Accounting for some similarities and differences among the Indo-Portuguese creoles

J. CLANCY CLEMENTS

Abstract

The present study appeals to the notions of social organization and Portuguese presence to account for some of the key difference among the creoles spoken in Diu, Daman, and Korlai (India). The concepts of frequency and perceptual salience are tapped to account for some of the similarities among these creoles.

1. Introduction

Within the group of Indo-Portuguese (IP) creoles, Ferraz (1987: 338-339), among others, distinguishes two general groups: the Gauro-Portuguese creoles and the Dravido-Portuguese creoles, the first having Indo-Aryan sub-/adstrate languages (Marathi and Gujarati) and the second Dravidian substrate languages (Kannada, Malayalam, and Tamil). The IP creoles to be compared in this study are Gauro-Portuguese creoles, specifically those spoken in Korlai, Daman, and Diu (see map 1).

In this study, four main questions will be discussed:

1. What do demographics and the presence of Portuguese have to do with the similarities and differences among the three creoles?
2. What does frequency have to do with the similarities and differences among the three creoles?
3. What do language acquisition and language processing have to do with the similarities and differences among the three creoles?
4. Where are these similarities and differences among the three creole languages manifest and why?

Map 1. Some locations in the Indian west coast where Indo-Portuguese creoles were or are spoken (adapted from Clements, 1996: xvi).
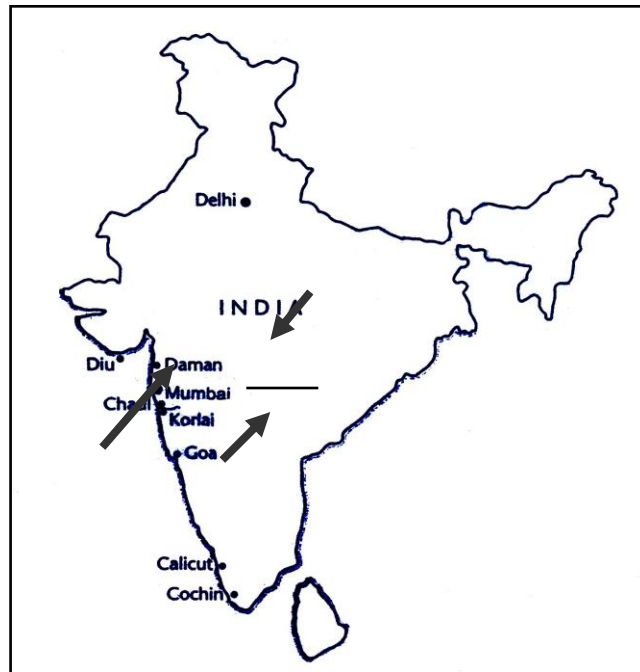
Within the evolutionary model of language change, the ecology of a given language is defined in terms of who speaks what with whom. In the case of the IP creoles, we can frame this in the following terms. In each of the communities under consideration, the demographics (especially community size) and the presence of Portuguese (which affects the linguistic feature pool) can be used to account for key differences among the three creoles. Within the evolutionary model, there are also certain factors that contribute to the selection and/or maintenance of certain lexical or structural features. In this paper, I discuss how frequency, processes involved in second language acquisition, and language processing account for key similarities in the three creoles. As will become clear below, the notion of language processing I use here refers specifically to perceptual salience of a lexeme or structure in the target language. All terms will be defined below. The discussion here is limited to the examination of the inventory and distribution of phonemes in the three creoles, the nominal and verbal morphology found in the creoles, as well as their pronominal systems, certain word order phenomena, and their respective core lexicons. Before discussing the linguistic phenomena, the relevant parts of the model and hypotheses stemming from them will be introduced.

2. Demographics, community size, and social stratification

The general model adopted for this study is the evolutionary model of language change proposed by Croft (2000) and Mufwene (2001). Specifically, Croft (2000) argues for the importance of social, cultural, and political considerations in defining a language. Both Croft and Mufwene define a language in population-theoretic terms, considering a language to be analogous to a biological species. In contrast to a biological species, however, which is defined as a group of interbreeding individuals, a language is defined by a speech community using a given language and the speech community is made up of individuals who are routinely engaged among themselves in communicative intercourse. The population-theoretic definition of a language implies that 'every speaker perceive[s] every other speaker as someone he or she should be able to communicate with by using what they perceive as the same language' (Croft, 2000: 18). An important part of this definition is the interaction between individuals in the community of speakers. As Croft states, '[c]ommunicative interaction depends not only on the degree of structural similarity of the varieties spoken, but also on the social behavior of the speakers. Serbian and Croatian are mutually intelligible to a high degree, but many speakers do not communicate with the opposite community due to the recent political changes in former Yugoslavia' (2000: 19). Thus, the analog of reproductive isolation is, in the case of language, communicative isolation, and interbreeding in a biological population equates to conversational intercourse in a speech community.

Croft also notes that in a speech community we find subdivisions analogous to those found within a species. For example, the biological notion of geographical race is a subpopulation of a species defined geographically, and typically has diverged structurally to some extent, though not enough to impede interbreeding. The notion deme is less inclusive than a geographical race and refers to a subpopulation within a species which, as with a geographical race, has a higher likelihood of interbreeding and a lower likelihood of breeding with members belonging to adjacent demes (Croft, 2000: 19). For Croft, race and deme correspond to the notion 'social network' within a given speech community: a social network is 'a group of people who are most likely to communicate with each other, and not so much with those outside the network' (2000: 20).

In this paper, I will argue that the size of a given speech community is important for the development of social networks. In order to discuss the role size of speech community plays in the development of language, I will tap other work by Croft, which in turn draws from a number of sources that deal with language change and social relations.

In this study, we deal in particular with creole languages, the creoles of Diu, Daman, and Korlai, India. All three formed, or began to form, some time in the 16th century, as a consequence of the contact between the Portuguese and the respective communities in which the Portuguese established settlements. Interestingly, all three creoles are spoken by somewhat isolated speech groups, communities that have maintained themselves due largely to religious and social barriers. In addition, Korlai and Diu were physically isolated, by rivers. This is an important aspect of each community and it needs to be kept in mind as we comment on the nature of each of the different speech communities in Diu, Daman, and Korlai.

In the general social evolution of groups, Diamond (1997: 268-269), in Croft (2003: 5), speaks of four types of social organization. These are shown in Table I, adapted from Croft's Table 2.

