) have been extensively discussed in the OT literature (cf., Sells, Rickford & Wasow (1994, 1996), Grimshaw (1994), Grimshaw & Samek-

Lodovici (1995), Smolensky et al. (1995), Bakovíc (1995), and Samek-lodovici (in preparation)).

6.1. Matrix (Direct) Questions

Beginning with the matrix questions in SIE and IVE, we need to deal with the problem of Inversion vs. Adjunction, i.e., whereas SIE allows subject-verb inversion, IVE does not. In the discussion of questions, OP-SCOPE will not appear in the tableaux because it is inviolable in both SIE and IVE. This constraint forces wh-phrase to move to a position from which it can c-command the entire proposition (=IP) at S-Structure. The constraints that need to be recruited to yield direct questions are: OP-SPEC and STAY. The interaction of these two constraints in the order given in (22) yields the categorical prediction of direct questions in SIE: the wh-phrase in CP-Spec followed by an aux in Comp. The tableau in (23) shows a competition between two candidates, an adjunction structure and an inversion structure. Both violate the low ranking constraint STAY, and therefore STAY remains inactive on the candidate set. Since adjunction violates OP-SPEC, a higher ranked constraint, while inversion does not, inversion structure is more harmonic, and wins.

(22)   SIE: **OP-SPEC >> STAY**

(23) Tableau: SIE

Candidates		OP-SPEC	STAY
adj	What you would like to eat *t* ?	*!	*
inv ⇒	What would you *t* like to eat *t* ?		**

Turning to IVE, recall that direct questions in IVE involve an adjunction structure, (12a) above; the wh-phrase adjoins to IP-Spec instead of moving to CP-Spec as it does in SIE. It turns out that both OP-SPEC and STAY yield the adjunction structure too, albeit with a different ranking. The IVE grammar ranks STAY over OP-SPEC (24, below), which gives the desired results (25). The tableau in (25) shows, again, two competing candidates, both violating the highest ranking constraint STAY. Notice however, the inversion structure incurs two violations of STAY — one by moving wh-phrase and the other by moving the Infl/Aux to Comp — as opposed to only one violation of STAY — moving wh-

phrase — in the adjunction structure. In this competition, inversion loses because it incurs more violations than adjunction.

(24)   IVE: **STAY >> OP-SPEC**

(25) Tableau: IVE

Candidates		STAY	OP-SPEC
adj ⇒	What you would like to eat *t* ?	*	*
inv	What would you *t* like to eat *t* ?	**!	

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

6.2. Embedded (Indirect) Questions

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

Consider first SIE. Since SIE does not permit inversion in indirect questions, OB-HEAD must have a lower prominence vis-a-vis OP-SPEC and STAY. We have already established that the grammar of SIE ranks OP-SPEC over STAY (22, above). OB-HEAD, given its diminished status in SIE, must be ranked below STAY; the relevant ranking is given in (26).

The tableau (27) shows two competing candidates, both deferential to OP-SPEC. Since OP-SPEC cannot distinguish between the two candidates, the evaluation is passed on to the next important constraint, STAY. Again both violate STAY, but it is the inversion structure that incurs two violations of STAY as opposed to non-inversion structure which violates STAY only once. In this competition, then, non-inverted structure is harmonic, and wins.

(26)   **SIE: OP-SPEC >> STAY >> OB-HD**

(27) Tableau: SIE

Candidates	OP-SPEC	STAY	OB-HD
no-inv ⟹ I wonder what *e* he is eating *t*		*	*
inv I wonder what is he *t* eating *t*		**!	

Turning to indirect questions in IVE, recall that these require inversion with wh-movement, suggesting that OB-HEAD is a constraint of high-prominence. Recall, too, that we have already established that in IVE STAY outranks OP-SPEC (24, above). By ranking OB-HEAD over STAY and OP-SPEC, as shown in (28), we get the desired output.

Once again, the tableau in (29) shows two competing candidates. The optimal output, given the dominance hierarchy in (28), is the inverted structure because the non-inverted structure violates OB-HEAD.

