PIDGINIZATION AND CREOLIZATION OF LANGUAGES

PROCEEDINGS OF A CONFERENCE HELD AT THE UNIVERSITY OF THE WEST INDIES MONA, JAMAICA, APRIL 1968

EDITED BY
DELL HYMES
Professor of Anthropology, University of Pennsylvania

CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE
LONDON NEW YORK NEW ROCHELLE
MELBOURNE SYDNEY
There was some question at one time as to whether or not pidgins were languages. This was an era when decisions were being made about the use of pidgins in business, government, and education in former colonial territories throughout the world. There was always a superfluity of languages, and in the keen competition that characterizes multilingual areas, the pidgins were outclassed; they hardly 'made the leagues' at all. In the contest with the natural languages, both indigenous and imported, they showed up very poorly, having no supporters with mother-tongue loyalties. Of course, the arguments given – when they were not simply disregarded – were mainly linguistic: pidgins have vocabularies and grammars so restricted as to be inadequate for communication. Seldom did anyone observe that they actually were in widespread and successful use. Without them communication would, in many areas, be severely hampered or almost impossible. But still, they were not languages.

The usefulness of pidgins is clear, but there seems to be very little more support for them now than there was ten or twenty years ago. This is certainly true of the Central African Republic which has been independent since 1960. Four months of close observations in 1967 (June–September), three of which were spent in the capital, Bangui, convinced me that although Sango was still the langue nationale (French being langue officielle) nothing official was being done to develop its use. There was a government-appointed Sango committee, but it has met only a few times since its establishment in 1962, and very little has come of the meetings. It made no recommendations for the use of the several thousand dollars earmarked by Unesco for developing Sango during the fiscal year 1966–7 and linguistic studies were sacrificed to musical instruments in the subsequent requests from the C.A.R. delegation. Some of the Central African members of the committee, for there are some European members as well, are mildly nationalistic about Sango, attributing to it value as a symbol of independence. The last meeting of the committee in August 1967 favored the adoption of a standard orthography and spelling before further work, meaning a grammar and dictionary (independent of the American publications), be taken up. Since then the chairman of the committee, whose official position was Directeur de l'Enseignement, and the Minister of Interior were removed, the President retaining this ministry's
portfolio himself. In any case, the committee could only act in an advisory capacity; decisions had to be made by the President. (Sango will be further characterized below.)

The situation is different now, but only partly. What makes the difference is that linguists have taken up the study of pidgins with the seriousness that hitherto characterized only the investigation of natural languages. The increase in the number and variety of publications on this subject is witness to this fact. (Not all of this has had scientific motivations, of course. The U.S. Office of Education, which has subsidized much of the work on pidgins, is concerned with the preparation of language teaching material. It obviously has nothing to say to the foreign countries about the use of pidgins and must rely on informed observers for information about the usefulness, for Americans, of this or that pidgin.)

**Distinctiveness of pidgins**

But have we proven that pidgins are indeed languages? No one has addressed himself to this task with greater vigor than R. A. Hall, Jr. *Hands Off Pidgin English!* (1955, and the exclamation point is part of the title) is not just a grammar of Melanesian Pidgin; it is also a brief for the social and linguistic legitimacy of the language. His *Pidgin and Creole Languages* (1966) does not alter his stand. It is one thing to say, however, that pidgins deserve linguistic investigation or that they manifest the kinds of structures that we find in natural languages and another thing to say that they are (real) languages in any very precise sense. Linguists clearly should investigate various types of semi-languages for the information they can give us about the place of languages in human communication.

It is clear that pidgins are different from natural languages in several significant respects. Even Hall recognizes the fact that the life history of pidgins is different (Hall 1962). Historically, that is, genetically, they may be even more different than he is prepared to admit, if relexification is accepted as the best hypothesis (Whinnom 1965). In Stewart's scheme, creoles and pidgins are also distantly removed from standard languages, because a pidgin, for example, has no 'codified set of grammatical and lexical norms which are formally accepted and learned by the language's user', no 'existing community of speakers', and lacks 'homogeneity' (its 'basic lexicon and basic grammatical structure [are not] both from the same pre-stages of the language') (1962).

Many other characteristics of pidgins have been mentioned, but they are as varied as the two kinds of observations referred to in the preceding paragraph. There has been little concern with distinguishing between superficial features and defining characteristics. This is the fundamental problem in taxonomy. Although most people can recognize a cow when they see one, they could not
Salient and substantive pidginization

tell what distinguishes a cow from a goat, and it is, in fact, very difficult to define cow without appeal to characteristics which are not immediately apparent, e.g. the dentition. In pidgins there may be a striking amount of reduplication or iteration, but it is doubtful, as far as I am concerned, that reduplication is a defining characteristic of these languages. This morphological characteristic is probably due to the source languages. In itself it helps us to recognize most pidgins, but not to distinguish them from other kinds of languages. Reduplication is therefore a salient feature of pidgins, not a substantive one.

Salient pidginization describes the striking concomitant changes that natural language $A$ ($A_n$) experiences in the process of becoming a pidgin ($A_p$). One of these is, or may be, a very high rate of borrowing - so high that a pidgin is sometimes automatically called a mixed language. (Many have used 'creole' with this meaning.) Substantive pidginization, on the other hand, is the set of changes that characterizes all pidgins. One of these, perhaps the most important one, is simplification. For example, the total lexical inventories of pidgins are known to be impoverished by comparison with those of natural languages.

If we can agree on what constitutes substantive pidginization, we will have agreed on what is a pidgin. We will also have agreed on whether or not pidgins are languages. For example, does Sango's restricted vocabulary make it a non-language? Haugen thought so (Samarin 1966a:209). There may have been some misunderstanding, however, about the total morphemic inventory of Sango. If the figure of 489 is given for 'Sango' words (Samarin 1966a:188), as opposed to French borrowings, this is not to ignore the many French words that have become fully naturalized - *merdée* 'to bother', for example. Taber's dictionary (1965) lists 186 such borrowings: 123 nouns, 40 verbs, 9 adjunctives, and 14 connectives and particles. Still, the total is somewhere between 700 and 1,000 morpheme-words. Although this is quite low when compared to the inventory of English or even Zulu, it is not so striking when compared with Caucasian languages (e.g. Kuipers 1960). (And it is of the same order of magnitude as the number of fundamental word-forming elements that Swadesh has found characteristic of all languages (including Mayan and Aztec) not enlarged with an exogenously-derived learned vocabulary (the world languages of modern science such as English and Russian especially).) This is to point out that it is not simply the size of a pidgin's inventory that may be significant, but its character.

The nature of pidgins

In trying to determine what kinds of languages pidgins are we can take two avenues of investigation. We can first proceed by examining the assertions that have been made about languages in general, sometimes without any
thought given to pidgins, to determine the measure of their applicability to pidgins. We can also proceed in the other direction, by comparing pidgins with natural languages.

