Studies in African Linguistics
Volume 14, Number 2, August 1983

GENETIC RELATIONSHIP AND THE CASE OF MA'A (MBUGU)*

Sarah G. Thomason
University of Pittsburgh

This paper addresses the general question of genetic vs. non-genetic language development, in the context of a structural and historical discussion of Ma'a (Mbugu), a language with Cushitic basic vocabulary that is spoken in Tanzania. The grammatical structure of Ma'a is compared to characteristic Cushitic and Bantu structures. The conclusion that emerges from this comparison is that Ma'a probably does not have enough Cushitic grammar to qualify as a Cushitic language in the full genetic sense; and if it does not, its origin must be nongenetic. The final section of the paper seeks to determine the particular route of nongenetic development that Ma'a has followed, using the direct evidence of published comments about its speakers' history and the indirect evidence of comparison with other languages whose origin is nongenetic or, like Ma'a, on the borderline between genetic and nongenetic.

1. Introduction

Ma'a (Mbugu)[1] is a favorite battleground for proponents and opponents of hypotheses of language mixture, since it is said to have Bantu grammar but not Bantu vocabulary. It is best known to non-Africanists from Goodman's description, and probably most people who know about the language would agree with his conclusion that "the development which Mbugu has undergone defies easy categorization; it remains a unique linguistic specimen" [1971: 253]. In this paper I will argue that Ma'a is indeed a "mixed language"

---

*This paper is a greatly expanded and extensively revised version of a case study originally contained in Thomason & Kaufman [1975]. The present draft has benefited greatly from both substantive and bibliographical information provided by Christopher Ehret, to whom I am very grateful indeed for his generous help. Any remaining errors of fact or interpretation are, of course, my own.

[1]Elderkin comments that the self-name *Ma'a* is preferable to the more familiar *Mbugu* as a designation for this language, since the name *Mbugu* is also used for a Bantu language spoken in the same region [1976:280].

which probably should not be classified genetically into any language family, and that its history can be partly inferred from its structure and from published information about the history and cultural traits of its speakers. I will also show that this type of linguistic mixture is not unique, but is found in a few other languages whose history is better attested.

Greenberg was apparently the first, and for some time the only, linguist to classify Ma'a genetically as a Cushitic language. In making this classification, he used his method of mass lexical comparison [Greenberg 1955, 1963], so his criterion for the grouping was strictly lexical. Since 1971, more extensive lexical studies have shown conclusively that the basic vocabulary of Ma'a is primarily of Cushitic origin (see especially Ehret [1980]), and as a result of these studies the genetic grouping with Cushitic now seems to be widely accepted. Welmers, for instance, observes that "the development of Ma'a, though certainly unusual, is within the familiar framework of continuous language history with extensive external influence", and that "the continuous or genetic history of Ma'a is Cushitic" ([1973:8]; see also Ehret [1974, 1976, 1980]).

Opposing views can be found even in the relatively recent literature, however. Dolgopolskij groups Ma'a with Bantu because of its Bantu morphology ([1973]; cited by Zaborski [1976:83]), and the older "mixed language" claim is favored by Whiteley [1960a]. Elderkin also seems to prefer the last treatment, and to conclude from the case of Ma'a that genetic classification of languages is in principle unscientific: "classification of languages rests on the selection of one part of a language to typify that language, and this selection is arbitrary" [1976:296-7]. He goes on to say that Ma'a is a crucial example because "no theory of linguistic relationships should have an exception" (p. 297).[2] Elderkin is quite right to suspect that something is wrong with a genetic linguistics that takes just one part of a language as

[2]This is a rather common view. Note, in the same volume, Grover Hudson's remark that genetic classification is 'exhaustive, since no languages can be left out' [1976:237]. The same position is reflected, I think, in the strenuous efforts that have been made by some creolists to justify genetic classifications for Caribbean creoles (see e.g. Meillet [1921:82] and Hall [1958:370f.]).

diagnostic for purposes of classification. I could, in theory, graft Russian lexical morphemes onto my English grammar, but I would not then be speaking Russian; and I would no longer be speaking English, either.

The flaw in Elderkin's reasoning is his assumption that genetic classification of languages must be exhaustive. The idea that some languages do not fit into the standard genetic picture dates from the nineteenth century, with Schuchardt's work on pidgin and creole languages. It remains controversial even for these classic examples of "mixed languages",[3] but the controversy arises, in my opinion, from a failure to take seriously enough the first principle of genetic relationship--namely, that a daughter language is a changed later form of its single parent language. On this view, we must surely assume that any given daughter language in a family tree arises through an unbroken series of generation-to-generation transmissions of an entire language, that is, a complex set of interrelated lexical, phonological, morphosyntactic, and semantic structures. Changes, both externally and internally motivated, accumulate gradually enough that (as can be shown for Indoeuropean, for instance) systematic reflexes of proto-language structures can be found in all linguistic subsystems of a daughter language even after five or six thousand years. But if the chain of transmission is broken at any point, then the resulting language no longer belongs in any family tree, because it is not a changed later form of any single parent language: it does not meet the conditions for genetic classification.[4] Such a break in transmission will always be reflected, as I have argued elsewhere [Thomason

---

[3]See Thomason [1980] for a discussion of this controversy.