Table I. Types of Social Organization

	BAND	TRIBE	CHIEFDOM	STATE
Membership population	dozens	hundreds	thousands	over 50,000
Settlement pattern	seasonal dispersion into families	sedentary: one village or pastoral	sedentary: many villages, or pastoral	sedentary: many villages and cities
Basis of relationship	kin	kin-based clan	class and residence	class and residence
Stratification	egalitarian	egalitarian (or 'big man')	stratified by kin (elite, commoner, slave(?))	complexly stratified, not by kin, slavery

The archeological record suggests that there is a link between the growth (in terms of size) of a societal group and emergence of more complex social organizations on certain measures of complexity (Croft, 2003: 4). In fact, the most important characteristic for defining the scale of society types in Table I is, as Croft (2003: 6) notes, increasing population size of the society. And the increasing size of a group can lead to fission into two smaller groups. When fission takes place, smaller social units form, which has as a consequence a decrease in complexity by the same measures as those that quantify complexity (Croft, 2003: 4).

Table II shows two types of fission. The difference between type 1 and type 2 would depend on the size of the social unit. A large band would arguably be still smaller, and thus less complex, than a tribe. However, the outcome of the fission of a tribe or a large band would be similar in that the result would in both cases be two (or more) smaller bands. This same argument would also apply, *mutatis mutandis*, to chiefdoms and states.

Table II. Two types of social fission

Type 1	one large band => two smaller bands one large tribe => two smaller tribes one large chiefdom => two smaller chiefdoms one large state => two smaller states
Type 2	one tribe => two or more bands one chiefdom => two or more tribes one state => two or more chiefdoms

Keeping in mind that an important defining characteristic of a group is size, I suggest that although the creole communities in question do not fit, strictly speaking, within the definitions of bands, tribes, etc., as speech communities that are isolated from their surrounding communities due to caste and religion, they function in this manner.[1] That is, I propose that we think of the creole communities in question as functional tribes or chiefdoms, as the case may be. If we accept this working assumption, the following prediction and a language-related corollary can be advanced: the smaller (or larger) the speech community is, the less (or more) complex it will be in terms of social units. The corollary to this prediction is: the smaller (or larger) community is, the less (or more) heterogeneous the creole will be.[2]

One major exception to the isolation in all three communities has been the presence of the Portuguese culture and language throughout their respective histories. As we shall see, this presence has made a difference in the development of all the creoles. The presence of Portuguese and its linguistics features, in the respective communities has an inverse relation to the influence of the respective substrate language in the communities (Gujarati for Diu and Daman, Marathi for Korlai). This is especially true starting in the 20th century

---

[1] As discussed in Ansaldo (2009), the practice of endogamy also is relevant for the IP creoles under consideration. We know that in Korlai, marriage has taken place within the community for the most part. Church records show that in the last 19th and early 20th century, brides were brought from Goa. Since the change of the church language from Portuguese/Latin to Marathi, there have been marriages between Korlai villagers and Christian community members from the general Mumbai area. In Diu and Daman, records indicate marriages between members from the two communities were not uncommon, but both communities spoke and speak varieties that are mutually intelligible.

[2] As pointed out by one reviewer, this model does not include the effect of political history. However, the model applies profitably to the development of the IP creoles discussed here.

when education started playing a more prominent factor in language use in the three communities. This will be addressed below.

Comparing the size of Diu, Daman, and Korlai, in terms of their respective populations, it is clear that Diu and Korlai, with their populations around 250 and 800 inhabitants, respectively, are more homogeneous than Daman, which has approximately 4000 inhabitants (Cardoso, 2009; Clements, 2001a; Clements & Koontz-Garboden, 2002). However, these population figures are the current ones. What were the populations like in these communities historically? As far as the records available to me allow, I attempt to answer this question now. Moreover, I address the questions of the Portuguese presence in each community, as well as look at the factors of social stratification, and relations between the inhabitants of each community.

2.1. Diu

By the 17th century, the creole language in Diu had already formed. Based on the demographic distribution reported by Cardoso (2009: 72), between the 18th and 20th centuries, Diu's Catholic population fluctuated between 230 and 600, and is currently around 250.[3] Thus, as a group, the creole-speaking community in Diu would be considered a functional tribe. Indeed, it shares various characteristics with those of a tribe: many of the members of the community are related to one another, either by birth or marriage, and the community is largely without stratification, as the great majority of its habitants are white-collar workers, employed in government and education. That is, it is largely a community without social stratification and most likely has been for 200 years or more.[4]

Table III contains information about employment, but also about the languages in contact and the presence of Portuguese in Diu since the mid-16th century. For 300 years (roughly 1550 to 1850), the presence of the Portuguese language, culture, and government was strong.

Since the 1850s, this has waned considerably and from the latter part of the 19th century onward there has been an increasing presence of English, both in government and the church. Gujarati, the regional language, still has a strong presence in Diu, as well. Today, Diu Portuguese is spoken in Diu, for the most part, by fewer than 250 speakers. Given the number of speakers and the lack of social stratification in Diu, we would expect to find little variation

---

[3] Cardoso (2009: 92) states that in 1792 Diu had 598 Catholic (and arguably creole--speaking) inhabitants; between 1842 and 1900, this population ranged between 300-350, between 1920 and 1950, the population range was 228-393. In 1991, the creole-speaking population was 250.

[4] One reviewer pointed out that in Daman, and possibly in Diu, there were slaves of African origin. When I speak of lack of social stratification in Diu, the reference is to the community of Indian origin.

relative to, say, Daman, which has around 4000 speakers. That is, in the language use of Diu Portuguese speakers, we do not expect to find class--sensitive language structure. That is, in the language use of Diu Portuguese speakers, we do not expect to find class-sensitive language structure.

Table III. Information on the Diu Christian community

| Languages in Contact	Presence of Portuguese	Principal means of employment
-Portuguese -Gujarati -English (starting in late 19th c.)	– strong governmental presence till the mid-19th century; – very weak governmental presence since then; – relatively strong cultural presence till 1961; – a diminished cultural presence up to the present	– government – education

2.2. Daman

Of the three areas communities, Daman was settled the last, becoming a part of the Portuguese Empire in 1581 (Clements & Koontz-Garboden, 2002: 197-198). As in the case of Diu, the Portuguese presence in Daman was strong till around the middle of the 19th century and the regional Indic language of the area has always been Gujarati, though as in Diu, English in Daman has become increasingly influential in government, education, and the church from the late 19th century onwards.