(28)   **IVE: OB-HD >> STAY >> OP-SPEC**

(29) Tableau: IVE

Candidates	OB-HD	STAY	OP-SPEC
no-inv I wonder what *e* he is eating *t*	*!		
inv ⟹ I wonder what is he *t* eating *t*		**!	

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

6.3. Pro Drop

The empirical facts of pro-drop are straightforward: SIE, like other standard varieties of English, does not permit pro-drop. IVE, on the other hand, allows pro-drop but it is restricted to those arguments (subject/objects) that are topic-connected. These different patterns of generalization can be expressed by letting three constraints — PARSE, DROP TOPIC, and SUBJECT — interact in different ways. Since SIE does not permit argument pro-

dropping, it must be the case that PARSE (an argument) and SUBJECT are ranked higher in priority than DROP TOPIC. The non-pro-drop phenomenon in SIE follow from the dominance configuration given in (30).

As shown in tableau (31), candidate (b), which satisfies PARSE and SUBJECT is preferred over both candidate (a), which violates PARSE, and candidate (b), which violates SUBJECT. Thus the ranking PARSE above SUBJECT above DROP TOPIC yields the non-pro-drop generalization in SIE.

(30)   **SIE: PARSE >> SUBJECT >> DROP TOPIC**

(31) Tableau: SIE

Candidates	PARSE	SUBJECT	DROP TOPIC
(a) left just now only	*!	*	*
(b) ⟹ he left just now only			*
(c) left just now only he		*!	*

Turning to pro-drop in IVE, we find evidence of different ranking of the three constraints. Earlier, in (15a-c), we provided evidence that the grammar of IVE does not require an overt subject (or object) when it is topic-connected, which means that the constraint DROP-TOPIC must dominate PARSE and SUBJECT. In fact, the ranking configuration in (32) gets us the desired results. In tableau (33), we find that candidate (a) is the harmonic choice since the other two candidates incur violations of the highest ranked constraint, DROP-TOPIC.

(32)   **IVE: DROP TOPIC >> PARSE >> SUBJECT**

(33) Tableau: IVE

Candidates	DROP TOPIC	PARSE	SUBJECT
(a) ⟹ left just now only		*!	*
(b) he left just now only	*!		*
(c) left just now only he	*!		*

With respect to the phenomenon of pro-drop, the difference between the grammars of SIE and IVE is reducible to different rankings of the same constraints.

6.4.   Null Expletive (*it*) Subjects

Turning finally to null expletive subjects, we noticed earlier in (18) that SIE requires expletives in finite clauses whereas IVE does not require expletives in finite clauses.   This difference follows from the interaction and satisfaction of two constraints, FULL INT and SUBJECT. We follow Grimshaw & Samek-Lodovici (1995) in assuming that an expletive is a regular pronoun with its lexical conceptual structure (at least partly) unparsed. Since SIE requires subject, even expletives in subject position, it must be the case that in this grammar SUBJECT outranks FULL INT, as shown in (34). In tableau (35), the candidate with the expletive in subject position (=IP-Spec) wins because it only violates FULL INT whereas the other candidate violates SUBJECT, a fatal violation, given the prioritized ranking in (34).

(34)   SIE: SUBJECT >> FULL INT

(35) Tableaux: SIE

Candidates	SUBJECT	FULL INT
pro is clear that he will not come	*!	
⇒ It is clear that he will not come		*

In IVE on the other hand, expletives can be dropped from subject position. This generalization can be captured by re-ranking the two constraints, SUBJECT and FULL INT in such a way that FULL INT outranks SUBJECT, as shown in (36). The tableau in (37) shows that in IVE an optimal candidate will satisfy FULL INT at the expense of violating SUBJECT.

(36)   IVE: FULL INT >> SUBJECT

(37) Tableaux: IVE

Candidates	FULL INT	SUBJECT
⇒ *pro* is clear that he will not come		*
It is clear that he will not come	*!	

To sum up, with respect to null expletives, the interaction of two constraints, FULL INT and SUBJECT, yields the distributio- nal differences between SIE and IVE.