[4]The notion of a "break in transmission" is necessarily vague, because there are borderline cases where transmission is neither clearly normal and continuous nor clearly abnormal and discontinuous. One example is Afrikaans, which is viewed by some linguists as a creole and by others as a direct outgrowth of Dutch; other examples are cases where two or more languages in an intimate contact situation are so closely related that the source of many structures cannot be determined. One such case is found in those areas of England where Old English and Old Norse were spoken, at a time when they still shared many lexical and grammatical features. Other possible examples of this type might be certain pidginized Bantu languages that arose in exclusively Bantu-speaking areas. These and other borderline cases are described in Thomason & Kaufman, Forthcoming.

1980], in a mismatch between the vocabulary and the grammar of the resulting language; it will not be possible to show that both have arisen from the same parent language.

The most obvious candidates for languages with nongenetic origins are pidgins, which do not arise through any sort of transmission. Instead, they are created as new languages in multilingual contact settings[5] in which a single lexicon is adopted--usually, though not always, taken almost entirely from a single language--and a new grammar evolves through intergroup communication, generally without a single-language target. The next most obvious nongenetic languages are creoles like those of the Caribbean, which developed rapidly among linguistically diverse groups of slaves who adopted the lexicon of the slavemasters and constructed a new grammar that apparently did not, at least in the beginning, involve any serious attempt to learn a single language's grammar.[6] These languages arose, therefore, outside of normal transmission processes. In most pidgins and in the Caribbean creoles, the vocabulary is taken from a single language, and the grammar is not derived from that language or from any other single language. The nongenetic historical development, though not evident in the single-source vocabulary of such a language, is reflected clearly in the grammatical structures: the least

---

[5]Or, much more rarely, in bilingual contact situations. Whinnom (1971) has argued that, for social reasons having to do with the availability of the target language, no pidgin can develop in a bilingual setting. But his model does not take into account the possibility that speakers of the vocabulary-base language might deliberately withhold access to their language. Such deliberate withholding is attested for the 17th-century Delaware-based Amerindian pidgin and for Mobilian Jargon, and it probably also accounts in part for the emergence of Tây Bồi between French and Vietnamese speakers in Vietnam.

[6]The Caribbean creoles are considered by some creolists, e.g. Alleyne, Bickerton, and Thomason & Kaufman, to have developed in a process of abrupt creolization--that is, without going through a fully crystallized pidgin stage. Most of them have remained in contact with the vocabulary-base language under social circumstances that encourage convergence toward that language, i.e. decreolization. As a result, these creoles may safely be assumed to be more like the European vocabulary-base language now than they were when they first crystallized as creole languages. Caribbean creoles like Saramaccan which have *not* remained in contact with the vocabulary-base language show more African, and fewer European, structural features.

decreolized Caribbean creoles and European-vocabulary pidgins like Tok Pisin (Neomelanesian) show few or no universally marked features characteristic of European languages, but they do have marked features characteristic of the relevant substrate languages (African, Melanesian).

The lesson to be learned from these cases is that lexical correspondences, no matter how numerous and systematic they are, cannot stand alone as sufficient evidence of normal transmission, and hence of genetic classifiability. From a retrospective viewpoint, in order to rule out the possibility of nongenetic development, we must show systematic correspondences in grammatical as well as in lexical structures, and cognation in grammatical as well as in lexical morphemes.[7]

If we look at Ma'a from this perspective, we will not focus on the lexicon, because the Cushitic origin of the basic vocabulary is no longer in doubt. Ma'a thus belongs either in the Cushitic group or in no genetic group. And we will not focus primarily on the mere fact that Ma'a has some Bantu grammatical features, because most languages that acquire foreign grammatical features do so without losing their genetic continuity. Instead, the crucial question has to do with Cushitic grammar: does Ma'a have enough of it to qualify as a Cushitic language in the full genetic sense?

This question is addressed in Section 2 below, in a systematic comparison of Ma'a structures with characteristic Cushitic and Bantu grammatical structures. The method of comparison is primarily typological. The reason for this typological emphasis is that even "hard-to-borrow" features like inflectional affixes might show regular phonological correspondence with comparable affixes in one language, but close functional and/or positional correspondence with affixes in another, as a result of interference. So, for instance, the so-called "second genitive" case in Russian, a partitive con-

---

[7]The insistence on grammatical correspondences in languages that are claimed to be genetically related is of course not new; many historical linguists have emphasized the importance of such correspondences at least since Gyarmathi's time (late 18th century). But this aspect of genetic linguistics has sometimes been neglected because vocabulary is easier to elicit, easier to compare, and certainly easier to quantify than grammar.

struction, has as its marker a native Russian suffix -u , but the grammatical distinction between partitive and non-partitive genitives entered Russian through the influence of neighboring Finnic languages. Examples of this sort are common, and they show the need for extreme caution in interpreting historically the products of the most intensive contact situations, like the one in which modern Ma'a arose: solid evidence of a particular origin must be sought in structural features in which the potential source languages disagree. Where the sources agree typologically, a definite origin can be assigned only to morphemes that agree in form *and function* with one source or another.