Since the mid-19th century, Daman has had Catholic 3000+ inhabitants. In the 1990s, there were roughly 4000 members in the community (Clements, 1991). As shown in map 2, Daman has two general neighborhoods. The most populated is the Badrapur-Big Daman area (henceforth the Badrapur area) with around 2500-3000 inhabitants; the lesser populated being the Small Daman area, with around 1000-1500 inhabitants. The two areas are separated by the narrow Damanganga River.[5]

---

[5] It is conceivable that the river may have contributed to the development of the different communities. However, the Damanganga is about 20-25 meters across and would have been easily crossed with launches and wading during low tide in the dry season.

Map 2. The Daman Area.
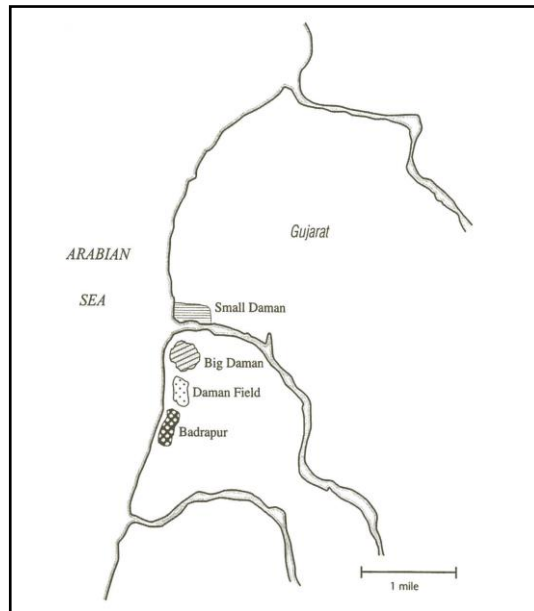
In terms of size, the respective creole-speaking communities in Daman and Diu are substantially different: Diu is tribe-sized at 250 inhabitants while Daman is chiefdom-sized at 4000+ members. In terms of social stratification, Daman is also different from Diu in that the Badrapur area consists largely of blue-collar workers (trades people, custodians, etc.), whereas in the Small Daman area we find a substantial white-collar work force (government workers, educators, lawyers, etc.). That is, Daman has social stratification, visible in terms of social class (employment), area of residence, and −as we shall see− in language use and language structure. This information is shown schematically in Table IV.

Table IV. Information on the Daman Christian community

DAMAN	Languages in Contact	Presence of Portuguese	Principal means of employment
	– Portuguese – Gujarati – English (starting in late 19th c.)	– strong governmental presence till the mid-19th century; – weaker governmental presence since then; – relatively strong cultural presence till 1961; – a weaker cultural presence up to the present	– government – education – law – skilled and unskilled blue--collar labor

2.3. Korlai

The area where Korlai is now located was settled the earliest of all three communities. The Portuguese first arrived in the area in 1505 and by 1523 a fortification had been built in Chaul, on the north side of the Kundalika River (see Map 3).

Map 3. The Korlai area.
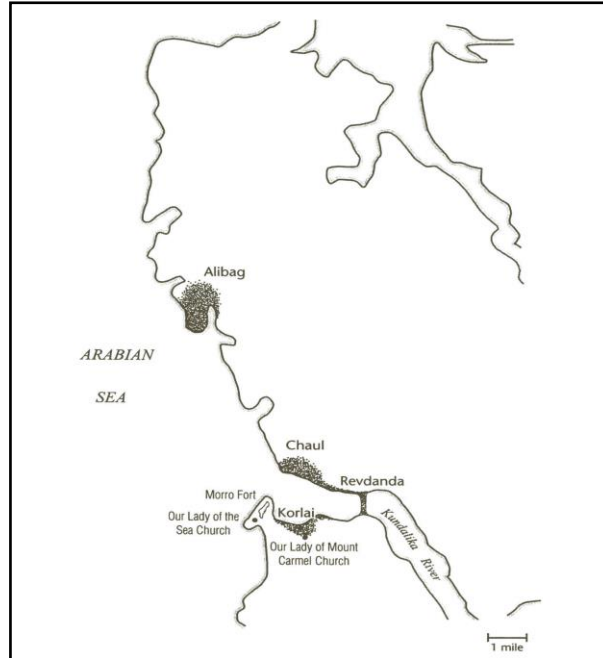
In the last 200 years, Korlai's population has increased gradually from 325 inhabitants in 1814 to over 600 in 1920, to more than 800 inhabitants in the 1990s (Clements, 1996: 220). Thus, in terms of size Korlai is closer to a tribe. With regard to social structure, Korlai inhabitants have always been agriculturalists, either as tenant farmers or land owners. Given that the great majority of them still work as agriculturalists, their community has little to no social stratification, as one would expect of tribes. This information is given in Table V.

As a result, we would expect little to no linguistic variation based on social stratification. However, due to the increased pressure on the part of Marathi, we might expect other types of variation.

Table V. Information on the Korlai Christian community

Languages in Contact	Presence of Portuguese	Principal means of employment
– Portuguese – Marathi	– relatively strong governmental presence till 1740; – since then presence has been strongly diminished (presence of parish priest) up 1960s; – since 1964, no presence of Portuguese	– agriculture

To summarize this section, I have described the creole-speaking communities of Diu, Daman, and Korlai within a typology of social organization, according to size, presence of Portuguese culture, government and language, and presence of stratification according to what we may call social networks. In the smaller, tribe-like communities of Diu and Korlai, we find little to no stratification. In Daman however, a community more than ten times larger than those in Diu and Korlai, there is class-sensitive stratification. This information is summarized in Table VI below.

Table VI. Social type of and Portuguese presence in Diu, Daman, and Korlai

	Size of creole-speaking community	Class-sensitive stratification	Portuguese presence
Diu	less than 250 (tribe)	little to none	relatively stronger
Daman	4000+ (chiefdom)	present	relatively stronger
Korlai	800 (tribe)	little to none	relatively weaker and minimal to non-existent since beginning in the 20$^{\text{th}}$ century

In the next section, I will discuss the impact of social type in Diu, Daman, and Korlai, on the language structure and language use of the respective creole languages.

3. Social organization, nature of contact, and the acquisition process

Above, I described the social type of Diu, Daman, and Korlai and the role of Portuguese presence and its linguistic features in language change, all within an evolutionary framework of language change. In this section, I propose an account for the variation among the creoles based on the factors just mentioned, but before undertaking this, I need to introduce other factors

involved in the evolutionary model of language change and the predictions they generate for the three creoles in question.

