Towards a Theory of Constraints in OT: Emergence of the not-so-unmarked in Malayalee English

Tara Mohanan and K. P. Mohanan
National University of Singapore
elltaram@nus.edu.sg ellkpmoh@nus.edu.sg
Version: May 2003

Contents

Towards a Theory of Constraints in OT: Emergence of the not-so-unmarked in Malayalee English*

Tara Mohanan and K. P. Mohanan
National University of Singapore

Abstract

This paper explores the phenomena of 'persistence' and 'emergence' in the patterns of contrast, distribution, and alternation in the phonology of Malayalee English, a transplanted second language system, and argues for the need to supplement the theory of constraint interaction in Optimality Theory with an explicit theory of constraint generation. The proposal involves two parts. First, when two or more constraints share a common core but vary in the details of their manifestation, their redundancy can be eliminated by deriving them from an underspecified *constraint core*. Adding further specifications of domain, locus, trigger, or outcome value, would yield the fully specified constraints. The constraints derived from the same constraint core would form a *constraint homologue*, providing a basis for the study of invariance and probable variations in language typology, language change, and language contact. Second, to capture cross-linguistic probabilities in the ranking of constraints, we propose that each markedness constraint homologue be paired with its faithfulness counterpart, with a universal default ranking relation. Such a constraint pair with default ranking expresses a universal tendency. In a weak tendency, faithfulness outranks markedness; given a language, reversing the default ranking would activate the tendency. In a strong tendency, markedness outranks faithfulness; reversing the ranking would deactivate the tendency. Strong tendencies correspond to the unmarked state of affairs in the SPE sense of 'unmarked'. Our analysis also reveals that, contrary to expectation in language contact, some patterns in ME that it shares with neither its substrate nor its superstrate sources illustrate the emergence of marked configurations of facts. We suggest that these marked structures can be explained as the resolution of conflicting pulls from the parent languages within the space provided by universal grammar.

* We are deeply indebted to Arto Anttila for detailed discussions and alternative accounts, and to Devyani Sharma for extensive and insightful comments. We gratefully acknowledge Joan Bresnan, Paul Kiparsky, and Peter Sells for conceptual clarification and useful suggestions, Bao Zhiming, Chin Seok Koon, Madalena Cruz-Ferreira, and Vivienne Fong for comments on an earlier version, and Sunita Abraham for also detailed editorial suggestions. We owe thanks to Joan Bresnan and the members of her Advanced LFG class (Winter 2003) for their significant contribution in directing our attention to important issues in OT, and stimulating the growth of the paper. The paper has benefited from questions and suggestions from the audience at the Stanford Linguistics Colloquium (April 2003), and from subsequent discussions with John Singler.

1 Introduction

1.1 The phenomenon

Instances of a target language contrast being lost in a second language system are well known. A classic example is the loss of the contrast between [r] and [l] in the English of many Japanese and Chinese speakers, and the loss of the contrast between [s] and [z] in the English of many South Indian speakers. Such loss of contrast is typically attributed to the influence of the first language system.

The converse of this phenomenon, however, is somewhat unexpected. Imagine a group of American speakers learning Japanese. It would be surprising to find their variety of Japanese to develop a contrast between [r] and [l]. Yet, such a phenomenon is found in Malayalee English (henceforth ME), a variety of English used by speakers of Malayalam,[1] which exhibits a contrast between alveolar and retroflex consonants:[2]

(**1**)	a.	*slowness*	[sloonas]	b.	*lawless*	[lɔɔlas]	c.	*very*	[weri]
		bonus	[booNas]		*folly*	[fɔɔLi]		*merry*	[meRi]

This contrast in ME can be traced to the contrast in Malayalam illustrated in (2):

(**2**)	a.	[pattam]	'herd'	b.	[n̲etti]	'forehead'
		[paTTam]	'kite'		[weTTi]	'cut down'
	c.	[puli]	'leopard'	d.	[weeli]	'fence'
		[puLi]	'tamarind'		[weeLi]	'marriage'
	e.	[muna]	'nib'	f.	[inam]	'type'
		[t̲uNa]	'help'		[iNa]	'pair'
	g.	[kara]	'shore'	h.	[kiiri]	'mongoose'
		[kaRa]	'stain'		[kiiRi]	'tore'

This paper examines the facts of the distribution and alternation of the alveolar and retroflex consonants in ME. While the contrast in (1) reflects the contrast in (2), the ME patterns of distribution and alternation intertwined with the contrast are not found in Malayalam. For instance, the contrast in (1a, b) is crucially tied up with morphology: in ME, nasals and laterals are never retroflex morpheme-initially, and they are never alveolar after front vowels morpheme-internally. The latter restriction is absent in Malayalam. Alveolars and retroflexes in ME also exhibit patterns of alternation not found in Malayalam, as illustrated by the pair *take* [Teek] and *intake* [inteek]. Such patterns are particularly fascinating because they

---

[1] Malayalam is a Dravidian language spoken in Kerala, in Southern India.

[2] While the persistence of substrate contrasts in a second language system may not be uncommon in syntax and semantics, they are not widely reported in phonology. Devyani Sharma (pc. April 2003), however, has pointed out the example of German English, which exhibits the persistence of the contrast, absent in English, between voiceless velar plosives and fricatives, their distribution being based on universal contextual preferences. This parallels the ME contrast between alveolars and retroflexes.

cannot be explained in terms of a theory that views patterns of a second language system as (i) the transfer of a substrate pattern; (ii) the retention of a superstrate pattern; or (iii) the direct reflection of a universal pattern.

1.2 The theoretical questions

The research program associated with SPE phonology (Chomsky and Halle 1968) sought to address two important questions:

A. What are the structural properties of human languages?
B. How do we develop a theoretical framework and model to formally express these structural properties so as to rule out logically possible but unattested properties?

Most subsequent research in phonological theory, all the way through Lexical Phonology (Pesetsky 1979; Mohanan 1982, 1986; Kiparsky 1982, 1985; Pulleyblank 1986), was pre-occupied with these questions. Chapter 9 of SPE introduced two other important questions:

C. What are the structural properties found repeatedly in human languages?
D. How do we develop a theory that correctly predicts the asymmetries between commonly and infrequently found structural properties?

The research programs associated with Natural Phonology (Stampe 1972) and theories of underspecification (Kiparsky 1982; Archangeli 1984, 1988; Steriade 1987, 1995) sought to address these questions under the rubric of markedness or naturalness. With the advent of O(ptimality) T(heory) (Prince and Smolensky 1993), two more equally important research questions emerged in the field:

E. What are the structural properties that distinguish one language from another?
F. How do we develop a theory that yields these typological variations in a way that would rule out logically possible but unattested variations?

OT is a theory of constraint interaction, which expresses the resolution of conflicting requirements from different constraints by specifying their relative ranks. In response to (F), OT claims that all structural properties that distinguish one language from another can be deduced from the combination of universal constraints, their language particular rankings and the idiosyncratic properties of morphemes. Alternative ranking assignments yield different grammars.

It goes without saying that redundancies in the constraint system (where constraints with approximately the same generalization may vary in their specifics) need to be eliminated, and the degrees of freedom in ranking assignments (so that UG correctly rules out unlikely/impossible grammars) must be restricted. One way to achieve this is to develop a theory in which (i) constraints with family resemblances are grouped together in terms of an underspecified common core and fully-specified variable manifestations, and (ii) pairs of faithfulness and markedness constraints are specified for their universal default (i.e., modifiable) ranking. Such a theory of constraints would complement the theory of

constraint ranking in OT.[3] Based on an analysis of the facts of the segmental phonology of ME, this paper sketches some aspects of such a theory.

1.3 Offspring systems in language contact

Our exploration of ME is also guided by an important theoretical question: how do novel patterns emerge in second language systems? This question may be relevant for theories not only of second language systems but also of contact languages in general (including pidgins and creoles), and can therefore be generalized as: how do novel patterns emerge in an 'offspring' language born through contact between substrate and superstrate languages? In this context, questions A-F in the previous section can be reformulated as follows:

I	What are the structural properties:			
	a.		of an individual contact language?	[grammar]
	b.	(i)	shared by a contact language and its substratum and/or superstratum?	[evolution]
		(ii)	unique to a contact language in relation to its substratum and superstratum?	
	c.	(i)	shared across contact languages?	[universals]
		(ii)	unique to a contact language in relation to other contact languages?	[typology]
II	How do we develop a theory that explains the structural properties that:			[linguistic theory]
		(i)	are shared across contact languages;	
		(ii)	distinguish a contact language from another; and	
		(iii)	distinguish a contact language from its substratum and superstratum?	

The bioprogram hypothesis (Bickerton 1981) seeks to respond to questions (Ici) and (IIi), and claims that these properties are instances of the emergence of default universals.[4] The substratist hypothesis seeks to respond to questions (Icii) and (IIii), and claims that (many of) these properties, if not part of the superstratum, are instances of the persistence of the

---

[3] We are grateful to Arto Anttila and Paul Kiparsky for helping us think along these lines. The intuition of the relationship between constraints that call for a mechanism of constraint generation and the need to minimize redundancy is implicit in most OT analyses. Our proposals in this paper are an explicit concretization of this intuition.

[4] See Kay and Sankoff 1974, Haiman 1985, and Mühlhäuser 1986 for other versions of the universalist hypothesis.

properties of the substratum. Various scholars (Mufwene 1986, 1991; Rickford and McWhorter 1997; Bao 1998; Kandiah 1998; DeGraff 1999; Singler 2000; Siegel 2002), recognizing that both the substratum and universals have a role in shaping the offspring system, have pointed out that these hypotheses complement rather than contradict each other. There also seems to be a general agreement among scholars that while the marked structures in contact languages are inherited from their source languages, the structures that diverge from the sources are unmarked universals (Singler 1988, Thomason and Kaufman 1988, Bresnan 2001, among others).

Our concern in this paper centers around the unique structural properties that distinguish the second language, ME, from both its substratum and the superstratum ((Ibii) and (IIiii)). We claim that at least some of these properties are instances of the emergence of the marked rather than the unmarked, and that they result from the tension between the substratum and the superstratum, resolved within the space provided by UG.

Early work on second language systems, a sub-type of contact systems, was associated with applied linguistics, in particular, the tradition of error analysis and contrastive analysis. As pointed out, for instance, in Köhler (1984), patterns in second language varieties that were different from those in the standard native varieties were depicted as errors or deviations (Bansal 1969; Tiffen 1974; Platt and Weber 1980; Wells 1982; among others). Second language research subsequently developed as part of sociolinguistics, in the study of non-native or "new varieties" (NV) of English (Kachru 1983, 1992; Williams 1987; Sridhar and Sridhar 1992; Kandiah 1998; among others). This tradition sought to free NVs (which are either stable second language systems or first language systems that historically grew out of second language systems) from the colonial dependence on old or standard varieties (OV), and viewed NV systems as socially equal to the OVs. The descriptions of NVs nevertheless often tend to document the ways in which NVs differ from OVs, continuing the methodology of the earlier tradition.

In contrast to the above approaches, the case study of ME in this paper is located alongside the study of languages in contact (e.g., Sebba1997; Broselow et al. 1998; Bao 1998), and second language studies that explore system internal patterns within the theoretical linguistic tradition (Warrier 1976; Vijayakrishnan 1978; Appa Rao 1978; Broselow 1992; Singh 1995; Lombardi 2000; Das 2001; Sharma 2001, in prep.).[5] Within this tradition, the OT research program allows us to frame questions about the relationship between offspring languages (whether pidgins and creoles or second language systems) and

[5] To take an example, Singh (1995) points out that Punjabi English has an asymmetry between the possible pronunciations of [s]+stop clusters and stop+sonorant clusters: while *school* may be pronounced as either [sə̀ku:l] or [isku:l], *glass* is pronounced only as [gila:s], not as *[igla:s]. He provides an explanation for the asymmetry in terms of a theory of syllable structure that appeals to sonority hierarchy. Now, while the applied linguistics and NV traditions would notice the vowel insertion in Punjabi English to break up complex onsets, the question of why Punjabi English does not have *[igla:s] as an option, and the search for an answer in terms of a general theory, are both characteristic of the theoretical tradition.

their parent languages (substratum and superstratum) from a theoretical typological perspective.

The emergent alveolar-retroflex pattern in ME exhibits idiosyncratic differences in the distribution of each of the coronal segments; these differences, while they reflect some of the patterns in Malayalam, are not simple transfers from Malayalam. Nor can they be construed as universals in their actual manifestation.[6] Our analysis shows the need to recognize the role of the parent languages and UG as influencing but not uniquely determining the properties of the offspring, thus allowing for the emergence of novel language particular patterns.

1.4 Outline of the paper

The paper takes the following path. After providing a general sense of the phonology of ME in section 2, we lay out the patterns involving alveolar and retroflex consonants in ME that are of special interest to us, and provide an analysis of the patterns within the broad outlines of OT in section 3. Section 4 outlines the consequences of the analysis for OT, and offers proposals for developing a theory of constraints. Drawing on the proposals in section 4, section 5 reformulates the analysis given in section 3. This is followed in section 6 by a discussion of the alveolar-retroflex patterns in ME from a diachronic perspective, offering an explanation for the patterns in section 3 in terms of the tension between the substrate and superstrate systems in the evolution of ME. We conclude by summing up the central claims and consequences in section 7.