In the case of Ma'a, we will see that in general its structures are similar to Cushitic structures only where Cushitic and Bantu are typologically similar. Where Cushitic and Bantu differ, Ma'a usually agrees with Bantu. Specifically, Ma'a corresponds to Cushitic in a few phonological units, syntactic structures, and derivational processes, and in one feature of the inflectional morphology. Otherwise Ma'a matches Bantu closely, and most strikingly in the inflectional morphology, where it has a complete and productive set of Bantu inflectional structures. Overall, few productive nonlexical structures in Ma'a can be shown to be of definite Cushitic origin, whereas, by contrast, many can be shown to derive from Bantu. Usually, as several authors have observed, the Bantu structures can be traced to the Bantu language Pare, a southern dialect of Asu, whose speakers have been in intimate contact with the Ma'a people since about the seventeenth century [Ehret & Nurse 1981:141-2]. The other Bantu source is Shambaa, whose speakers are now neighbors of the remaining Ma'a speakers in the northeastern corner of Tanzania.

2. Ma'a, Cushitic, and Bantu Structures

In the discussion that follows, I am basing my statements about Ma'a primarily on the five most useful published sources available to me: Ehret [1980], Copland [1933-34], Green [1963], Tucker & Bryan [1974], and Elderkin [1976]. The first four sources make use of primary data; Elderkin's analysis is based on secondary sources. All five sources together provide only a

fragmentary sketch of Ma'a grammar, but additional information has been provided by Christopher Ehret [personal communication, 1982].[8] For Cushitic grammar I am relying heavily on the very useful sketches in Bender [1976].[9] I will emphasize Iraqw and Dahalo, the two Southern Cushitic languages for which I have descriptions, since Southern Cushitic is the branch that Ma'a matches lexically.

2.1. Phonology. As far as phonology is concerned, the Ma'a inventory consists mostly of phonemes that are common in both Cushitic and Bantu (and, for that matter, in other language groups around the world). Given the fact that the mix of structures in Ma'a must reflect extensive contribution from both sources, the presence of such common phonemes in Ma'a cannot be ascribed definitely to either source. In particular, though these phonemes correspond regularly to identical or similar phonemes in other Southern Cushitic languages, their phonetic representations in Ma'a might just as well be due to the fact that Bantu languages also have such sounds as to direct inheritance from Cushitic. Evidence for phonological inheritance from Cushitic, or for interference from Bantu, must therefore rest on the presence or absence in Ma'a of phonemes that occur in only one of the two groups, and

---

[8]The only major sources that are not available to me are Meinhof [1906] and the source called FILE by Tucker & Bryan [1974]. But since the sources I do have make frequent references to these, it is unlikely that information crucial to my argument is missing. The Ma'a data that Ehret used in his 1980 study came from his own field work in 1967 and 1973 and from Bernd Heine's field work [Ehret 1980:11f.].

[9]This book contains sketches of the following Cushitic languages: Iraqw and Dahalo [Elderkin 1976]; Beja [R. Hudson 1976]; Highland East Cushitic [G. Hudson 1976]; Werizoid [Black 1976]; Afar [Bliese 1976]; Oromo (Galla) [Gragg 1976]; and Dasenech [Sasse 1976]. Other sources on Cushitic languages that I have used are Welmers [1952] and [1973] on Saho; Bender et al. [1976] on Hadiyya (Highland East Cushitic) and Oromo; Whiteley [1960b] on Iraqw; R. Hudson [1974] on Beja; and Tucker & Bryan [1966] on Cushitic languages in general (especially Galla, Somali, Awiya, Bilin, and Beja).

the strongest evidence will be phonemes that are universally marked.[10] Ma'a does have a few phonemes that provide such evidence, as we will see on examining the phonemic inventory in Table 1 (data from [Ehret 1980:113]).

Table 1. Ma'a phonemes

Consonants						Vowels	
p	t		č	k	ʔ	i	u
b	d		ɟ	g		e	o
^{m}b	^{n}d		^{n}ɟ	^{ŋ}g		a	
f	s	ɬ	š	x	h		
v	z			ɣ			
m	n		^{y}n	ŋ			
	r	l				Tones	
w			y			HIGH	LOW

The Ma'a inventory has one phoneme, the voiceless lateral fricative /ɬ/ , which is rather highly marked in universal terms and which is clearly a Southern Cushitic (SC) inheritance. Lateral obstruents occur in other SC languages, but such phonemes apparently do not occur in other branches of Cushitic, and they do not occur widely in Bantu. So in this respect, at least, Cushitic words are pronounced in Ma'a with a characteristic SC sound. Ma'a has two other phonemes which, though not universally marked, are more common in Cushitic than in Bantu: /ʔ/ and /x/ . Of these /ʔ/ is far more common than /x/ as a phoneme in Cushitic. Other than these, no Ma'a phonemes look like promising candidates for specifically Cushitic inheritances.

One subphonemic phonetic feature of Ma'a seems likely to be due to Bantu influence. The voiced stop phonemes /b d ɟ/ have implosive pronunciation

[10]By "universally marked" I mean, here, a phoneme or phoneme type that is uncommon in languages of the world--so uncommon that it can probably be considered relatively unlikely to arise spontaneously, and thus more likely to be present as a result of inheritance or convergence.