3.1. Social organization and Portuguese contact

In examining language change, an important question to ask is: Which components of the grammar of a language are more (or less) susceptible to undergo language change? A historical linguistics textbook such as Trask (1996) suggests that, typically the morphology and phonology of a given language are less susceptible to change than the syntax or the non-core vocabulary of the same language. The core lexicon is more stable than the non-core lexicon.

3.1.1. The lexicon

Thus, if there is a change in the core lexicon in any of the creoles in question, we hypothesize that it may be due to pressure from the surrounding communities, which would be more probable if there was a lack of a Portuguese presence, and a concomitant greater influence of the surrounding substrate languages. Thus, I hypothesize that there would be less likelihood of lexical replacement in the core vocabulary in those creoles with a relatively greater presence of Portuguese in their community. Given what we saw above, it would be expected that, in their respective core vocabularies, Daman and Diu would have replaced fewer lexical items (if any) than Korlai. The hypothesis seems to be borne out: Cardoso (personal communication, 4 June, 2009) states that in Diu only between 1% and 3% of the Diu core vocabulary is from a source other than Portuguese.[6] In the case of Daman, only two nouns from the core vocabulary have been replaced (1%): /tænk/ 'lake' /dãntʃ/ 'horn'. In contrast, in Korlai 11% (23 of 207 core vocabulary items) has been replaced. We list the lexical items replaced in Table VII below, ordered according to the categories used by Pagel *et al*. (2007).

The categories in which lexical replacement is found in Korlai are not random. Using the frequency of a word in discourse, Pagel *et al*. (2007) found that not all words are replaced at the same rate. In the languages they examined, they found that "[f]or a given frequency of meaning-use, prepositions and conjunctions evolve most quickly, followed by progressively slower evolution for adjectives, verbs, nouns, special adverbs ['what', 'when', 'where', 'how', 'here', 'there' and 'not'], pronouns and finally numbers" (in descending order) (Pagel *et al*., 2007: 719).[7]

---

[6] The core vocabulary list, taken from Comrie and Smith (1977), comprises 207 items.

[7] Pagel *et al*.'s lexical replacement scale is roughly suggestive of the lexical portion of Thomason and Kaufman's (1988: 74-95) five-level, contact intensity-based, borrowing scale, according to which non-basic vocabulary is first borrowed, followed by conjunctions and certain adverbials, then adpositions, then inflectional

   The prediction for the creoles in question is the following: if the presence of Portuguese influenced lexical evolution in the creoles, we would expect to see the more change in Korlai and the less in Daman and Diu, given that Korlai ceased to have a significant Portuguese presence from 1740 onwards, whereas in Diu and Daman a strong Portuguese presence lasted up to the mid 1800s, and still today there is some presence of Portuguese, most notably in the church liturgy and in certain cultural ties. After all, Diu and Daman were Portuguese possessions until 1961. In Daman, only two non-Portuguese nouns are found in the core vocabulary; in Diu arguably nouns are also found, though no more than 3% at most. As Table VII shows, in Korlai the conjunctions/preposition category is the most highly affected, where 33% (2/6) of the items in the category have been replaced. At the other end of the spectrum, no numbers have been replaced, which follows from Pagel *et al.*'s ordering.

Table VII. Lexical replacements to the core vocabulary in Korlai

Part of Speech	
Nouns **(8/85 [9%])**	*bərəf* 'ice', *ɖʱəg* 'cloud', *dhopa* 'knee', *ful* 'flower', *reti* 'sand', *ʃiŋ* 'horn', *təla* 'lake', *pənkhə* 'wing'
Adjectives **(2/36 [6%])**	*bhoʈəl* 'dull', *gol* 'round'
Verbs **(8/55 [14%])**	*suskar tuma* 'breathe', *tarta hika* 'float', *ʋahun anda* 'flow', *bərəf hika* 'freeze', *loʈu* 'push', *dabu* 'squeeze', *bəsku* 'pierce', *ʋitʃar hedze* 'think')[8]
Pronouns **(1/8 [12%])**	*ye* 'this'
Numbers **(0 [0%])**	
Conjunctions and Prepositions **(2/6 [33%])**	*ani* 'and', *karən* 'because'
(Special) Adverbs **(1/12 [8%])**	*daʋri badzu* 'leftside'
TOTALS	22/207 = 11%

   In the middle are, in order of percentage of replacement, verbs (14% [8/55]), pronouns (12% [1/8]), nouns (9% [/85])), (special) adverbs (8%

---

   affixes in native words and, finally pronominal elements. The scale makes no mention of numerals.

[8] The verbs *suskar tuma* 'take breath', *tarta hika* 'become floating', *ʋahun anda* 'go flowing ', *bərəf hika* 'become ice', *ʋitʃar hedze* 'make thoughts' are complex verbs, the first element of which is from Marathi and the second from Portuguese

[1/12]), and adjectives (6% [2/36]). This ordering is not predicted: we would not expect to see so much replacement in the verb category but not in the noun and adjective categories. Moreover, we would not expect a pronoun to have been replaced. In sum, while the categories in the middle do not conform to Pagel *et al.*'s ordering, at either end, their ordering is correct: most replacement is found in the conjunction/preposition category and no replacement at all is found in the number category. In using Pagel *et al.*'s ordering in the manner in which I have, a caveat is in order. Pagel *et al.* looked at lexical evolution in general using large corpora and concentrating on languages that have been evolving thousands of years. Of course, this is not the case with the creoles in question. We know that they formed after 1520 and that, at least with Korlai, there has been a substantial amount of contact--induced language change. Yet, even with this caveat, Pagel *et al.*'s ordering is applicable to the languages in question.

Not only is the lexicon proper affected by the social organization of the communities and the presence/absence of Portuguese. We find a comparable effect in the question-word and pronominal systems, as well. In Table VIII, the question words are listed for the three creoles under discussion. Table IX contains those question words that display variation among the creoles. All the creoles share 'who, what, where, why, how much/many', and 'which'. Where there is variation (see Table IX), it is found in the largest community, Daman, and the use of the forms is sensitive to sociolinguistic factors such as education and socio-economic class. This is predicted by the nature of bigger chiefdom-like community.

In the two smaller communities, the forms found reflect the socio--economic nature of the community. That is, the more basilectal forms *kɔr* 'when' (< Ptg. *que hora* 'what hour') and *kilɛ* 'how' (< Ptg. *que laia* 'what manner') are originally bimorphemic wh-words and found in Korlai, which consists almost entirely of agriculturalists.