Let us look first at four assertions about natural languages. Two have to do with language use and two with characteristics of language systems.

Chomsky believes, along with Humboldt, that language is primarily a ‘means of thought and self-expression’ (1966:21). This being true, ‘The purely practical use of language is characteristic of no real human language, but only of invented parasitic systems’ (22), adding in a footnote, ‘For example, the lingua franca of the Mediterranean coast’, probably referring to Sabir (for which, see Whinnom 1965). A pidgin is therefore a parasitic form of speech whose primary function is not the expression of thought but that of practical interchange. (The paraphrase is my own.) Chomsky is led to this assertion by his commitment to the Cartesian principle, so named by him, that language is a human species-specific endowment that liberates man from contextual restraints and stimuli and permits him to freely encode his own experience. He does not suggest how thought and self-expression are to be measured, and he surely could not deprive my Central African assistant of these gifts when he and I discuss cultural relativity and the importance of saying ‘Thank you’ in Western society! Still, we know what Chomsky is driving at.

There are two assertions by Joos that apply to pidgins, only the second of which I will consider in any length. First, ‘... all languages are as complex as is biologically tolerable, and hence all are equally complex ... It is to the advantage of every community to make its language as complex as biologically possible ...’ (1964:191). This is patently inapplicable to pidgins, for they are far less complex than normal languages. It might, in any case, be more appropriate to say, as suggested by Gleason (in conversation), that complexity in language, whose function is the resolution of conflicts between various functions of language, is limited only by the ability of the pattern-setting speaker to cope with it.

The second of Joos’s claims has to do with *hapax legomena*. He believes that the number of types in any natural-language text long enough to deserve the title ‘text’, say 500 words long or more, that occur only once ‘always falls between 46% and 48%’ (personal communication, but cf. Joos 1936 and 1937–8 and Hanley 1937). The corollary of this assertion is that figures that diverge greatly from these would belong to ‘pseudo-texts’ or ‘pseudo-languages’. In James Joyce’s *Ulysses* the *hapax legomena* are 50–5% of its vocabulary and one Esperanto text of about 20,000 tokens had a figure of 63%. But a Basic English text of a size similar to the Esperanto one had 38%. Joos concludes: ‘Neither Basic English nor Esperanto is a language.’

When applied to other languages for the present study the test produced
Salient and substantive pidginization

ambiguous results. Only in the case of Sango, where 31.5% of the types are hapax legomena, is it validated. One would have expected similar results for Police Motu, another pidginized lingua franca, but the percentage is 42.94%. The figure for Jamaican Creole is as high as one expects for natural languages (46.48%), but this may only show that the text is closer to English than other varieties of JC are. On the other hand, the figures for Ngbandi (32.09%), Gheya (40.89%), and Cashinawa, a language of Peru (34.21%), are lower than one expects for natural languages. The skewing factor here may be in the size rather than in the kind of texts that were obtained, but since the Ngbandi text is a portion of the Bible translated by a missionary, we may have to deal with inadvertent pidginization. The figure is low in the highly inflected Cashinawa language perhaps because all morphemes are lumped together.

In any case, Gleason suggests (in conversation) that such gross calculations are not going to prove as diagnostic as some other kinds. It may be just as important, if not more important, to see the curve of the percentages and their dispersion. The hypothesis put forth by him is that a pidgin will have a sharper drop in a text of a given size than a natural language. [That is, an element is not only more likely to recur, but also to do so in a shorter interval.] He experimented with Kâte, a natural language from New Guinea. Five passages of 200 words each from five stories in Wowose Tikhata Buk I and one passage of 600 words from Harijcek Ere Binay were counted. The percentage of hapax legomena falls as the sample gets larger and the dispersion appears to become less. Nothing is said about how this compares with a pidgin language.

Only a few comprehensive statements of the above-mentioned kind are necessary to illustrate both their value and their limitation. On the one hand, they set off pidgins in a class of languages different from natural ones. On the other hand, they do not contribute to our understanding of the pidgins themselves. This remaining class of unnatural languages is really no class at all. It is simply comprised of what is left over after other forms of speech have passed the test of normality. Since this is true, the comprehensive statements (if not the ones referred to, then others) are weaker than they ought to be. The inclusion of the term 'normal' is only a device to save the statements from invalidation. Implied is the following: 'If one finds a language where this generalization does not apply, it is probably not normal.'

The term 'normal' occurs in the fourth statement considered in this discussion. Hymes, in writing about the varieties of language that humans use, asserts that 'No normal person, and no normal community, is limited in repertoire to a single variety of code [implying that a code could have more than one variety], to an unchanging monotony which would preclude the possibility of indicating respect, insolence, mock-seriousness, humor, role-distance, etc., by switching from one code-variety to another' (1967:9). Does this apply to pidgins? No, since Hymes is talking, not about particular kinds
of linguistic systems, but about language use. This is to say that a speaker of a pidgin, as a normal human being in a normal society, can be expected to have more than one code-variety for different uses. The pidgin, on the other hand, is not normal, and when a person is speaking a pidgin he is limited to the use of a code with but one level or style or key or register, to cite some terms used for this aspect of the organization of language. (One might speak here of a ‘monostratic’, ‘monoclaval’ or ‘monotonic’ code.) In other words, he does not have the rich variety of language styles from which to choose whatever is appropriate to the context, situation, or person (or people) to whom he is talking. (What happens to pidgin — or for that matter, an artificial language — in a normal society where a plurality of functions is universally characteristic of language is, of course, another matter. Is it inevitable, for example, that people who are limited to a pidgin somehow contribute to the evolution of a language that will respond to socially significant differences in human behavior? And even if this is inevitable, there are many who would like to know what the behavioral and psychological consequences are of being restricted to a pidgin or other type of ‘monoclaval’ language.)

The crucial term in the immediate discussion is code. We mean, of course, a clearly defined and integrated array of linguistic features, opposed to another similar array. Precisely what is needed to produce a ‘code-variety’ is unclear. It is probably something more than a single feature. Thus, it would seem ill-advised to suggest two code-varieties for Sango just because of the use of the second/third person plural pronoun (ála) in politely addressing another individual or referring to him. Deference to an elder is not indicated in any other ways in the language itself. (The use of babá and mamá, with the obvious meanings, is not really parallel.) In fact, it would be possible to describe ála as being the morphemic realizate of the semons of person and number/age-with-respect-to-speaker. (Politeness is an attitude, not necessarily a part of the semantic structure of language.) It should further be noted about ála that not all Sango-speaking people use it in the way described here. It is the Gbaya-Ngbaka-Manza who use ála deferentially, with interference from their own languages (cf. Gbeya wi ‘second person plural’ and wa ‘third person plural’). It is very significant that although the President is called by many lofty titles in Sango and French on the Sango radio broadcasts, he is never referred to by ála.7

In summary, the pidgin-speaking community is not normal from a socio-linguistic point of view. Neither is the language normal. A pidgin is a language, but a different kind of language.