[Ehret 1980:130], and even /g/ is pronounced with weak implosion [Ehret, p.c. 1982]. Now, in many Cushitic languages implosives (usually only /ɗ/) and ejectives occur as phonemes. Dahalo, for instance, has several glottalized phonemes of each type, and Ehret [1980] reconstructs implosives for Proto-SC. The fact that the implosive pronunciation is not distinctive in Ma'a would therefore suggest that Ma'a matches Bantu in this respect, since Bantu languages often have implosives as allophones of pulmonic voiced stop phonemes. But Ehret comments that the implosive feature is not "attributable to Bantu and, where it occurs in East African Bantu languages, it can be laid to pre-Bantu habits of articulation" [p.c. 1982]. Nevertheless, the loss of glottalization as a distinctive feature in Ma'a still requires an explanation, and of course it is quite possible for a Bantu language to have acquired allophonic implosives from earlier interference and then later to have influenced Ma'a to lose glottalized stops as a distinctive phoneme type. Since both Pare and Shambaa have allophonic implosive pronunciation of voiced stops [David Odden, p.c. 1982], this seems the most likely source of the Ma'a implosives.

Several features of Ma'a phonology can definitely be ascribed to Bantu influence. The most striking one is the highly marked series of four prenasalized voiced stops, /ᵐb ⁿd ⁿɟ ᵑg/ , which entered the language first in Bantu loanwords [Ehret 1980:113].[11] The other non-SC phonemes in Ma'a are not universally marked, but the Bantu influence is clear. Ma'a acquired the phonemes /v ɟ ɣ/ in Bantu loanwords, and Bantu /ɟ/ and /ɣ/ caused the original Ma'a allophones [ǰ] and [ɣ] to assume phonemic status. Moreover, the *absence* in Ma'a of several marked phoneme types is noteworthy, because Bantu also lacks them. Since all of these occur in other SC languages and are reconstructed by Ehret for Proto-SC, their absence in the Ma'a cognates

[11]Although Ehret [1980] reconstructs a series of prenasalized voiceless stops for Proto-SC, the prenasalized voiced stops in Ma'a do not correspond to these. The only SC language with prenasalized voiceless stops is Dahalo (though Ma'a has corresponding nasal + stop clusters in some environments), so the reconstructed set seems rather dubious for Proto-SC, especially since Dahalo, like Ma'a, has undergone heavy interference from neighboring languages--with, however, very different results.

strongly suggests elimination under Bantu influence. Ma'a has lost pharyngeal fricatives, which are common in Cushitic in general and occur in Iraqw and Dahalo in particular. The labialized dorsal stops that are characteristic of Southern, Northern, and Central Cushitic do not occur in Ma'a. The Proto-SC ejectives and retroflex stops are gone. Besides these marked phoneme types, Ma'a has lost the contrast between long and short vowels that characterizes most Cushitic languages, including Proto-SC. I cannot judge the probability of Bantu influence on this loss, since, though many Bantu languages lack phonemic vowel length [Welmers 1973:25], I have no specific information about this feature in Pare or Shambaa. As for prosodic features, Ma'a has two phonemic tones and predictable stress on the first stem syllable [Ehret, p.c. 1982]. According to Ehret, Bantu loanwords contributed to the development of tone in Ma'a; he does not reconstruct lexical tone phonemes for Proto-SC, and he considers Iraqw to have neither phonemic tone nor phonemic stress [p.c. 1982]. Dahalo, like Ma'a, has lexical tone distinctions. In general, Cushitic languages have phonemic stress, while Bantu languages have phonemic tones.[12] Finally, at least one phonotactic feature of Ma'a may be due to Bantu influence. Quite recently, to judge by the chronology indicated by other sound changes, Ma'a lost all word-final and verb-stem-final consonants; this change, according to Ehret, may have been motivated by "the Bantu pattern in which no word could end in a consonant" [1980:110].

In sum, the Ma'a phonemic inventory has two phonemes that are characteristic of Cushitic but not of Bantu, namely /ɬ/ and /ʔ/ . But in most phonological features in which Bantu and Cushitic differ, Ma'a matches Bantu rather than Cushitic. These features include the presence of prenasalized voiced stops and of phonemic tones, and the absence of pharyngeal fricatives, labialized dorsal stops, ejective and retroflex stops, and final consonants. Ma'a also differs from typical Cushitic structure in its lack of distinctive

[12]Welmers [1973:78] observes that, though Cushitic languages are not in general tonal, some are analyzed as having phonemic tones. But, he says, Saho pitches are predictable if stress is treated as phonemic, and he believes that this is likely to be true for at least some other Cushitic languages too.

vowel length, but I do not know the status of vowel length in the relevant Bantu languages. In addition to these general structural divergences from Cushitic, Ma'a has acquired the phonemes /v ɟ ɣ/ under Bantu influence.

2.2 Morphology. When we turn to the morphology, we find a sharp distinction between inflectional and derivational patterns as far as their historical sources are concerned. The inflectional system of Ma'a is almost entirely of Bantu origin, but the derivational affixes seem to be about evenly divided between Bantu and Cushitic suffixes. Typologically, however, Bantu and Cushitic agree in the particular kinds of derivational processes attested in Ma'a. So, in the morphology as in the phonology, where the two groups differ typologically Ma'a almost always resembles Bantu rather than Cushitic.

