THE CREOLIZATION OF PIDGIN MORPHOPHONOLOGY

William J. Samarin
Department of Anthropology, University of Toronto

Meeting of the Society for Pidgin and Creole Linguistics
University of Amsterdam, The Netherlands, 10-11 June 1993

Abstract

In support of the growing conviction that creolization, as a linguistic process, probably begins before a 'creole' with native speakers emerges, this paper examines one of the alleged characteristics of creolization: morphophonological condensation: e.g., *tɔŋɔnà > twà* 'when' (Sango). Two pidgins of similar ages, one with a European lexifier language, the other with an African lexifier language, are compared: Tok Pisin and Sango. After examining seven explanations for condensation (intention, nativization, fluency, tempo, naturalness, style, child-language errors) and finding them wanting, we consider the agents of change (children and adults), evidence for condensation in the pidgin stage, the linguistic 'stuff' that is changed, the etiology (the linguistic processes) of condensation in Sango, and the linguistic consequences of condensation. The latter in Sango is, among other things, the beginning of what may become subjectival prefixation and possessive suffixation. (Key words: Creolization, Condensation, Tok Pisin, Sango, Language change, Historical linguistics, Phonology)

Introduction

It would help to advance the study of pidgins and creoles if there were established a list of assertions, claims, generalizations, and the like, about these kinds of languages. With this list we could judge for ourselves how many of them were credible. We might all accept some, but some of us would judge others as being moderately or outrageously ill-founded and premature. Such a list would also be a focus for research and discussion.

Before addressing ourselves to one, let us consider sixteen examples of these kinds of assertions, a few noted only in recent reading.

"... pidgins do not have a stable phonology." (Romaine 1992:178)

"... the sound inventory of pidgins, particularly when compared to their lexifier and substratum languages" is small (Mühlhäusler 1986:148)

"The general characteristics of pidgins [or perhaps only pidgins based on European languages] are well known ..." (Labov 1990:14)

"... pidgins [or perhaps only pidgins based on European languages] do not provide all of the features which native speakers seem to demand in a language." (Labov 1990:7; quoted in Sankoff and Laberge 1974:74 and Sankoff 1980:197)

"... pidgins and creoles are as different as chalk and cheese; ..." (Bickerton 1983:236, with qualifying footnote)

"When pidgins [or perhaps only pidgins based on European languages] become creoles, and acquire native speakers, they are spoken with much greater speed and fluency" (Labov 1990:14; cf Romaine 1992:172, where the statement is qualified)

"Normally creolization happens first with a generation of boys, and only later spreads to girls." (Foley 1988:181)
"... 'creolization' is a process extending over a number of generations of native speakers." (Carden and Stewart 1988:51, cf 45)

"... it is well documented that stylistic diversification can set in well before creolization" (Mühlhäusler 1983:69)

"There is no concordial agreement between subject and predicate in pidgins and creoles" (Todd 1974:14)

[In pidgins] "... there is a lack of obligatory marking of grammatical distinctions ..." (Sankoff and Mazzie 1991:1)

"... pidgins show extreme 'multifunctionality,' i.e. the same lexical items occur as nouns, verbs, and adjectives." (Sankoff and Mazzie 1991:1)

"... pidgins exhibit a high degree of motivation and transparency in compounding" (Romaine 1992:155)

"... pidgins ... have widespread homophony" (Romaine 1992:151)

"Pidgins ... have little room for synonymy" (Romaine 1992:161)

Pidgins and creoles "do not constitute a formal type of languages which may be contrasted with other traditional types .... [They] have so far revealed no property that has not been attested in any non-PC [non-pidgin and non-creole] language" (Mufwene 1988:265)

The last of any such list might be the following: "... for most views on pidgins and creoles there is also an opposite view" (Mühlhäusler 1989:37).

In this paper I address myself to one belief about pidgins and creoles, as languages, and about creolization as a process, demonstrating that explanations for its occurrence are not scientifically founded. Positively, I hope to demonstrate that morphophonological change, while predictable on the basis of what little we know about creolization, can be only partly explained and that the factors are several. The conclusion must be drawn that creolization is not as dramatic a departure from a pidgin state as has been assumed.

My method might be described as scientific bricolage. I rummage through the literature looking for bits and pieces of things that I might be able to use in creating something that might help in leading to an understanding of pidgins and creoles. My predecessors, pioneers in this venture, whose statements are cited below, did not address themselves boldly and directly -- as one would expect in scientific discourse -- to the matter that concerns us and therefore should not be held responsible for failing to provide us with a complete and well-argued brief. If what follows has any merit, even while failing to be complete, it is because they were willing to play around with certain ideas which we now can examine and evaluate.

Phonological condensation in creolization

It would be generally if not universally held that in creolization -- that is, when pidgins become creoles -- they experience morphophonological change. To be more precise, to a greater or lesser degree some of their morphemes or words are at least reduced in form: they are shortened or condensed. (Reduction, condensation, shortening, erosion, and contraction have been used more or less synonymously; here we use condensation.) This belief might be expressed in the following words, not yet attested by me but implied or assumed in the literature: Morphophonological condensation characterizes creolization as a process of language change.

This view, whether true or not, has been made of only one creolizing pidgin, namely, Tok Pisin, because an attempt was made to explain it in that language. The process of reduction may, of course, have characterized some or all of the creoles in existence today, but we do not have for them the kind of data
that are available for Tok Pisin and, only recently, Sango. A well-known example in Tok Pisin is the reduction of the future marker bai. Sango can also be cited to prove this assertion, and it will be frequently cited in this paper.\(^2\) (The kinds of changes are too numerous to be fully discussed here, but some are mentioned.)\(^3\) In this language we now find fásó or fásó 'now' presumably from fádesó, where a syllable is deleted, as also in koró 'village' from kodoro. (However, intermediate forms still in existence constitute evidence for a gradual erosion rather than the real 'deletion' of a syllable: fátsó and kodro or kotro (to which we return below).\(^4\)

With the demonstration of reduction taking place in our lifetime -- and in the case of Sango mostly since I began learning it in 1952 -- we have learned something significant about pidgins and creoles, on the one hand, and, on the other, about language change. But we cannot be satisfied with that knowledge; in itself, it is not understanding. We must ask ourselves why the shortened forms of 'now,' for example, have arisen in Sango and why, apparently, they arose, as seems the case, mostly in the last couple of decades, when the language became much more widely used, in more contexts, more frequently, and by many more people, both adults and children, as either their primary or only language of daily use. It is not enough to say that without variation -- in this case between long and short forms -- a pidgin is not a natural linguistic system, whereas a creole is, because the latter has inherent internal variation (cf Labov 1990:26; Sankoff explicitly characterizes a creole as a natural language by comparison with a pidgin [1980:198]). How, we must ask, did this alleged 'naturalness' come into existence?

In separate sections we consider the explanations that have been proffered for condensation (why does condensation take place?), the agents of change (who is responsible for it?), evidence for condensation in the pidgin stage, the kinds of 'stuff' that undergo change (what forms are most vulnerable?), the etiology of condensation in Sango, and the consequences of condensation in Sango.

Explanations for condensation

Seven explanations have been given for the emergence of reduced forms and their concomitant variation. One can assume that they are not necessarily mutually exclusive.

1. Intention. The first explanation is teleological. It attributes purpose to the speakers of a creolizing pidgin. In the quotation cited above we are told that "pidgins do not provide all of the features which native speakers seem to demand." Others similarly hypothesized that "native speakers in general tend to have little patience with or respect for a language with few grammatical categories, limited lexicon, virtually no morphophonemic reduction rules, insufficient redundancy, and little opportunity for stylistic manoeuvre" (Sankoff and Laberge 1974:76, emphasis added here and in the whole of this paragraph). Purpose seems to be implied also in the following statement: in or along with creolization there emerge "sociological contexts that put speakers of pidgin ... in need of a 'universal language': a language that would serve all purposes of communication in the ongoing life of a stable community ..." (Jourdan 1991:195). Purpose is certainly expressed in the assumption that "relexification ... only occurs when a new language is invented for social-boundary marking or similar purposes" (Bickerton 1988:281, cf 278, 280). The extreme to which this view can be taken, but not necessarily by any of the authors just cited, is illustrated by this statement in general: "Since language is an instrument, it follows that a language can be evaluated, altered, corrected, regulated, and improved, and new languages can be created at will" (Tauli 1968:9; reviewed by Haugen [1969] unfavorably).