2 Malayalee English and its Phonology: Lay of the land

Our object of study in this paper is the nature and evolution of the phonological system of present day ME, a transplanted second language variety of English spoken by speakers of Malayalam. As in the case of other linguistic systems, ME exhibits a great deal of regional and social variation across grammars. As is to be expected, there also exist variations correlating with differences in education (level, medium of instruction, and so on), and the degree of exposure to native English speech (through television and cinema). The data discussed in this paper are culled from the common features in the speech of the authors' large extended families, fairly representative of the variety of ME spoken by fluent speakers from central Kerala, educated in Malayalam medium schools. The patterns we describe are found across varieties of ME, and are typical or representative of ME to the extent that if we were to hear the clustering of these features on the radio, or when walking down the street, we could correctly identify the variety as ME, and the speaker as a Malayalam speaker.

---

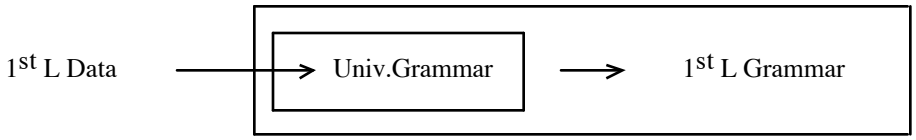
[6] The appearance of such novel patterns in creolization inspired the universalist hypothesis (Kay and Sankoff 1974, Bickerton 1981, Mühlhäuser 1986, among others). Such patterns are acknowledged in second language systems as morphogenesis (Mohanan 1992), fulguration (Kandiah 1998), and emergence (Broselow et al 1998).

2.1 Malayalee English as a second language system

Before we proceed to discuss the structure of ME, we spell out some of our basic assumptions, and outline a perspective within which we can raise relevant questions in the domain of second language systems.

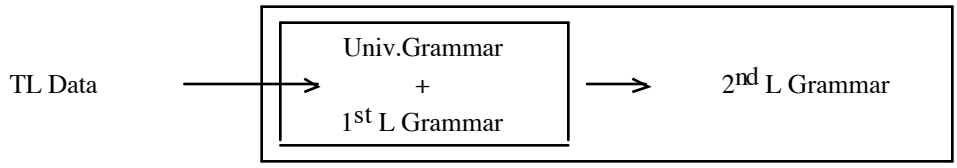
We use the term *second language* from an acquisitional perspective to refer to a system that an individual acquires after having acquired a first language. The differences between the acquisition of the first language and that of the second language lie in the nature of the initial state. The initial state for first language acquisition is the system of the language faculty that has not yet been exposed to linguistic data:

Figure 1: First Language Acquisition



For second language acquisition, the initial state exposed to data from the target language (TL) is a system that includes the language faculty *and* the grammar of first language:

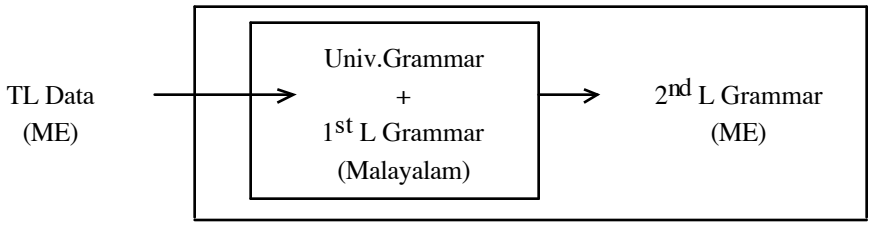
Figure 2: Second Language Acquisition



We assume that the two grammars of a bilingual speaker constitute interacting overlapping subsystems of a single *bilingual system*. A theory of code switching (whether involving two first languages or a first and a second language), for example, is concerned with the nature of the interaction between the two linguistic subsystems in a bilingual system. Second language formation within this view is a process involving the growth of a monolingual system into a bilingual one.

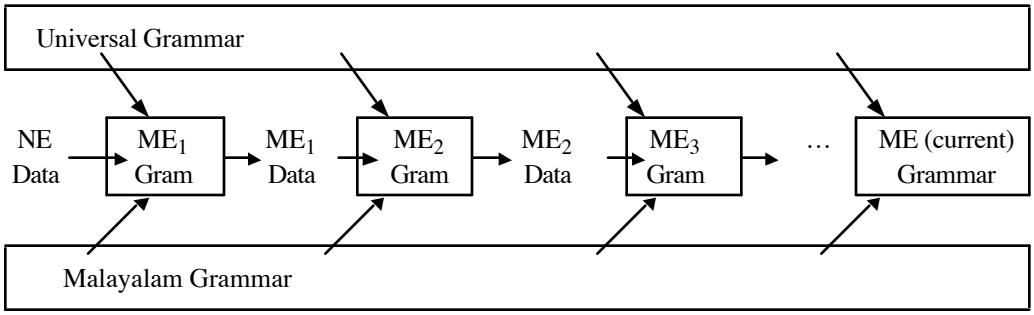
As mentioned earlier, ME is a transplanted system, not an interlanguage system. An *interlanguage* (Selinker 1969, 1971; Trudgill and Hannah 1982; Davis, Criper & Howatt 1984) is an unstable system in which second language learners, exposed to data from a native variety, continually approximate the native target. In contrast, a *transplanted* system (Kachru 1983) is a self-replicating stable system, in which learners speak the variety they are exposed to, namely, the second language variety. In the case of ME, young learners successfully learn the target variety, namely, that spoken by the adult community:

Figure 3: The acquisition of ME, a transplanted system



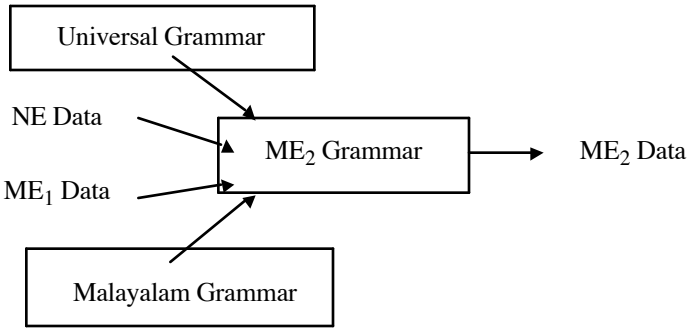
An interlanguage system can be viewed as gaining the status of a transplanted system when it feeds itself and stabilizes, more or less the way a pidgin becomes a creole. The questions of the evolution of ME, then, are not questions of language acquisition but of language change: how did contact between the grammar of Malayalam and data from native varieties of English (NE) (specifically varieties spoken in Britain), in the course of successive generations of second language learners, give rise to the current system of ME?

Figure 4: The evolution of ME (language change)



It must be noted that this picture of ME was true of the situation when the authors were learning English, but has changed in the recent past, because of widespread availability, through cable television, of data from native varieties of English:

Figure 5: The acquisition of ME: the current situation



While figure 5 is the appropriate model for the ME spoken by the younger generation today, figure 4 remains appropriate for our purposes. The variety of ME we describe is that used by speakers typically above the age of 30, not exposed to the media in their formative years.

While ME largely remains a second language system, a number of non-native varieties of English are now evolving into first language systems. Take Singapore English, for instance, which for a substantial number of speakers is a first language. These speakers are either monolingual, speaking no language other than English, or symmetrically bilingual, speaking English and, say, Mandarin.[7] The model in figures 4 and 5 can be extended to understand the historical evolution of such nascent first language systems as well.

2.2 Patterns of contrast in ME

The vowel system The surface contrasts among vowels in ME are listed in (3a). Given in (3b) are examples illustrating the contrasts listed in (3a):

(**3**) a. Pure vowels

	front	central	back
high	i ii		u uu
mid	e ee	∂∂	o oo
low	ææ		a aa ɔɔ

Diphthongs

ai au

b. Vowels with examples:

i	*bit*	[b**i**t]	*beer*	[b**i**yar]	**ii**	*beat*	[b**ii**t]	
e	*about*	[**e**bauT]	*balloon*	[b**e**luuN]	**ee**	*bait*	[b**ee**t]	
a	*idea*	[aiDiy**a**]	*mutter*	[m**a**TT**a**R]	**aa**	*balm*	[baam]	*art* [aaRT]
u	*book*	[buk]			**uu**	*boot*	[buuT]	
o	*occur*	[okkaR]	*collide*	[koLaiD]	**oo**	*go*	[goo] *coat*	[kooT]
					∂∂	*girl*	[g∂∂L]	
ai	*find*	[faind]			**ææ**	*bat*	[b**ææ**t] [8]	
au	*found*	[faund]			**ɔɔ**	*cot*	[kɔɔT] *caught*	[kɔɔT]

The vowel inventory of ME is clearly different from the inventories of any of the native varieties of English. Yet, the phonological system of ME originally arose from exposure to

---

[7] By 'symmetric bilingualism,' we mean the co-presence of two linguistic systems acquired simultaneously, as opposed to 'asymmetric bilingualism,' where one system is acquired after another.

[8] The reason for treating [ææ] and [ɔɔ] as long vowels in ME is that in phonological patterns that appeal to the distinction between simple and complex nuclei, these nuclei pattern like complex nuclei. One such pattern involves a geminate/non-geminate alternation among consonants that is dependent on whether the preceding nucleus is simple or complex.

(the British varieties of) native English (NE). How did exposure to NE data result in the birth of the ME vowel inventory?

As may be expected, the ME inventory appears to be a modified version of the inventory of Malayalam, the substratum. Compare the pure vowels in the two inventories:[9]

(**4**) a. Malayalam

i ii		u uu
e ee	ə əə	o oo
	a aa	

b. Malayalee English

i ii		u uu
e ee	əə	o oo
ææ	a aa	ɔɔ

In (4b), we find elements that are not present in standard NE varieties. For instance, corresponding to the NE diphthongs [ei] and [ou] are the long [ee] and [oo] in ME (*bait* [beet], *boat* [booT]). We may view this as an instance of the adaptation into ME of the Malayalam nuclei in order to accommodate the required target contrasts.[10] Alternatively, this may be an instance of an overlap with the targets [ee] and [oo] in, say, Irish English, which is likely to have served as the input to the historical origins of ME. In other words, the contrasts available in the first language are sufficient to meet the demands of the target data. These patterns are consistent with the substratist hypothesis.

Now consider the long vowels [ææ] and [ɔɔ] in ME, which appear to be reanalyzed versions of the NE short vowels [æ] and [ɔ] (see footnote 8). Note that Malayalam does not have a low front vowel or a low back rounded vowel, whether short or long. Hence, an explanation in terms of the adaptation of a substrate contrast is not available for these ME vowels. A possible explanation for the reanalysis is that they are phonetically long in NE, and are hence treated as phonologically long in ME.

The consonantsystem We now turn to the distinguishing properties of the consonant system of ME. To begin, consider the surface contrasts among the consonants:

(**5**) Consonants

	lab	dent	alv	retr	pal.alv	pal	vel	glott.
stop	p b	t̲ d̲	t	T D	c j		k g	
fric	f		s	S				h
lateral			l	L				
nasal	m		n	N			ŋ	
tap			r	R				
glide	w					y		

[9] As we will see, the phonetically central vowels [əə], [a], and [aa] in ME pattern phonologically vowels, as reflected in the table in (4b).

[10] Such adaptation involves the phenomenon of *transfer* or first language *interference* (Weinreich 1953; Lado 1964) generally found in interlanguage systems.

Examples illustrating each of the segments in the inventory in (5) are given in (6):

(**6**) Consonants with examples:

p	*pen*	[pen]			**b**	*bet*	[bet]	
t̲	*thick*	[t̲ik]			**d̲**	*this*	[d̲is]	
t	*sit*	[sit]	*meeting*	[miittiŋ]				
T	*put*	[puT]	*retake*	[riiTTeek]	**D**	*good*	[guD]	
c	*check*	[cek]			**j**	*judge*	[jaDj]	
k	*mystic*	[mistik]	*sticky*	[stikki]	**g**	*got*	[gɔɔT]	
s	*sell*	[sel]	*rice*	[rais]	**f**	*fist*	[fist]	
	zoo	[suu]	*rise*	[rais]				
S	*shell*	[Sel]	*measure*	[meSaR]	**h**	*hen*	[hen]	
m	*must*	[mast]			**n**	*name*	[neem]	
ŋ	*singing*	[siŋŋiŋ]			**N**	*gun*	[gaN]	
l	*lick*	[lik]	*slowly*	[slooli]	**r**	*carom*	[kæræram]	
L	*pull*	[puL]	*holy*	[hooLi]	**R**[11]	*theorem*	[t̲iyaRam]	
			flit	[fLit]				
w[12]	*wine*	[wain]	*vine*	[wain]	**y**	*wire*	[wayaR]	*yes* [yes]

Categoriesofsecondlanguagecontrasts Most traditional descriptions of the phonology of second language varieties highlight the contrasts in the target language that are absent in the second language (e.g., the absence of [z] in ME: *rice*/*rise* [rais]). Left unnoticed are the reverse cases, namely, contrasts in the second language that are absent in the target language. What we find fascinating in ME are instances of the latter kind, as exemplified by the near minimal pairs in (1) (as well as examples like *sourness* [sawaRnes] vs. *harness* [haaRNes] ([n]/[N]), and *slowly* [slooli] vs. *holy* [hooLi] ([l]/[L]). No native variety of English has a contrast between alveolar and retroflex segments. This contrast, however, is one of the most distinguishing characteristics of ME.

