1. Introduction

Geographical regions of the world may show linguistic distributions of great diversity (such as, for example, the state of California for American Indian languages, the Jos Plateau in Nigeria, the highlands of Papua-New Guinea, etc.) or of relative uniformity (such as the contemporary United States or Japan) or these regions may lie somewhere in-between, as Greenberg [1956:109] suggests. Diversity may of course obtain both individually and intra-individually. Greenberg [1956] and later Lieberson [1964] have developed quantitative measures to "... render such impressions more objective, allow the comparing of disparate geographical areas, and eventually to correlate varying degrees of linguistic diversity with political, economic, geographic, historic, and other non-linguistic factors" [Greenberg:109]. Greenberg proposed eight measures of linguistic diversity which were designed to determine the possibilities of communication within a delimited area. These measures permit us to quantify diversity along a continuum which ranges from a situation where no code is shared by any two persons to the opposite situation where all persons share a common code.

Greenberg's references are primarily to the linguistic diversity of spatially delimited populations, however as Lieberson points out, such measures may be applied to "any socially meaningful population delineated..."
on a non-areal basis" [1964:526]. In the present paper, we will look at the question of multilingualism by a comparison of successive generations using some of the quantitative measures elaborated by Greenberg.

2. The Population

The Mijikenda (Swahili 'nine villages') live in an area of over 5,000 square miles along the Kenya coast and the immediate hinterland. These people speak nine contiguous mutually intelligible dialects. They number about 300,000. The Giriama form the most populous component of the Mijikenda, numbering over 160,000. The Northern Giriama, on whom this discussion of multilingualism is focused, practice a variety of forms of cultivation ranging from subsistence swidden to cash crop agriculture on government settlement schemes. This research was conducted in the Gede Settlement Scheme, located between Mombasa and Malindi on the Kenya coast. The population there is primarily Northern Giriama (representing about 75% of the total) with an admixture of other Mijikenda peoples, the Cushitic Waata Galla, upcountry peoples such as the Bantu Kikuyu, Kamba, Luhya and Nilotic Luo. The Giriama dominate the area of the settlement scheme culturally, if not linguistically.

Adjacent to the settlement scheme are nucleated Swahili villages. The Swahili sometimes settle within Mijikenda population areas, occupying positions either as tradesmen or less frequently, as farmers in single or extended family compounds.

The following pattern of multilingualism characterizes the population of the settlement scheme. Most Mijikenda men speak Swahili in addition to the vernacular language. The typical variety of Mijikenda Swahili is fluent but obviously non-native to a native speaker. Among the women, monolingualism in the vernacular is still not uncommon, but an increasing number of them are becoming more than incipiently bilingual in this

---

2Diebold [1964:504] offers the following definition of minimal bilingual skill: "contact with possible models in a second language and the ability to use these in the environment of the native language." In the Mijikenda case, "possible models" may include not only native speakers, but also speakers of Swahili as a second language, who may deliberately assist coethnics in acquiring limited proficiency in a non-native language.
language. The Waata follow a pattern similar to the Mijikenda but they are fluent in their own vernacular as well. Upcountry people of different ethnic groups tend to follow a similar pattern in their language repertoires. They are usually less fluent than Mijikenda in Swahili, but more so than their coethnics at home. Primary school graduates among the above groups may also speak English. A typical member of an upcountry ethnic group living on the coast speaks better English than does a typical Mijikenda. A possible reason for this is the occupational status of upcountry groups, employment in the 'European' hotels. Their occupational status is in turn a function of educational attainment which may in turn be a product of their ability to achieve a high level of proficiency in English, the medium of the instruction in the school system.

With regard to interintelligibility among Mijikenda dialects, my own observations of cross-ethnic group communication indicate that intelligibility is high. Members of these groups claimed a lesser degree of interintelligibility than their language behavior indicated. Interintelligibility appeared to be positively correlated with lexicostatistical estimates of subgrouping. Because language variety was considered an important marker of ethnic membership and group solidarity, such a claim about mutual intelligibility might have been predicted.

As a result of increased exposure to education and increased contact with outside cultures among these people, there has been a significant change in the type of multilingualism prevalent on the settlement scheme. This change has been in the direction of expanded individual language repertoire.