This attempt -- or this attempt in its present unsophisticated form -- to explain changes in creolization by intention on the part of speakers of a pidgin is subject to criticism in spite of the qualifying or hedging use of "seem to," "in general," and "virtually." There are at least three things that one can say about this position. (1) It specifies nothing about the ostensible creators of new forms. Of all the demographic characteristics perhaps age is the most important. Are the innovators adults who have spoken the pidgin for whatever number of years, or are they children? It is hardly likely that young children would experience any need (cognitive, psycholinguistic, or whatever) that adults would not experience. (About children more is said below.) (2) It specifies nothing about the social context in which innovation occurs. (3) It does not attempt to explain the creation of multiple forms in a speech community and their use intersubjectively. Although one might assume that a number of experimentations with the word for 'now' would be attempted by speakers of Sango, fulfilling an unconscious or conscious need, only a few of which, of
course, are accepted and learned by others, how do people agree on a social meaning for all the variants? Young people just do not say (or have not yet been observed in our texts to say) as people in their parents' and grandparents' generation would have said -- fàdesô -- but anything phonologically close to it in the proper syntactic context appears to be acceptable: e.g. fàsô, fàsàsô, fàtsô (pronounced fâtsô, a consonant cluster, not an affricate), fôsô, etc. And, as Romaine has also documented for Tok Pisin (1992:XXX), variants occur in the speech of a single speaker and even in one speaker's discourse, as can also be abundantly illustrated in Sango. (4) It is not reinforced with references to other cases of intention in language change.

The next three causative factors -- nativization, fluency, and speech tempo -- are not clearly or consistently and rigorously separated by those who invoke them explicitly or implicitly. It would appear that some believe that nativization -- having the pidgin as one's first language -- results in more fluent and faster speech than that which characterized the speakers of a pidgin as a second language. (Further discussion follows.)

2. Nativization. With respect to nativization (when people acquire a language as their first language, one that Calvet calls une langue grégaire [1992]) we might cite the statement that "... the most striking differences between native speakers [of Tok Pisin] and non-native speakers are probably to be found in the morphophonemics ..." (Sankoff and Laberge 1974:83; cf Romaine 1992:31). (The qualifying adverb "probably" appears to be applied to the part of language that is most affected, not to the difference between native and non-native speakers.) It has been found, however, that in Tok Pisin nativization is not as important a factor as having the language as the primary means of communication (when polylinguals use one or more languages more than others). This conclusion is drawn from the fact that some changes in that language took place before nativization began (discussed below). This seems to be true of Sango also, as will be demonstrated below. In any case, if we want to seriously consider nativization as a causative factor, we ought to investigate social factors as well. Would we expect the same kinds of changes in an isolated rural village, whose culture is still largely traditional, as in a city, whose culture is largely westernized or 'modernized'? Such ideal or hypothetical cases may not easily be found because of schools, where pidgin Tok Pisin is the lingua franca among pupils but English is taught. Something close to the ideal is the village of Gapun, studied by Kulick recently (1992), where shift to Tok Pisin is taking place.

3. Fluency. Fluency as a causative factor would naturally follow from nativization, but it can also characterize second-language speakers. In the following statement fluency is linked with tempo, which we look at next: "... it is only in a community of fluent second- and first-language speakers that rules which greatly enhance rate of production will have a chance of acceptance" (Romaine 1992:172, cf 54; see also Foley 1988:179). (In the vernacular of English this might be read as, "You don't have to say fàdesô. Say fàsô: you can talk faster that way.") Rules in the strict sense of the term, however, are not provided. What we get are lists of variants of, for example, bai, the future marker, and save, an aspect marker meaning 'to be accustomed to' (Romaine 1992:173, 175). It is not clear for these words how speakers knew what to do (i.e. what rules to follow). In any case, since there are so many variants, different speakers were following different 'rules,' unless, of course, all we mean by rules in phonological reduction is "Condense this word any way you can." (With respect to 'rules' in Sango see below.) Assuming that something like this is meant, we are led to paraphrase this statement by saying that (a) condensed forms help people to speak faster and that (b) fluent speakers of a language can produce condensed forms. The claim is implausible on at least two counts. (1) There is no evidence that only those who speak languages with short words can speak fast (or generate discourse quickly). If rate of production were a desired goal for human beings, there would not be any, or very few, languages with long words. (2) Fluency itself is arguable as a scientific concept. Does it, for example, apply only to performance? If so, it cannot simply mean 'speak a language rapidly,' for one can do something badly but rapidly. If, then, it characterizes linguistic competence, it must mean something like 'speak a language like a native.' And a person may speak slowly as a matter of habit or when motivated sociolinguistically (e.g., by context, topic, interlocutors, etc.) Since the competent person has been called the "maximally gifted person," one of whose abilities is to be able "to talk in coherent, reasoned and semantically dense sentences, mastering the syntactic/semantic resources of the language ..." (Romaine 1984:5, citing Fillmore 1979:93), this can surely characterize many second-language speakers, whether the languages be pidgins or not.

If fluency means something like the competence of a native speaker, can we get speakers of a pidgin (or creolizing pidgin) to distinguish between fluent and non-fluent speakers? I doubt it. If I asked a
number of persons for whom Sango is their primary language to identify among a sample of tape-recordings of first-language and second-language speakers of Sango, which ones spoke as well as they did (or just which ones spoke Sango well), I expect that I would find them unable to distinguish the cohorts. One might, on the other hand, expect them to distinguish between urban and rural on the basis of the fact that adult inhabitants on several occasions told me that they could do so, being able, they said, to identify even a person's ethnicity. One might also expect some rural Sango to be old-fashioned. For example, forty years ago the common word for 'uncle' was àu; today in Bangui it is kốvà; the urban form of the French loanword for 'to begin,' in one of its variants, is kùmá́sèè, whereas the older and rural one is kùmá́sì. However, this expectation was not confirmed by a test of 171 subjects in Bangui in 1992, who were selected because they knew no other Central African language than Sango. Only 52% of the subjects could identify those who were urban (Samarin 1993). (This test might be faulted for the brevity of the excerpts from recordings that subjects heard: 7 seconds.)

I do not know what to do with the phrase "have a chance of acceptance" in the quotation cited above (Romaine 1992:172). One interpretation of this sentence would make it possible for anyone to produce shortened forms, but that only in a community of fluent speakers would they be accepted. If that is so, I cannot imagine why it would be true.

4. Tempo. An "expanded pidgin," possibly the same as what I am calling a creolizing pidgin, has been characterized as having, among other features, "an increase in speech tempo" (Romaine 1988:138). And being able to talk faster is something, apparently, that people desire, for we are informed, as we saw above, that rules that "greatly enhance rate of production," presumably by morphophonemic condensation, are the ones that are more likely to be accepted (Romaine 1988:138). Condensed forms have indeed been attributed to the alleged fact that native, and especially young (presumably native), speakers talk faster than non-native speakers: for example, it is said that "The children [native speakers of Tok Pisin] speak with much greater speed and fluency, involving a number of morphophonemic reductions ..." (Sankoff and Laberge 1974:76); "young children speak a faster and structurally more advanced variety [of Tok Pisin] ..." (Mühlhäusler 1980:57). Although one person has publicly admitted that "systematic experimental comparisons of the speech tempo of 'natural' languages and pidgins" have not been done, she claims that Labov had "established that there was a significant difference in the rate of articulation of fluent Tok Pisin and Buang speakers ..." (Romaine 1988:132). Labov indeed claims that in spite of the fact that Tok Pisin has been used as a second language for a hundred years, it "is still spoken markedly slower than the native languages" (1990:14). As an "example," he says, not evidence (I point out), for this allegation, he cites the case of a Buang speaker, about 40 years of age, who had been "observed," and presumably recorded, at a village meeting by G[illian] and D[avid] Sankoff. In switching from Buang to Tok Pisin, we are told, the man took as long to say seven words (with what I count to be nine syllables) in Tok Pisin as in the immediately preceding eleven words (and sixteen syllables by my count) in Buang.

At this juncture we do not have any evidence that those who speak Tok Pisin as a second but primary language or as a native language, whether young or adult, talk faster -- usually, characteristically, in more contexts, etc. -- than those for whom it is a second language. All statements but Labov's are not supported by any kind of evidence. And Labov's is implausible, because the discourse from which two examples are taken is not situated sociolinguistically. That is, the contextualization is inadequate. Nonetheless, it is suggestive, and suggestive of a different deduction. It is to be noted, first, that the speech was made in a village meeting. One would expect a certain style of talk at such a meeting, and if we knew more about the speaker, we might predict that if he were a big man (the kind of personage identified as such in the literature on Papua New Guinea), he would indeed give a 'speech' rather than simply 'talk.' The fact that he switched from his native language with, presumably, co-ethnics (or wantoks < one talk) to Tok Pisin suggests the kind of language use that has been documented for Papua New Guinea. Competence in oratory characterizes big-men (Kulick 1992:142-149; Lederman 1984; Merland and Rumsey 1990; Watson-Gegeo and White 1990), and the use of other languages is part of this competence. This being the case, it is reasonable to suppose that in switching from Buang this particular person deliberately spoke in a different manner. (At times possibly measured, cadenced, sententious, or whatever.) That is all that would take to explain a measurable and consistent difference in one discourse between these two languages, in which, we might speculate, most, if not all switches were from Buang to Tok Pisin, not in reverse order. But we are not told. It should be further noted as relevant that Romaine herself observes that in her experience
switches from vernaculars to Tok Pisin in sermons had an "emphatic function" (1992:325). We are justified in supposing that 'emphasis' was characterized by a more deliberate pace.