To place the ME contrasts in the perspective of language contact, consider the inventories of the superstratum (NE) and the substratum (Malayalam) in (7). The NE segments are in bold face, and the Malayalam ones in regular font underneath them.

---

[11] The segment [R] is phonetically post-alveolar, not retroflex: its articulation does not involve the curling of the tongue blade. However, like retroflexes, [R] is accompanied by tongue body retraction, in contrast to the tongue body fronting of [r] (Warrier 1976). It also forms a natural class with the retroflexes in phonological patterning, as we will see below. Hence we group [R] with retroflexes.

[12] The segment we represent as /w/, though a bilabial glide in some contexts, is realized as a labio-dental frictionless continuant in many other contexts. Phonologically, however, the segment behaves like a glide, particularly in patterns involving syllable structure.

(**7**) *English and Malayalam consonants*

	lab.	dent.	alv.	retr.	pal-alv.	pal.	vel.	glot.
plosive	**p b** p b	 t̪ d̪	**t d** t	 T D	**c j** c j	 k' g'	**k g** k g	
fric	**f v**	**θ ð**	**s z** s s'	 S	**ʃ ʒ**			**h** h
nasal	**m** m	 n̪	**n** n	 N	 ñ	 n'	**ŋ** ŋ	
lateral approx			**l** l	 L				
central approx	**w** w		**r** r	 R		**y** y		

First consider the fairly straightforward correspondences between NE and Malayalam: [p, b, c, j, k, g, s, h, m, ŋ, w, y] exist in both languages, and are part of the ME system. In other words, the three systems exhibit an overlap with respect to these segments.

NE [θ, ð, ʃ] are reanalyzed as Malayalam [t̪, d̪, S], illustrating the adaptation of the superstrate contrasts to the nearest available substrate structures.[13]

ME, like Malayalam, has no voiced fricatives, despite the presence of [v, θ, z, ʒ] in NE. The NE [z] and [ʒ] are reanalyzed as ME [s] and [S] respectively, thereby neutralizing the voicing contrast. The NE /v/ and /w/ neutralize in ME as a glide (/w/), one of whose allophones in Malayalam is the frictionless continuant [ʋ]. And although the voicing contrast between NE [θ] and [ð] is maintained in ME, these segments, as mentioned earlier, are adapted as dental stops.

The segment [f], which does not exist in Malayalam, is introduced in ME, illustrating how the target data occasionally forces the introduction of a structure (configuration of phonological features as a segment) to preserve a superstrate contrast. [14]

---

[13] Observe that French speakers of English reanalyze the non-distributed alveolar [t] and [d] of NE as the dental/distributed [t̪] and [d̪] of French, and the dental fricatives [θ] and [ð] of NE as alveolar [s] and [z]. Thus, while ME preserves the dentality of [θ] and [ð] and abandons the fricativity, French English preserves the fricativity and abandons the dentality. One way to understand this difference is to assume that in Malayalam, the dental~alveolar contrast is 'stronger' (i.e., has higher functional load in the segment inventory) than the fricative~plosive contrast, while in French, the reverse holds.

[14] One might ask why ME introduces the new voiceless fricative /f/ into the system in order to maintain a target contrast, but despite the target contrast of voicing among fricatives, does not introduce voiced fricatives. A plausible answer to this question may lie in the strength of contrast in the substrate system (see footnote 13). Voicing contrast among fricatives is not part of the Malayalam system, and hence the NE contrast is neutralized in ME. On the other hand, Malayalam does have place contrasts among fricatives; hence an additional fricative does not disturb the overall pattern congruity of the segment inventory.

The cases discussed above illustrate (i) overlap between the substrate and superstrate systems, (ii) adaptation of superstrate contrasts, (iii) loss of superstrate contrasts, and (iv) retention of superstrate contrasts absent in the substratum. We characterize these categories as follows:

I. A unit x is present in the superstratum and the substratum.
   It is present in the offspring. *Overlap*

II. A unit x in the superstratum is absent in the substratum.
   (a) x is adapted in the offspring as x' from the substratum. *Adaptation*
   (b) x is lost in the offspring. *Loss*
   (c) x is introduced in the offspring to retain contrast. *Retention*

The category in I requires no discussion; studies of contact systems have focused on the three categories in II. It must be noted that a particular phenomenon may involve more than one category. For instance, the neutralization of NE /v/ and /w/ to the frictionless continuant [□] in ME involves the loss of the NE contrast, as well as the adaptation of both segments to [□]. In contrast, the neutralization of [s] and [z] involves only loss, not adaptation, and the appearance of [t̲, d̲] for the dental fricatives [□,□] involves only adaptation, without loss of contrast.

We are now left with the alveolar and retroflex consonants. While the alveolar-retroflex contrast clearly originates in the substratum, what is interesting is that they are unmotivated from the point of view of approximating the superstratum. They may therefore be seen as the persistence of a substrate contrast in the offspring system. The alveolar-retroflex contrast in ME thus illustrates a largely overlooked fifth category:

III. A unit x' in the substratum is absent in the superstratum.
   x' persists in the offspring. *Persistence*

The persistence of substrate regularities such as phonotactic restrictions and allophonic processes, which limit the options available to a phonological system, are well recognized in second language studies. But the persistence of a substrate contrast, which expands the available options in a system, is somewhat unusual. How do we explain such apparently unmotivated persistence? Before we search for clues for an answer, we will look briefly at some patterns of distribution and alternation in ME, to see if the patterns yield further surprises that do not fit the categories above.

2.3 Patterns of distribution and alternation

Having looked at the patterns of contrast in ME in the light of the substratum and superstratum, let us look at three phonological patterns of distribution and alternation that give ME a distinctive flavor. The first two patterns — post-nasal voicing and the fronting of velar stops — are modified versions of corresponding patterns in Malayalam. Generalizing our five categories of contrast to cover patterns of distribution and alternation as well, we

may view post-nasal voicing and velar fronting as instances of the persistence of a substrate pattern in the offspring. However, the third pattern — intervocalic gemination of voiceless stops — is found neither in the substratum nor the superstratum. Therefore, it cannot be construed as an instance of the persistence of a substrate pattern. Nor could it be a matter of adaptation, loss, or retention. We will suggest that this innovation is the result of what we call emergence, arising out of the tension between the substratum and the superstratum.

Post-nasal voicing In most representative varieties of ME, the voiceless stop has a gap in its distribution: it does not occur morpheme-internally after a nasal (*pump* [pamb] *[pamp]; *limping* [limbiŋ] *[limpiŋ]). Across morphemes, a voiceless stop gets voiced after a nasal (Warrier 1976). The alternation commonly applies within words (8), but not across words (9):[15]

(**8**)	a.	*possible*	[pɔɔsibL]	*impossible*	[imbɔɔsibL]
	b.	*power*	[pawaR]	*empower*	[embawaR]
	c.	*distract*	[DistRææ kt]	*contract*	[kɔɔNDRææ kt]
	d.	*discrete*	[DiskRiit]	*concrete*	[kɔɔŋgReet]
	e.	*kind*	[kaind]	*mankind*	[mææŋgaind]

(**9**)	a.	*become possible*	[bikampɔɔsibL]	* [bikambɔɔsibL]
	b.	*one trick*	[waNTRik]	* [waNDRik]

The distributional gap as well as the pattern of alternation can be expressed as: Post-nasal obstruent stops are voiced in ME. This pattern is identical to post-nasal voicing in Malayalam (Warrier 1976, Mohanan & Mohanan 1984): [t̲aamara] 'lotus'/[ce<u>nd</u>aamara] 'red lotus'; [kuTTi] child]/[aa□guTTi] 'boy'; [pacca] 'green'/[iLambacca] 'light green'. This then is a simple instance of the persistence of a substrate pattern in the offspring.

The pattern, however, exhibits an interesting twist. In Malayalam, it holds within as well as across word, whereas in ME, it is restricted to a word internal environment. The regularity, as well as the difference, is expressed by the pre-theoretic generalization in (10):

(**10**)	*Post-nasal obstruent stops are voiced.*	Malayalam:	in a phonological phrase
		ME:	in a phonological word

Velar fronting ME exhibits an alternation between regular and fronted velar stops, as illustrated in (11):

(**11**)	a.	*mystic*	[mistik]	*mystical*	[mistik'k'al]
	b.	*critic*	[krittik]	*critical*	[krittik'k'al]
	c.	*chick*	[cik]	*chicken*	[cik'k'an]

[15] We have not indicated post-nasal voicing in our transcription in this paper except in this sub-section. For instance, while the word *intake* in ME is [indeek], we have transcribed it as [inteek], so as to avoid confusion between the [d] in the [T-t-d] pattern and the [D-d] pattern.

We express this pattern of alternation as: Intervocalic voiceless velar stops are front after high front vowels. This generalization also explains a distributional pattern: unfronted voiceless velar stops do not occur intervocalically after [i]: *America* [ameerik'k'a] (*[ameerikka]); *Africa* [aafrik'k'a] (*[aafrikka]).

Velar fronting is another instance of the persistence of a substrate pattern. Intervocalic velar stops become fronted in Malayalam after front vowels (Mohanan & Mohanan 1984), yielding alternations like [awal**kk*ð***] 'to her' ([awal] 'she') and [kuTTi**k'k'*ð***] 'to the child' ([kuTTi] 'child').[16] Once again, the pattern in ME shows an interesting difference. In Malayalam, the fronting applies to both voiced and voiceless stops (k and g), after all front vowels ([reek'ha] 'line', [meeg'ham] 'cloud'). In ME, it is restricted to the voiceless stop ([k]) after a high front vowel ([i]).[17] We state the generalization in Malayalam as (12a), and that in ME as (12b):

(**12**) a. Intervocalic velar stops are front after front vowels. (Malayalam)
b. Intervocalic voiceless velar stops are front after high front vowels. (ME)

Intervocalic gemination of voiceless stops Intervocalic voiceless stops in ME are always geminate word internally (*packet* [pæækket], *report* [RippooRT]), unlike voiced stops (*figure* [figaR], *baboon* [bææbuun]). The following examples illustrate the single-geminate alternation: [18]

(**13**) a.	*bakes*	[beeks]	*bake eggs*	[beek egs]	*baker*	[beekkar]
b.	*gallop*	[gæælap]	*gallop again*	[gæælap egeen]	*galloping*	[gæælappiŋ]
c.	*leaped*	[liipD]	*leap outward*	[liip auTwððD]	*leaping*	[liippiŋ]
d.	*eats*	[iits]	*eat it up*	[iit it ap]	*eating*	[iittiŋ]

The pattern can be described as: Intervocalic voiceless stops are geminate word internally. This pattern, while rooted in the structure of Malayalam, is not found in Malayalam in this form. Thus, in Malayalam, (i) all consonants exhibit a single vs. geminate contrast (voiced

---

[16] Mohanan and Mohanan (1984) refer to this phenomenon in Malayalam as palatalization, using the term 'palatal' to refer to segments with palatal articulation without involving the blade of the tongue, thus excluding the palato-alveolar [tʃ] in English.

[17] Rubach (1984) proposes that first language rules are not transferred to the second language unless they are automatic and noncyclic. A similar position is taken in Singh (1991). But palatalization in Malayalam has lexical exceptions (Mohanan & Mohanan 1984) and hence is not automatic. The facts of palatalization in ME therefore present a problem for this claim.

It may turn out that the hypothesis holds in non-transplanted second language systems where the input is the native variety, and that non-automatic rules are transferred only as part of the language change in transplanted second language systems. We do not have sufficient data to investigate the issue here.