7.   Conclusions

The success of (socio)linguistic theory depends largely, I believe, in its ability to demonstrate the systematic nature of language variation and use. In this paper I have argued that the mechanism of constraint interaction and satisfaction, as conceptualized in OT, allows for a straightforward account of English language variation in India.   In OT, UG is conceptualized as a set of potentially conflicting constraints holding in all languages, with cross-linguistic variation arising from the fact that different languages, language varieties resolve the conflicts among these constraints differently.

I have presented evidence to claim that the differences between the observed patterns of generalization in SIE and IVE are best accounted for in a conceptualization of grammar that is based on a general notion of priority.   This OT-theoretic conceptualization allows us to capture the intuition that the grammatical constraints that govern the syntactic behavior of IVE are *not* unique to it. Specifically, in section 6, I have shown that the difference between the grammars of SIE and IVE is reducible to different rankings of the same constraints, which is only expected given that OT appeals to variation in ranking to provide different grammars.

Given the logic of the argument, that variation in constraint ranking yields different grammars, and the evidence presented in sections 4 and 6 to support it, it does not seem plausible to maintain the "deviation from the norm" hypothesis of Quirk (1988, 1989, 1990; cf., also Prator 1968) to account for variation in Indian English.

Finally, I believe that studies on language variation, such as this one (cf. also Sells, Rickford and Wasow 1994, 1996, Rickford et al. 1994, Mesthrie 1992), show ways in which sociolinguistic theory and current syntactic theorizing can inform and enrich each other.

References

Anttila, Arto (1995). Deriving variation from grammar: A study of Finnish Genetives. ms. Optimality Archives.

Bakovic, Eric. (1995). A Markedness Subhierarchy in Syntax. Presented at OT Workshop at MIT, May 19, 1995.

Bamgbose, A. (1992). Standard Nigerian English: Issues of Identification. In B. Kachru (ed) *English Across Cultures.* Urbana, IL: University of Illinois Press. 148-161.

Bhatt, Rakesh (1995a). Prescriptivism, creativity, and world Englishes. *WE,* 14:2. 247-260

Bhatt, Rakesh (1995b). The Uprooted, The Indentured, and The Segregated:South African Indian English. *JPCL,* 10:2. 381-396.

Bokamba, E. (1992). The Africanization of English. In B. Kachru (ed) *English Across Cultures.* Urbana, IL: University of Illinois Press. 125-147.

Chomsky, Noam. (1981). *Lectures on Government and Binding.* Dordrecht, Holland: Foris.

Chomsky, Noam (1986). *Knowledge of Language.* New York: Praeger.

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

Déprez, V. (1991). WH-Movement: Adjunction and Substitution. In D. Bartes (ed.) *WCCFL 10.* Stanford: CSLI. 103-114.

Duff, A. (1837). *New Era of the English Language and English Literature in India.* Edinburgh: J. Johnstone.

Goodall, G. (1991). On the Status of SPEC of IP. In D. Bartes (ed.) *WCCFL 10.* Stanford: CSLI. 175-182.

Grant, C. (1831). Observations on the State of Society among the Asiatic Subjects of Great Britain, particularly with respect to Morals, and the Means of Improving it. In *General Appendix to Parliamentary Papers 1831-1832.* London.

Grimshaw, Jane (1991). Extended Projection. Ms. Brandeis University.

Grimshaw, Jane (1994). Projections, Heads, and Optimality. Ms. Rutgers U (To appear: Linguistic Inquiry).

Grimshaw, Jane (1995). Optionality and Optimality. .Presented at OT Workshop at MIT, May 19, 1995

Grimshaw, Jane & Vieri Samek-Lodovici (1995). Optimal Subjects. *UMOP 18.* 589-606.

Hancin-Bhatt, B. and R. Bhatt (1995). Optimality, Crosslanguage Transfer, and L2 Phonology. Paper presented at SLRF '95, Cornell University, Ithaca, New York. Sept 29-Oct 1.

Hosali, P. & J. Aitchison (1986). Butler English: A minimal pidgin? *JPCL* 1:1. 51-80.

Jaeggli, O. & K. Safir (eds.) (1989). *The Null Subject Parameter.* Dordrecht, Holland: Kluwer.