Table VIII. Question words in Diu, Daman, and Korlai
(Cardoso, 2009: 130; and Clements & Mahboob, 2000: 469-471)

	'who'	'what'	'when'	'where'	'why'	'how'	'how much/ many'	'which'
Diu	*kẽ*	*(u)ki*	*kwɔn*	*ɔn*	*pərki/ purki*	*kom*	*kwõt*	*kwɔl*
Daman	*kẽ*	*ki/ ki koiz*	*kwan/ kɔm/ kyɔr*	*un*	*parki*	*kilay/ kɔm*	*kwant*	*kal*
Korlai	*kɛ̃*	*ki*	*kɔr*	*un*	*pəri*	*kilɛ*	*kãt*	*kal*

Table IX. Question words with form variation in Diu, Daman, and Korlai

	'when'	'how'
Diu	*kwɔn*	*kom*
Daman	*kwan/kɔm/kyɔr*	*kilay/kɔm*
Korlai	*kɔr*	*kilɛ*

The more acrolectal form *kwɔn* 'when' (< Ptg. *quando* 'when') and *kom* 'how' (< Ptg. *como* 'how') are found in Diu, where the population consists largely of white-collar workers. Finally, in Daman, which is a larger community and has both blue- and white-collar jobs, we find both the basilectal and acrolectal forms.

In the respective pronominal systems, we find variation distributed in a similar way. In Tables X and XI, the respective subject and object pronominal systems found in the three creoles are given. In the larger community of Daman, two clearly distinguished pronominal systems are found:

Table X. Subject pronominal systems in Diu, Daman, and Korlai (Cardoso, 2009: 131--132 for Diu; Clements & Koontz-Garboden, 2002: 207-209 for Daman and Korlai)

	1s	2s fam.	2s form	3s	1p	2p	3p
Diu	*yo*	---	*use*	*el, ɛl*	*nɔs*	*usez*	*e(l)z*
Daman (Sm D)	*yo*	---	*use*	*il, ɛl*	*nɔs*	*usez*	*ez*
Daman (Bg D/Bad.)	*yo*	*ɔs*	*use*	*il, ɛl*	*nɔs*	*usez*	*ilot*
Korlai	*yo*	*wɔ*	*use*	*el*	*nɔ*	*udzo*	*elo*

Table XI. Object pronominal systems in Diu, Daman, and Korlai (Cardoso, 2009: 131--132 for Diu, Clements & Koontz-Garboden 2002: 207-209 for Daman and Korlai).

	1s	2s fam.	2s form	3s	1p	2p	3p
Diu	*a mi*	---	*a use*	*a el, a ɛl*	*a nɔs*	*a usez*	*a e(l)z*
Daman (Sm D)	*ami*	---	*aʋse*	*ayil, ayɛl*	*anɔs*	*aʋsez*	*ayi(l)z*
Daman (Bg D/Bad.)	*pa(r)mi*	*pɔrɔs*	*puse*	*pirel, pirɛl*	*pənɔs*	*pusez*	*pilot*
Korlai	*par(m)i*	*pɔrɔ*	*puse*	*pel*	*pɔnɔ*	*pudzo*	*pelo*

One system is largely used in Small Daman that coincides almost exactly with that found in Diu; the other is largely found in Big Daman/Badrapur that shares many features with that found in Korlai.

The key features that characterize the systems found is Diu and Small Daman are: no 2s familiar forms (the 2s formal forms are the only ones found), gender distinction in the 3s, and incorporation of *a* into the object pronouns.[9] In these communities the forms ɔs and pɔrɔs are stigmatized and either avoid, as in Daman, or non-existent, as in Diu.[10]

The pronominal systems found in Big Daman and Korlai are characterized by: distinction between 2s familiar and formal forms, originally bi-morphemic 3p form *ilot* (< Ptg. *eles outros* 'they others') and *pilot* (< Ptg. *para eles outros* 'for them others') in Big Daman and *elo* (< Ptg. *eles outros* 'they others') and *pelo* (< Ptg. *para eles outros* 'for them others') in Korlai, and the use of Portuguese *para* as the marker incorporated into the object pronominal system. Moreover, the Korlai pronominal system has traits not shared by either system found in Diu and Daman: gender non-distinction in 3s, originally bi-morphemic 2p form *udzo* (< Ptg. *vocês outros* 'you-plural others').

The differences between the systems in Diu and Small Daman, on the one hand, and in Big Daman and Korlai, on the other, fall out naturally from the distinctions in social organization and in social class. The fact that Daman is a larger community lends itself to fission into two smaller speech communities and, indeed, this is what we find. The difference in social class in Small Daman (predominately white collar) and Big Daman (predominately blue collar) accounts for the acrolectal traits in the pronominal system in the former and the basilectal traits in the latter. Given the small size of the communities in Diu and Korlai, we would not expect there to be lectal differences, and expectation is borne out: Diu is largely white collar and its pronominal system exhibits only acrolectal traits. Korlai, by contrast, is almost exclusively agriculturalist and exhibits forms from what in Daman would be the basilectal variety.

3.1.2. Sound systems

Another aspect to discuss involves the respective sound systems in Diu, Daman, and Korlai. Above, it was noted that the sound system is typically more stable than, say, the syntax of a language (e.g. word order, etc.). Given the variable presence of Portuguese in the three communities, one would

---

[9] The pronoun *a* marked and continues to mark the indirect object forms in canonical Portuguese, both spoken and written, but as Baxter (1988) notes, *a* also marked animate direct objects in 16th century Portuguese. This could be why *a* is employed as a marker for all object pronouns in Diu and Small Daman, without distinction between direct and indirect objects.

[10] Clements (2001c) discusses in detail the basilectal and acrolectal verb and pronominal forms in Daman and the stigmatization of the former.

expect that if we found change in the sound system in one or more of these creoles, we would be more likely to find it in Korlai because of the lack of a significant Portuguese presence there since 1740. For present purposes, I will only look at the consonant inventories of the creoles because that is where we find the differences. In Table XII, the respective consonant systems of Diu, Daman, and Korlai are shown.