[18] We have found it useful to ask ME speakers to transliterate English words in Malayalam orthography. They systematically use double consonant characters for intervocalic voiceless stops (as in *groping* or *reaper*). This is in contrast to non-intervocalic voiceless stops (as in *pen* and *reaps*), and intervocalic voiced stops (as in *robing* and *liberal*). As far as we know, speakers of other varieties of Indian English do not exhibit this behavior when transliterating English words in their respective regional scripts.

and voiceless plosives, as well as fricatives, nasals, laterals, and glides); (ii) single and geminate consonants exhibit patterns of alternation in several morphosyntactically governed environments both within and across words; and (iii) intervocalic voiceless stops are systematically voiced if single, preserving the voicelessness if geminate.[19] However, Malayalam does not have a purely phonological alternation between single and geminate intervocalic voiceless stops. How did the ME pattern arise?[20]

The emergence of the alternation in (13) between single and geminate voiceless stops may be viewed as resulting from the different interpretations of the intervocalic and non-intervocalic voiceless stops. At the phonetic level in Malayalam, as noted above, intervocalic stops are geminate if voiceless; if single, they are voiced. In ME, voicing intervocalic stops to meet this substrate restriction would result in neutralizing the superstrate voicing contrast. Since NE requires maintaining the contrast, we can explain the intervocalic gemination in ME as a reanalysis of the voiceless intervocalic single stops in English as geminates, regardless of their phonetic duration. This strategy allows ME to preserve the intervocalic contrast between voiced and voiceless stops in the superstratum, while obeying the substrate prohibition against voiceless intervocalic single stops.[21]

Viewed this way, the pattern of intervocalic gemination in ME in (13) emerges from the interaction between the properties of the substratum and the demands of the superstratum, and is not a form of mere persistence. One way to explain the emergence of this pattern is from an OT perspective, in terms of constraint interaction. Following a suggestion from Arto Anttila (pc. March 2002), we can account for the ME pattern in terms of the three universal constraints in (14a). The different rankings of the three constraints in (14b) yield the grammars of Malayalam, English, and ME:

---

[19] For instance, /waaṯam/ [waaḏam] 'paralysis'; /koTi/ [koDi] 'flag'; /makan/ [magan] 'son'. These intervocalic stops are also lenited. For details, see Mohanan and Mohanan (1984); Mohanan (1986).

[20] The emergence of such patterns in second language systems is not unique to ME. For instance, coronal palatalization and non-local labial assimilation (Bao 2003) and [sp] metathesis (Mohanan 1993) in Singapore Colloquial English are not found either in Standard English or in any of the substratal sources.

[21] One might ask why gemination is not adopted to preserve voicing contrast among post-nasal stops (section 2.3.1) in ME. The answer lies in the constraint in Malayalam against post-nasal geminates, which persists in ME. Since the sequence N+CC constitutes an ill-formed representation, an NE form such as [siŋkiŋ] cannot be interpreted as /sinkking/. The only alternative is to analyze it as ME /siŋkiŋ/, and allow the constraint in (10) to yield ME [siŋgiŋ].

(**14**) a. (i) Single voiceless intervocalic stops are prohibited.
(ii) Moras not present in the input are prohibited in the output.
(effect: gemination is prohibited.)
(iii) [+voice] not present in the input are prohibited in the output.
(effect: voicing is prohibited.)

b.

ranking	outcome	language
(i) >> (iii) >> (ii)	voiceless geminate stop	ME
(iii) >> (ii) >> (i)	voiceless single stop	English
(i) >> (ii) >> (iii)	voiced single stop	Malayalam

This analysis derives the facts of ME gemination from universal constraints by specifying distinct rankings for Malayalam, English, and ME. [22]

2.4 Processes in the formation of contact languages

In section 2.2, we identified five categories in the description of the contrasts in an offspring language in terms of those available in the substratum and the superstratum, namely, overlap, adaptation, loss, retention, and persistence. Having looked at patterns of alternation in ME, we must ensure that they are covered by these categories, and generalize the characterization of the categories if necessary. Furthermore, the phenomenon of gemination in ME discussed above suggests the need to acknowledge a sixth category, namely, emergence.

We state the categories below in a generalized form to apply to patterns of not only contrast but also distribution and alternation:

I. A pattern x is present in the superstratum and the substratum.
It is present in the offspring. *Overlap*

II. A pattern x in the superstratum is absent in the substratum.
(a) x is adapted in the offspring as x' available in the substratum. *Adaptation*
(b) x is lost in the offspring. *Loss*
(c) x is introduced in the offspring to retain contrast. *Retention*

III. A pattern x' in the substratum is absent in the superstratum.
x' persists in the offspring. *Persistence*

IV. A pattern x'', present in neither the superstratum nor the substratum
emerges as a novel pattern in the offspring. *Emergence*

The categories described above are summarized in the table below, where x, x', and x'' are patterns (of contrast, distribution, or alternation) in the offspring. x represents a pattern in the superstratum that may or may not be present in the substratum, x' is a pattern in the

[22] Note that to get the result of voicing instead of gemination after nasals in ME (section 2.3.1), the prohibition against post-nasal geminates must be ranked higher than (14aii).

substratum absent in the superstratum, and x'' is absent in both the superstratum and the substratum:

Table: Processes in the formation of a second language system

	OVERLAP	ADAPTA-TION	LOSS	RETEN-TION	PERSIST-ENCE	EMERG-ENCE
SUPERSTRATUM	x	x	x	x	—	—
SUBSTRATUM	x	x'	—	—	x'	—
OFFSPRING	**x**	**x'**	**—**	**x**	**x'**	**x''**

The analysis of intervocalic gemination in the previous section shows that at least some instances of the emergence of novel patterns in contact languages, motivated by the tension between the substratum and the superstratum, can be explained in terms of a ranking of universal constraints distinct from that of the substratum and the superstratum. A similar strategy may yield the other categories of patterns — overlap, adaptation, loss, retention, and persistence — in terms of differences in constraint ranking, but we will not attempt to derive them in this paper.

The emergence of the pattern of gemination in (14) is a relatively simple one. In order to better understand the nature of emergence in offspring systems, and to see if language particular ranking of universal constraints can yield all instances of emergence, we must look at more complex cases. In section 3, we explore the alveolar-retroflex contrast in ME as a case study of such emergence of patterns in a contact situation.

3 Alveolar and Retroflex Consonants

3.1 Backness and retroflexion

Subminimal pairs like those in (15) illustrate the contrast between alveolar and retroflex consonants in ME:

(**15**) a. (i) *slowly* [sloo**l**i] (ii) *sourness* [sawa**R**n**n**es]
b. (i) *holy* [hoo**L**i] (ii) *harness* [haa**R**N**N**es]

Notice a significant factor in the contrast: the words in (15a) are bi-morphemic, while those in (15b) are mono-morphemic.[23]

[23] The appearance of the retroflex [N] in the word *governess* [gawa**R**N**N**es] in ME appears to contradict this statement. However, interpreted as 'caretaker for children of rich families', the word is not seen as semantically related to *govern* or *government* by speakers of ME. In the absence of semantic motivation for treating the word as bi-morphemic, the phonological contrast between *sourness* and *governess* suggests that *governess* is best treated as mono-morphemic.

Morpheme internally, alveolars and retroflexes are in complementary distribution in ME. A crucial conditioning factor affecting their morpheme-internal distribution is the vowel that precedes them: alveolars occur after front vowels, and retroflexes after back vowels, as illustrated in (16)-(18):

(**16**)	a.	*many*	[me**n**i]	*ban*	[bææ**n**]	*coin*	[kɔɔyi**n**]	*dine*	[Dai**n**]
	b.	*money*	[ma**N**i]	*prawn*	[pRɔɔ**N**]	*loan*	[loo**N**]	*frown*	[fRau**N**]
(**17**)	a.	*valley*	[wææ**l**i]	*pallor*	[pææ**l**aR]	*coil*	[kɔɔyi**l**]	*bell*	[be**l**]
	b.	*volley*	[wɔɔ**L**i]	*colour*	[ka**L**aR]	*pole*	[poo**L**]	cool	[kuu**L**]
(**18**)	a.	*atom*	[ææ**tt**am]	*bet*	[be**t**]	*write*	[Rai**t**]	*sit*	[si**t**]
	b.	*autumn*	[ɔɔ**TT**am]	cut	[ka**T**]	*wrote*	[Roo**T**]	*put*	[pu**T**]

This is an extremely robust pattern: mono-morphemic nonce forms like *[bani], *[sali], *[gɔɔli], *[mææLi], *[miT], and *[geNar] are regarded by ME speakers as strange or unacceptable.

When morpheme-initial and post-consonantal, the different coronals — nasal, lateral, plosives, and tap — exhibit different patterns of distribution. We proceed directly to look at these patterns in some detail.

3.2 Variations in the pattern

Coronal nasal Morpheme initial coronal nasals in ME are alveolar (e.g., *nose* [noos]) without exception, even when preceded by a back vowel in the same word (e.g., *slowness* [sloones], in contrast to *bonus* [booNas]). When preceded by a consonant other than /R/, the coronal nasal is again alveolar:

(**19**)	a.	*kidney*	[kiD**n**i]	*hackneyed*	[hææk**n**iiD]	*Kitchner*	[kic**n**aR]		
	b.	*corner*	[kɔɔR**N**aR]	*furnace*	[faR**N**as]				

In short, retroflex nasals occur after a back vowel or /R/ within the same morpheme. We can account for these facts in OT by assuming the constraints in (20a-c), and their ranking specified for ME:

(**20**) <u>Coronal nasal:</u>

a. A coronal nasal is alveolar.
b. Coronals agree in [back] with the preceding V within a morpheme.
c. A coronal nasal is retroflex after [R] within a morpheme.

RANKING: ME: (20b), (20c) >> (20a)

The domain specification of (20b, c) as morpheme internal is needed to explain the surface contrast between bi-morphemic *slowness* [sloones] and *sourness* [sawaRnes] on the one hand, and the mono-morphemic *bonus* [booNas] and *harness* [haaRNes] on the other. [24]

Lateral As with nasals, a lateral in ME is invariantly alveolar morpheme-initially: *low* [loo], *shoelace* [Suulees], *clueless* [kLuules]. But unlike the nasal, the lateral is retroflex after a wider range of consonants, including all retroflexes, and non-coronal obstruents:

(**21**) a. *apply* [appLai] *blue* [bLuu] *fly* [fLai]
aglow [egLo] *clear* [kLiyaR]
b. *cutlary* [kaTLaRi] *butler* [baTLaR]
c. *rattler* [ræætlaR] *athletic* [at̲lettik] *hamlet* [hææmlet]

In other words, the lateral behaves exactly like the nasal post-vocalically and morpheme-initially, but after consonants, its differs in behavior. These facts follow from the constraints and their ranking in (22):

(**22**) Coronal lateral:
a. A lateral is alveolar.
b. Coronals agree in [back] with the preceding V within a morpheme. (=(20b))
c. A lateral is retroflex after obstruents within a morpheme.
d. A lateral agrees in [back] with the preceding coronal within a morpheme.

RANKING: ME: (22d) >> (22c), (22b) >> (22a)

Pairs like *early* [ððRLi]/*dearly* [diyaRli], and *couplet* [kapLet]/*hopeless* [hooples], illustrate that the constraints in (22b-d), like (20b-c), are restricted to the morpheme internal domain.

Coronal voiced stop In the 'typical' or representative varieties of ME, the voiced alveolar stop [d] occurs only when preceded by [n]; in all other environments, the voiced coronal plosive is [D]:[25]

(**23**) a.	*fend*	[fend]	*grind*	[graind]	*bandage*	[bæændeej]
b.	*fond*	[fɔɔND]	*ground*	[grauND]	*wonder*	[waNDaR]
	cold	[kooLD]	*Rushdie*	[RaSDi]	*board*	[booRD]
	weld	[welD]	*seldom*	[selDam]	*abduct*	[abDakt]
	divide	[DiwaiD]	*feed*	[fiiD]	*ride*	[raiD]

This generalization holds across morphemes as well:

---

[24] There exists a handful of proper nouns such as *London* [laNDan], and *Joan* [joon] (cf: *John* [jɔɔN]), where contrary to (20), the final nasal is alveolar, and *British* [briTTiS], where the plosive after a front vowel is retroflex. Such examples must be stipulated as exceptions.

[25] We must note that while intervocalically, the voiced coronal plosive is invariantly retroflex (*ready* [reDi]), word-initially, many educated ME speakers are increasingly using the alveolar (*debt* [det]).

(**24**)	[d]		[D]	
a.	*debt*	[Det]	*indebted*	[indetteD]
b.	*direct*	[Dayerekt]	*indirect*	[indayerekt]
	nondirect	[nɔɔNDayerekt]		
c.	*define*	[Difain]	*indefinite*	[indefinit]
	undefined	[aNDifaind]		
d.	*groaned*	[grooND]	*cleaned*	[kLiind]
	dreamed	[DRiimD]		
	missed	[misD]		
	picked	[pikD]		
	saved	[seewD]		
e.	*kingdom*	[kiŋDam]	*christendom*	[kRistyendam]
	freedom	[fRiiDam]		
f.	*outdoor*	[auTDooR]	*indoor*	[indooR]
	sundance	[saNDaans]	*fan dance*	[fændæens]

In the case of coronal nasals and laterals, we stated the general constraint relevant for them in terms of alveolar ((20a), (22a)). Given (23)-(24), the general statement for the [d]/[D] pair must be that it is retroflex:

(**25**) Voiced non-dental coronal plosive:

a. A voiced coronal plosive is retroflex.

b. A voiced coronal plosive is alveolar after [n] within a word.