Exposure to the formal education process began historically with the advent of colonialism in East Africa in the late nineteenth century. The impact of the European-style school system and what was perceived by most Mijikenda as 'its language' (Swahili/Giriama: Kizungu 'European language, i.e. English') became deeply entrenched in the cultural and linguistic

---

3 For lexicostatistical evidence and comparative word lists on the sub-grouping of these dialects, cf. Sedlak [1975].
milieu of the masses of the Mijikenda only after the introduction of universal education following the achievement of political independence by the Republic of Kenya. Since the medium of instruction for the primary schools even at the village level was English for the most part, this factor contributed to the development of a population multilingual to various degrees. Before the coming of the British, it is likely that the great majority of the Mijikenda were monolingual. The few who were bilingual acted as cultural intermediaries with neighboring ethnic groups such as the coastal Swahili and Pokomo, or the upcountry Kamba, Taita and Taveta (all Bantu peoples) in trading relationships.

With the advent of the British came another factor which contributed to the expansion of individual language repertoire. Subpopulations of present-day Kenya became increasingly mobile as a result of new economic and political configurations imposed by the colonial system. This juxtaposition of a great number of Niger-Congo Bantu, Nilotic and Cushitic speakers produced a need for a channel of communication among these groups. Because Swahili was already rooted firmly on the coast and had served this function even before the entrance of the British, it took on a greater role. English was at that time and, for the most part, is today, a language of the elite, in spite of its spread through the school system.

Seventy-five individuals were selected from the total settlement scheme population of 16,000. These seventy-five individuals were representative of the settlement scheme population ethnically, linguistically, and economically, insofar as this was possible.

3. Expanded Repertoire

The influence of both the English language and the Swahili language factors has resulted in an increase in the mean number of languages spoken by the present generation compared to that of their parents' generation. An index of claimed language repertoire referring to this mean number for the respondents' generation (cf. (1) below = 2.72) exceeds that for the parents' generation (cf. (2) below = 1.69) by a factor of 1.03, the real increase in the number of languages in the average repertoire of the respondents. Parents' generation will be defined as the set of respondents'
putative fathers and biological mothers. The index of claimed language repertoire will be defined as the mean number of languages spoken by each generation. For the respondents' generation it is simply the summation of languages spoken by each respondent, divided by the total number of respondents:

\[ \text{RESPONDENTS' INDEX} = \frac{\sum_{i=1}^{n} \text{language}}{n} = 2.72 \]

For the parents' generation, the computation of this index is somewhat more complicated: the summation of the number of languages claimed by the respondent to be present in the repertoire of each parent \((F + M)\) divided by two. This results in an average for one parent. This average is then multiplied by a correction factor of 1.46. The result is divided by \(n\), 75, the total number of respondents:

\[ \text{PARENTS' INDEX} = \frac{\sum_{i=1}^{n} \frac{F + M}{2} \cdot 1.46}{n} = 1.69 \]

The correction factor of 1.46 represents a weighting to compensate for under-reporting of parents' language repertoire by the respondents. This adjustment is reflected in (2). Table 1 shows the contrast in the language repertoires of the parents as reported by the respondents versus their language repertoires as they themselves reported it. The derivation of the correction factor is also shown here.

In the course of the interview, respondents were required to provide less information on the repertoires of their parents than on their own repertoires. Informal discussions with several of the parents of the respondents to the survey brought the following pattern to light. Some respondents failed to report either Swahili, the lingua franca, or another vernacular as components of their parents' repertoires.