Table XII. Consonantal Phonemes

Diu (19) **(Cardoso, 2009)**	/p, b, t, d, k, g, m, n, ŋ, ɾ, f, ʋ, s, z, ʃ, tʃ, dʒ, j, l/
Daman (22) **(Clements &** **Koontz- Garboden,** **2002)**	/p, b, t, d, k, g, m, n, ñ, ɾ, r, f, v, s, z, ʃ, ʒ, tʃ, dʒ, j, l ([l, ʟ]), ʎ/
Korlai (25, **excluding sounds in** **parentheses)** **(Clements &** **Koontz-Garboden,** **2002)**	/p, b, (bh), t, th, ʈ, ʈh, d, dh, ɖ, k, (kh), g, (gh), m, n, ɾ, f, ʋ, s, ʃ, ts, dz, tʃ, dʒ, j, l, h/
Portuguese (15th- **-16th c.)** **Clements, 1996)**	/p, b, t, d, k, g, m, n, ñ, ɾ, r, f, v, s, z, ś, ź, ʃ, ts, dz, tʃ, dʒ, j, w, l, ʎ/

A full discussion involving an in-depth comparison of the consonant systems of our three creoles would take us too far afield from our present purpose of examining the degree of change in the creoles. I note, however, that such a study needs to be undertaken. In Table XIII, the new phonemes for each creole are included.

Table XIII. New Phonemes in Diu, Daman, and Korlai

Diu	/ŋ, ʋ/
Daman	/v, ʒ, ([l, ʟ])/
Korlai	/(bh), th, ʈ, ʈh, d, dh, ɖ, (kh), (gh), ʋ, h/

In Diu, neither /ŋ/ or /ʋ/ is found in 15th-16th c. Portuguese. In Daman, we find /v/, which existed in Portuguese but in neither of the other creoles. This /v/ may have been maintained by acrolectal Daman, but it may also have been reintroduced. This is a distinct possibility since /ʒ/ did not exist in 15th-16th c. Portuguese but developed later from /dʒ/. Thus, /ʒ/ would have to be introduced into Daman. This same reasoning applies to the velar lateral [ʟ], which is a realization of /l/ in Daman.

The creole that exhibits the highest degree of change in its consonant system is Korlai. All the phonemes listed without parentheses are found in minimal pairs. With the exception of /gh/, all the phonemes in parentheses are found in words of Portuguese origin, but not in minimal pairs (cf. Clements, 1996: 61-80). If we just compare the number of new consonantal phonemes, Diu has two, Daman also has two, plus one new phonetic realization of /l/, while Korlai has 11 new consonantal phonemes. This situation in Korlai again is accounted for by lack of substantial Portuguese presence there from 1740 onwards.

3.1.2. Syntax

The last phenomenon to examine in this section is the basic word order and adposition order in the three creoles. Since the 15th century, Portuguese has exhibited a relatively free word order, with a marked preference for SVO order and exclusively prepositions. Based on the differential presence of Portuguese in the three communities and the fact that both Gujarati and Marathi are SOV and postpositional languages, we expect Diu and Daman to show far fewer signs of contact-induced change than Korlai. In fact, this is the case; both Diu and Daman have default SVO order, although both allow OV order in focus constructions. As for adposition order, both Diu and Daman have prepositions with one exception: in both languages the adposition *junt* is postposed with pronouns and preposed with full NPs.[11]

The situation in Korlai is significantly different. As shown in Clements (2001b), the older population in Korlai is in a transition, in which object order relative to the verb is split fairly evenly. In the younger generation, OV order is at 73%, such that Korlai can now be said to exhibit OV order. Regarding adpositions, all adpositions with CV structure have remained prepositions. The rest of the adpositions have become postpositions by means of a conventionalized construction. So, for example, *baʃ* 'under' is found as in *baʃ alb* 'under (the/a) tree' but much more often is used as *alb su baʃ* 'under (the/a) tree'. Similarly, *dẽt* 'in' is more often used as *kadz su dẽt* than as *dẽt kadz* 'in (the/a) house'.

So far, I have discussed various differences among the three creoles and how to account for them. The main factors that account for the changes are the size of the speech communities and the extent of the Portuguese presence, which, as mentioned above, is inversely related to the influence on each creole from the respective substrate languages.

In the next section, I address similarities among the creoles and suggest a way to account for them that involves primarily the nature of processing in

---

[11] Actually, the situation is somewhat more complex. In Diu, Cardoso (pers. comm. 21 November, 2009) notes that *jũt* marginally appears as a postposition with full NPs. In my elicitation work for Daman, an informant reported that all prepositions can appear as postpositions with pronominal elements, although in actual discourse and in the stories I collected, only *jũt* appeared as a postposition.

naturalistic second language acquisition, but I will also appeal to the feature pool configuration (extent of Portuguese presence vs presence of substrate influence) in order to account for particular choices in the formation of the creoles under consideration.

3.2. Naturalistic Second Language Acquisition and Language Processing (Perceptual Salience)

In this section, I argue that the uniform way in which L2 learners process input largely accounts for key similarities we find in Diu, Daman, and Korlai. I appeal to notions of naturalistic second language acquisition, frequency of use of items in discourse, perceptual salience of items that are processed by L2 learners. I now define these terms briefly.

By naturalistic L2 acquisition I mean untutored L2 acquisition. Frequency refers to the number of times in a given corpus that a certain item or form appears. The notion of perceptual salience is more complex. I will discuss only two definitions of perceptual salience, both of which are proposed by me, though other scholars have probably proposed similar definitions. The definition I propose is relative and is based on the ubiquity of CV structure, as well as the general assumptions that stressed syllables and free-standing morphemes are more easily perceived in the speech chain than unstressed syllables and clitics/affixes, respectively. The definition is relative, and is expressed in (1) below.

(1) Definition of Perceptual Salience
   a. $C_i V_j$ is more perceptually salient than $V_j C_i$, V_j
   b. stressed syllables are more perceptually salient than unstressed syllables
   c. free-standing morphemes are more perceptually salient than clitics/affixes

An additional comment is necessary for (1a). What is stated in (1a) is that, in comparing a CV to a V or VC structure, the V or C remain constant in this statement. That is, the sequence *ta* would be more perceptually salient than *at* or *a*. (1a) would be made stronger by claiming that, independent of what vowel or consonant is involved, any CV is more perceptually salient than any VC or V. For present purposes, I will assume the weaker definition.

What I have discussed so far assumes the notion of a 'landing site'. That is, I am assuming that in naturalistic L2 acquisition, forms and structure are incorporated (aka transferred and accommodated) into a second language based on there being a landing site, i.e., somewhere to transfer to. Assuming a

landing site, then the frequency of occurrence of a given item and its perceptual salience become factors.[12]

With specific regard to frequency, two aspects of frequency of occurrence are distinguished by Bybee (1985): the frequency of a single item (lexical strength) and the frequency of a sound sequence within a paradigm (lexical connections). For example, in the present-tense copula paradigm of Portuguese *ser* 'be', shown in (2), the by far most frequently occurring single form, *é*, has a V structure. However, the second most frequently used form, *são*, has a CV structure, which it shares with the two other forms *sou* and *somos*.[13] The prediction is that if any of the creoles have any reflexes of the present-tense form of 'be' as copula, it will be either *é* or *são*.[14]

(2)   <u>C V(CVC)</u>   <u>V(C)</u>
        s o u   é s
        | | |
        s o m o s   **é**
        | |
        | |
        | |\
        s ã o

For the past, the most frequently used form by far in the corpus by Davies and Ferreira (2006-) is *era*. I therefore would expect that, if there were any forms in the creoles from the past copula paradigm, the form *era* would be selected.