RANKING: ME: (25b) >> (25a)

Though (25b) holds across morphemes, including the stems of a compound ((24f)), it does not hold across words. Thus, the retroflex stop is preserved in phrases like *when days pass* [wen Dees paas], *in direct conflict* [in Dayarekt kɔɔNfLikt], and *children dance* [cilRen Dæens], requiring the domain specification of (25b) as word -internal.[26]

Coronal voiceless stop If we look at the examples in (18), repeated below for easy reference, we find that morpheme-internally after vowels, [t] and [T] exhibit the same distribution as the coronal nasal and lateral, [T] appearing after back vowels, and [t] after front vowels:

(**18**)	a.	*atom*	[æættam]	*bet*	[bet]	*write*	[Rait]	*sit*	[sit]
	b.	*autumn*	[ɔɔTTam]	cut	[kaT]	*wrote*	[RooT]	*put*	[puT]

[26] It is quite possible that (25b) applies only when /n/ and /d/ are in the same syllable. If so, it can apply postlexically, given syllabifications like the word-internal *indirect* /i.nda.ya.rakt/, and *in direct* /in. Da.ya.rakt/ across words, which follow the Malayalam syllabification pattern (K P Mohanan 1982, 1986; T Mohanan 1989). For lack of evidence on syllabification in ME, we will not pursue this possibility here.

After consonants, [T] occurs only after retroflexes ((26a) vs. (26b, c)):

(**26**) a. *cult* [kaLT] *heart* [haaRT] c. *built* [bilt] *dreamt* [DRemt]
b. *soft* [sɔɔft] *actor* [æktar] (cf. *build* [bilD] *dreamed* [DRiimD])

As (27) shows, the morpheme-initial voiceless coronal plosive (like its voiced counterpart) is systematically retroflex word-initially (column A), and after front vowels (column B). In other words, it is alveolar after a front vowel, as in (18a), only if the vowel is in the same morpheme. After non-retroflex consonants, it is alveolar across morphemes as well (column C), though not across words ((27e)):

(**27**)

	A	B	C	
a.	*take* [Teek]	*retake* [riiTeek]	*mistake* [misteek]	*intake* [indeek]
b.	*trust* [TRast]		*distrust* [DistRast]	
c.	*train* [TReen]	*detrain* [DiiTReen]	*well-trained* [weltRaind]	
d.	*tint* [Tind]		*rose-tinted* [roostindeD]	
e.	*tips* [Tips]	*free tips* [frii Tips]	*ten tips* [Ten Tips]	

As with the nasal, lateral, and voiced plosive, we can derive the facts in (26)-(27) from the following constraints and their ranking:

(**28**) Voiceless non-dental coronal plosive:
a. A voiceless coronal plosive is retroflex.
b. Coronals agree in [back] with the preceding V within a morpheme. (=(20b))
c. A voiceless coronal plosive is alveolar post-consonantally within a word.
d. A voiceless coronal plosive agrees in [back] with the preceding coronal within a word.

RANKING: ME: (28d) >> (28c), (28b) >> (28a)

The characterization of the voiceless plosive in the constraint in (28a) as retroflex groups it with its voiced counterpart (d/D). Now, given the extremely restricted distribution of the alveolar [d], it is entirely reasonable to assume (25a). However, one might wonder why the constraint in (28a) cannot be formulated to group the voiceless plosive with the nasal and lateral instead, as alveolar. Such a move would replace the constraints in (28) with those in (29):

(**29**) Voiceless non-dental coronal plosive:
a. A voiceless coronal plosive is alveolar.
b. Coronals agree in [back] with the preceding V within morphemes.
c. A voiceless coronal plosive is retroflex morpheme-initially, except when preceded by alveolar within the same word.
d. A voiceless coronal plosive agrees in [back] with the preceding coronal within words.

RANKING: ME: (29d) >> (29c), (29b) >> (29a)

Notice that (29a), (29b), and (29d) are equivalent to (28a), (28b), and (28d) respectively. However, (29c) requires a statement of exception, unlike its counterpart in (28c). We therefore choose the more parsimonious statement in (28) to express the patterns of distribution and alternation between [t] and [ɾ].

Tap Of the pair of taps [r] and [ɾ], the alveolar [r] appears in an extremely restricted environment: it cannot occur anywhere except (i) intervocalically after a front vowel ((30a)), and (ii) after /g/ ((31a)):

(**30**)	a.	*very*	[weri]	*marry*	[mæəri]	*ariel*	[eeriyal]
	b.	*worry*	[waɾi]	*sorry*	[sɔɔɾi]	*orient*	[ooɾiyant]
		color	[kaLaɾ]	*park*	[paaɾk]	*lord*	[lɔɔɾD]
(**31**)	a.	*grow*	[groo]	*degree*	[digri]	*grand*	[græənd]
	b.	*bray*	[bɾee]	*pray*	[pɾee]	*tree*	[tɾii]
		fry	[fɾai]	*three*	[t̠ɾii]	*shrink*	[sɾiŋk]
(**32**)		*red*	[ɾeD]	*rabbit*	[ɾæəbit]	*rob*	[ɾɔɔb]

In short, the coronal tap is alveolar after front vowels, and after [g].

The tap exhibits a difference in behavior, however, between the post-vocalic and post-consonantal environments. After [g], the tap is exceptionlessly alveolar, but after front vowels, the retroflex [ɾ] does sometimes occur, as illustrated in (33):

(**33**)	a.	*erosion*	[i'ɾooSaN]		b.	*variety*	[we'ɾaitti]
		erupt	[i'ɾapt]			*various*	['weeriyas]
		therapy	[t̠e'ɾæəppi]			*vary*	['weeri]

In (33a), the intervocalic tap is retroflex despite the preceding front vowel. Notice, however, that in every instance, the tap is at the beginning of a stressed syllable. In other words, [r] occurs after front vowels only if it is not foot-initial. This pattern is confirmed by the example of alternation in (33b).

Even with the above caveat, the generalization has counterexamples. In forms like *terror* [teɾaɾ] and *irritate* [iɾitteet], the tap is retroflex despite being in an unstressed syllable after a front vowel. For an explanation for these apparent anomalies, we must look to orthography. Double *rr* in spelling, even when preceded by a front vowel, appears as [ɾ] in ME if the vowel is short. Consider the contrast between (34a) and (34b):

(**34**)	a.	*carriage*	['kæəreej]		b.	*merry*	['meɾi]
		marry	['mæəri]			*terror*	['teɾaɾ]
		serious	['siirias]			*irrigate*	['iɾigeet]

Taking into account all the factors involved in the distribution and alternation patterns of [r] and [ɾ], we formulate the constraints that yield the facts as (35):

(**35**) Coronal tap:

a. A tap is retroflex.
b. Coronals agree in [back] with the preceding V within morphemes. (=(20b))
c. A tap is alveolar after [g] within a morpheme.
d. A tap is retroflex if it is foot-initial.
e. A tap is retroflex when double in spelling after a short V.

RANKING: ME: (35e), (35d) >> (35c), (35b) >> (35a)

3.3 Pulling the pieces together

In sum, we find that alveolar and retroflex segment pairs in ME are systematically in complementary distribution. Ignoring some of the idiosyncrasies in the individual statements, the broad generalizations that emerge from the statements in (20), (22), (25), (28), and (35) are as follows. The voiced coronal plosive is special: it is retroflex in all environments except after [n]. For the remaining coronals, there are three relevant environments. Morpheme initially, the nasal and lateral are alveolar, while the plosive and tap are retroflex. Post-vocalically, the coronals are alveolar after front vowels and retroflex after back vowels. Post-consonantally, all but the tap are alveolar after alveolar, and all but the nasal are retroflex after retroflex. After other consonants, the preferred manifestation in this position is alveolar for nasal and plosive but retroflex for lateral and tap.

In the light of these generalizations, the statements in (20), (22), (25), (28), and (35) can be unified as a single set of ranked constraints as given in (36):

(**36**) Non-distributed coronals:[27]

a. (i) Nasals and laterals are alveolar.
   (ii) Voiceless plosives and taps are retroflex.
   (iii) Voiced plosives are retroflex.

b. Coronals agree in [back] with the preceding V within a morpheme.

c. Taps are retroflex when (i) foot-initial; or (ii) double in spelling after a short V.

d. (i) Laterals are retroflex after obstruents within a morpheme.
   (ii) Voiceless plosives are alveolar post-consonantally within a word.

e. (i) Laterals agree in [back] with the preceding coronal within a morpheme.
   (ii) Voiceless plosives agree in [back] with the preceding coronal within a word.

f. (i) Nasals are retroflex after [R] within morphemes.
   (ii) Voiced plosives are alveolar after [n] within morphemes.
   (iii) Taps are alveolar after [g] within morphemes.

RANKING: ME: (36f), (36e) >> (36d), (36c), (36aiii) >> (36b) >> (36ai), (36aii)

---

[27] The specification 'non-distributed' is meant to exclude dental plosives, fricatives, and nasals.

4 Theoretical Consequences

4.1 Expanding the analysis to Malayalam and English

Optimality Theory is committed to the position that all structural differences in linguistic patterns across languages are describable in terms of differences in the ranking of universal constraints. Given this position, we must assume that the constraints in (36) are universal. If so, they need to be supplemented by additional constraints to yield the right results for other languages. Take (36ai-ii), for instance. Given that coronals are not retroflex in English, we need a constraint that non-distributed coronals are alveolar, and rank it higher than (36a) for English. Given that in Malayalam, alveolars and retroflexes are not in complementary distribution, we need a faithfulness constraint that requires the input specifications of [back] in coronals to be preserved in the output, and rank this constraint higher than the other constraints. This involves the addition of (iv) and (v) to (36a), and their incorporation into the ranking specifications:

(**36**) a. (iv) Non-distributed coronals are alveolar.
(v) Input specifications of backness in coronals are retained in the output.

RANKING:	ME:	(36aiii) >> (36b) >> (36ai), (36aii) >> (36aiv), (36av)
	English:	(36aiv) >> (36ai), (36aii), (36aiii), (36av), (36b)
	Malayalam:	(36av) >> (36ai), (36aii), (36aiii), (36aiv), (36b)

4.2 Domain specification

Notice that the only difference between (36ei) and (36eii) is in the specification of the domain (morpheme/word) of the constraints. Suppose we think of the two constraints as variant manifestations of the same CORE, forming a constraint HOMOLOGUE, analogous to homologous organs in biology (or archetypes in mythology).

As is well-known, homologous organs like limbs, digits, and eyes exhibit an incredibly complex range of variation both across species (legs of humans, cows, spiders and birds), and across organs within a species (arms vs. legs, toes vs. fingers, thumb vs. little finger). Yet, they have a deep fundamental unity of form and function. Viewed this way, we can think of voicing assimilation, consonant lenition, vowel epenthesis, glide insertion, and so on, as constraint homologues that have the same core, and variable manifestations within and across languages. Given that the homology of biological organs is rooted in genetic structure, modifications of which give rise to evolution, the concept of homologous constraints or constraint homologues provides for a unified conception of synchronic as well as diachronic invariance and variability in language structure.

Viewing groups of constraints as constraint homologues requires factoring apart the shared properties of the member constraints from their differences. In the case of (36ei) and (36eii), the only difference lies in the specification of the domain:

(**36**') e. Laterals and voiceless plosives agree in [back] with the preceding coronal.
DOMAIN: (i) Lateral: morpheme-internal.
(ii) Plosive: word-internal

The issue of domain specification also appears in the voicing of post-nasal plosives. The generalization in (10) would be formulated in OT as in (37):

(**37**) a. Post-nasal obstruent stops are voiced within a phonological phrase.
b. Post-nasal obstruent stops are voiced within a phonological word.
c. Specification of [-voice] in the input must be preserved in the output.

RANKING: ME: (37b) >> (37c) >> (37a)
English: (37c) >> (37a, b)
Malayalam (37a, b) >> (37c)

The constraint core and its different manifestations of (37a, b) can be stated as the constraint homologue in (38a):

(**38**) a. Post-nasal obstruent stops are voiced. DOMAIN
(i): phonological phrase
(ii): phonological word
b. Specification of [-voice] in the input must be preserved in the output.

RANKING: ME: (38ai) >> (38b) >> (38aii)
English: (38b) >> (38ai, ii)
Malayalam (38ai, ii) >> (38b)

Notice that ranking the faithfulness constraint *above* the markedness homologue has the effect of suppressing or de-activating the pattern of the voicing of post-nasal plosives, as in English. Ranking it *below* the markedness homologue has the effect of fully activating the pattern, as in Malayalam. Ranking it *between* two variant constraints of the markedness homologue provides for partial (de)activation of the pattern, as in ME.

Given the role of faithfulness in activating and deactivating the markedness patterns, it would be useful to think of faithfulness and markedness as forming a CONSTRAINT PAIR. It would also be useful to specify a universal default ranking for each constraint pair, such that a grammar that deviates from the default needs an additional stipulation, while one in which the ranking conforms to UG needs no stipulation. For instance, if the universal default ranking of (38a, b) is (38b)>>(38a), then ME and Malayalam would need additional stipulations to activate (38a), while English doesn't. If, on the other hand, the universal ranking is (38a)>>(38b), then English and ME would need to deactivate (38a). We will pursue the advantages of this suggestion in subsequent sections.