If there are no experimental studies comparing the rate of speech of fluent speakers of a pidgin and fluent speakers of its creolized counterpart, assuming that such studies could be pursued in a scientifically acceptable way, and if the one alleged piece of "objective evidence" (REFERENCE) can be explained in other and better ways, do we have any reason for believing that pidgins are indeed spoken more slowly than creoles? It has been suggested (Romaine 1992:172) that they would be, since they are second languages. But this argument is insubstantial, because no evidence is cited to support the claim that second languages are always or more commonly spoken at a slower rate than first languages.

My doubts about the role of speech tempo in explaining the origin of shortened forms of words is supported by Don Kulick who has worked extensively in Papua New Guinea:

I am completely confident in stating that there is absolutely no difference [in the rate at which a vernacular and Tok Pisin as a creole] are spoken -- certainly none that could be held to be significant or even interesting across all contexts, speakers and speech genres.

... [The older men of Gapun who learned Tok Pisin in the 1930s and have been speaking it since then] are absolutely fluent in the language, and it would be ridiculous to assume that they speak it at a slower rate than they speak their vernacular. One of the aspects of language use that I document in my book [Kulick 1992] is the dexterity with which speakers in Gapun continually oscillate between Tok Pisin and the vernacular language. There is no sense whatsoever that tempo is a distinguishing feature between these languages.

Furthermore, the children in Gapun who have acquired Tok Pisin as their first language do not speak the language at a faster rate than older bilinguals (personal communication, 30 January 1993, used with permission).

Before leaving this topic we must comment on one word that is used instead of 'fast,' sometimes as a synonym -- a matter too important to put in a footnote. The word is 'allegro,' as in 'allegro speech' and 'allegro style.' If this word is used as a perfect equivalent of 'fast,' there might not be a problem (all other things being equal). But when it implies, or, as in 'allegro style,' identifies a particular manner of speaking, it may be, and usually is, misleading, for writers about pidgins and creoles characterize neither the style nor the contexts of the use of this alleged 'style.' Whereas news announcers on radio and television might be said to have an allegro style different from the one they use in normal conversation, can something similar be said in a scientifically rigorous manner of the speech of young native speakers of a pidgin? It is very doubtful. Like all users of language, their speech tempos are not just allegro and lento but range on a continuum in a dynamic way (that is, in the same discourse with the same interlocutor, etc.).

5. Naturalness. A creole has been explicitly defined as a natural language (Sankoff and Laberge 1980:198), and has been considered to be such by others. (This belief would necessarily make pidgins in some way less than natural.) It is said, for example, that in Tok Pisin b|o| 'theirs' is more natural than the original pidgin form bilong o| (Romaine 1989:217). One apparent explanation for this belief is that "natural linguistic systems show inherent internal variation ...") (Labov 1990:26). Presumably when people only said bilong o| and every other word in Tok Pisin in only one way, it was not really natural. Aphoristically, this belief might be expressed as "People don't talk in citation forms." The latter statement was put to me by a colleague as we discussed the topic of this paper. This extemporaneous statement in casual conversation might be nuanced in the following way: "In natural languages most people, when speaking the vernacular, that is, in casual speech, in everyday contexts, will use some forms that are reduced or condensed by comparison with other forms used in other registers." What appears to be axiomatic has, however, not been demonstrated comparatively, that is, cross-linguistically. And it would appear just as reasonable to suppose that a number of as yet undetermined factors, some linguistic and some nonlinguistic, determine a typology of phonological condensation. I know very little about the languages of the world in this connection, but no language that I do know has as much as urban Sango now has and as little as Gbeya has (Samarin 1967a). For the latter I will say that having spoken the latter for forty years, and in the 1950s every day of the week for five years, I know of nothing in the language other than 'citation forms.' Gbeya does not have inherent variation. There are, however, forms that are morphologically constrained. For example, rip 'eye, face' will be used in certain constructions such as rip-ri (face of water) 'the surface of the water' but...
rifim ‘my face.’ It is quite probable that rif is both historically and grammatically derived from rip, but it is not the case that the latter is the citation form for the word, unless we now define ‘citation form’ in such a way that it would accommodate the argument. (The fricative, incidentally, cannot occur at the end of a word in contemporary Gbeya.) Certainly, one would not get it in response to the question, "Whose face (did you say)?" Although *rifim may once have been possible, it is not today. Moreover, neither vowels nor consonants are replaced, modified, or lost in certain environments. Phonologically, Gbeya is, with the possible exception of ideophones, which are already in a class by themselves (Samarin 1967b, 1970), a monostylistic language (using ‘style’ rigorously), no matter how incredible that might appear to some. And if it is true of Gbeya, it probably is true of related Gbaya languages and to some degree of other Ubangian languages.

But Gbeya only illustrates the fact that citation forms, as linguists know them, are the result of language change introduced by literacy or by some comparable cultural focus on language form in a preliterate society. If the speakers of Sango (or at least some of them, whose social characteristics have not yet been determined) consider yèkè to be the ‘correct’ form of the verb ‘to be,’ even though in extemporaneous discourse it is used rarely and always when its use can be explained sociolinguistically -- e.g., in church, by Protestants, etc. -- it is certainly because this was the form in which it was codified in the literature of Protestants and Catholics. Among children the full form might be considered kè; altogether, there are 18 known variants of the verb (Samarin 1993a). And although vanga dà (mouth [of] house) has been used for ‘door’ (as with nù-tuwà in Gbeya and other Ubangian languages [Boyeldieu 1987]) at least since the 1950s along with vanga ti dà, many Central Africans consider the latter the correct form. Thus, after I counted four instances of the shorter form in a young man’s narrative to the shoemaker with whom I was sitting, I asked him, pretending relative ignorance of the language, which of these two phrases was better. As would be expected in our own society, unaware of which one he had used, he told me categorically that the longer form was better.

Although no survey of literacy in Sango has been undertaken in Bangui, my own involvement in literacy projects in the Central African Republic in the 1950s leads me to believe that the proportion of literate children would be very small by comparison with adults. Among the latter I would further suspect that reading skills were acquired before coming to Bangui.

6. Style. Woven into the attempts to explain the differences between varieties of Tok Pisin (largely on the basis of second-language vs first or primary language, on the one hand, and between pidgin vs creole, on the other hand) has been the concept of style, as has been noted. It is used in connection with rate of speech production: there is, it is said, a “rapid speaking style” (Romaine 1992:175) and there are “allegro speech styles” (Romaine 1992:172). Consequently, the shortened forms are called “stylistic variants” (Mühlhäuser 1986:70). In no discussion of morphophonemic change in Tok Pisin, however, as we have already noted, has an author rigorously correlated the shortened forms with anything that might be properly called style: as related, for example, to formal vs informal speech, to one topic of discourse as opposed to another, to one set of interlocutors as opposed to another, or even to speed of utterance. In the latter case we have not learned that any one speaker has more short forms when speaking rapidly than when speaking slowly. It would appear, therefore, that the word ‘style’ is being used loosely only to mean ‘a way of talking’: in other words, children (or urban children) use short forms; adults don’t or do to a lesser degree.

The emergence of short forms does not necessarily mean that one immediately has a style; it only means that the material is now there for speakers to use in the creation of a style. (See note 21.)

7. Child-language ‘errors.’ It seems to be implied that short forms in Tok Pisin appeared in early language learning on the part of children. They figure in the next section.

The agents of change

Change can take place only as people use a language. With creolization what kinds of persons are we concerned with? Or, we might ask, which of their social characteristics can be linked to linguistic innovation? Age is the one exclusively invoked. "Children," it is said of language change in general, "play an active role in elaborating, constructing and (re)-creating their language." Indeed, they are "the prime instigators of change" (Romaine 1989:218, 200). Elsewhere we are simply told that "it is in phonology in
general that children appear most significantly to differ from adult speakers of Tok Pisin at present ...” (Sankoff and Laberge 1980:269; cf Romaine 1990:191-192).