In the case of a regular verb such as *falar* 'speak', the most frequently occurring finite form is the 3s *fala*, which also happens to be the one that shares its sequence of sounds with all the other forms in the paradigm. In the preterit, the most frequently found form is *falou*, again the 3s form and, thus, we would expect this form to be selected by the creoles. The frequency of the present and preterit forms of *falar* are given in Tables XIV and XV. In Table XIV, I also include one non-finite form, that of *falar*. Given that it is by far the most frequently occurring form of all the forms so far discussed, the chances of it being selected as a base present-tense form would be high.

---

[12] Siegel (2008: 156) identifies six factors: markedness, perceptual salience, transparency, simplicity, frequency, and congruence. We concentrate here only on: frequency and perceptual salience, assuming with Haspelmath (2006) that markedness, in all its complexity, is best subsumed under frequency.

[13] The frequency counts for the Portuguese present-tense forms are found in Clements 2009: 22-23.

[14] The non-finite form *ser* would also be a candidate. However, in no Portuguese- -based creole is this form found. For this reason it is disregarded here.

(3)   <u>CV´ CV(C)</u>   <u>CV CV´</u>
       f a l   -o   f a l a m o s
       | | | | | | |
       f a l a s   (f a l a i s)
       | | | | | | | |
       f a l a   f a l a r
       | | | |
       f a l a m

Taking content verbs first, such as *falar*, what we find in Diu, Daman, and Korlai follows directly from our expectations based on frequency of occurrence of the forms in the paradigms. In Daman and Korlai, the base form of the verb is based on the infinitival form (*falá* and *halá*, respectively). For the present tense, at least, the default form *fal* in Diu is not derived from the infinitival form, as in Korlai and Daman, but rather on the 3SG form *fala*, i.e. the most frequently-occurring form of the present-tense paradigm. If only frequency were involved, in Diu we would also expect *falá* as the base form, even in the present tense, but this is not the case. One way to account for the difference is to appeal to the nature of the community in Diu: it has always been a small population and thus would be homogeneous. It has also been historically a white-collar community. Between the forms *falá* and *fal* as a base form, the latter, present-tense form would be closer to an acrolectal form referring to present situations.

(4)   CV CV´ …
       f a l ei
       | | |
       f a l aste
       | | |
       f a l ou
       | | |
       f a l á mos
       | | | |
       f a l a stes
       | | | |
       f a l a ram

Table XIV. Frequency of present-tense forms and the infinitival form of *falar* in spoken Portuguese, according to Davies and Ferreira (2006-)

	Form	Frequency in a corpus of 2,083,296 words	Tokens per million
1s	falo	116	55.68
2s	falas	9	4.32
3s	fala	455	218.40
1p	falamos	75	36.0
2p	falais	---	---
3p	falam	106	50.88
Infinitive	falar	1108	531.85
Gerund	falando	329	157.92

In the past tense, the preterit base form is derived from *falou*, as would be predicted given the frequency of occurrence of the form relative to the others in the same paradigm

One sees the strength of frequency of occurrence in the retention of the three verb classes in all three creoles. These were retained because these creoles, even Diu in this case, kept the reflexes of the infinitival forms: *falá* 'speak, talk' *bebé* 'drink', and *subí* 'go up'.

Table XV. Frequency of preterit-tense forms in spoken Portuguese, according to Davies and Ferreira (2006-)

	Form	Frequency in a corpus of 2,083,296 words	Tokens per million
1s	falei	183	87.84
2s	falaste	3	1.44
3s	falou	328	157.44
1p	falámos	27	12.96
2p	falastes	---	---
3p	falaram	27	12.96

3.2.1. The copula

The last point has to do with the copula forms we find in Diu, Daman, and Korlai. Given that the question of the different copulas in the three creoles is complex, I will only discuss the more general aspects of the copula in the three creoles.[15]

---

[15]   The presence of reflexes from Portuguese *estar* is seldom encountered and arguably part of the most acrolectal speech only and thus will not be considered here.

First, all creoles have the reflex of *tem/têm* 'she/he has; they have' in the present and *tinha* 'she/he had' in the past (see Cardoso, 2009: 330; Clements & Koontz-Garboden, 2002; Clements, 1996). The fact that all creoles have reflexes of Portuguese 'have' is most easily accounted for by appealing to the substrate languages, Marathi and Gujarati, in which no difference is made lexically between 'have', 'become', and 'be'.

While Korlai only has one copula (*tɛ, ti*), both Diu and Daman have, as well, reflexes of the 3SG Portuguese copula forms in the present and past: *é* 'PRS.be' and *er* 'PRT.be'. The fact that we find 3SG forms follows directly from the frequency argument mentioned above: *é* and *era* are the most frequently used in their respective paradigms. What is interesting about the distribution of the copulative verbs in Diu and Daman is that, overall, it patterns as does *ser-estar* in Portuguese, where reflexes of *tem/têm* and *tinha* replace the functional equivalent of the *estar* forms (cf. Clements & Koontz--Garboden, 2002).

4. Concluding Remarks

I started by framing this study within an evolutionary model of language change. Specifically, I sketched out some points about the effect of size on the nature of the communities of Diu, Daman, and Korlai and the effect of the Portuguese presence in these communities. I have suggested that the size of the community affects the degree of linguistic variation in these communities. It was found that that we have more than one variety of a creole in Daman, precisely the community with 4000+ inhabitants. Where we find a more homogeneous linguistic situations is in Diu and Korlai, each of which has fewer than 1000 inhabitants. I also discussed the nature of employment in the three communities and how this affects selection and maintenance of variants in the creolization process. Diu has been and continues to be largely white collar and displays many traits that would be considered acrolectal when compared to, say, Korlai at the other end of the spectrum and where the vast majority of inhabitants are agriculturalists. The biggest community, Daman, is where we find both an acrolectal variety, spoken in Small Daman, and a basilectal variety, spoken for the most part in Big Daman. Revealingly: Big Daman is largely blue collar more comparable to Korlai, and Small Daman white collar and more comparable to Diu.