4.3 Loci and triggers

The need to eliminate redundancy across constraints by deriving a group of constraints with family resemblances from a single constraint core comes up again when we look at our treatment of the fronting of intervocalic velar stops in (12). The constraints in (12a) and (12b) are identical, except for the additional specification of the locus of the constraint in ME as voiceless, and its trigger as high:

(**12**) a. Intervocalic velar stops are front after front vowels. (Malayalam)
b. Intervocalic voiceless velar stops are front after high front vowels. (ME)

An OT analysis of these generalizations would be along the following lines:

(**39**) a. Intervocalic velar stops are front after front vowels.
b. Intervocalic voiceless velar stops are front after high front vowels.
c. Specification of fronting in velar stops in the input is preserved in the output.

RANKING:	ME:	(39b) >> (39c), (39a)
	English:	(39c) >> (39a), (39b)
	Malayalam	(39a) >> (39b), (39c)

Once again, we can factor out from the core the additional specifications for the locus and trigger, as in (40):

(**40**) a. Intervocalic velar stops are front after front vowels.

LOCUS		TRIGGER	
(i)	—	(iii)	—
(ii)	voiceless	(iv)	high

b. Specification of backness in velar stops in the input is preserved in the output.

RANKING:	ME:	(40aii, iv) >> (40b) >> (40ai, iii)
	English:	(40b) >> (40a)
	Malayalam	(40a) >> (40b)

Ranking faithfulness above markedness deactivates the assimilation, as in English, and the reverse ranking fully activates it, as in Malayalam. And interleaving faithfulness with markedness again yields partial activation, as in ME.

4. 4 Universal default ranking

It is reasonable to assume that given a universal default ranking of a faithfulness-markeness pair of constraints, conformity with the default would be more common across languages, and deviation from it less common. If so, the universal default ranking of a pair as Mx>>Fx predicts that the markedness pattern would be found in a majority of languages, and Fx>>Mx predicts that the markedness pattern would be found only in a minority of languages. We take it that the fronting of velar stops after front vowels is cross-linguistically uncommon (unlike the fronting of velar stops before front vowels). The universal default ranking (40b)>>(40a) would reflect this asymmetry: English, then, would

need no additional stipulation, while Malayalam and ME would require a stipulation that reverses the default ranking to activate velar fronting. In the case of post-nasal voicing in the previous section, we left open the choice of the universal default ranking. The choice between the two possible rankings would depend on whether or not post-nasal voicing is found in a majority of languages.

Now, the need for balancing the tension between minimizing articulatory effort and maximizing discriminability (Lindblom 1986; Bosch et al. 1987; Mohanan 1993; Boersma 1998; Flemming 2002; Padgett 2003) in the phonetic module may carry the burden of explaining many of the cross-linguistic asymmetries of probability in segment contrasts (for instance, the ubiquitous contrast between nasal and non-nasal stops, the less common contrast between nasal and non-nasal vowels, the relative rarity of the contrast between voiced and voiceless nasals, and the absence of a contrast between voiceless and voiced nasal vowels). If so, phonology need say nothing about such asymmetries.

However, this source of explanation does not extend to asymmetries in patterns of distribution and alternation across segments. Without universal default ranking of faithfulness-markedness pairs, the theory would yield a factorial typology in which the fronting and not fronting of velars after front vowels are equally probable. It would also incorrectly predict that a grammar that requires coronals to be retroflex (prohibiting alveolars) after back vowels is as likely as one without the requirement: we are not aware of any grammar other than that of ME in which alveolars are prohibited after back vowels. Universal default rankings allow us to avoid such incorrect predictions of cross-linguistic asymmetries of probability in factorial typology.

The basic idea suggested in the above discussion is that each constraint pair that carries the default ranking expresses a *universal tendency*. A strong tendency is one in which markedness outranks faithfulness, and a weak tendency, one in which faithfulness outranks markedness. Language particular reversing of the default ranking deactivates a strong tendency and activates a weak tendency. To support the claim that a given constraint pair is a universal tendency, we must show that it is repeatedly found in languages; to show that it is a strong tendency, we must show that it is found in a majority of languages.[28]

4.5 Markedness

The theoretical concept of markedness involves an asymmetry between the marked and the unmarked, and makes a set of empirical claims about certain cross-linguistically observed asymmetries in segment inventories, cross-segmental patterns of distribution and alternation, patterns of acquisition and language change, and so on. For instance, we find languages

---

[28] One may think of each of these tendencies as an attracting or repelling magnetic field, as it were, placed in the trajectory of a moving pendulum. In the course of its historical evolution, the pendulum of each grammar would be a non-linear dynamical system: caught by multiple fields of attraction, its trajectory would form a chaotic attractor, continuously evolving without settling down to any fixed point (K. P. Mohanan 1993).

with voiced and voiceless fricatives (English), and languages with only voiceless fricatives (Malayalam), but not languages with only voiced fricatives. Likewise, there are languages with voiced and voiceless nasals (Burmese), and languages with only voiced nasals (English), but not languages with only voiceless nasals. Theories of markedness aim to predict these asymmetries observed across languages, and do so by postulating the categories of 'marked' and 'unmarked', and the assumption that the choice of the marked entails the presence of the unmarked. Given this assumption, [–voice] as unmarked in a [–sonorant, +continuant] segment, and [+voice] as unmarked in a [+nasal] segment, correctly predict the cross-linguistic implicational asymmetries in segment inventory.

Asymmetries of feature values are also found in patterns of alternation. For instance, there are languages in which coronals assimilate in place to labials and velars, but not labials and velars to coronals (English), languages in which coronals assimilate to labials and velars, labials assimilate to velars but not coronals, and velars do not assimilate (Korean), and languages in which coronals, labials and velars freely assimilate to one another (Malayalam). But there are no languages in which labials and velars assimilate to coronals, but coronals do not assimilate. These asymmetries can be captured by a markedness scale of place in which coronals are most unmarked, and velars are most marked. Similar asymmetries are found in other cross-segmental patterns such as harmony, neutralization, epenthesis, and deletion.

An implicational statement of the form, "A language that has X also has Y, but a language that has Y need not have X," derives a cross-linguistic asymmetry of probability, namely, that Y is more common than X. Such asymmetries of probability also fall within the scope of markedness theory.

Theories that seek to capture the substance of markedness, including SPE and underspecification theories, and their OT descendents (Kiparsky 1982; Archangeli 1984, 1988; Steriade 1987, 1995, De Lacy 2002, among others) involve claims such as the following: the output of neutralization and epenthesis is the unmarked member; only the marked member can resist spreading constraints (assimilation, harmony); the unmarked member is likely to appear earlier than the marked one in language development; the unmarked member is likely to be the overtly unexpressed one in a paradigm; and so on. At the heart of markedness, then, is the convergence of these various cross-linguistic asymmetries (Mohanan 1993).

This concept of markedness is distinct from that used in the context of the *emergence of the unmarked* in OT. To quote McCarthy and Prince (1994:1) "In Optimality Theory, forms are marked with respect to some constraint C if they violate it — indeed, as Smolensky (1993) emphasizes, they are *literally* marked in that they incur violation-marks for constraint C as part of their grammatical derivation." Under this meaning, 'unmarked' in a cross-linguistic context refers to all the optimal forms that follow from language-particular ranking specifications. This yields a factorial typology of the possibilities and impossibilities of grammatical structures within the universal space for grammars. However, it does not attempt to predict the relative cross-linguistic probabilities of possible structures.

Thus, it predicts that there exist languages that prohibit voiceless nasals or voiced fricatives, and languages that allow them. But they fail to predict that the former types of languages are more common, or that languages that require coronals to be retroflex after back vowels are extremely rare.

Depending on whether or not the theory has provisions for (a) specifying universal ranking, which individual grammars may (or not) be permitted to modify, and (b) allowing constraints to be unranked with respect to each other, we have the following scenarios of cross-linguistic predictions for constraints C1 and C2.

(**41**)

Provision in the theory for:		Prediction	
universal ranking	partial ordering in grammars	probability of languages with ranking	
x	x	equal:	{C1>>C2 }, {C2>>C1}
x	√	equal:	{C1>>C2 }, {C2>>C1}, {C1, C2}
√ {C1>>C2}	x	all:	{C1>>C2}
√ {C1>>C2}: default	x	most: some:	{C1>>C2} {C2>>C1}
√ {C1>>C2}: default	√	many: some:	{C1>>C2} {C2>>C1}, {C1, C2}

As these scenarios show, specific assumptions about universal default ranking are crucial for deriving predictions on asymmetries of cross-linguistic probabilities.

5 Fleshing out the analysis

5.1 A revised account

Drawing on the proposals argued for in the previous sections, we now reformulate the account of ME in (36) as in (42) below, along with the corresponding specifications for English and Malayalam. In the table, each item (42a, b, c, d, e, f) forms a constraint homologue. Each homologue is composed of an underspecified core, and specifications of domain, locus, trigger and outcome value which yield the fully specified constraint. We assume that each markedness homologue is paired with a faithfulness constraint, and may be ranked relative to it by a universal default ranking (UDR). The specifications ACT(ivated) and DEACT(ivated) indicate signal that the UDR is reversed: when faithfulness outranks markedness (F>>M), the reversal activates an inactive markedness constraint; when markedness outranks faithfulness (M>>F), the reversal deactivates an active markedness constraint.

Notice that (36ei, ii) and (36fi, ii) all have to do with the assimilation of backness in adjacent coronals. Factoring away the domain specification from the constraint allows us to capture this essential aspect of (36ei, ii) in a single constraint, (42e). Factoring out the locus-

trigger information from the constraint allows us to incorporate into (42e) the substance of (36fi, ii) as well:

(**42**) Non-distributed coronals:

	UDR	DOMAIN	LOCUS	TRIGGER	OUTCOME	
a *Coronals are [□ back].* (M>>F)						
ME			nasal/later.		□: –	**i**
			vl. plos./tap		□: +	**i i**
			vd. plos.		□: +	**i i i**
Eng.					□: –	
Mal.	DEACT					
b *Coronals agree in [back] with the preceding V.* (F >>M)						
ME	ACT	morph.-int.				
Eng.						
Mal.						
c *Taps are retroflex (i) foot-initially; or (ii) when double in spelling after a short V.* (F >>M)						
ME	ACT					
Eng.						
Mal.						
d *Coronals are [□ back] after consonants.* (F >>M)						
ME	ACT	morph.-int.	lateral	obstruent	□: +	**i**
		word-int.	vl. plos.		□: –	**i i**
Eng.						
Mal.	ACT	morph.-int.	lateral	obstruent	□: +	**i**
e *Coronals agree in [back] with a preceding coronal.* (F >>M)						
ME	ACT	*word-int.*	nasal	[R]		**i**
		morph.-int.	lateral			**i i**
		word-int.	vl. plos.			**i i i**
		word-int.	vd. plos.	[n]		**i v**
Eng.						
Mal.	ACT	word-int.	nasal	[R]		**i**
		word-int.	lateral			**i i**
		word-int.	vl. plos.			**i i i**
		morph.-int.	vd. plos.	[n]		**i v**
f *Taps are alveolar after [g].* (F >>M)						
ME	ACT	syl.-int.				
Eng.						
Mal.	ACT	syl.-int.				

RANKING: ME: (42f), (42e) >> (42d), (42c), (42aiii) >> (42b) >> (42ai), (42aii)

5.2 Compositionality of universal constraints

The idea of universal constraint cores being underspecified for locus, trigger, domain, and outcome value leads to the possibility of decomposing the constraint cores themselves into their components. To illustrate, the constraints in (43) can be unpacked as shown below:

(**43**) a. Adjacent obstruents within a syllable agree in voice.

DOMAIN	LOCUS	TRIGGER	OUTCOME
syllable	obstruent	obstruent [voice]	[voice]

b. Sonorants are voiced.

DOMAIN	LOCUS	TRIGGER	OUTCOME
segment	sonorant		[+ voice]

Likewise, the universal constraint cores in (42a-f) may be unpacked, and given as in (44) along with their variable manifestations:

(**44**) Coronals in Malayalam, English, and ME:

		UDR	DOMAIN	LOCUS	TRIGGER	OUTCOME
a	**CORE**	**M>>F**		**non-distr. cor**		**[back]**
	ME: **i**			nasal		= –
				lateral		= –
	i i			tap		= +
				vcless plosive		= +
	i i i			vcd plosive		= +
	Eng.					= –
	Mal.	DEACT				
b	**CORE**	**F>>M**		**non-distr. cor**	**$V_{[\ back]}$ —**	**[back]**
	ME	ACT	morph-int			
	Eng.					
	Mal.					
c	**CORE**	**F>>M**		**tap: foot-initial or (C)V— (spelt rr)**		**[+ back]**
	ME	ACT				
	Eng.					
	Mal.					
d	**CORE**	**F>>M**		**non-distr. cor**	**C —**	**[back]**
	ME: **i**	ACT	morph-int	lateral	obstruent	= +
	i i		word-int	voiceless plosive		= –
	Eng.					
	Mal.: **i**	ACT	morph.-int.	lateral	ostruent	= +

e	CORE	F>>M		non-distr. cor	cor$_{[\alpha\,back]}$ —	[α back]
	ME: **i**		word-int	nasal	[R]	
	i i	ACT	morph-int	lateral		
	i i i		word-int	vcless plosive		
	i v		word-int	vcd plosive	[n]	
	Eng.					
	Mal.: **i**		word-int	nasal	[R]	
	i i	ACT	word-int	lateral		
	i i i		word-int	vcless plosive		
	i v		morph-int	vcd plosive	[n]	
f	**CORE**	**F>>M**		**tap**	**[g] —**	**[–bk]**
	ME	ACT	syl-int			
	Eng.					
	Mal.	ACT	syl.-int.			