Two explanations have been given for what children do. First, in a pidgin community, ignoring the possible differences between a rural and an urban one, children “have a greater scope to act as norm makers due to the fact that a great deal of variability exists among the adult community” (Romaine 1989:216-217). This explanation is implausible, because it has not been demonstrated that the kinds of changes that have taken place in Tok Pisin and Sango are linked in any way with variability in the adult population of second-language speakers. Furthermore, it is not true, at least for Sango, that “a great deal of variability” exists — only some kinds of variability. (Vocalic variation was cited in note 2. One goal of the project that led to the grammar of Samarin 1967 was to determine how much uniformity there was throughout the country.) Second, “It may be,” we are told, “that a certain amount of this reduction is part of the normal tendency at this stage in children’s acquisition of Tok Pisin, which accelerates this natural part of the expansion phase” (Romaine 1992:175). It is an interesting statement, but it is not an unambiguous one. Why, for example, is only “a certain amount” of reduction explained in this way, and what kinds of reduction might be explained? There is, moreover, a “normal tendency” to do what? Do all children in acquiring language for the first time reduce words? To that question the answer is, of course, yes (at certain ages). But the reduction of child language at some early stage of acquisition seems to be different from that with which we are concerned here. Urbanized Sango has the same indigenous phonemes (i.e. ignoring those that have recently come from French) as any other variety of Sango; it is the shapes of words that differ. For Tok Pisin we are ill informed about the age of the children who ostensibly are introducing the new forms. The role of children in language change certainly cannot be denied. Bhat (1970), H. A. Gleason (see Samarin 1967c:61), and Gumperz and Wilson (1971, according to my understanding of what may have happened) have documented their role. What is needed, however, is an explanation of why children (at whatever age) create reduced forms rather than adults and why this is more true (if it is) of creolizing pidgins than of natural languages. We come back to this topic below.

If there is now in some places a variety of Tok Pisin that morphophonologically characterizes children, we are informed that “as they grow older” they “revert to a more conservative norm” (Mühlhäuser 1980:57; also 1986:69, 70). We would have to conclude, therefore, that among adolescents, for example, there are fewer short forms than among younger children. No evidence for Tok Pisin has yet been provided for this reversion to an earlier form, and no evidence can be provided for Sango, although one would expect it. Not only do adolescents use short forms but adults as well. One older woman of Gbanu ethnicity, whom I found selling snuff, had come to Bangui as a married woman, yet she said very calmly, leisurely, and politely to someone asking for change for a hundred-franc note, ày.d.apè (dots separating ‘words;’ -y is one of the variants of the verb ‘to be’ that follows the subject marker à:) instead of àkè dà apè (is/are there not) ‘There isn’t any’ (i.e. ‘I don’t have any small change’). In other words, if short forms were introduced and are being used notoriously by young people, they are being learned by some adults. (I do not remember having heard in 1988, 1991, and 1992 short forms in the speech of radio announcers. I have not recently recorded samples of adult speech in Bangui.)

Creolization in a pidgin

The agents of creolization in Tok Pisin are not, however, unambiguously identified as children, for creolization seems to have begun some time ago. Children may, of course, have been some or all of the innovators, but writers do not speculate. It is about Tok Pisin presumably that the following is said: “it is well documented that stylistic diversification [which we have identified as the emergence of short forms] can set in well before creolization” (Mühlhäuser 1983:69). The only evidence we have appears to be the following. (1) Both the phonological reduction and the syntactic positioning (from sentence external to preverbal position) of bai (also babai, ba, and b’ from balmbai [Romaine 1992:173, 269]), the marker of futurity, are said to have been “well under way prior to the existence of a large number of native speakers ...” (Sankoff and Laberge 1980:208; this belief seems to be shared by Jourdan [1991:192, 194]. But Romaine argues that change (from sentence external to preverbal position) probably took place before phonological change (1992:253-254); as an option, it may have been available “for at least a century” (1992:257). (2) The reduction of word stress “first became apparent in the speech of the last generation of Tok Pisin speakers in one environment and has now become much more pronounced in the speech of adolescents and is
spreading to new environments" (Farascl 1989:137). (We return to the topic of grammatical role and phonological change below.)

Precreolized reduction can also be attested for Sango. Six examples are cited.

(1) Although one of the common forms of the noun tene ‘word, matter, affair’ among children is te (sometimes nasalized with short or long vowel), there was something at least sixty years ago that Protestant missionaries codified as <tèti> (arrows bracketing a written form), where the grave accent indicates the vowel quality, not tone. In 1952 I heard it only as tenti, where I am marking tone. Its meaning is ‘because’ and followed by nyè ‘what?’ (i.e. because of what?) it means ‘why?’ This is clearly derived from tene and the preposition ‘of.’ (I do not recall having heard or found in my recent transcriptions something like te [for tene] ti mò nyè [word of you what?] ‘What’s your problem?’ but I would expect it.) A synonym with the same syntactic behavior is ngbâŋâ tî, now reduced variably to ngbââti or ngbâti, the noun means something like ‘judgement, decision.’ The early date for the reduced form is not the only interesting thing about <tèti>. The other is that it must be a calque on one or more Ubangian languages other than the lexifier language, because this construction, while absent in the latter (if we take Lekens [1958] seriously, which we have every reason to do), is paralleled by an identical construction in Gbeya. (This is not to say that Gbeya is the source.)

(2) Although go is frequently found in transcriptions of children’s speech for gwè ‘to go,’ in the 1950s I certainly heard it in the nominalized form gongo as an alternate for gwengo.20

(3) Similarly, it was common in the 1950s to hear kwé ‘all’ reduced in a reduplicated form: kôkôkwé ‘absolutely everything.’ Today in Bangui one hears the abbreviated form also when followed by the perfective marker àwè (subject marker.finish): lô tô kôow < lô tô kwé àwè (3s.pm eat all.finish) ‘he’s eaten [it] all.’ (The abbreviation of gwè and kwé was one reason I analyzed these words in the 1950s as goè and kôé."

(4) As noted above, there is an intermediate form between kodoro (the only one cited in Lekens’ dictionary) and koro; namely, kodro or kotro. Here it needs to be said that this was available to speakers in at least the 50s. Although Lekens (1958) lists only the long form kodoro, there are some dialects in which the vowel before r is commonly deleted. (It is quite possible that several dialects contributed originally and have contributed subsequently to Sango.) Even words from other Ubangian languages experience vowel deletion: kârâko > kârko or krâko ‘peanut,’ alleged to be from Banda. (We have not yet found *kâko.)

(5) One type of reduction was not at all unusual in 1962: that was the elision of either the final vowel of one word or the first vowel of another when found juxtaposed: tôt apè < tô tô apè (cry not) ‘don’t cry; lô gwè nà gâla pè (or gâl apè) (3s.pm go to market not) ‘she didn’t go to the market.’ (See Samarín 1967:45 for more examples. It should be noted that of the 17 cited, the negative is involved 7 times; of the two conjoining vowels in 8 other cases the second is again a; in one case a is the first vowel; in one case the phrase is bâle ókó [unit-of-ten one] ‘ten.’ Furthermore, 8 examples are from radio broadcasters; two other speakers are adolescent girls who claimed that they knew no other Ubangian language than Sango; one speaker was an adult female Protestant Bible school student of Kanre ethnicity; and one an adult female of Gbanu ethnicity, about whom it was noted at the time of taping, [speech] shows influence of travel)."

(6) Some instances of the deletion of final vowels even when followed by a consonant were found in recently examined data from 1962: sârà ten[es] nà mbî apê (do speech prep. me not) ‘(she) didn’t speak to me’ (one of the two creole speakers just mentioned); âsâarâ [a] ten[es] nà keke (subj.marker.do speech prep. tree) ‘said to the tree’ (a rural Protestant woman of Manza ethnicity over 50 years of age; and other examples of ten[es] ‘to say’ come from radio broadcasts.26

It would be wise to conclude, therefore, that morphophonological reduction may begin in the pidgin, when still spoken for the most part only by adults (and, in the case of Sango in those days, by men more than women) and that children ‘take the ball and run with it.’

We will not here attempt to explain why children (and presumably those for whom the pidgin has become the primary if not first language) should be more responsible than adults for reduction. We might nonetheless mention that on the basis of what apparently has been observed in Bislama it is believed that
children regularize and streamline irregular forms and those that are not “functionally fixed” (Jourdan 1991:195). This does not seem to apply at all to Tok Pisin and certainly not to Sango.

Let us now look at what kinds of words in a pidgin experience reduction.

The stuff that is changed

Only some words undergo change, and it is curious that little has been made of this fact. In Tok Pisin we find these kinds of words:

1. Fillers when thinking of the name of a person or thing (like our ‘Whachamacallit?’)
2. Quotation frames to introduce reported speech
3. The preposition bilong (and certain phrases in which it is found)
4. Certain three-syllable demonstratives and articles
5. Certain three-syllable pronouns
6. The connective sapos ‘if’
7. Only some verbs (e.g. ‘give,’ ‘follow’)
8. ba, already mentioned
9. save as marker of habitual, not, apparently, as the verb ‘to know, to know how to’

By contrast one notes that whereas an 11 year-old boy used many shortened forms in Tok Pisin, lukautim pikinini ‘take care of child’ did not experience shortening (Romaine 1992:174). By contrast, in Sango mérangé ‘offspring [animal or human], child, kid [as an appellation for child], young person, son/daughter’ has several reduced alternants. A child can refer to a peer with this word.