2.2.1. Nominal inflection. In the nominal subsystem the most important grammatical categories are noun classification, number, pronominal possession, and adjectival agreement. Almost everyone who has written about Ma'a has emphasized the presence of Bantu noun-class prefixes, both on nouns and as agreement markers on adjectives, verbs, and certain particles. These prefixes, usually from Pare but occasionally (judging by the phonological shape) from Shambaa, fall into the standard Bantu classes 1/2, 3/4, 5/6, 7/8, 9/10, 11, 13, 14, 15, and 17 [Ehret 1980:131]. Most writers have also remarked on the inconsistent use of the prefixes. Bryan reports, for instance, that one informant gave mu-haraza as the citation form for 'river', but later used the unprefixed form haraza in conversation [Tucker & Bryan 1974:192]. Moreover, some adjectives are invariable, i.e. they do not agree with the nouns they modify , and some nouns are also invariable.[13]

---

[13]Although the type of data is limited--only words of Cushitic origin, and (presumably) only citation forms given--a count of the noun classes represented in Ehret's published Ma'a data [1980] gives a rough idea of the level of attestation for the various classes: 134 nouns have no prefix (but some of these would be prefixless in the singular in Bantu too); 239 nouns have class prefixes, including m(u)- (class 1 or 3; 60 nouns), mi- (cl. 4; 1 noun), i- (cl. 5; 72), ma- (cl. 6; 17), ki- (cl. 7; 35), N- (cl. 9/10; 17), lu- (cl. 11; 20), ka- (cl. 13; 4), (v)u- (cl. 14; 10), and ku- (cl. 15; 3). This last class contains all verb infinitives, so the actual number of Ma'a nouns in ku- is no doubt very large, even though Ehret's

But none of the Ma'a sources highlights the point of major significance for the genetic question. The crucial fact about the Ma'a system of noun classification is not that it resembles Bantu and has morphemes of Bantu origin, but rather that it is so *unlike* Cushitic that it cannot possibly be viewed as a continuation, or even as a partial remodelling, of an earlier Cushitic classificatory system. Typologically, Bantu and Cushitic languages differ in every respect in the ways in which they classify nouns.

First of all, the semantic content of the Bantu noun classes is quite varied. Typical semantic features used for classifying nouns are human; trees and other plants; animals; long, thin objects; paired objects; items of material culture; diminutives; augmentatives; abstract nouns; and the verbal infinitive. Cushitic and other Afroasiatic languages, by contrast, have just two noun classes, based on the semantic feature of biological gender: masculine and feminine.[14] Second, the formal expression of noun classification in Bantu is entirely prefixal, with paired singular/plural prefixes on nouns and agreement prefixes (sometimes different from the noun prefixes) on adjectives and other modifiers, on verbs, and on some particles, e.g. the associative particle used in genitive constructions. Cushitic languages do not have a uniform set of gender affixes on nouns themselves, though in some of the languages the gender of at least some nouns is reflected in the form of the noun. Examples are found in Saho, in which stressed nouns and unstressed nouns ending in a consonant are regularly masculine, while unstressed nouns ending in a vowel are feminine [Welmers 1973:222]; Dasenech, in which distinct masculine and feminine singulative suffixes are added to collective nouns [Sasse 1976:203]; and Afar, which has a masculine vocative suffix opposed to a feminine vocative suffix [Bliese 1976:150]. In some Cushitic languages modifiers agree with head nouns in gender, usually by the presence

---

data did not contain many. Ehret observes that prefixless nouns belong in class 9/10 for purposes of concord [1980:131].

[14]As in most languages with noun classification systems, Bantu and Cushitic noun classes contain many nouns which do *not* meet the semantic criterion, but which are classified arbitrarily or merely by their phonetic shape. For instance, not all feminine Cushitic nouns refer to female creatures, and Bantu languages have inanimates in the "animals" class.

or absence of a feminine suffix. Many of the languages distinguish masculine and feminine third person singular verb forms in combined tense-aspect/subject suffixes, and most have distinct masculine and feminine third singular pronouns. Cushitic languages clearly do not all treat gender in the same way, and, according to Greenberg [1963:45], some of the western languages lack grammatical gender entirely. Nevertheless, all but one of the languages for which I have data have a masculine/feminine distinction at least in third singular free pronouns.[15] Cushitic languages that distinguish gender elsewhere as well typically have suffixes that vary according to gender. As far as the Southern Cushitic languages are concerned, Elderkin mentions a masculine/feminine distinction in Iraqw only in the free pronouns, while Dahalo is said to have natural gender which is marked at least in free pronouns, adjectives (by suffixes), and demonstratives. Proto-SC nouns, according to Ehret, were marked for gender by suffixes attached to the noun stem [1980:48].