**Diagnostic tool for normality**

What is needed is a clearer understanding of what linguistic abnormalcy is,
More specifically, how can we best describe pidginization? (The use of this process noun is deliberate. I think that pidginization is a common phenomenon in human language and that pidgins are only special cases of it. The study of both will obviously go on together.)

There might be considerable value, for example, in the establishment of a recognition procedure for pidgins (or pidginization). On the basis of a certain amount of connected discourse which is subjected to a particular kind of analysis, we would hope to determine, with a reasonable margin of error, whether or not the text was normal. There should be no a priori reason why texts from marginal languages (in Stewart's term) should be excluded. The processing of a great variety of data would have the value of breaking us of some of our prejudices and expanding our knowledge of what constitutes pidginization.

Another value of a pidgin-diagnostic tool would be to free us at least in part from dependence on accidents of historical information. We are at the moment concerned with contemporary pidgins and creoles, cases known today to be such. This is understandable, for there is still a great deal to learn about them. For example, I should very much like to know more about the origin of Sango; the amount of multilingualism in the Ubangi basin, the human relationships that existed between the speakers of vernacular Sango and other people of the area, the identification of the carriers of Sango and the channels that characterized its use, the relative importance of African (both indigenous and foreign — e.g. from Senegal, Congo, etc.) and European (French, Portuguese, Greek) carriers of the language, etc. But the time ought to come when other languages will be examined in the light of these cases. This may not be easy, for if pidgins survive, they apparently adapt themselves to the needs of their speakers; they become naturalized, although not necessarily to the same degree. If we examine a language at a very great time after its pidginized era, we might have to be less dogmatic about its past than we would like to be. (See Southworth's study of Indo-Aryan in this volume.) I therefore agree with Le Page when he writes that 'Many of the world's languages have probably undergone some degree of creolization at one time or another; by studying what is happening under our noses at the present day we should get a much better idea of what has happened... in the past' (1966:vii). Investigators should be encouraged to suggest some possibilities. There has perhaps been too much timidity. How widely is it known among pidginists, for example, that 'Oceanic Melanesian represents a pidginized form of Austronesian grafted onto basically non-Austronesian languages' (Capell 1966:537)?

If a new approach to pidgin studies can free linguists from historical constraints, thus contributing to the rigor of their methodology, it may at the same time make a contribution to historical reconstructions. The new approach may help us to understand better what we observe and to reveal what we do
not even suspect. Thus, if we can determine a set of linguistic characteristics that evince rapid (or traumatic) social changes, we may be able to increase our ability to reconstruct the histories of societies. If, for example, among related languages A, B, C, and D, A reveals some of our diagnostic traits, we would be justified in seeking evidence for the kinds of social changes they are associated with. This at least, is a task to which sociolinguistics should seriously address itself.

A linguistic diagnostic tool for 'pidginness' would be applied to a corpus, as was mentioned above. Therefore the selection of this corpus must be pursued with the utmost deliberation. As in all linguistic investigation we must be concerned with the sampling technique as well as the kind and quantity of the corpus collected.

**Simplification**

Let us now return to comparison of pidgins with natural languages. We need to examine the concept of simplification. The following represents the common belief:

From a structural point of view, the essential characteristic of a pidgin language is that it is sharply reduced in its pronunciation and grammar and in its vocabulary. In general, this reduction is in the direction of whatever features are common to the language of all those using the pidgin, for mutual ease in use and comprehensibility, thus arriving at a kind of greatest common denominator (Hall 1966:25).

Two things are wrong with this characterization of pidgins: the notions of commonality and reduction are both imprecise.

As for *commonality*, what is implied here is that the pidgin AB results when speakers of A seek to learn B but fail to acquire the use of some features of B, certain grammatical categories, for example (and conversely, of course, for speakers of B learning A). The following grid shows the set of theoretical possibilities. (+) marks the presence, (−) the absence, of a feature. Rows 1a, b are for features present in both A and B, rows 2a, b for features present only in A, rows 3a, b for features present only in B, and rows 4a, b for features present in neither. In these pairs (a) represents the case in which the outcome is the presence of the feature in AB, (b) the case in which it is not.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1b</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>2a</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>2b</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>3a</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3b</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>4a</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>4b</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>
Hall's statement implies that a pidgin language will have features (here, grammatical categories) that reflect the possibilities represented by 1a, 2b, 3b (and probably 1b), as well as of course 4b. The features positively present in the pidgin will be those represented by 1a, features present in both sources (the common denominator). From this point of view, a pidgin is not expected to be as complex as A or B (even though the grid shows four plus marks in each column), though balancing features from both (2a, 3a); and the possibility that the formation of a pidgin could involve innovation (4a) is not even considered. But it is possible for the pidgin to have a category which was absent in the two languages in contact. I believe that there are examples of it (4a) in Sango. The most notable one is the existence and use of *ekë* 'to be'. In Sango it is used both as a kind of auxiliary verb to mark progressive aspect (*lo ekë te kôbe* 'he is eating') and as a copula (*lo ekë makanzi* 'he is a village-headman'). No such verb has yet been found in Ngbandi (in the absence of data on vernacular Sango or Yakoma). Moreover, in that language, identification is expressed by the juxtaposition of two elements.

**Ngbandi**

\[
\begin{align*}
\text{ê ãzi} & \ 'we are people' \ ('lit. \ 'we people') \\
\text{lô nà tâ tî nkpc} & \ 'she had a pot of oil' \ ('lit. \ 'she with pot of oil')
\end{align*}
\]

**Sango**

\[
\begin{align*}
i \text{ekë dzo} & \ 'we are people' \ ('lit. \ 'we people') \\
\text{lo ekë na tî mafuta} & \ 'she had a pot of oil'
\end{align*}
\]

In Sango juxtaposition is limited to only certain kinds of constructions: e.g. *sô zo wa* 'who is this person?' *sô makanzi* 'this is the village-headman' or *lo ekë makanzi*. Let this example illustrate the much-overlooked innovative powers of a pidgin.