The glance at the kinds of words that are reduced in Tok Pisin would suggest, as it does for Sango, that it is the most commonly used words that are more susceptible to reduction than others. Although Romaine’s corpus from research undertaken in 1986 and 1987 consists of 705,121 words (1992:136), we cannot learn anything in her work about the frequency of the vulnerable words. She mentions it only with respect to the shift from [p] to [t] as a result of increasing knowledge of English in the country, a feature of decreolization (Romaine 1992:190-191).

A concordance count of my own recent corpus of Sango has not yet been made, but all the affected words that I have come across are certainly ones whose token count in texts would be high. For example, in my 1962 corpus of non-French tokens of words (N = 33,990) in addition to bound morphemes, the latter represented by only three types, the prepositions tf ‘of’ represented 8.99% and nà ‘at, to, for, etc.’ 7.64% of the corpus. And in a microcorpus of 100 words, in which there are 32 types, the two of them together represent 22% of all tokens (Samarin 1971). It is not surprising, therefore, that the first has given rise to several reduced forms, one of which is the loss of the consonant and vowel, leaving only the tone (Samarin 1992b), and the second to à : e.g. dà tf là > dàlà là (house of 3s.prm) ‘his house; àgowè (or àgò) nà nga > àgowà à nga (subject-marker.go prep. river) ‘she went to the river.’ However, I do not mean to say that frequency alone explains the emergence of reduced forms of tf.

It stands to reason, therefore, that grammaticalized words of frequent occurrence would be more vulnerable to erosion than others. This is illustrated by bai, save, and laik in Tok Pisin. In Sango, however, kë, the verb ‘to be,’ which is also used to mark the continuative or irrealis, is at the present time no more
frequently eroded as an auxiliary than as an independent verb (Samarin 1993a): e.g. lò kè kà (3s.prn is there) 'she's over there,' lò kè ga (3s.prn is come) 'she's coming' or 'she's going to come.'

To look at some other data, let us compare the frequencies of nouns and verbs in my corpus of 1962. There were 231 nouns of non-French origin, the tokens of which come to 5,497. The frequencies ranged from 1 to 374 in a gradual manner. If we start at 183 tokens, since the preceding one is 151, we have 8 words that represent 33.12% of all noun tokens. Every one of them occurs today in some condensed form: some, like mèrenqé 'child,' in several different forms. Of verbs there were 135. We can subtract wè 'to be finished,' whose frequencies (N = 258) included the perfective marker àwè. We can also subtract kè 'to be' (N = 1,058), one of whose uses is a marker of continuative, imperfective, irrealis, future, and the like. If we start with 208 tokens, since the preceding one is 165, we find that 6 verbs represent 32.71% of the tokens (N = 4,876). Of these, bëqà 'to see' is reduced to bà, as are the other few words with falling tone; gwè 'to go,' as we have seen, to gô; tènè 'to say' to tèè (as well as other forms), and sârà 'to do, make' to sàa and sâr (not yet to *sâ)

It can be seen, therefore, that there is in Sango a correlation between frequency of occurrence and condensation.

A more thorough study of frequencies might reveal that some infrequently used words are condensed by analogy because of phonological similarity. Indeed, we may already have an example in the text discussed in the following paragraph. There we find the pronoun of reported speech, nî (e.g. 'He said he would [= I will] do it'), occurring as n with mid tone. This pronoun does not occur very frequently, and certainly much less frequently than nî, which is commonly reduced to n, at least when used with its original function as an anaphoric determinant: e.g. mbò nî 'the dog.'

It is appropriate here to give some idea of what happens in discourse. We take one text by way of example, transcribed first by a Central African but checked by myself variably at normal to half the normal speed on a Lanier VoiceWriter P-148. It is an improvised argument (pàpà) on the part of two boys, 2'27" in duration. Their ages were given as three and four, but my notes indicate that I did not believe this. They were, in any case, too young to have begun going to school. In spite of the high percentage of condensed forms -- 9.59% of 323 words in the discourse of the ‘plaintif' (N = 31) and 11.83% of 169 words in that of the ‘defendant' (N = 20; 18 if 3 questionable ones are deleted, therefore 10.05%) -- these were children who had never been to Bangui. Both were Gbeya, one lived in Bossangoa, the other in the small village of Bongam, 12 km away, where the recording was made. They were, of course, speakers of Gbeya. It should be noted that although rural and very young, their Sango is good, as far as I am able to judge it (except for a possible mistake discussed below). They use three levels of pitch correctly instead of two, which Gbeya has, and the only thing that reveals their ethnicity is the use of the Gbeya sentence-final particle ndé. (It would appear that they learned Sango along with Gbeya as first languages.) As can be seen in Table 1, there are only 10 words that were shortened, whose tokens in the whole text were 52, but 35 tokens were produced by only three words: jàâ 'topicalizer or existential particle,' ngbângàti 'because,' and tî 'of.' However, the frequency of the first two might be a function of the genre of discourse: the first (N = 13) because the particle is used as a topicalizer (e.g. 'You're the one who stole my mango'); the second (N = 9) because the boys are challenging each other with the words 'why?' (ngbângàti nàqè) and 'because.' As for nouns and verbs, it is seen that only 'word,' 'to say,' and 'to want' are found.

The 'anomalous' forms (only because I had never encountered them before) need to be commented on. How are we to explain mángô ní mbi 'my mango'? (The child certainly knew the use of tî, for it was correctly used 14 times, followed by nouns, pronouns [1s, 2s, 3s], and verbs.) Four explanations are suggested. (1) The form nî might be described as being the result of assimilation (by voicing and nasalization) of the short form tî (where the vowel is lost; see below). It should be noted that the tone is correct: the phrase has high-low-mid tones just as in mángô mbi. (2) If there were only one such token of nî, one would be inclined to describe it as a slip-of-the-tongue, an error of performance, like the Canadian announcer's saying that there would be snow "aroun' min..., midnight." This is highly unlikely given the fact that it occurs three times -- without any self-correction -- in a single sentence of 40 words in a conversational 'turn' of 62 words. (3) It might be considered an error resulting from the conflation of the preposition and the determinant nî. It should be noted, however, that the child knows the use of this word: it occurs in the full form in nà plaàs (< French place) nî 'at the place' and twice as in ngù nî 'the water.'
(Something like *mângô n’ lì mbi [mango det. of 1 s.prn] is ungrammatical; the determinant must follow the noun phrase.) Similarly, although the forward-moving tonalization of lì occurs three times, with the nouns lì ‘head’ and ngbangba ‘cheek’ in phrases meaning ‘his ...’ (e.g. lì lôô ‘his head’), they also occur with lì with 1s, 2s, and 3s pronouns (4 times).

Furthermore, this child says mângô lôô ‘his mango’ instead of either mângô tf lô or mângô tô. Instead of the tone’s moving backward, which is the usual case when the preposition is reduced, it moves forward. Did the child do this because he had already internalized a ‘rule’ that permits one to move the tone forward, as in kôl lô ‘her husband,’ where the high tone occurs with lô? Or is this form constructed on the analogy of lô tènê, which used to be a function of intonation, meaning ‘What he SAID was ...’ Today, it is just a variant of tènê preceding reported speech. I would accept the second but not the first innovation, but that is only because I can explain it better linguistically. Certainly in the exchange between the two boys there is no apparent insistence, as if the boy were saying ‘His mango.’

It should be noted in this section that as far as Sango is concerned, the size of the word is not at all a factor. Four- and three-syllable words (of which there are few in the language, possibly all borrowed from one or more Bantu languages) can be reduced: e.g. màfùtà > mààftà ‘oil, fat, grease,’ sùkûlà > sukwà and skàà ‘to wash’ but not màkùnzi ‘chief, headman.’ And, as has already been seen, monosyllabic words are also affected.

It would be wrong to leave the impression that only single words are affected by condensation, for as we saw above, in Tok Pisin sometimes two words are fused into one. In Sango fàdesô ‘now,’ tôngasô ‘thus, like this,’ tôngànà ‘when, if’ are probably phrases in origin in which só and nà are separate words; nà ndô wà (at place what?) > nándwà ‘where?’ is certainly a phrase. But other common phrases are not condensed. I have yet to find a condensation of lâkwé Qâ ‘sun,’ kwé ‘all’) ‘always,’ which is frequently repeated for emphasis, whereas other repeated words are transformed if not reduced: hjo hfo > yôôyôô ‘quickly.’

The consequences of condensation in Sango

On the basis of an analysis of a carefully checked sample of transcriptions of urban, including creolized, Sango (Samarin and Walker 1992c), we can make the following statements about the consequences of condensation, grouped according to whether they are phonological, morphological, syntactic, or stylistic.