Ma'a noun classification follows the Bantu pattern faithfully. It has prefixes of Bantu origin attached to Cushitic noun stems that are "associated with the same classes as the semantically comparable Bantu stem would be" [Elderkin 1976:289]. Most adjectives take the appropriate noun-class prefixes for the nouns they modify, and verbs take concordial prefixes to agree with full-noun subjects and objects. In genitive constructions, the 'of' particle -a is combined with the noun-class prefix appropriate for the head (possessed) noun, as in afá yá mɔhɛ́ 'a person's goat' (lit. 'goat class=9-of person') vs. afá ɟá mɔhɛ́ 'a person's goats' (lit. 'goat class=10-of person) [Tucker & Bryan 1974:200]. In all, as noted above, fifteen noun classes are attested in noun and concordial prefixes. Gender markers of Southern Cushitic origin do occur in Ma'a; compare, for instance, iŋi 'brother' and iŋinta 'sister' or i'alú 'sheep' and i'alé 'ram'. But, though common, they are not productive [Ehret, p.c. 1982]. The language's current pattern of noun classification is Bantu, and it has replaced an older

---

[15]The exception is Dasenech, which does distinguish masculine and feminine gender in nouns, adjectives, and verbs. So, all the Cushitic languages described in my sources have grammatical gender.

Cushitic pattern, which is now attested only in relic word pairs. Even the personal pronouns, which are of Cushitic origin, lack the usual Cushitic masculine/feminine gender distinction.[16]

The number category also reveals a deep typological division between Bantu and Cushitic, and here again Ma'a is Bantu in type. As mentioned above, number is marked in Bantu nouns by the noun-class prefixes, which occur in paired singular/plural sets. It is therefore an obligatory category, both in the noun itself and in other words that agree with the noun. In Cushitic, by contrast, the category of number is not obligatory, at least for some nouns. When plural is marked, a wide variety of markers, apparently with lexically governed distribution, is used in many of the languages, including Iraqw and Dahalo. For a given language, these markers may include several suffixes, an infix, accentual alternations, reduplication of the final consonant, and a change in vowel pattern. More striking still is the marking of number in some nouns by adding a singulative affix instead of a plural one. Compare,

---

[16]Ehret [p.c. 1982] observes that "there is no specific evidence for attributing the loss of gender in Ma'a to Bantu influence," and that "in the pronouns the masculine forms were generalized as might be expected with normal processes of language change." It is true, of course, that a Bantu-style system of noun classification is not inherently incompatible with a Cushitic system based semantically on biological gender. A language *could* have both. But though exceptions certainly exist, most languages in the world do not mix biological-gender (or animacy) classification with other kinds of noun classification. Bantu, in any case, does not. So the circumstantial evidence for the loss of Cushitic noun classes in Ma'a *because of* the rise of Bantu noun classes is very strong, particularly in light of the fact that Ma'a agrees typologically with Bantu, and differs from Cushitic, in so many other respects: if, as Ehret believes, Ma'a was once an ordinary SC language, it has shifted typologically toward Bantu in all its grammatical systems. And if it has shifted toward Bantu in other subsystems, why not assume that Bantu interference was the causal factor in this instance as well as in the more obvious instances (like the agreement patterns or the phonemic tones)? Similarly, "normal processes of language change" may not demand explanations as dramatically as apparently abnormal ones do, but that does not mean that we should not seek explanations for them. Since even the most natural changes often fail to occur, it is never inappropriate to ask why a particular change happened when it did; and if a reasonable explanation is available, it should not be rejected merely because similar changes have occurred under different antecedent conditions.

for instance, Dahalo ʔúšò 'male elephant' : plural ʔùšàšè and kì:dzò 'old man' : plural kí:dzò:mà , but á:dʒù 'lung fish' : singulative á:dʒùmè [Elderkin 1976:292].

Ma'a marks number by the paired Bantu noun-class prefixes, both in nouns and in other words that agree with nouns. Examples (from Green [1963]) are mu-'o m-gititu 'small mouth' (-'o 'mouth' and -gititu 'small') : plural mi-'o mi-gititu . Ehret [p.c. 1982] cites two Ma'a nouns of Cushitic origin that have typical SC number marking: nihi 'animal' : plural nihena , and tambala 'snail' : plural tamba (the singular of 'snail' has a suffix -a ; stem-final -l has been lost in the suffixless plural by a regular sound change). Ma'a has few, if any, other traces of Cushitic number marking [Ehret 1980:48], though my sources list several nouns with a Cushitic suffix -no that indicates mass quantity [Ehret, p.c. 1982] and apparently functions sometimes as a quasi-plural: iɬare (Bryan) or ɬare (Green) 'cloud' : plural ɬareno (Green), maɬareno (Green, Bryan), or maɬare (Bryan); ŋgile 'bee' : ŋgileno 'swarm of bees' (Green); kunge 'Kweme nut' : plural kungeno (Green); and 'i'alú 'a sheep' : plural ale:no (Meinhof [1906]) or ma'alú (Bryan) (Green [1963:185]; Tucker & Bryan [1974:207]). Even here it is noteworthy that the variant forms for 'clouds' and 'sheep', with only the Bantu plural prefix ma- , occur in the most recent source;[17] Ehret [p.c. 1982] confirms that the -no suffix is losing ground in the language. (The hybrid form maɬareno 'clouds', with both the Bantu plural prefix and the Cushitic plural suffix, suggests that one mechanism for the replacement of Cushitic patterns by Bantu ones may have involved double-marking of nouns at one stage.)