(44a-f) are the constraint cores of the alveolar-retroflex patterns in ME, English, and Malayalam, together with the fragments that derive the fully specified constraints from the constraint cores, and the language particular stipulations of ranking.

The picture in (44) may be viewed as a partial theory of the cross-linguistic manifestation of coronals in terms of a core and variable manifestations. It is a partial theory for two reasons. First, its scope is restricted to non-distributed coronals: to be complete, it should include distributed coronals as well. Second, its scope is restricted to three grammars, namely, those of English, Malayalam, and ME: to be complete, the typology it covers should include other languages as well. We believe, however, that it could be a promising starting point for a theoretical typology of coronals in human languages that addresses not only questions of language structure, but also language acquisition, language change, and language contact.

5.3 Language particular constraints?

OT does not employ language particular constraints to derive cross-linguistic variation. Given this strong position, we must assume that the constraints in all the constraint homologues in (44) are universal. Now, recall the methodological requirement that, in order to postulate a constraint pairing, the pair must express a universal tendency, whether strong or weak, repeatedly found in languages. (44a) is not particularly interesting, since all that it does is provide the choice between alveolars (non-distributed coronals with fronted tongue body) and retroflexes (non-distributed coronals with retracted tongue body). (44e) is straightforward: it requires adjacent coronal consonants to agree in backness, which we can take as a universal tendency. The suggestion that (44b) is a universal tendency, however, may be problematic. (44b) requires vowels preceding the retroflexes to be back, as is attested in Kodagu (Bhat 1973, cited in Flemming 2003). But it also requires vowels

preceding alveolars to be non-back, a pattern we have not seen attested in any language or variety other than ME.

The assumption that (44c, d, f) represent universal tendencies is even more problematic. As far as we know, (44c) is not surface true in any phonological system other than ME, and (44d, f) are not surface true in any system other than Malayalam and ME. We could, of course, postulate (44c, d, f) as universals and not have them activated in any grammar other than (Malayalam and) ME. But this would be explanatorily equivalent to saying that (44c) is a language particular constraint in ME, and (44d, f) are language particular constraints in Malayalam and ME. Occam's razor requires that we do not postulate a constraint or entity in UG unless it significantly reduces the stipulations in grammars, that is, unless it serves an explanatory purpose. Assuming that (44c, d, f) are universals, activated only in (Malayalam and) ME, serves no explanatory purpose. This suggests that they be treated as language particular constraints, and that we relax the OT prohibition against such constraints.

6 Morphogenesis of the Alveolar-Retroflex Patterns

6.1 The diachronic question

What this paper has done so far is to provide a theoretical synchronic account of the patterns that distinguish ME from the substratum and the superstratum, and to articulate and defend the theoretical assumptions underlying the analysis. We have not, however, answered the diachronic question: how did the somewhat uncommon pattern of surface contrast of alveolar and retroflex consonants and their distribution and alternation emerge in the structure of ME? Granting the minor differences in domain specification in (42e), the language particular stipulations for ME in (42di, e, f) are clearly carried over from the substratum. The stipulations in (42ai) are identical to those of the superstratum. Those in (42aii, aiii, b, c), however, are unique to ME. How did they arise in the ME system?

At the heart of our exploration of the phonology of ME is the general question: "How does structure arise in linguistic systems?"[29] Given the perspective outlined in section 2.1, we may identify three forces that shape the structure of an offspring in a contact situation:

---

[29] It may be useful to bear in mind that the term 'structure' has two distinct meanings in the context of language, namely, (i) the structure of a particular linguistic expression (e.g., feature geometric structure of the segment /p/, or the syllable structure of the word *spray*); and (ii) the structural properties of a language or one of its aspects (e.g., syllable structure in Malayalam; morphological structure in Sanskrit). In our questions about the emergence of structure, we mean structure in the second sense. This sense of structure is the opposite of randomness, same as the concept of order in theories of complexity. If human languages had universal units of representation, but no laws governing their combination (i.e., if the atomic units of representation could freely combine in any random manner to form larger representations), languages would have structure in the first sense but not the second. Complexity in language structure is reflected in linguistic patterns, constraints being a formal device to express them.

UG, the substratum, and the superstratum. Let us state the influence of the parents as the following general principles:

(**45**) a. The offspring retains the structure of the substratum.
b. The output of the offspring system must match the output of the superstratum.

When the pull from the two forces results in a conflict, the offspring system has three options: (i) be loyal to the substratum (adaptation, loss, persistence); (ii) conform to the superstratum (retention); and (iii) find a stable compromise, drawing on UG (emergence). [30]

6.2 Levels of representation in language contact

The distinction between the underlying, lexical, and phonetic levels of representation in Lexical Phonology (Pesetsky 1979; Mohanan 1982, 1986; Kiparsky 1982, 1985; Pulleyblank 1983, among others) may allow us to better understand the role of principles (45a, b) in shaping the evolution of (42a-c). The theories of lexical phonology and morphology as developed in the early eighties contained two sets of leading ideas. One had to do with a *model* of phonology, with specific hypotheses about levels of phonological representation, their relationships, and the interface between phonology and morphology. The other set of hypotheses had to do with the *interaction between phonological laws* (rules/constraints).

Stratal OT (Kiparsky 1997; McCarthy and Prince 2001) combines the OT theory of constraint interaction with the stratal model of Lexical Phonology, where each stratum can specify its own ranking of universal constraints. Each stratum also has its own input-output pairs, the input to a given stratum being the optimal output candidate of the previous stratum. This architecture of OT naturally yields the underlying, lexical, and phonetic levels of representation in Lexical Phonology illustrated in the examples given below:

		racial	*baron*	*tenth*	*king*
(**46**)	underlying	reis iæl	bæron	ten □	king
	lexical	reiʃəl	bærən	ten□	kiŋ
	phonetic	reiʃl	bærn	ten□	kiŋ

In Stratal OT, the 'underlying representation' would be the input to the lexical (word formation) component, the 'lexical representation' being its output, and input to the phrase level organization.

---

[30] John Singler (pc. April 2003) has pointed out to us that although found in stable second language systems like ME and Singapore English, emergence of the marked is unlikely in pidgins and creoles. This intuition makes sense, as pidgins and creoles, unlike second languages, do not typically involve the acquisitional asymmetry of a linguistic system with a stable grammar incorporating a new grammar. If this is correct, we would expect emergence of the marked in creoles that evolve from second language systems. If Singapore English is a creole, or is in the process of becoming one, this expectation is borne out (Mohanan 1992, Bao 1998).

The distinction between underlying and lexical levels of representation leads to a distinction between the inventories of segments at each of these levels. For instance, there is evidence to assume that underlying representations in English distinguish labial and alveolar among nasals, but not velar. But lexical representations of English distinguish labial, alveolar, and velar nasals. Thus, the distinction between *kin* and *Kim* would be represented in terms of the coronal and labial nasal in both the underlying and lexical levels. The distinction between *kin* and *king*, on the other hand, would be represented in terms of the coronal and velar nasal only at the lexical level; at the underlying level, the words would be represented as /kin/ and /king/, without a velar nasal.

The set of segments needed to distinguish each morpheme from the rest of the morphemes in the language constitutes the underlying inventory. The set of segments needed to distinguish each word from the other words (lexical level) constitutes the lexical inventory. The inventories at the two levels in Malayalam and NE help us understand the patterns of persistence and emergence in ME.

6.3 Persistence of contrasts

Alveolar and retroflex consonants are distinct in the underlying as well as lexical inventories of Malayalam (Mohanan & Mohanan 1984):

(**47**) *Malayalam inventories of non-dental coronals*

underlying/input	:	–	T	–	D	n	N	l	L	r	R
lexical	:	t	T	–	D	n	N	l	L	r	R

Native varieties of English do not require the contrast at either level:

(**48**) *Native English inventories of non-dental coronals*

underlying/input	:	t	–	d	–	n	–	l	–	r	–
lexical	:	t	–	d	–	n	–	l	–	r	–

Given that the distribution of alveolar and retroflex segments in ME crucially refers to morphological information, the lexical inventory of ME must include both alveolars and retroflexes. However, since their distribution is predictable, the contrast is not required in the underlying inventory. The regularities in the three inventories can be summarized as below:

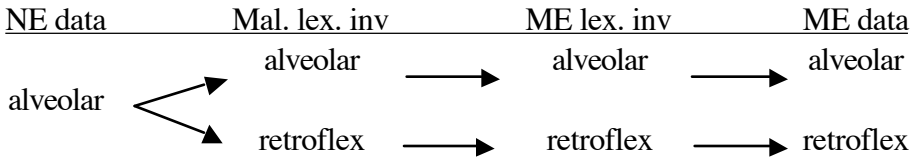
(**49**) *Alveolar and retroflex*

	Malayalam	ME	Native English
Underlying	contrastive	non-contrastive	non-contrastive
Lexical	contrastive	contrastive	non-contrastive

It now becomes evident that at the underlying level, ME expresses the same number of contrasts as NE. At the lexical level, the number and feature composition of ME segments (with the exception of the extra [d]) is identical to that of Malayalam. The question is: how come the ME system followed the superstratum in its underlying inventory, and the substratum in its lexical inventory?

Given the assumption that it is the lexical inventory that guides the conscious perception of segment distinctions in a language (Mohanan 1986), principle (45a) would guide speakers of Malayalam to analyze the NE phonetic segments in terms of the grid provided by the Malayalam lexical inventory:

(**50**) *Persistence of contrast: alveolars and retroflexes in ME*



Since morphemic contrasts are expressed at the underlying level, the absence of the contrast at the underlying level in ME follows from the absence of such a contrast in the superstratum.

A parallel situation is found in Malayalam syllabification, where the coming together of the Sanskrit and Dravidian systems of syllabification through language contact result in conflict (T. Mohanan 1989). The conflict is resolved by using Dravidian (substrate) syllabification at the lexical level, and Sanskrit (superstrate) syllabification at the underlying level. The proposed assumption to explain this phenomenon is that "the foreign input is assigned a surface representation that is compatible with the native system, and an underlying representation is borrowed from the foreign grammar." (T. Mohanan 1989:615) In ME, the foreign (NE) input is assigned a lexical representation that is compatible with the Malayalam system, while the underlying representation is "borrowed" from NE.

A natural question that arises at this point is: "The persistence of the alveolar-retroflex contrast in Malayalam results in the splitting of the NE alveolars into alveolars and retroflexes in ME. How come the contrast between the palatalized and non-palatalized alveolar fricatives (/s/ and /s'/) in Malayalam did not yield a similar split of the English /s/ into [s] and [s'] in ME? A plausible answer may lie in the concept of 'strength of contrast' (see notes 10 and 12): while the alveolar-retroflex contrast covers six pairs of segments (t/T, d/D, r/R, s/S, n/N, l/L), and is therefore strong, the palatalized-non-palatalized coronal contrast, involving just one pair (s/s'), is not strong enough to persist in the offspring.

6.4 Emergence of the marked

Given the above account of the persistence of contrast, let us now look at the unique patterns of the distribution of alveolars and retroflexes in ME. (42d-f) reveal the persistence in ME of the Malayalam specifications. What remains to be explained is the emergence of the specifications in ME in (42a-c).

The emergence of (42a) is largely grounded in the distributional patterns in Malayalam, where retroflex nasals and laterals are not found morpheme-initially, nor are alveolar plosives. This is reflected in (42ai, iii), and part of (42aii). We have no explanation

for why taps are retroflex morpheme-initially in ME (the other part of (42aii)), since Malayalam has both the alveolar and retroflex taps in this environment.

What explains why the coronals are alveolar after front vowels, and retroflex after back vowels in ME (42b)? The answer, we suggest, lies in the phonetic quality of alveolar and retroflex consonants. As noted by Steriade (2000) and Flemming (2003), retroflexes and alveolars have retracted and advanced tongue body articulations respectively, anticipated in the preceding back and front vowels. It would be reasonable to surmise therefore that this gradient coarticulation is phonologized as a non-gradient distributional restriction in ME. Though phonetically grounded, however, the prohibition of alveolar consonants after back vowels does not appear to be attested in any system other than ME. Hence we must view it as a highly marked, if not language particular, pattern.