Of much more interest is the concept of reduction. We are generally led to believe that what is involved in reduction is primarily a decrease in the number of linguistic elements, whether phonological, morphological, or lexical. And Le Page, agreeing with Hall in principle, further restricts the concept of reduction by adding that 'since the inflectional structures of two languages in a contact situation rarely coincide, inflection is the commonest casualty in a contact situation' (1967:86). Such a narrow view of simplification is unfortunate although it is understandable. Both Le Page and Hall are thinking too much about pidgins they are most familiar with, the ones that are closely linked to European languages. This view ignores the possibility that pidginization could occur without the drastic reduction of an inflectional system. Moreover, pidginization can certainly occur when the source language has very little inflection. Sango is a good example of this kind of development. In the source language what are commonly called verbal categories are marked by co-occurrence patterns of tone in verbs and pronouns (and the subject marker *a*-). Thus, future time or anticipatory aspect is marked by pronoun set 3 (with high tone) and verb tone set B (but there are eight classes
WILLIAM J. SAMARIN

of verbs): e.g. \textit{é fono} 'we will walk' (Nelson 1952). In Pidgin Sango, verbs and pronouns are invariable, but some of the inflection remains with subject marker \textit{a-}: \textit{fadé i fono} 'we will walk,' \textit{i fono awe} 'we have walked' (literally 'we walk it-is-finished').

Both Hall and Le Page seem to have forgotten (or rejected?) the notion of purposeful simplification on the part of B-speakers (that is, of the language being learned). It has generally been held that B-speakers both simplified their language to make it easier for the learners and also imitated the learners in the simplifying changes they were making. A single term covers both kinds of behavior, but they are fundamentally different. Simplification on the part of A-speakers involves the inability (a) to infer linguistic structure from random utterances, and (b) to remember both the elements and the rules. Simplification on the part of B-speakers, on the other hand, is self-motivated: It assumes either some kind of notion, albeit extremely naïve, of what is difficult in their language or enough acquaintance with language A to realize what the learning difficulties are going to be, or both. I would not expect the second kind of knowledge on a large enough scale to account for pidginization. (But for its possible role, see Ferguson's paper in this volume.) The first kind, if it were entirely self-motivated, would be extremely interesting. Would we find, perchance, that there was a universal intuitive notion of simplification? Would the speakers of vernacular Sango, for example, have deliberately eliminated tonal inflection?

A broader concept of reduction will comprehend more linguistic phenomena, thereby establishing relationships that have gone unnoticed. I should like to suggest that there is something in pidgin languages, imperfect learning of a second language, loss of one's own language, and restricted codes, that is common to them all. These are only examples; the list can undoubtedly be lengthened, and with the larger number will come the need for establishing a more refined typology of pidginization than I can offer at this time.

\textit{Reduction in functions}

The fundamental characteristic of pidginization is indeed reduction or simplification, and it need not be 'drastic' or 'sharp' at all. Moreover, it is not necessarily a purely linguistic phenomenon. Pidginization should be seen as \textit{any consistent reduction of the functions of language both in its grammar and its use}. The key word here is 'functions'. With reference to the various uses to which language is put, this characterization means that a language is used to talk about less topics, or in fewer contexts, to indicate fewer social relations, etc. Imagine the whole gamut of uses to which any specific language is put, and a pidginized form of that language would have fewer such uses. This function reduction is what is indicated in the term 'trade language', but many
other restricted uses are obviously possible. A speaker whose proficiency in a second language is primarily in some specialized area of learning or human activity (sports, religious ritual, hunting, farming) speaks a pidginized form of the language. Boys who learn Yiddish to talk about the Talmud and Molokan young people in America who learn to talk Russian to take part in Molokan religious life are sometimes examples of such speakers. And most of us who only occasionally speak a second language probably have a limited world of discourse. This kind of pidginization therefore comprehends the limitations implied in the statements by Chomsky and Hymes referred to above.

Reduction in language use probably always has some repercussions on language output. When they are massive, a pidgin language results. Limited pidginization is much more common. It is common, for example, in linguistic field work and in the making of bilingual dictionaries of preliterate languages. In the initial stages of field work one chooses to elicit utterances which will reveal just as much information as can be controlled by the investigator. One permits the linguistic structure to drip out, so to speak; one does not turn on the faucet wide open. This means use of the same constructions again and again. Even if one attempts to vary them, as generative grammarians claim to do, they will vary in only a few ways. The same kind of stereotypy is found in dictionaries, even those as good as Lekens' *Dictionnaire Ngbandi* (but not his *Ngbandi Idioticon*). This dictionary illustrates the meanings of words with very few kinds of syntactic constructions, showing a preference for pronoun subjects and one of the past tenses. Unwittingly much of the language structure is amputated.

It is choice that both requires and permits a language-learner to pidginize. He is faced with options between which he has to decide. Some of these are, for the native speakers, stylistic choices, others are to avoid ambiguity (Gleason 1965:44off). To the learner they are just a headache. Thus, if he has difficulties with arranging pronouns in a French sentence, he may avoid constructions like *je ne lui en ai pas donné* 'I didn’t give him any of it’ for those like *je ne lui ai pas donné de ça* (or the named object). American speakers of Sango have further pidginized this pidgin by ignoring the (stylistic?) option between sentences like *lo gá ape, téné ake ape* (‘he come not, word is not’) and *tongana lo gá ape, téné ake ape* both of which mean ‘If he doesn’t come, it’s all right’. They use only the second one, presumably because it is closer to English, which has a subordinating conjunction. The statistical differences in the incidence of *tongana* in African and American texts probably also indicates other significant differences in encoding. So striking are the differences that I can on the basis of the use of *tongana* alone unerringly identify a Central African who has lived in a very dependent relationship on English-speaking missionaries.
The case of *tongana* in Sango is evidence for options in a pidgin, but Sango has fewer of them than one expects in a natural language. (The use or nonuse of the copula in *só aské yɛ* or *só yɛ* ‘what is this?’ is another example.) Thus it should be possible to identify pidginness on internal evidence alone. Be that as it may, we have in the reduction of options a means for characterizing historical pidginization. By way of illustration let us look rather closely at the Sango connective *ti* (*na* could also be used). Both of these occur in the source language with the same general functions they have in Sango, but their uses are not identical. Moreover, the total number of connective types is greater in the source language than in Sango.\(^{11}\)

Ngbandi has two connectives comparable to Sango’s *ti*. They are *té* and *ti* with varying tone. Both of them mark possession:

<table>
<thead>
<tr>
<th>Ngbandi Words</th>
<th>Sango Words</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>zàmà ti mbi</em> ‘my knife’</td>
<td><em>dà té lò</em> ‘his house’</td>
</tr>
<tr>
<td><em>ágbia ti kbólo</em> ‘village chiefs’</td>
<td><em>Lo-ğbia tè Nsàpa’ kingdom of God</em></td>
</tr>
</tbody>
</table>

(Mark 14:53) (Mark 14:25)