Swahili was often omitted by respondents since it was almost universally a component of their own repertoires and therefore could be assumed to be a component of their parents' repertoires. Swahili, in fact,
Table 1

PARENTS' LANGUAGE REPERTOIRE

<table>
<thead>
<tr>
<th>Respondent</th>
<th>Respondent's Report</th>
<th>Parent's Self-report</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>31</td>
<td>DKm</td>
<td>G</td>
</tr>
<tr>
<td>41</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>53</td>
<td>Km</td>
<td>Km</td>
</tr>
<tr>
<td>56</td>
<td>W</td>
<td>W</td>
</tr>
<tr>
<td>59</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

Total 13 19

CORRECTION FACTOR = \(\frac{19}{13} = 1.46\)

Column
A = father's language repertoire: respondent's report
B = mother's language repertoire: respondent's report
C = father's language repertoire: self-report
D = mother's language repertoire: self-report
(underlining = languages not reported by children)

Language Key
A = Omani Arabic
D = Duruma
G = Giriama
Ki = Kikuyu
Km = Kamba
S = Swahili
W = Waata
was less widely known by the parents and was less often a component of
the parents' repertoire.

Other vernaculars were also frequently omitted by respondents be­
cause the period of acquisition and use by the parent may have come and
gone before the respondent was old enough to notice. Vernacular languages
acquired were usually those of large ethnic segments of the Kenya popula­
tion, such as the Kikuyu, the Kamba or the Luo. According to the respon­
dents who had learned vernaculars other than their own, the following
conditions of the contact situation favored acquisition. These are listed
in order of importance: most of the other participants in the contact
situation spoke a vernacular different from that of the respondent; the
contact was intense and over a period of at least months; most of the
other members of the group spoke limited Swahili; the respondent spoke
limited Swahili. Other conditions may also have contributed to producing
a situation favorable to the acquisition of other vernaculars, but the
above were the primary ones proposed in the respondents' 'ethnosociology
of language.'

The typical repertoire of the parents' generation consists of ver­
nacular and Swahili, while that of the respondents' generation consists
of vernacular, Swahili, and possibly English, rather than a second African
language. Swahili, if known, is acquired largely through contact with
native speakers, while English, if known, is acquired almost solely
through the Kenya school system where it is the medium of instruction
and there is no contact with native speakers.

4. Greenberg's Diversity Measures

Greenberg has proposed a number of language diversity measures which
may be of value in the analysis of the sociology of the language situation
in this Mijikenda community and in the analysis of expanded individual
repertoire.

The simplest model proposed by Greenberg is called the monolingual
nonweighted method, and is referred to as method A. It can be described
as follows: "If from a given area we choose two members of the population
at random, the probability that these two individuals speak the same lan­
guage can be considered a measure of its linguistic diversity. If everyone
speaks the same language, the probability that two such individuals speak the same language is obviously 1, or certainty. If each individual speaks a different language, the probability is zero. Since we are measuring diversity rather than uniformity, this measure may be subtracted from 1, so that our index will vary from 0, indicating the least diversity, to 1, indicating the greatest" [1956:109].

The total probability of selecting two speakers of the same language is the sum of the probabilities of such an event for each individual language $M$, $N$, $O$, etc. The probability for each of these is $m^2$, $n^2$, $o^2$, etc., where $m$ is the proportion of speakers of $M$ to the total population, etc. Summing up, the formula becomes:

(3) Method A = $1 - \Sigma (i^2)$

Method A is referred to as a monolingual method, because the compilation of each implies counting each individual as a speaker of a single language. This method, then, would reflect only mother tongue complexity within the multilingual society and not the fact that the language repertoires of individual speakers have expanded. Presumably, modifications could be made in each of the methods which would allow them to account for this fact.

Greenberg [1956:111] suggests that this method may be altered by the adoption of a split personality method, in which a speaker of two languages would be counted as two speakers, a speaker of three languages as three speakers, etc. For the 75 individuals now living in the Gede Settlement Scheme in Coastal Kenya and for 75 of their parents who have their origins in the same area, let us examine the results using Method C, the split personality method. This method is not weighted for linguistic similarity. For the children's generation, $C = .766$ and for the parents' generation, $C = .460$. Using this method, the probability that any member of the children's generation speaks the same language as any other member is .766 and the probability that any member of the parents' generation speaks the same language as any other member is .460.