I then surveyed some key differences among the three creoles: the lexicon, question words, pronoun systems, consonant inventories, object order relative to the verb, and adposition order. The differences are summarized in Table XVI.

The large degree of difference between Korlai, on the one hand, and Daman and Diu, on the other, follow from, not only the nature and size of each community, as mentioned above, but also from the variable presence of Portuguese language and its linguistic features. In Korlai, this ceased to be a

factor starting in 1740, whereas in Diu and Daman, the strong presence of Portuguese stopped around the mid-19th century, but Diu and Daman were Portuguese possessions until 1961 and Portuguese influence continues on till today, although it has been strongly reduced.

Table XVI. Summary of the differences in Diu, Daman, and Korlai

	Diu	**Daman**	**Korlai**
Lexicon	1-3% new core lexicon	1% new core lexicon	11% new core lexicon
Question words	no bi--morphemic forms	no bi-morphemic forms	2 bi-morphemic forms
Pronoun systems	acrolectal system	acrolectal and basilectal systems	basilectal system
Consonant inventories	2 new consonants	3 new consonants	11 new consonants
Direct object order relative to verb	SVO	SVO	SOV
Adposition order	Prepositions	Prepositions	Postpositions

The similarities among the three creoles were argued to be a result of frequency of occurrence of forms and the effect of perceptual salience in process of creole formation, which are also parts and parcel of the evolutionary model of language change. This is most apparent in the verbal forms that found their way into the three creoles: the base forms of the verbs, both present- and preterit-tense, are all either the most frequently occurring form of the finite verb paradigm or the most frequently occurring non-finite form.

In the copula forms, we find the same pattern: the forms in the creoles are Portuguese 3SG forms, i.e., the most frequently occurring forms of the respective paradigms *tem* 'she,he, it has' *tinha* 'she, he, it had', *é* 'she, he, it is', and *er* 'she, he, it was'. In an interesting development we find that Daman and Diu have both copulas, whose distribution is, roughly, analogous to that of *estar* and *ser* in Portuguese today. In Korlai, however, only one copula verb is found, with *tɛ* in the present tense and *ti* in the past tense. While the distribution of the copula is accounted for by appealing to the variable presence of Portuguese in the communities, the actual forms of the copula follow from the frequency. Finally, the fact that all three creoles have retained all three formal verb classes from Portuguese can be accounted for by appealing to frequency: all creoles have the infinitival forms of each verb class (*kata* 'sing', *bebe* 'drink', *subi* 'go up'), precisely the most frequently used form of the respective paradigms in spoken language.

In future research, the complexities of copula forms and copula use needs to be examined, especially the presence (infrequent in Daman and more frequent in Diu) of reflexes of Portuguese *estar* forms. However, I believe that the evolutionary model, which incorporates notions of social organization, the presence of linguistic features (due in this case to the presence or absence of the Portuguese), frequency, and perceptual salience, is useful in accounting for these and other similarities and difference among the Indo-Portuguese creoles.

References

Baxter, A. (1988) *A grammar of Kristang (Malacca Creole Portuguese)*. Canberra: Pacific Linguistics.

Bybee, J. (1985) *Morphology*. Amsterdam: John Benjamins.

Cardoso, H. (2009) *The Indo-Portuguese language of Diu*. Utrecht: LOT [Doctoral dissertation, University of Amsterdam].

Clements, J. C. (1996) *The genesis of a language: the formation and development of Korlai Portuguese*. Amsterdam: John Benjamins.

Clements, J. C. (2001a) The Indo-Portuguese Creoles: Languages in Transition, *Hispania* (1991), **74**, 637-646.

Clements, J. C. (2001b) Word order shift and natural L_2 acquisition in a Portuguese creole. In *Romance syntax, semantics and L_2 acquisition* (C. Wiltshire & J. Camps, editors), pp. 73-87. Amsterdam: Benjamins.

Clements, J. C. (2001c) Acrolect evolution and stigmatized speech in Daman Creole Portuguese. Paper delivered at *Colóquio sobre os crioulos de base lexical portuguesa*, Coimbra, Portugal, June 29-30.

Clements, J. C. (2003) The tense-aspect system in pidgins and naturalistically learned L2, *Studies in Second Language Acquisition*, **25**, 245-281.27.

Clements, J. C. (2009) *The legacy of Spanish and Portuguese: Colonial expansion and language change*. Cambridge: CUP.

Clements, J. C & Koontz-Garboden, A. (2002) Two Indo-Portuguese creoles in contrast, *Journal of Pidgin and Creole Languages*, **17**, 191-236.

Clements, J. C. & Mahboob, A. (2000) Wh-words and question formation in pidgins/creole languages. In *Language change and language contact in Pidgins and Creoles* (J. McWhorter, editor), pp. 459-497. Amsterdam: John Benjamins.

Comrie, B. & Smith, N. (1977). Lingua descriptive studies: Questionnaire, *Lingua*, **42**, 1-72.

Croft, W. (2000) *Explaining language change*. London: Longman.

Croft, W. (2003) Social evolution and language change. Manuscript. University of New Mexico.

Davies, M. & Ferreira, M. (2006-) *Corpus do Português* (45 million words, 1300s--1900s). Online at: http://www.corpusdoportugues.org.

Diamond J. (1997) *Guns, germs, and steel: a short history of everybody for the last 13,000 years*. London: Vintage.

Ferraz, L. I. (1987) Portuguese creoles of West Africa and Asia. In *Pidgin and Creole Languages – Essays in Memory of John E. Reinecke* (G. Gilbert, editor), pp. 337--360. Honolulu: University of Hawaii Press.

Haspelmath, M. (2006) Against markedness (and what to replace it with), *Journal of Linguistics*, **42**, 25-70.

Milroy, L. (1987) *Language and social networks*, 2nd edition. Oxford: Blackwell.

Mufwene, S. (2001) *The ecology of language evolution*. Cambridge: Cambridge University Press.

Pagel, M, Atkinson, Q. D. & Meade1, A. (2007) Frequency of word-use predicts rates of lexical evolution throughout Indo-European history, *Nature*, **449** (October 11), 718-721.

Siegel, J. (2008) *The emergence of pidgin and creole languages*. Oxford: Oxford University Press.

Trask, R. L. (1996) *Historical linguistics*. London: Arnold.

J. Clancy Clements
Indiana University
(Department of Linguistics and Spanish & Portuguese)
Dept. of Linguistics, MM322, IUB
1021 East Third Street
Bloomington, IN 47405
USA

clements@indiana.edu