Examples like these suggest that there is no restriction on the use of either *té* or *ti*, but the corpus gives some indication that there may be a semantic difference; if not, I would assume a stylistic one. In other constructions there are clear (as far as our corpus is concerned) differences of use. Only *té* is used to ‘intensify’ the subject: *là kpé té* ‘là ‘they fled’ (literally ‘they flee of them’). Only *ti* is used with the interrogative substitute *nè* ‘what?’ and to introduce verbs or nominalized verbs: *tene tene ti nè* ‘say what?’ (literally ‘say word of what?’), cf. *tene ti nevè* ‘falsehood’ (‘word of lie’), *tene té* ‘là ‘their word’; *tongo lo mè ti lìa be* ‘raillerie (qui lèche le cœur) amère’, *à lèmbì ti vàŋgo* ‘can sell’ (‘equal to the selling’) (Mark 14:5). However, *ti* is absent in some such cases for no apparent reason, marked by (): *à no () 1bâ nyi lo* ‘il est allé en visite chez sa fille’, *lò bàbà () lingo kwà* ‘il continue à travailler (‘he continued doing work’). On the other hand, *ti* (or *té* or both) is consistently omitted in constructions marked by what one can call inalienable possession:

<table>
<thead>
<tr>
<th>Ngbandi Words</th>
<th>Sango Words</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ngù lì</em> ‘tears’ (cf. <em>ngù tí le keke</em> ‘fruit juice’, Mark 14:25)</td>
<td></td>
</tr>
<tr>
<td><em>là lò</em> ‘his head’</td>
<td></td>
</tr>
<tr>
<td><em>pè lò</em> ‘after him’</td>
<td></td>
</tr>
<tr>
<td><em>hà lò</em> ‘her husband’</td>
<td></td>
</tr>
<tr>
<td><em>sà ndo</em> ‘forest animal(s)’</td>
<td></td>
</tr>
<tr>
<td>*ngó <em>da</em> ‘axe handle’</td>
<td></td>
</tr>
</tbody>
</table>

It is possible to have three and even four nouns in a possessive series: *olo mène tilè mbi là* ‘this is my blood’ (‘blood of body of me’).

In Sango the picture is considerably more simple. There is only one unvarying connective which marks possession, and it is used with none of the restrictions found in Ngbandi. There is still, however, some option with inalienable possession. I should guess that the farther one got away from
Bangui, where Sango–Yakoma still exerts a little influence on the lingua franca, the fewer would be the options. The following phrase illustrates not only the single option that is possible but also the reason for the high frequency of *ti: méréngé (ti) kili ti ita ti mbi ti kili ‘my brother’s son’ (‘child of male of sibling of me of male’).

What is significant about *na* and *ti*, both of which have a high incidence, is not simply their frequency. Yet it was their frequency that attracted my attention and led to the formulations being presented in this paper. I started with very different goals, however. I had assumed that they might lead me to evidence of significant semantic and syntactic reorganization. I was looking for something characteristically pidgin about paraphrasis, for example. Instead, something else was found. So striking are the statistics that it is worth saying something about them.

In a corpus of 37,217 words and morphemes there were 11 connective types (i.e., different words), making up a total of 6,940 tokens or occurrences of all connectives. This represented 20.4% of the non-French part of the corpus, itself representing 33,990 tokens. But *na* and *ti* together account for 81.5% of all the Sango (non-French) connectives, the other 9 accounting for only 1,281 tokens. As is indicated in Table 1, these two words represent 15.19% of the total Sango corpus. The figure can be higher than this in random portions of text. For example, in the 100-word microcorpus from text N80 (from Samarin 1966b), *ti* represents 16%, and *na* 6%. In other words, in a running microtext where there were only 32 types, 2 of them together accounted for 22% of all the tokens.

**Table 1. Incidence of ‘na’ and ‘ti’ in a Sango corpus**

<table>
<thead>
<tr>
<th></th>
<th>Tokens</th>
<th>Non-French tokens (33,990)</th>
<th>All tokens, including French (37,017)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ti</em></td>
<td>3,056</td>
<td>8.99%</td>
<td>8.20%</td>
</tr>
<tr>
<td><em>na</em></td>
<td>2,603</td>
<td>7.64%</td>
<td>6.99%</td>
</tr>
<tr>
<td><em>Totals</em></td>
<td>5,659</td>
<td>16.63%</td>
<td>15.19%</td>
</tr>
</tbody>
</table>

The lesson that these facts teach is clear: we must refrain from assuming that the salient features of a particular pidgin are a characteristic of pidginization itself. Even if we found that connectives had a higher incidence in pidgins than in natural languages, we would be obliged to seek the causes. The effects of pidginization might be similar and the causes quite different. And while we are suggesting the exercise of caution, we might add that we must distinguish what is the result of pidginization from what is accidental in the text, whether
because of the way it was obtained, the topic, the genre of discourse, the style, etc. For example, the rather uninteresting narrative style of Sango, with clauses strung along in a close parallel to the actual events that took place, is very much like the narrative style of Tonga, another African, but Bantu and natural language. What characterizes Tonga is a redundant style which uses four devices: staging, overlapping, repetition, and synonymous expressions (Jones and Carter 1967). For what is translated as 'The whites followed us to the Bembezi and there was a great battle, Tonga has the following:

... again we ran away, we came here to the Bembezi. When we had come here to the Bembezi, the whites followed us, and came and found us at the Bembezi. They found us at the Bembezi, there was a very great battle (Jones and Carter 1967:112).

Other kinds of pidginization

In illustrating the concept of reduction in pidginization, we have been talking primarily about language learning, but two other instances of pidginization were earlier mentioned. The first is loss of memory, not that kind that is treated medically or psychiatrically, but the more normal kind, particularly when it is associated with 'disculturation', that is, losing the knowledge of and feeling for one's former existence. I should imagine that it is common among immigrants and people whose former way of living has been destroyed (e.g. American Indians). The only documentation of pidginization by memory loss known to me is that of an Auca Indian. The following is a description of her 'language' based on data that were elicited only with great difficulty. For 17 years she had not spoken her language; while living among the Quechua, to whom she had fled when she was 15 years old, she had deliberately tried to repress all she knew of her Auca existence.

... the verb suffixal system was impoverished, but supplemented by occasional Quichua [sic] affixes. The native pronominal set (apart from the honorific and the first-person exclusive plural forms) seems to have been complete. Most of the tense or aspect group was found, but only three occurred with normal frequency. The modal system was impoverished. Modal suffixes indicating uncertainty, doubt, probability, intention, negation, and obligation, were used so rarely as to be outside her effective control. Modal enclitics indicating derision, disgust, satire, extreme emotion, and emphatic negative seem to have been completely missing. Of the imperative system, only one of several suffixal combinations was used (Saint and Pike 1962:28-9).