Case inflections do not occur in Ma'a or, in general, in Bantu. It is hard to say whether or not this lack puts Ma'a in significant contrast with Cushitic, however. Most Cushitic languages seem to have at least a two-case distinction, between a subject case and an accusative or "absolute" case; but

[17]It should also be noted, however, that Bryan collected her Ma'a data in 1959, and Green does not say when he collected his, so that these two sources might represent contemporaneous usages. The Meinhof data, of course, is the oldest of the three.

I have no information about the category of case in Southern Cushitic. In one construction involving case Ma'a does differ significantly from the Bantu languages of the area. Shambaa, like Swahili, uses a locative suffix -ni beside a locative construction with the associative particle -a (see below) to express location. For instance, 'on the mountain' is mwima-ni or zuu ya mwima (literally 'aboveness of the mountain'). By contrast, Ma'a uses a preposition in locative expressions, as in aná longorí 'on the mountain'.[18] I have little information about locatives in Cushitic, so I cannot tell whether the Ma'a construction agrees with those in Cushitic languages or not (though most branches in the group have postpositions rather than prepositions).

Pronominal possession is marked on nouns in most Cushitic languages by a set of suffixed possessive pronominals which is etymologically related entirely or in part to the set of independent pronouns. In Iraqw these suffixes are added to the noun after a suffixed class marker (a syntactic class marker that determines the person and number of verb concord). In Bantu, pronominal possessors are full pronouns which, like other adjectives, follow the noun and take noun-class agreement prefixes. Ma'a pronominal possessors are of Cushitic origin and are apparently suffixed, as in Cushitic. Goodman states that they do not take concordial agreement [1971: 245], but Tucker & Bryan observe that this is true only when the possessives are used attributively. They do take the appropriate Bantu concord prefixes when used predicatively, as the following examples show: yá iŋí ni mu-ɣɔ́ 'this child is mine' (DEM child BE class=1-my) and yá va-iŋí ni va-kánù 'these children are ours' (DEM class=2-child BE class=2-our). Compare attributive use in mu-harega go 'my arm' (class=3-arm my), i-ɬɛ́ ɣɔ́ 'my name' (class=5-name my), and ki-kire go 'my stool' (class=7-stool my) [Tucker & Bryan 1974:202]. The Ma'a attributive possessive construction thus lacks the concord prefix and connective we would expect in a Bantu language, as e.g. in Swahili wa-toto wa-a-ŋgu 'my children' (class=2-child class=2-

[18]This information about Ma'a locatives, and the Ma'a and Shambaa examples, were provided by Christopher Ehret [p.c. 1982].

CONNECTIVE-my) vs. vi̇-su vi-a-ŋgu 'my knives' [Gleason 1955:48]. The attributive construction in Ma'a is similar to analogous constructions in Cushitic and is, in fact, the only inflectional pattern in Ma'a whose origin is clearly Cushitic. Moreover, when attached to kin terms pronominal possessors beginning in a vowel are frequently preceded by a Cushitic connective morpheme -r- [Ehret, p.c. 1982], so that even the morphophonemic behavior of the possessives is Cushitic.

2.2.2. Verbal inflection. Verbs in both genetic groups, Cushitic and Bantu, inflect for the same general categories: tense/aspect, person, number, and noun class. However, the two groups differ sharply in the actual verb construction. Typologically, Cushitic verbs tend to be flexional, with partially or wholly unsegmentable tense-aspect/subject affixes, while Bantu is agglutinative. Thus, the Iraqw second person plural marker is composed of the second singular suffix -t plus -a (plural + present/future) or -e (plural + past); similarly, the Lowland East Cushitic language Oromo (Galla) has a third singular masculine imperfective suffix -a opposed to a third singular masculine perfective suffix -e .[19] Most Cushitic languages, like Iraqw and Oromo, have suffixed tense-aspect/subject markers. In Bantu, by contrast, tense/aspect and subject agreement are expressed by separate affixes--subject prefixes, tense/aspect prefixes and often suffixes too. Another difference is that object marking is rare in Cushitic verb morphology, and it does not occur, apparently, in Southern Cushitic,[20] but Bantu languages have full sets of object prefixes as well as subject prefixes.

The Bantu third person agreement prefixes occur in paired singular/plural sets according to the class of the noun referent; these are prefixed to nouns and adjectives. First and second person prefixes do not inflect for

---

[19]Bender et al. use the traditional terms 'imperfect' and 'perfect' to refer to these two aspects in Oromo, and probably they correspond in the typical Afroasiatic way to Elderkin's tense-labeled categories in Irawq.

[20]*Pace* Elderkin [1976:294f.], who speaks of a particle which either precedes the verb or is suffixed to it. I take it that this is not a true suffix when it follows the verb, but rather still a particle. An example is ʔákùwàté lùβò ~ lùβókùwà 'I hit(past) (someone) for you'.

noun class. Cushitic verbs typically show a masculine/feminine distinction in the third singular only, not in the plural, and even this minimal gender distinction seems to be lacking in Iraqw and Dahalo. Number agreement with the subject is marked either flexionally, as in Dahalo third singular -i vs. third plural -e: , or agglutinatively, as in Iraqw third singular -/i/ vs. third plural -/i + r/ .