1. Phonological.

(a) Word-initial and word-medial consonant clusters.

(b) Word-final consonants, especially nasals and r.

(c) Tone-bearing syllabified nasals.

(d) Loss of high-low and low-high tone sequences on long vowels in some words but the emergence of new glides phonologically and morphologically (mérengé > méenqé ‘child,’ dà tf ôô > dàé ôô ‘his house.’ (There never was a high-mid glide before.)

(e) Homophony, of which there was very little in the pidgin stage: tè (old word) ‘to eat,’ tè < tènê ‘to say.’

(f) A loss, or potential loss, of redundancy. On the basis of what has already been confirmed, the following is possible: à à à à à à < àà à à à à à ‘they copula at belly of house’ ‘they’re in the house.’ However, the following is documented: kôgâá tâà nà à tè nà à à <
kògàra ti àlà n' àtènè nà àlà (in-law of them determiner subject-marker say to them) 'their in-law said to them.'


(a) Zero forms of words, of which there is as yet only one example and only in this kind of construction, but quite common: mbi 0 ti lango < mbi yí ti lango (I want of sleep) 'I want to sleep.'

(b) Affixation may be developing in sequences like the following: lòò ga < lòò ke ga (3s.prn copula come) 'she's gonna come;' blìkm (the nasal carrying mid tone) < blìk (Fr. bic) ti mbi (pen of me) 'my pen;' mu ūndà < mbi ūndà (1s.prn ask) 'I ask,' which I assume will eventually become, if it has not already become, mūndà, like mìngà < mbi ūngà 'I know.'

(c) If some or all of the current ways for expressing possession are preserved, and if others emerge, this part of Sango’s grammar will be complex. In addition to the two that have just been cited, we can look again at two others: màngò lòò 'his mango,' màngò n mbi 'my mango.'

(d) It follows from the above observations that morphemic variants, some occurring as free forms and some in construction with other morphemes, are already numerous in the language. Another example might be cited: kir.à n gè < kin nà ni gè (return with determinant here) 'give it to me here.' The speaker is male, 17 years of age, and came to Bangui from Nigeria ten years before the recording was made.

3. Syntactic. Although nì, the determinant, today varies with ŋ following a noun, the time may come when the latter will be the only form used in casual speech. Should that happen, there would still be the variant in other positions: e.g. nì láà ‘that’s it.’

4. Stylistic. There is some evidence already that what might be called formal discourse uses fewer, if any, contractions. This was true of a radio broadcast in Sango by President André Kolingba in 1992. Only one contraction was heard, and that at a point where, in my opinion, he was trying to express solidarity with the nation’s citizens, many of whom were demanding political changes: tâlà ‘your.’ An analysis of different kinds of radio broadcasts by employees of Radio Bangui might reveal more stylistic correlations. (Contractions heard in 1962 were mentioned above.)

The etiology of condensation in Sango

Having found explanations for morphophonological change in creolization less than adequate or convincing, I am obliged to offer some alternative explanation. It is to be noted, first, that attempts to explain what has happened in Tok Pisin were not based on the foundation already laid in historical linguistics, even though one of the authors identifies herself as a historical linguist (Romaine 1992:vii). One might be justified in concluding, therefore, that my predecessors found nothing in other languages that would help explain Tok Pisin’s evolution. In dialects in general one would expect similar changes, if not explanations. The vernacular (as opposed to a literary standard) is, of course, what we are interested in, but whereas there has been a considerable amount of study of the distribution and use of nonstandard forms in different speech communities, the emergence of these forms has not been explained. Today why should a fourteen year-old boy say pràly for probably? And why should jus’sec have arisen for just a second?

I do not have a comprehensive theory or hypothesis that would account for the emergence of shortened forms in all languages and varieties throughout time. But before I venture a possible explanation for Sango, let us note by way of summary some things -- factual and speculative -- about this language.

(1) Morphophonological change was taking place before the language began to be used by members of a speech community as first or even primary language.
Change seems to have begun with quite ‘natural’ phonetic processes (viz. vocalic elision across word boundaries; a geolinguistic substratal or adstratal explanation has yet to be found).

Recent change, although apparently intense by comparison to what was found thirty years ago, is also largely ‘natural.’ The processes that will be used in explaining morphophonological change in Sango will be assimilation, syncope, and apocope. In tônsô ‘thus, in this way’ (where ‘n’ marks a nasalized vowel) can be traced from tôṅgásô via forms where the homorganic prenasalized voiced stop is replaced by a velar nasal; similarly there are a number of intermediate forms between tôṅgânhà and twà ‘if, when.’ The loss of one of two identical vowels across word boundaries (i.e., elision) occurs frequently, but even some unlike vowels suffer deletion: e.g., tf àlà > tàlà (of 2p.pron/3p.pron) ‘yours/their.’ However, not all elisions are thus explained. The final vowel of apè, the negative, for example, is frequently dropped before the sentence particle mà, which, at the same time is abbreviated: ... apm, the nasal carrying low tone. It is possibly because of this pattern that ap occurs before other words with initial nasals, even when these occur in what would appear to be a new sentence, sometimes separated by slight pause. In one improvised argument of 5'15" duration, containing 65 tokens of the negative, there were 9 instances of this pattern. Because the forms ap and p occurred four times without a following nasal (once before a non-nasal consonant, twice before vowels, and one at the end of the speaker’s turn), we must observe here the liberation of the negative from phonological constraints, leading to the incipient generalization of p as one of its forms. The first vowel of the negative was dropped (revealed by the absence of its mid tone) when preceded by a word ending with the same vowel: e.g., hîngà pè (know not) ‘don’t know.’ While not accounting for shortened forms as such, assimilation explains a large number of variants. The most striking example, in this case of discontinuous assimilation, is one I heard from a young woman in her 20s, because it was the only time I ever heard it in this very common word: ànîng.apà for ànîng apè (subject marker.be long not) ‘it won’t be long.’ Once a word has changed because of phonological constraints, it can be used freely. Thus, we find in the following sentence from the speech of a 12 year-old Muslim girl the word ‘all’ followed by a consonant: zô ko nà tf lò (person all [copula deleted, as permitted by the grammar] prep, of 3s.prn) ‘everyone has her thing (i.e., belonging).’

Change is sometimes ‘explained by’ or at least correlated with phonological (and possibly syntactic) constraints. For example, tf is not reduced when it occurs in a highlighted phrase: e.g., Ti mbi, mbi ké àwè (of me, I refuse finish) ‘As for me, I have refused.’ It would be advisable to study the use and distribution of various degrees of amplitude (stress) in Sango discourse with respect to the emergence of condensed forms. It is my impression that there is something different about sentence stress in the speech of urban young people from what it was years ago. In Tok Pisin also one might profitably study this topic. In one study of 1500 words of tapesrecorded discourse from each of 30 speakers it was found that dispela and prepositions (not named) were almost never stressed in both first- and second-language speakers (Faraclas 1989:133).

Change affects frequently used words more often than others. I would base my explanation for change in Sango on two factors that would probably be accepted by all. (1) Speech production, while guided by the matching of sound to meaning in a patterned way, is never perfect: the bull’s-eye is not always hit. (2) Speech is always constrained by some social model. Where no clear model exists, there will be greater variation. I would like to believe that forty years ago, when I learned Sango and when this language, in the absence of competence in French, served as a symbol of upward mobility, adults (who made up the majority of speakers) were careful to talk like others. (Protestants, for example, spoke and wrote -- as they continue to do -- a variety of Sango that was different in ways commonly recognized by many people: e.g., using ànj instead of j, or the less frequent q, for the first person plural pronoun.) In recent years children, who are perhaps in all speech communities the major innovators, either do not have a clear model or, as I prefer to believe, do not care about the adult model (or models) they are exposed to. (This view seems to be shared by one other. See Foley 1988:190-181. See also Romaine [1992:26], who says that linguistic diversity may result from the absence of "pressures towards convergence.") On the basis of my research in 1962 and three subsequent research tours, totalling 30 weeks, in the Central African Republic I would doubt the view that the children of Bangui have an adult Sango model that is less clear than it once was. Among themselves, of course, it is different. If, or since, that is true, we would like very much to know how and why the model broke down.)
The children of Bangui make up a subculture very different from that of rural children. It is my impression that they are, for example, much less supervised by adults than are rural children: most of them go to school; although the girls have some chores at home, they have more free time than rural children do, and the boys are largely free to do what they want. Moreover, urban children stay up later into the night and roam around at night much more than rural children do, a fact attested by me over the years. One piece of evidence of the change in culture is the fact that very few children in Bangui, as I learned in 1992, heard at home traditional folktales in either an ethnic language or Sango. Urban children are also less likely to use the 2p and 3p. pronoun ali in deferential speech. This happens even in Gbeya when Bangui children try to speak to me in their ethnic language. It is probably true for children that in Sango almost anything goes. I would suspect that children would accept apà for ape ‘not’ in the clause cited above, and, since the two girls in conversation were well dressed, might even imitate them. Adults, on the other hand, I would predict, would consider this a mistake: one made by a child or by someone still learning Sango.