The second is in restricted codes, as opposed to elaborated codes. These are constructs of speech forms suggested by the sociologist Basil Bernstein to account for linguistic variables correlative with different forms of social relations or qualities of social structure (see Bernstein 1967 with bibliography). In one sense the speech forms are determined by the roles the speaker assumes,
and a role is a set of shared and learned meanings through which a person interacts with others. Bernstein asserts that the role is therefore a 'complex coding activity' that controls the production of speech. The codes thus generated are discoverable in terms of the probability of predicting what form of speech a speaker will use. Of these two codes it is the restricted one that seems to reveal features of pidginization. Linguistically, it is characterized by verbal planning that is 'confined to choice of sequence rather than selection and organization of the sequence' (128). This apparently means that the speaker has fewer options. He has less freedom in organizing his utterances on the basis of greater vocabulary and syntactic patterns. A restricted code is best comprehended in its pure form 'where all the words, and hence the organizing structure, irrespective of its degree of complexity, are wholly predictable for speakers and listeners' (127). From a sociological point of view, in this code speech 'is not perceived as a major means of presenting inner states to the other' (130; cf. the statement by Chomsky above).

If, as Bernstein suggests, the pure form of restricted codes would be ritualistic modes of communication, then glossolalia ('speaking in tongues') is the epitome of such a code. What he says of relationships regulated by protocol and cocktail party routines is eminently true of this form of nonsense language. He writes:

In these relations individual difference cannot be signalled through the verbal channel except insofar as the choice of sequence or routine exists. It is transmitted essentially through variations in extraverbal signals. Given the selection of the sequence, new information will be made available through the [with following emphasis added] extraverbal channels, and these channels are likely to become the object of special perceptual activity (1967:127-8).

Glossolalia must be characterized both as a restricted code and as a highly pidginized form of speech because of its predictable form, using Bernstein's criterion, and because it represents the nearly maximal amount of reduction that language can experience and still sound like language. With a very restricted number of consonants and vowels a limited number of syllables is produced and arranged in a highly predictable order. Imposed on this segmental phonology, however, is a suprasegmental structure that is equally expressive as that of the native language of the speaker.12

There are, of course, very important differences between glossolalia and the kinds of restricted codes set up by Bernstein. For one thing, they are, or can be, extemporaneously generated; yet all the glossas -- glossolalic 'languages' -- so far studied have revealed a remarkable similarity of general structure (Samarin 1968). Moreover, there is no cognitive semantic encoding, at least of the normal type. That is to say, there is no correlation of the segmented stream of speech with units of meaning. There is meaning of a very different kind -- the meaning of function; glossolalia provides phonological material that
is the vehicle of affect without the distraction of semantic content. This is not far from the extemporaneous pidgins that one devises when, knowing less than the rudiments of Portuguese or Serbo-Croatian, he seeks to establish rapport with peasants who speak these languages.

Closer to the speech forms Bernstein is focussing his attention on are jargons and secret languages. They too are pidgin languages or are pidginized forms of other languages. Having collected some samples of the 'secret' language of the Gbeya grass-hunting Ngaragé society (Samarin 1959), for whom the speaking of Sango and Gbeya is proscribed during the hunting season, I would suggest that it is similar to Sango in having a limited inventory of lexical units and syntactic patterns, with one exception: it does not borrow from French as Sango does.

Another kind of 'exotic language' that displays features of pidginization is the special language used in Australia by the Dyirbal in the presence of certain in-laws. According to Dixon (1971), this language, called Dyalŋuy, has the same grammar and phonology as the everyday language, but is lexically more parsimonious. There is as a rule a one-many correspondence between items from the two vocabularies. For example, whereas the normal language (Guwal, or Dyirbal) has names for specific kinds of grub, Dyalŋuy (which Dixon dubs 'mother-in-law language') has only one generic term; and whereas there are three verbs in Guwal to denote different kinds of cutting, there are but two in Dyalŋuy. If one must be specific in the latter about a certain kind of grub, one uses a relative clause with the generic term.13

All of this is paralleled in pidgins. What is particularly interesting about the Dyirbal in-law language is that it is a pidgin (in my opinion) that, in the first place, is not used as a second language by a society, and, in the second place, does not appear to be derived by process of pidginization, as ordinarily understood. It is even possible, Dixon suggests (personal communication) that the language evolved from a collection of lexical items borrowed from neighboring languages. Guwal and Dyalŋuy are not 'cognate' lexically, and multilingualism within groups is common in Australia. The in-law language would thus appear to be reduced, both in structure (lexicon) and use, and a mixture as well (Dyirbal grammar and phonology, exotic lexicon), not through exigencies of contact, but as a deliberate creation; not as an unrespected compromise, but specifically to express respect. In any case, Dyalŋuy and languages like it may force us to redefine pidgin, and certainly require the distinction between pidgin and pidginization.

The pidginized nature of a restricted code is seen in the extent to which the full potential of language, as revealed in an elaborated code, is contracted. Bernstein sees this as a result of the major function of the code, which is to 'define and reinforce the form of the social relationship by restricting the verbal signalling of individual experience' (128). But we need not accept his
interpretation of the linguistic facts. He may profit as much from a consideration of pidginization as we from a consideration of restricted codes sociologically defined. I believe that we are looking at similar phenomena from two very different points of view.

As compared to pidgins, restricted codes have a different genesis. Pidgins result from language learning situations whereas restricted codes are part of the shared and learned behavior of a social group. There is no suggestion that an elaborated code was reduced, no matter how slowly, to the level of a restricted code. It could very well be that the latter is a linear descendant of a certain form of speech. If it proves to be very much like other forms of speech in other societies, we may have to conclude that it is not pidginized at all but that it is normal language. If this be the case, then the elaborated code is the innovation, like — in a weak simile — the improved fruit on a grafted tree. One would expect, for example, that as a society's culture becomes more complex (i.e. ‘richer’), as the language is forced to serve more functions (e.g. oratory or art), and as individuals acquire more and more functions, there would be a concomitant elaboration of the code. But would the contrary occur? If not for whole societies, at least for segments? This is what may produce restricted codes in complex societies and pidgins in multilingual ones.

Koinés

One type of language that has been excluded from the discussion to this point is the koiné. The reason is that koinés, like the Common Greek (so called by the Greeks themselves) lingua franca of the Mediterranean basin from which our technical term is derived, have never (at least to my knowledge) been considered pidgins. Unlike pidgins, koinés are not drastically reduced forms of language in spite of the fact that some simplification can be expected in them (Samarin 1962b). For example, Cohen (1956:310) talks about the ‘élimination d'un nombre de ses traits propres’ in Attic as koiné developed from it. What characterizes them linguistically is the incorporation of features from several regional varieties of a single language. This kind of amalgamation (or dialect mixing) can lead to a certain amount of heterogeneity. That is, a koiné, caught at an early stage of its history, might consist of many kinds of speech that are not easily correlated with non-linguistic factors like region, function, social status, etc. In time, however, the mix might jell, not without varieties of speech like those characteristic of any normal speech community: the gelatin is more like head-cheese than a dessert jelly. Some such development may be part of the history of Standard Macedonian with a base of what Lunt calls 'interdialects' (1959:23; see Samarin 1966a:195). Another feature that distinguishes koinés from pidgins, a feature that is implied in what has just been said, is that they are never detached from the languages they issue
from. That is, Common Greek as spoken at the beginning of the Christian era (but already 300 years old) in Jerusalem, Corinth, and Rome, would not be discontinuous from Attic Greek which was its base. (Those who insist on the continuity of, say, Jamaican Creole with English and Haitian Creole with French, might want to reexamine their positions with koinés in mind.)