Kachru, Braj (1983). *The Indianization of English.* New Delhi: OUP.

Kachru, Braj (1990). *The Alchemy of English.* Urbana, IL: University of Illinois Press.

Kachru, Braj. (1996). English as lingua france. In Hans Goebl et al. (eds.) *Contact Linguistics.* New York: Walter de Gruyter. 906-913.

Labov, William. (1970). The logic of nonstandard English. In Alatis, James (ed.) Linguistics and the teaching of standard English to speakers of other languages and dialects. Georgetown University Round Table on Languages and Linguistics 1969. Washington, DC: Georgetown University Press.

Labov, William (1972a). *Sociolinguistic Patterns.* Philadelphia: University of Pennsylvania Press.

Labov, William (1972b). Some Principles of Linguistic Methodology. *Language in Society,* 1. 97-154.

Labov, William (1982). Objectivity and Commitment in Linguistic Science. *Language in Society,* 11. 165-201.

Law, N.N. (1915). *Promise of Learning in India by Early European Settlers.* London: Longman.

May, Robert (1985). *Logical Form.* Cambridge: MIT Press.

McCarthy, J. (1995). Some Misconceptions about Optimality Theory. Ms. UMass, Amherst.

Mesthrie, R. (1992). *English in Language Shift.* Cambridge: CUP.

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

Pollock, Jean-Yves (1989). Verb Movement, UG and the Structure of IP. *Linguistic Inquiry,* 20. 365-424.

Prince, Alan & Paul Smolensky (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms. Rutgers U

Quirk, R. (1988). The question of standard in the international use of English. In Peter Lowenberg (ed.) Language spread and language policy: issues, implications, and case studies. Proceedings of GURT 1987, Washington DC.

Quirk, R. (1989). Language varieties and standard language.. *JALT II,* 1, 14-25.

Quirk, R. (1990). Language Varieties and Standard Language. *English Today,* 21, 3-10.

Richter, J. (1908). *A History of Missions in India.* New York: F. H. Revell

Rickford J., T. Wasow, N. Mendoza-Denton & J. Espinoza. 1994. Syntactic Variation and Change in Progress: Loss of the Verbal Coda in Topic-Restricting *as far as* Constructions. *Language,* 70.

Rizzi, Luigi (1986). Null Objects in Italian and the Theory of *pro*. *Linguistic Inquiry*, 17, 501-58.

Rizzi, Luigi (1990). Speculations on Verb Second. In J. Mascaró and M. Nespor (eds.), *The Henk van Riemsdijk Festchrift, Grammar in Progress*. Dordrecht: Foris. 375:386.

Samek-Lodovici, Vieri. In preparation. Topic and Focus Effects on Clause Structure: An Optimality Theory Approach. Ph.D Dissertation, Rutgers University.

Sells, P. J. Rickford & T. Wasow (1994). An Optimality Theoretic Approach to Variation in Negative Inversion in AAVE. Ms. Stanford University.

Sells, P. J. Rickford & T. Wasow (1996). An Optimality Theoretic Approach to Variation in Negative Inversion in AAVE. *Natural Language and Linguistic Theory* 14: 591-627.

Sey, K. (1973). *Ghanian English*. London: Macmillian.

Smolensky, Paul et al. (1995). When is Less More? Faithfulness and Minimal Links in wh-Chains. Presented at OT Workshop at MIT, May 19, 1995.

Sridhar, S. N. (1992). The Ecology of Bilingual Competence: Language Interaction in the Syntax of Indiginized Varieties of English. *World Englishes*, 11:2/3, 141-150.

Sridhar, K. K. & S. N. Sridhar (1992). Bridging the Paradigm Gap: Second Language Acquisition Theory and the Indiginized Varieties of English. In B. Kachru (ed) *English Across Cultures*. Urbana, IL: University of Illinois Press. 91-107.

Trudgill, P. & J. Hannah (1985). *International English*. London: Edward Arnold

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

Wolfram, Walt (1993). Ethical Considerations in Language Awareness Programs. *Issues in Applied Linguistics*, 4. 225-255.

Department of English
University of Tennessee
Knoxville, TN 37966

rakesh@utk.edu