The emergence of (42c), as far as we can see, is neither phonetically nor phonologically motivated. We seriously doubt the legitimacy of claiming this constraint to express a universal tendency, let alone an unmarked tendency.

7 Summing up

One of the important findings of our case study of coronals in ME is that while some of the patterns that distinguish a contact language from its substrate and superstrate sources may be instances of the emergence of the unmarked universal, we also find instances of the emergence of marked patterns, some of which may even be language particular. In this paper, we have outlined an Optimality Theoretic analysis of such patterns in ME, and explored its consequences for theories of language contact and for a theory of constraints in phonology.

Our central claim regarding the emergence of structure in contact offspring is that an important source of such emergence is the conflict between the existing substrate system and the pull to fit the superstratum. Caught in the conflict, the offspring compromises by choosing the underlying representations of the superstratum and the lexical representations of the substratum. Connecting the two mismatching representations results in the emergence of novel patterns that exist in neither of the parent languages.

In our endeavor to describe the facts of ME phonology, we discovered that constraints postulated as universals often contain redundancies when a number of constraints express the same core generalization with variations in its manifestation. We have demonstrated that these redundancies can be eliminated if we assume a theory of constraint generation that derives the variants in a constraint homologue from an underspecified constraint core, by filling in the fragments of locus, trigger, outcome value, and domain.

As the linguist's analogues of homologous organs in biology, constraint homologues form important anchors of invariance, variation, and asymmetries of probability that lie at the intersection of language typology, language change, and language contact. Constraint homologues like place assimilation, rounding harmony, epenthesis, consonant lenition, and vowel reduction in phonology could be thought of as counterparts in syntax-semantics of

constraint homologues like anaphora, control, case marking, and word order. Like constraint homologues, we may also think of construction homologues that exhibit cross-linguistic invariance, variation, and asymmetries of probability in morphology-syntax, constructions such as the passive, causative, applicative, cleft, topic, and relative clause being classic examples.

As in the case of homologous organs, constraint homologues exhibit a hierarchical structure. The homologue of spreading constraints, for instance, may be viewed as being composed of assimilation on the one hand, and harmony on the other. As a sub-homologue, assimilation may be viewed as being composed of place assimilation, voicing assimilation, nasal assimilation, and so on. Likewise, the homologue of lenition may be viewed as being composed of the sub-homologues of vowel reduction and consonant lenition ('constraint assembly' in Mohanan 1993).

To express asymmetries in the cross-linguistic probabilities of ranking markedness and faithful constraints, we have suggested that each markedness constraint homologue be paired with its faithfulness counterpart, with the provision to specify their relative default ranking as a universal (UDR). Such constraint pairs express universal tendencies that, depending on which member of the pair is ranked higher, can be activated or deactivated by language particular reversal of the default ranking. Such universal defaults would express the substance of markedness universals that the Prague School, SPE, and underspecification theories have sought to express.

References

Appa Rao. 1978. *Phonology of Telugu English*. MLitt Dissertation. CIEFL, Hyderabad, India.

Archangeli, Diana. 1984. *Underspecification in Yawelmani Phonology and Morphology*. Doctoral dissertation. MIT, Cambridge, MA.

Archangeli, Diana. 1988. "Aspects of underspecification theory." *Phonology* 5:183-208.

Bansal, R. K. 1969. *Intelligibility of Indian English*. CIEFL, Hyderabad, India

Bao, Zhiming. 1998. "Theories of language genesis." In Foley, J. et al (eds) *English in New Cultural Contexts: Reflections from Singapore*. Oxford University Press.

Bao, Zhiming. 2003. "Social stigma and grammatical autonomy in non-native varieties." *Language and Society* 32.1.

Bhat, D. N. S. 1973. "Retroflexion: an areal feature." *Working Papers on Language Universals* 13:27-67.

Bickerton, Derek. 1981. *Roots of Language*. Ann Arbor, Mich.: Karoma Publishers.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Bresnan, Joan. 2001. "Pidgin genesis and Optimality Theory." In Jeff Siegel (ed) *Processes of Language Contact: Case studies from Australia and the Pacific*. 145-173.

Broselow, Ellen. 1992. "Parametric variation in Arabic dialect phonology." In E. Broselow, M. Eid, and J. J. McCarthy (eds) *Perspectives in Arabic Linguistics*. 7-45. Amsterdam and Philadelphia: John Benjamins.

Broselow, Ellen, Su-I Chen, and Chilin Wang. 1998. "The emergence of the unmarked in second language phonology." *Studies in Second Language Acquisition* 20:261-280.

Chomsky, Noam. & Morris, Halle. 1968. *The Sound Pattern of English*. Harper & Row.

Das, Shyamal. 2001. *Some Aspects of the Prosodic Phonology of Tripura Bangla and Tripura Bangla English*. Doctoral dissertation. CIEFL, Hyderabad. [ROA 493-0202]

Davis, A., C. Criper & A. Howatt (eds). 1984. *Interlanguage*. Edinburgh University Press.

DeGraff, M. 1999. *Language Creation and Language Change: Creolization, diachrony, and development*. Cambridge, MA: The MIT Press.

De Lacy, Paul. 2002. *The Formal Expression of Markedness*. Doctoral dissertation. University of Massachusetts, Amherst, MA.

Flemming, Edward. 2002. *Auditory Representations in Phonology*. New York: Routledge.

Flemming, Edward. 2003. “The relationship between coronal place and vowel backness.” Ms. Stanford University.

Haiman, John. 1985. *Natural Language: Iconicity and erosion*. Cambridge: Cambridge University Press.

Kachru, B. 1983. *The Indianization of English*. Oxford: Oxford University Press.

Kachru, B. (ed.) 1992. *The Other Tongue: English across cultures*. Urbana: University of Illinois Press.

Kandiah, T. 1998. “The emergence of New Englishes.” In Foley, J. et al (eds) *English in New Cultural Contexts: Reflections from Singapore*. Oxford University Press.

Kay, Paul and Gillian Sankoff. 1974. “A language-universals approach to pidgins and creoles.” In *Pidgins and Creoles: Current trends and prospects*, 46-84. Washington D.C.: Georgetown University Press.

Kiparsky, Paul. 1982. “Lexical Phonology and Morphology.” In *Linguistics in the Morning Calm*. Seoul, 3-91.

Kiparsky, Paul. 1985. “Some Consequences of Lexical Phonology.” *Phonology Yearbook* 2, 85-138.

Kiparsky, Paul. 1997. “LP and OT.” Handout from LSA Summer Linguistic Institute, Cornell University.

Köhler, K. 1984. “Contrastive phonology and the acquisition of phonetics skills.” In S. Eliasson, ed. *Theoretical Issues in Contrastive Phonology*. Julius Groos Verlag Heidelberg.

Lado, R. 1964. *Language Teaching*. New York: McGraw Hill.

Lindblom, Björn. 1986. “Phonetic universals in vowel systems.” In J. Ohala amd J. Jaeger (eds) *Experimental Phonology*, pp. 13-44.Orlando: Academic Press.

Lombardi, Linda. 2000. “Second language data and constraints on manner: explaining substitutions for the English interdentals.” [ROA-418]

McCarthy, John J. and Alan Prince. 1994. “The emergence of the unmarked: Optimality in prosodic morphology.” In Merce Gonzalez (ed). *Proceedings of the North East Linguistic Society* 24. pp. 333-379. Amherst, MA: GLSA Publications. [ROA-13-0594]

McCarthy, John J. and Alan Prince. 2001. “Prosodic morphology.” [Available on Rutgers Optimality Archive, ROA-482-1201]

Mohanan, K. P. 1982. *Lexical Phonology*. Doctoral dissertation, MIT. [Distributed by Indiana University Linguistics Club]

Mohanan, K. P. 1986. *The Theory of Lexical Phonology*. Reidel, Dordrecht.

Mohanan, K. P. 1992. "Describing the Phonology of Non-native Varieties." *World Englishes* Vol. 11. No.2/3 Pergamon Press. 111-128.

Mohanan, K. P. 1993. "Fields of attraction in phonology." In J. Goldsmith (ed). *The Last Phonological Rule*. pp. 61-116. Chicago: The University of Chicago Press.

Mohanan, K. P. and Tara Mohanan. 1984. "Lexical phonology of the consonant system in Malayalam." *Linguistic Inquiry* 15, 575-602.

Mohanan, Tara. 1989. "Syllable structure in Malayalam." *Linguistic Inquiry* 20(4): 589-626.

Mufwene, Salikoko. 1986. "Creolization and grammaticization: what creolists could contribute to research on grammaticization." In P. Baker and A Syea (eds). *Changing Meanings, Changing Functions*. pp. 3-28. London: University of Westminster Press.

Mufwene, Salikoko. 1991. "Pidgins, creoles, typology, and markedness." In *Development and Structures of Creole Languages. Essays in honor of Derek Bickerton,* 123-143. Amsterdam: Benjamins.

Mühlhäusler, Peter. 1986. *Pidgin and Creole Linguistics*. Oxford: Basil Blackwell.

Padgett, Jaye. 2003. Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21:39-87.

Pesetsky, David.1979. "Russian morphology and lexical theory." Ms. MIT.

Platt, J. & H. Weber. 1980. *English in Singapore and Malayasia*. Oxford University Press.

Prince, Alan and and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. http://www. ruccs.rutgers.edu/roa.html, ROA-537.

Pulleyblank, Douglas.1983. *Tone in Lexical Phonology*. Doctoral Dissertation. MIT. (Published by Reidel Publishing Co. 1986.)

Rickford, John and John McWhorter. 1997. "Language contact and language generation: Pidgins and creoles." In Coulmas F. (ed.) 1997. *The Handbook of Sociolinguistics*. pp. 238-256. Oxford: Blackwell.

Rubach, Jerzy. 1984 *Cyclic and Lexical Phonology* , Foris, Dordrecht

Sebba, M. 1997. *Contact Languages: Pidgins and creoles*. New York: St. Martin's Press.

Selinker, L. 1969. "Language Transfer." *General Linguistics* 9:67-92.

Selinker, L. 1971. "Interlanguage." IRAL 10, 209-231 (Reprinted in J.C. Richards (ed.) 1974. *Error Analysis: Perspectives on Second Language Acquisition*. 31-54, Longman.

Siegel, Jeff. 2002. “Pidgins and Creoles.” In Robert B. Kaplan (ed.) *The Oxford Handbook of Applied Linguistics*. Oxford: Oxford University Press.

Sharma, Devyani. 2001. “The pluperfect in native and non-native English: A comparative corpus study.” *Language Variation and Change* 13:343-373.

Sharma, Devyani. In preparation. Doctoral dissertation. Stanford University.

Singh, Rajendra. 1991. “Interference and contemporary phonological theory.” *Language Learning* 41:157-175.

Singh, Rajendra. 1995. “Rethinking the relationship between phonology and loan phonology.” *Folia Linguistica* 29:395-405.

Singler, John V. 1988. “The homogeneity of the substrate as a factor in pidgin/creole genesis.” *Language* 64:27-51.

Singler, John V. 2000. “Optimality Theory, the minimal-word constraint, and the historical sequencing of substrate influence in Pidgin/Creole genesis.” in John McWhorter (ed). *Language Change and Language Contact in Pidgins and Creoles*, 335-351. Amsterdam: John Benjamins.

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Handout from talk presented at Rutgers Optimality Workshop I, Rutgers University.

Sridhar, Kamal, and S. N. Sridhar. 1992. Bridging the paradigm gap: Second language acquisition theory and indigenized varieties of Engish. In Kachru (ed) 1992.

Stampe, David. 1972. *How I Spent my Summer Vacation [A dissertation on Natural Phonology]*. Doctoral dissertation. Chicago: University of Chicago. [Published 1980. New York: Garland.]

Steriade, Donca. 1987. Redundant values. In A. Bosch, B. Need, and E. Schiller (eds). *CLS 23: Parasession on Autosegmental and Metrical Phonology*. 339-362. Chicago: CLS.

Steriade, Donca. 1995. Underspecification and markedness. In John Goldsmith (ed.) *Handbook of Phonological Theory*. pp. 114-174. Cambridge, MA: Blackwell.

Steriade, Donca. 2000. Directional asymmetries in place assimilation: a perceptual account. In E. Hume and K. Johnson (eds) *Perception in Phonology*. Academic Press.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Tiffen, M. 1974. “Intelligibility of African English.” ELT Documents 74/2, The British Council.

Trudgill, P. & J. Hannah. 1982. *International English: a Guide to Varieties of Standard English*. Edward Arnold.

Vijayakrishnan, K. G. 1978 *Word Stress in Tamil English*. MLitt. Thesis, CIEFL, Hyderabad, India.

Warrier, (Mohanan) Tara. 1976. *The Phonetics and Phonology of Malayalam*. MLitt. Thesis, CIEFL, Hyderabad, India

Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Weinreich, Uriel. 1953 *Languages in Contact*. New York: Mouton.

Williams, Jessica. 1987. Non-native varieties of English: A special case of language acquisition. *English World-Wide* 8:161-199.