Macedonian has just been mentioned as illustrating the process of koineization; but other examples should also be cited. (For information about the following I am indebted to J. Edward Gates for access to the file of the "Dictionary of Linguistic Terminology" at the Centre for Linguistic Studies, University of Toronto, formerly at the Hartford Seminary Foundation.) Among the five 'style variations' of Spoken Arabic is 'the koincized colloquial', 'a plain colloquial into which leveling devices have been more or less liberally introduced' (Blanc 1960:85). This leveling is described by Blanc (82) in the following way:

In certain situations, usually interdialectal contact, the speaker may replace certain features of his native dialect with their equivalents in a dialect carrying higher prestige, not necessarily that of the interlocutor . . . Moreover leveling devices may be called into play without the speaker's actually stepping out of his native dialect, but by selecting from among a number of equivalent features available to him those which are more general or more urban and suppressing those which sound local or rustic . . . In general, leveling often takes place not so much in imitation of a specific dialect as in an attempt to suppress localisms in favor of features which are simply more common, more well known.

Although this description might give the impression that koineization is an impromptu thing, Blanc makes it clear elsewhere (1953:2-3) that the koiné is widely spread, found even among village dwellers. (He also mentions a 'pidgin' Arabic that is 'almost always easily recognized' which is also used when one wants to avoid using one's local dialect with strangers.) This colloquial koiné should not be confused with the one discussed by Ferguson.

. . . Most modern Arabic dialects descend from the earlier language through a form of Arabic, called here the koine, which was not identical with any of the earlier dialects and which differed in many significant respects from Classical Arabic but was used side by side with the Classical language during early centuries of the Muslim era (1959:616).

The Italian of late fourteenth-century Naples might also have been a kind of koine based on the speech of other regions Hall (1949:154). No such qualifications are made by Karlgren (1949:45) about the language of northern China which was a 'real' koiné in the seventh to the tenth centuries. For a contemporary koiné we can go to Yugoslavia where one of the two standard varieties of Serbo-Croatian is a Belgrade-based koiné (Bidwell 1964:532).

The koinés mentioned in the preceding paragraphs all seem to be languages that evolved through one channel only, the spoken one. Others have evolved
through the medium of writing. Atkinson cites the Ionic idiom in which Herodotus wrote as a kind of literary koiné (1933:248). 'Standard Yoruba', a koiné based on the Oyo dialect, might be another example since it is the language used in education, writing, and, of course, contact between people speaking different dialects of Yoruba (Bamgbose 1966:2). Since writing implies, at least for some cases, a certain amount of conscious 'engineering' of language (like the creation of 'union languages'), it is not surprising to find that the koiné which arose from the Tupi-Guarani group, the Lingua Geral, 'was the work of missionaries' after the conquest (Entwistle 1936:233). Vachek goes so far as to suggest that the important function of a literary language 'has always been to serve as a kind of koiné' (1966:98). However, we must remember the motivations suggested by Lunt and Blanc for the emergence of koinés, a topic about which Marouzeau is explicit:

La manie de copier le bourg, la petite ville, le chef-lieu, la capitale, tout cela contribue à fixer et à répandre une langue-modèle; les habitudes de langage se communique et s'imposent comme les idées et les préjugés, comme la mode et la façon de s'habiller; ainsi se crée ce qu'on appelle une langue commune . . . (1950:96).

The distinction between pidgins and koinés presented here differs from implications made in an earlier paper (Samarin 1966a). There I said that 'Town Bemba must certainly be considered a kind of pidgin language' (201n.) on the basis of Richardson's characterization of Town Bemba as having a 'tendency to indeterminacy in most departments of the language' (1963:145). Richardson, however, did not challenge this assertion in his comment on my paper, and in fact he said 'the parallel drawn... between Sango and Town Bemba in no way disagrees with the sense of my articles on the linguistic effects of urbanization' (Samarin 1966a:207). But he went on to say (208) that the real difference between Town and Village Bemba was in lexicon and style, implying that there was no reduction of the kind that I now use in characterizing a pidgin. (My purpose in that paper, of course, was something other than trying to characterize pidginization.) I would now say that Town Bemba might be a kind of koiné with pidgin features. If not, it might prove in any case to have features of salient pidginization with only a minimal amount of substantive pidginization. What kinds and degrees of pidginization occur in koinés still remain to be determined.

**Conclusion**

The terms 'pidginized', 'restricted', and 'elaborated', are stative terms by grammatical definition, but they are dynamic by implication. That is, something was pidginized, restricted, or elaborated. The terms need not be taken too literally. There is value in using them in a purely descriptive way in
formulating a taxonomy of language varieties and language functions. But they also have the value of reminding us that language is not static, no more so than societies are. It is the responsibility of linguistics therefore to address itself to the sociological dimensions of both synchronic and diachronic dynamism.

NOTES

1 It was just eight years ago that an African linguist of European extraction said to me, 'You talk as if you like Fanagalo.' In our conversation, I thought that I was simply asking the kinds of questions any linguist might ask about another language.

2 Recent investigations of glossolalia show clearly its value, though it is not a language in any ordinary sense of the term. See below.

3 This hypothesis is being attributed to several people. If someone deserves the honor of having first suggested it, we ought to have the detective work done once and for all. But perhaps there is convergence as well as diffusion in the explanation of the distribution of this idea.

4 Sango is one of the exceptions to the putative generalizations. Although repetition is used in the language, it is not found as a characteristic of a large number of words, as one finds in Jamaican iteratives (De Camp 1967). When one compares Sango with the source language, he sees what has happened: most of the ideophones have been lost in the process of pidginization. It is in these words that most of the reduplication occurs in Adamawa-Eastern languages. (Thomas (1963) seems to believe otherwise, but see my review (1965).) The result is that it is the loss of reduplication which is characteristic of the pidginization of vernacular Sango, not its presence.

5 We shall no longer pretend that pidgins are not languages. Up to now we have indulged in a bit of rhetoric. That pidgins are languages is beyond question; what we have to determine is what kind of language they are. There may be some forms of speech that are outside the range of linguisticality, but they are not illustrated by pidgins.

6 Le Page might disagree with this statement, for he writes that 'the grammar of a language is the sum total of all its relational devices and this sum total must be roughly constant. In other words, if you don't manage to get something across by one means, you will by another' (1967:86). He does not seem to exclude pidgins from this generalization. In fact, the next sentence limits grammatical simplification in pidginization to inflectional structures.