Contemporary, but especially urban, Sango is so different from what it was forty years ago that it would be difficult to say what a grammar of the language of children and young people would be like, apart from the fact that it would be a variable one. (What kind of pedagogical grammar would one write?) But to what extent can we say that it is a language – in the sense that one can make mistakes? Indeed, in the variation that we find, how can we identify mistakes? (A mistake, of course, would be something that some group of speakers would consider wrong.) In 1988 Central Africans told me, when I informed a group at a lecture at the University of Bangui of my discovery of the tonalization of ti ‘of,’ that this was imperfect children’s speech; they hadn’t yet learned the language. The members of the audience simply did not realize how frequent this reduction was and how widespread it was through the young population. (Assuming, most reasonably, that they had heard children speak in this way, we explain their reaction on the basis of subconscious editing: they had simply filled in what they would have said themselves.) If the tonalization of the preposition is not a mistake, what does one do with the following? Taperecordings reveal that the prefixal subject marker is sometimes used with high tone instead of the usual low tone and that it is omitted where it would appear to be obligatory (by the structure of earlier Sango). 30

Many if not all ‘mistakes’ have some cause. The two just cited, for example, might have an adstratal explanation. Speakers of the lexical-base language -- Yakoma and its closely related dialects -- may have introduced models that others adopted. The speakers may have been adults competent in Yakoma, for example, who spoke a variety of Sango influenced by their ethnic language, or by children who, because they spoke the ethnic language poorly, imitated what they did not really understand. For the present, however, there is no evidence to support the hypothesis. Indeed, an examination of discourse in Sango by children of Yakoma ethnicity does not reveal anything that would differentiate them from other children. 31 And there is at present no study of the place of the Yakoma, past or present, in the life of Bangui that would help explain their sociolinguistic importance as models. Yakoma-like changes, if that is what they are, must surely have arisen before the ascendancy of André Kolingba, of Yakoma ethnicity, to the presidency.

‘Doing the right thing’ when many people are ‘doing their own thing’ leads to a variable language. In fact, some results are so variable and unmotivated that they appear to be random. This is the case with the use of fade and andè to mark futurity in the speech of educated inhabitants of Bangui. Although both words are Yakoma adverbs meaning ‘right away’ and ‘soon’ that have been in Sango for a long time, young people used them in translating sentences from French into Sango almost as if they did not know how to use them, as if sprinkling them in a sentence was all that was needed. (A report was made in Samarin 1992a.)

It has been said that “From a very early age children monitor and correct their own and others’ linguistic production ‘errors’” (Romaine 1984:9). Does that apply to either Tok Pisin or Sango? For the latter I have no information, neither data based on my own observations or on what others may have told me. However, given the high degree of variation found in children’s Sango today, the implication appears to be that the statement does not apply.

In concluding this section we have to observe that the social contexts of creolization have to be considered in explaining any of its manifestations or consequences. Urbanization, we have observed, is undoubtedly one factor. Of course, invoking urbanization does not immediately give us an explanation. As Kulick (1992) correctly reminds us, such concepts do not provide us with the immediate causes for language
behavior. In his case -- explaining shift to Tok Pisin in a remote village -- he argues for an ideology: something that the people believe in and how that belief is related to Tok Pisin. There are, I should imagine, some socio-psychological or cultural dimensions to creolization that we ought to try to identify. This search is motivated by what we already know about language changes in an urban setting. (REFERENCES)

**Conclusion**

As a result of this étude, we might be led to the following observations about the 'creolization' of pidgin Sango morphophonology.

1. Several factors, not all of which have been identified and argued, appear to be working in changing the phonological shapes of words.

2. Many if not all changes start with certain words and then spread by something we might want to call analogy. As evidence we can refer to the condensation, cited above, of the infrequently used pronoun ni to n (with mid tone) on the pattern, I should think, of the very frequently used determinant ni, which becomes n (with high tone).

3. Change occurs with frequently used words first.

4. Change begins before (and possibly even long before) there is a large number of native speakers so distributed geographically that they have access to each other. Therefore, the linguistic differentiation between a pidgin and a creole is not one of drastic but gradual change. This statement with respect to Sango only supports what others have hypothesized or demonstrated (for example, Arends 1989, according to Kouwenberg 1991:130; Alleyne 1980:12).

5. Children innovate and create a variable language, because they are indifferent for socio-psychological reasons not yet determined to the adult model (or models).

6. Because children growing up retain some of their innovations, because they learn equivalent forms from their seniors, and because their seniors may adopt some of the forms used by the younger generation, an extremely variable language emerges.

7. Although the reduction of falling and rising tones on long vowels to a single vowel with a level tone (e.g. bàà > bà 'to see,' làà > là existential particle) might be seen as an instance of phonological drift, it should be observed that glides are being reintroduced by the reconstruction of the preposition tì 'of.'

**Acknowledgements**

Research that led to the writing of this paper was funded in part by the Social Sciences and Humanities Research Council (Canada), whose help is gratefully acknowledged. The Department of Anthropology of the University of Toronto provided research space and other facilities. Grace Brethren Foreign Missions (known as Mission Evangélique des Frères in the Central African Republic) also helped in many ways, not the least of which was allowing me to live at the guest house in Bangui. Without the authorization of the Central African Republic the research, of course, would not have been possible. The help and kindness of Mr. Germain Aoungbi-Kolzo, Directeur du Développement de la Recherche Scientifique et Technique, is much appreciated. For having commented on an earlier version of this paper, manifesting their virtues in collegial scholarship, Jack Chambers and Don Kulick are also thanked.
Notes

1. This work by Labov is being cited in its most recent form. Singler (1990:1) says that it was written in 1971 and had "never before been published." Romaine, on the other hand, gives its original date as 1970 and a publication date of 1977, for which see References. For Romaine 1992:172-174 one should consult also Romaine 1987 and 1988:138-140, where less qualified statements are made.

2. In Samarin 1966 Sango was called a creole not for any linguistic reason, but only because of field work in 1962 I had come to the conclusion that there was a considerable number of people for whom it had become their first language. Citations in Sango do not distinguish between the mid and low front and back vowels. Although phonemic, they appear randomly, as far as can be determined, in many pairs of words: e.g. pòko 'back.' The three tones are written in the commonly accepted manner.

3. For one study see Samarin and Walker 1992. It should be noted, however, that any one of the above-named terms used to describe the process is not in itself adequate. To understand the process we must first arrive at a description of its characteristics. In mààftà from màfùtà 'oil,' for example, there is compensatory lengthening. In kéeékè two consonants are dropped from kékèrékè 'tomorrow' without going all the way to something like *kékè. And a whole new range of canonical forms for words now exists, with syllable-final consonants, consonant clusters, etc.

4. The word for 'now' appears to be an innovation in Sango, consisting of the Yakoma adverb fàde 'soon' and one of the several homophonous words só. No convincing explanation for its origin has yet been determined. Because Lekens (1958) cites several variants of the indigenous word in which d is replaced by l or r, and because the latter two consonants are known to be lost in at least Ngabaka-Gbaya languages (Samarin 1959), it reasonable to suppose that the d-less forms like fàeso are derived from a word like fàle in a dialect of Ngbandi-Yakoma.

5. The adverb 'virtually' adds the notion of quantity, as if to say that a pidgin may have some reduction but not enough for native speakers.

6. This categorical statement is preceded on the same page by a qualified one, Labov 1970 having been cited: "This could be true for several reasons. One is simply that a pidgin is a second language for its speakers." Later (p. 176), Romaine acknowledges that "These results suggest that more than fluency is involved." For a statement similar to the one cited above see Mühlhäusler (1986:70): "... such changes promote optimization of production ... such stylistic variants can only emerge once the majority of pidgin speakers are reasonably proficient."

7. Romaine (1992:172) cites Mühlhäusler (1986:177) as having characterized an expanded pidgin (or expansion of a pidgin): "a steady increase in phonological distinctions, the use of what were formerly free variants for stylistic purposes, and the emergence of phonological rules which separate underlying from surface forms." He does not link any of these with fast speech.

8. Elsewhere [1992:172] she seems to take responsibility for this assertion: viz. "morphophonological condensation, which gives rise to short forms in allegro [i.e. fast], casual speech ..."

9. In citing this anecdote Romaine makes this one "speaker," who in Labov's opinion was "much more fluent" than someone whom he had heard speak "halting Hawaiian pidgin" (1990:15) into "fluent speakers" (1992:172). In other words, there was only one person and he was not characterized by Labov as fluent in Tok Pisin, only comparably more fluent. Mühlhäusler (1986:151) also cites Labov as having "established some significant differences between fluent Buang and fluent Tok Pisin."