However, if we look at the actual language repertoires of the successive generations, it will be seen that these probabilities may perhaps indicate
something other than diversity. In the case of the children's generation, 75 individuals speak a total of 21 different languages. Of these 75 individuals, 69 or 92% speak Giriama. Using Method C, counting each language of each individual as one unit, there are then 204 individuals. In the case of the parents' generation, 75 individuals were reported to have spoken only 11 languages. Of these 75 individuals, 63 or 84% of them spoke Giriama. Counting each language of each individual as one unit for this group, 127.02 (87 x correction factor of 1.46) individuals result. Knowing that the majority of both groups speak Giriama in addition to other possible languages, it is clear that the higher diversity figure for the children's generation must be due to the addition of other languages to their repertoires. This is indeed the case. Method A would show only mother tongue diversity and would therefore more accurately represent this aspect of the situation.

An additional measure, H, perhaps more accurately reflects the true sociolinguistic situation. This is the index of communication, the probability that if two members of the population are chosen at random, they will have at least one language in common. "This measure is obtained by first dividing the population into proportions of speakers of any one language only or any particular combination of languages and then calculating the products of each pair ..." [Greenberg 1956:112]. Products of pairs with at least one language in common are multiplied by 1; products of pairs with no language in common are multiplied by zero. The summation of the products yields the index of communication, which is approximately the same for both children's and parents' generations in the Kenya data.

Certain inferences might be drawn from a comparison of the index of communication and the split personality method across generations regarding redundancy of communicative form to perform a communicative function. For the moment, we shall assume that any communicative function may be filled by any communicative form, although we know that in the real social world this is not the case. Swahili and English, while expanding the repertoire of communicative forms, often compete to perform a communicative function necessarily filled by one or the other. For example,
in the home setting, two lower-level government officials (location chief and traffic bureau police lieutenant) may be discussing national politics. Because both of these officials received at least a secondary education, English may be used. Because both are also coastal people and are relatively fluent in Swahili, this language may also be used. In fact, code-switching is a characteristic feature of this setting; the switch is often triggered by the use of an English loan word in Swahili or vice versa. In the home setting, language choice for national politics may be random. The same two government personnel conversing on the same topic at the chief's office would most likely prefer Swahili. In this setting, the choice of Swahili, the lugha ya taifa, 'national language' over English, the lugha ya wageni, 'foreign language', suggests national political allegiance, while choice of English in this setting may connote at worst disloyalty, or at least Anglophilia.  

5. Discussion

In a multilingual situation such as that of the Gede Settlement Scheme, redundancy of communicative form is reflected in a high value for Measure C, a result using the split personality method, for the respondents' generation, with the index of communication remaining constant or nearly equivalent with that of the parents' generation. Efficiency of communicative form is reflected in the lower value of Measure C for the parents' generation, with the index of communication remaining constant or nearly equivalent with that of the respondents' generation.

At first sight, it appears that a similar value for the index of communication across generations has no disadvantages in terms of decreasing the actual communication potential. On closer analysis, a case could also be made for fragmentation of communication potential in a setting like the Gede Settlement Scheme where individual language repertoires have expanded. Certain topics, for example, may require a specific code, and thus exclude those without access to this code. In an earlier, more monolingual generation,
such may not have been the case. Thus the spread of two competing linguae francae (albeit the socioeconomic groups who have access to these two may not entirely overlap) within such a geographically delimited area which already has a high index of communication may be of questionable value. Only when the extension of the geographic and linguistic bounds is considered will an increase in the index of communication (Greenberg's $H$, the probability that any two speakers share one common language) become relevant and apparent.

Perhaps more satisfying from the point of view of the multilingual situation which we observed would be the application of Greenberg's Measure $G$, the random speaker-hearer method. This measure attacks the problem of the spread of a lingua franca, since it is the probability that "if one individual is chosen from the population who if polylingual (or multilingual), is equally likely to speak any of his languages, then a second member of the population will be able to understand him." By simple observation of the mother tongue percentages mentioned above for our data, it is obvious without calculation that this factor would be high.\(^6\)

---

\(^6\) In the discussion period following the presentation of this paper, C. A. Ferguson suggested that the analysis of linguistic diversity measures might now be developed and expanded to look at language situations in multilingual societies from a finer-grained and more discrete perspective. Diversity measures which would look at the allocation of linguistic codes according to roles of speakers, participants, topic, setting, etc., could be established. The establishment of ideal speaker-hearer folk categories might also be proposed in such an analysis.
REFERENCES