7 The monostratal nature of Sango seems to have a strong equalizing function. It certainly had on one particular occasion. I paid a call one day in New York City on a very high-ranking member of the C.A.R. delegation to the United Nations. The language we used was French, and our behavior was appropriately formal. We were naturally using vous with each other, and I, not certain what protocol required under the circumstances, was tossing in votre excellence from time to time. The introduction of Sango into the interview, but not by myself, completely altered our relationship. In Sango it was inappropriate to refer to 'his excellence'. For the remainder of the appointment we became casual, if not a bit intimate, with each other. And then, on parting, we switched back to the formulas that French provides for such occasions.
This linguistic diagnostic tool would obviously not tell us more than it was designed for. It could hardly tell us, for example, whether or not Afrikaans owes its existence to the use of Portuguese Creole in South Africa in the seventeenth and eighteenth centuries, as is suggested by M. F. Valkhoff (1966, and Samarin 1967a). But if we had a good sample of local and colloquial seventeenth century Dutch spoken by the indigenous population and foreign residents, his interests and ours would both be served. (On the general point, however, cf. Hymes (1959), discussing Taylor (1956).)

For what immediately follows I profited from a stimulating conversation with my colleague, H. A. Gleason, Jr.

Le Page also may be talking about non-linguistic (or sociolinguistic) pidginization in the following statement. The important expression is 'mode of behavior': '. . . one should . . . think of a pidgin as a mode of behavior which results when speakers of A [the learners of B] meet speakers of B [the language being learned] under certain social conditions . . .' (1967:86).

A diagnostic device that relies on measuring the kinds and number of options in a language obviates the difficulty of measuring redundancy. Although one might like to talk about redundancy being reduced in a pidgin in comparison with the languages in contact, this is difficult to maintain in the absence, as Le Page admits (1967:86), of any means to measure redundancy itself.

For the sake of convenience we shall call the source language Ngbandi, partly to avoid confusion between the names of the vernacular and the pidginized languages and partly because our best source of information about the Ngbandi–Sango–Yakoma dialect cluster is on Ngbandi itself. See the works by Lekens, Nelson, and the Gospel of Mark published by the British and Foreign Bible Society (1959). Nelson is the translator of Mark. Differences between Lekens and Nelson seem to be their differences, not the language's. I have tried with little success to get a corpus from vernacular Sango. I do have three 8-inch (78 r.p.m.) records of a religious nature produced by Gospel Recordings Incorporated (146 Glendale Boulevard, Los Angeles 26, California). They are labeled 'Sango: Ubangi', 1408-1A-LA3865 to 3870 and are not of high quality. The recording was made at Libenge, Republic of Congo (Kinshasa), around 1957. Rev. Richard B. Anderson, missionary of the Evangelical Free Church, who was there at that time, informs me that Sango fishermen lived on the sand-banks during the dry season, some of whom came from as far away as the confluence of the Uele and the Mbomu rivers. There was also a permanent Sango village on the banks of the Ubangi river not far from Libenge. Sango is spoken by about a million inhabitants of the Central African Republic and the adjoining parts of the Chad, Cameroun, and Congo (Kinshasa) republics, mostly as a second language. There is an increasing number of people for whom it is a first language, and for these it would be considered a creolized language, using the term in one of its senses. Whatever one says about the classification of Sango, Ngbandi is an Eastern (or as I prefer to call it, Ubangian) language of Joseph Greenberg's Adamawa–Eastern family, itself a part of Niger–Congo.

One person, eventually a glossolalist himself, characterized the tongues he heard at a religious retreat in the following way: '. . . I listened intently to the beautiful, expressive flow of language. Although I could not understand a word, it was unmistakably reverent and adoring, and something within my own spirit responded to that adoration' (Donald W. Basham, 'Keeping a divine appointment', in Full Gospel Business Men's Voice, December 1967, p. 26).
Dixon accounts for the similarities and differences in terms of nuclear and non-nuclear words, a taxonomy of the lexical inventory of languages that he suggests for incorporating the advantages of both the componential and definitional approaches to semantic description. Nuclear words are those that cannot be defined in terms of other lexical words and must be defined componentially; non-nuclear words, on the other hand, are defined in terms of the nuclear words utilizing the full grammatical possibilities of the language in the formulation of these definitions. Thus, nuclear terms are fewer in number but generally more frequent in occurrence than the non-nuclear ones. Since this is true of all natural languages, it is true of Guwal, but it is not true of Dyalñuy: the latter has equivalents for Guwal’s nuclear words but only a handful of non-nuclear items; and the non-nuclear words of Guwal are ‘defined’ in Dyalñuy in terms of one or more nuclear words. Dixon finds the ‘mother-in-law language’ very helpful as a source of clues to semantic relations within the everyday language.

REFERENCES


Blanc, Haim 1953. ‘Studies in North Palestinian Arabic’, *The Israel Oriental Society*, no. 4


British and Foreign Bible Society 1959. ‘Nɔɔ pà ã sù na ti Malako’ [The Gospel according to Mark], London


Hall, Robert A., Jr. 1949. ‘Nasal plus homorganic plosive in Central and Southern Italy’, *Archivum Linguisticum* 1:151–6

1955. *Hands off Pidgin English!* Sydney, Australia, Pacific Publications


1964. ‘Discussion’ [of Bastian et al.], *Linguistics and language study* (Fifteenth round table meeting), ed. by C. I. J. M. Stuart, 189–91 (Georgetown University Monograph Series on Languages and Linguistics, 17), Washington, D.C., Georgetown University Press


Saint, Rachel, and K. L. Pike 1952. ‘Aucu phonemics’, *Studies in Ecuadorian Indian languages* 1, ed. by Benjamin Elson, 2–30, Norman, Summer Institute of Linguistics of the University of Oklahoma


1961. 'The vocabulary of Sango', Word 17:16–22
1962a. 'Une lingua franca centrafricaine', Colloque sur le multilinguisme (Brazzaville 1962), 257–65, Brazzaville, Commission de Coopération Technique en Afrique
1966b. A grammar of Sango, The Hague, Mouton
1966c. The Cheya language: grammar, texts, and vocabularies (University of California Publications in Linguistics, 44), Berkeley and Los Angeles, University of California Press
1968. 'The linguisticality of glossolalia', The Hartford Quarterly (Hartford Seminary Foundation) 8(4):49–75
1970. Sango, langue de l'Afrique Centrale Leiden, E. J. Brill
1965. A dictionary of Sango, duplicated at Hartford Seminary Foundation for U.S. Office of Education
Taylor, Douglas 1956. 'Language contacts in the British West Indies 1; A case of intimate borrowing II; On the classification of creolized languages', Word 12:399–414
Vachek, Josef 1966. The linguistic school of Prague: an introduction to its theory and practice, Bloomington, Indiana University Press