10. One wonders if this person is not the same whom Sankoff (1980:44) refers to in discussing the use of Tok Pisin as the appropriate language for a person with power when he is in contexts relating to the broader colonial administration.
11. It should be noted that in an attempt to devise a model for speech production for the study of slips-of-the-tongue by simulating speech with a computer, it was found that speaking rate was an important determiner of the number of errors, but at fast rates most of the errors created nonwords (Dell and Reich 1980).

12. Another attempt to correlate speech tempo and reduction might also be discredited. Although it is said that "in rapid speech" Kwaio, spoken on central Malaita, "commonly uses an abbreviated na-ku" as an "I" subject-referencing pronoun, it is also said that the neighboring Kwara'e language has bound the cognate na and ku into a single form with characteristic metathesis: i.e., nauk (Keesing 1991: 223, 227 fn 14). It would seem that if rapid speech explained some forms in Kwaio, it would have to explain all forms in Kwara'e, suggesting that the speakers of the latter talked faster than speakers of the former, something which I am not inclined to accept.

13. An attempt was made in 1992 in Bangui to elicit from subjects imagined situations (scenettes they were called) where people would talk fast in Sango. Although three different techniques were tried, the assistant failed to get anything interesting. In this case I suspect that it was because the assistant, although educated in France and the United States, proved incompetent.

14. Gbeya, as pronounced by the people living in and around Bossangoa, around 350 km north of Bangui, is a variant of the more widely known name, Gbaya. It belongs to Joseph Greenberg's Eastern group of Adamawa-Eastern or Ubangian, as known among French linguists. The family of languages is widely distributed in the western part of the Central African Republic. On the languages of the Central African Republic see Samarin 1971b and Boyd 1989.

15. A colleague reminded me of the following words that I had read about two decades ago: "As far as we can see, there are no single-style speakers. ... every speaker we have encountered shows a shift of some linguistic variables as the social context and topic change" (Labov 1972:180). It has been suggested that because "we" has no plural antecedent, it can be considered an editorial reference to the author and that the passage can be rendered something as follows: "As far as I can see ... every speaker I have encountered." Perhaps the persons referred to were speakers only of English, for it has been said that Labov has worked on no other language than English (Calvet 1992).

16. Although a field worker should be ready, as in the anecdote cited, to undertake johnny-on-the-spot field linguistics, it may have unexpected results. On another occasion I asked a woman who was wheeling a barrow full of firewood down the road in Bangui (after greeting her, of course) if it was keke-wā or keke tf wā (wood [of] fire) 'firewood' that she was taking home. Probably having quarreled with others in gathering the small mango branches and having the prospect of walking several kilometers with her burden, she wasn't interested at all in telling me. In a querulous manner she responded with something like "I don't know!" meaning "I don't care!" A sociolinguistic study was done of attitudes towards phrases with and without tf in 1992. The data have not yet been analyzed.

17. Muhlhausler seems to share this view or have a similar one. He says that one source for what he calls "diversification" in a pidgin is backsliding, defined by him as "recourse to developmentally earlier stages" of the language (REFERENCE).

18. In Samarin (1971a:126fn9) I give credit to H. Al Gleason, Jr., my colleague at the Hartford Seminary Foundation, with whom sometime in 1967-1968 I discussed the paper that I was working on, in connection with characterizing pidginization as reduction in grammar and use.

19. Apparently shared by Romaine [1992:175], although I am puzzled by her saying that "short forms would expand in conformity with the norms of adult speech." Expansion was elsewhere characterized in the language with respect to reduction; here it seems to suggest the opposite.

20. (Here and with the following example, I consider it unlikely that the sequence CwV, or perhaps only Cwe, is actually an instance of pidginization, the simplification of the phonology of the lexifier language, on the grounds that some Ubangian languages, like Gbaya, do not have words with the sequence CW.)
21. Pitch glides are analyzed as occurring on two vowels. Phonetically, they appear to occur on a long vowel which is about twice as long as a short one; it seems to be that to me, and I familiar with phonemic short and long vowels in Gbeya.

22. There may be a lot more evidence of reduction in my tape recordings of 1962. They will have to be monitored with this matter in mind. The transcriptions, made in 1962-1963 by a person who was educated only in Sango with the Protestant orthography, therefore probably influenced by what he thought was the correct way to write Sango, undoubtedly paint over a great deal of the relevant data. However, all that was used in the grammar of 1967 was checked by myself.

23. (It has come to be used as a pronoun for inanimate objects, of which there was originally none, possibly under the influence of French. Another use of this word, for the first time encountered in this text, is discussed below.)

24. This is the kind of place where one might profitably look for a stylistic (although possibly unconscious) use of more or less contracted forms. However, I would have predicted, given a choice, of more contractions in the speech of the plaintiff.

25. It was common in recorded disputes for speakers to change from second person address to third person reference, as if addressing some other person in the audience. Because disputes are commonly witnessed or mediated by others, it is understandable that there could be this pronoun shifting. However, it was never accompanied by eye contact with me or any of the bystanders, evidence, I should think, that they were not dramatizing the event.

26. The argument was taped in 1991 at a Koranic school in Bangui in the quarter known as Sénégalais. The girls arguing were 12 and 11 years old (plaintiff and defendant respectively). The mediator was also a girl of 11. The defendant claimed to know what she called 'Sénégalais,' possibly Wolof or Fulani, the plaintiff a little Arabic. It is interesting that although this latter girl's Sango, with respect to condensation, is like that of other Bangui children, she told me that all her friends were 'Arab,' presumably meaning Muslim. Here, as in the argument cited above, the defendant used fewer contractions: she used the full form 76.19% of the time, the plaintiff 64.10% of the time. The role of mediator, those interested in field methods might like to know, was always introduced by Central Africans themselves. Sometimes one of the two participants would select the third party before the argument began, sometimes, when this had not been done, a third person would simply step forward. Frequently also when I was recording insults (attaques, they are called), a third party would intervene with àwè (finished) 'That's enough!' It was clear that the purpose of intervening was to prevent real trouble.

27. But we must not be categorical in saying that this phenomenon is a characteristic of urban or creolized Sango. The adult Kanre woman whose speech in 1962 was cited above said màfàtà instead of màfùtà for 'oil.' In Sango, as in Gbeya, and possibly other Ubangian languages, there are many words with identical vowels: e.g., in Gbeya, tänä 'turtle.'

28. Because of these words -- "in the course of acquisition, children go far beyond any models available to them" -- we might believe that Bickerton (1992:313) concurs, noting however what follows in the same sentence: "whereas adults do not and presumably cannot."

29. There is no contradiction here about models. It is possible that there is a clearer model in the adult or older population than there is among the younger population, closer in age to those whose habits are being formed. It should be understood that in Bangui and all of the Central African Republic, if not throughout Africa, the care of children and their early social behavior are quite different from that with which we are familiar in the west. It is empirically rather easy to document variation in a speech community, but much more difficult to study and describe a linguistic model in a community as dynamic and changing as Bangui's.

30. Marcel Diki-Kidiri, the Central African Linguist, claims (1977) that changing the tone from low to high on the subject marker makes the verb conditional, resulting in something equivalent to a clause with initial tôngànà: e.g. bàba tì mbi ága ... (father of 1s.prm subj.marker.come) 'if my father should come ...'
have never heard it, and when he told another Central African studying linguistics in France, the latter also
said that she had never heard it. Since tone is used in the morphology of the verb in Yakoma and related
dialects, it is understandable that some speakers might introduce this use into Sango.

31. Two studies undertaken in 1992 might shed some light on this topic. One is a comparative study
of children’s knowledge of ethnic languages among Banda, Gbeya, and Yakoma. The other is on the
ethnicity of one’s best friend.
Table 1. Contractions in a Dispute

<table>
<thead>
<tr>
<th>Word</th>
<th>Speaker A</th>
<th>Speaker B</th>
</tr>
</thead>
<tbody>
<tr>
<td>làà existential particle</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>tene 'word' noun</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>tènè 'say' verb</td>
<td>1</td>
<td>1 (?)</td>
</tr>
<tr>
<td>ni determinant</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>mbéni 'some, a' adjunctive</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>tì 'of' preposition</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>nà 'at' preposition</td>
<td>1</td>
<td>1 (?)</td>
</tr>
<tr>
<td>ni pronoun</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>mbi 'I' pronoun</td>
<td>0</td>
<td>1 (?)</td>
</tr>
<tr>
<td>tòngànà 'when' connective</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>ngbàngàtì 'because' connective</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>TOTALS</td>
<td>31</td>
<td>17 (20?)</td>
</tr>
</tbody>
</table>