Tense/aspect in Cushitic generally follows the familiar Afroasiatic pattern in which the major division is between the imperfective and perfective aspects. Subcategories of tense and, especially, aspectual functions are expressed by derivational processes like suffixation and initial-syllable reduplication. Bantu tense/aspect systems also distinguish completive from incompletive aspect [Welmers 1973:350,384], and they have in addition a paradigmatically related set of prefixes expressing features like simple present, present continuous, future, immediate past, remote past, stative (or perfect), and mood (e.g. conditional). Tense is in general a more important inflectional category in Bantu than in Cushitic.

Cushitic has a number of other characteristic morphological features, mostly derivational ones, in its verb system. Among these are prefixed reduplication as an intensive or frequentative formation, verb negation by means of a suffix or a prefix, and passive and causative formation by suffixation. Bantu verbs often share the last two features, e.g. negation in Kinyarwanda by means of a prefix, as in /nhi-ba-geénd-â/ 'they are not going' (NEG-they-go-ASPECT) [Kimenyi 1978:313], and passive formation by means of a suffix -wa , as in Venda funa 'love' : fun(i)wa 'be loved'.

In verb morphology, Ma'a patterns with Bantu wherever Bantu and Cushitic differ. Its verb morphology is agglutinative, and the inflectional morphemes are arranged in the common prefixal Bantu pattern:

(NEG +) Subject + Tense + (Object +) ROOT (+ extension)

(In Copland's text [1933-4:243,245], however, there is no object prefix if a full-noun object is present in the sentence.) Examples are ve-ne-tu-ifi 'they (the Masai) will destroy us' (class=2-FUT-us-destroy) and te-tu-ta-zaxo 'we will not hold' (NEG-we-FUT-hold). This ordering

contrasts sharply with the predominantly suffixing Cushitic patterns. Moreover, the inflectional affix morphemes themselves are all Bantu in origin, and most of them can be identified with Pare and/or Shambaa affixes. Ma'a even has some nonautomatic morphophonemic alternations characteristic of Bantu verb morphology, most notably the distinction in the negative between first person singular si- , e.g. si-'ánthù 'I do not cook', and the other persons with a separate negative prefix, e.g. tɛ-tú-'ánthù 'we do not cook' (NEG-we-cook; the present tense marker is zero) [Tucker & Bryan 1974:204]; compare Swahili si-anguki 'I do not fall' and ha-tu-anguki 'we do not fall' [Loogman 1965:200]. The class 1 (personal singular) verb agreement marker also varies as in Bantu. Ma'a has a- (or e- , in an alternation also found in Asu) as subject, as in Mwa-gilu a-ka-ba' 'Elder said' (class=1-elder class=1-PAST-say) [Copland 1933-4:244], and m- as object, as in n-áà-m-ma 'I have hit him' (I-PERFECT-class=1-beat) [Tucker & Bryan 1974:201]. Compare Swahili a-me-m-piga 'he has beaten him' (class=1-PERFECT-class=1-beat) [Gleason 1955:26].

The inflectional morphemes of Bantu origin in Ma'a verbs include subject/object markers for all three persons, both singular and plural, with a variety of noun-class prefixes attested for the third person; tense prefixes ta- (present or future), ne- (future), aa- (perfect), ka- (past); a tense/aspect suffix -ye (past NEG and conditional); a conditional prefix ku- (in the tense/aspect position); and a negative prefix si-/te- .

2.2.3 Derivational morphology. As mentioned above, derivational affixes in Ma'a seem to be divided about evenly between suffixes of Bantu origin and suffixes of Cushitic origin. But although the *morphemes* come from both sources, the productive derivational *patterns* in Ma'a are all, as far as I can tell, patterns in which Cushitic and Bantu agree, at least in ordering: all the processes involve suffixation. Cushitic derivational processes that are not found in Bantu, most notably prefixed reduplication and infixation, are apparently not productive in Ma'a, though lexicalized frequentative reduplicatives are rather common in Ehret's Ma'a data [1980], e.g. fufu 'to catch breath, rest' (from Proto-SC *fook'- , with prefixed reduplica-

tion and loss of the stem-final consonant).

The most common causative formative in modern Ma'a, according to Ehret, is probably the Cushitic suffix -ti [1980:63], e.g. -gugúlu 'run' : -gugulúti 'drive away'. The productivity of this suffix is indicated by (among other things) the fact that it has been added to stems since the recent and probably Bantu-influenced sound change that eliminated final consonants. Another Cushitic verb extension, -'V , is also still quite productive, and three others, -mu , -u , and -au , were productive until fairly recently [Ehret, p.c. 1982]. But a number of Bantu verb extensions are also quite productive in Ma'a and are used with Cushitic as well as with Bantu verbs. For instance, in Ehret's Ma'a data [1980] I find seventeen Cushitic verbs with the causative suffix -ija , ten with the stative -Vka , seven with the passive -wa , four each with the reciprocal -ana and the intensifier -ya , and a few other suffixes in one or two verb forms each. (Note that the entries in this data list are generally citation forms only, not sets of inflected and/or derived forms; still, the occurrence of the Bantu suffixes in the list gives some indication of their penetration into the Cushitic vocabulary.) Finally, Ehret [p.c. 1982] observes that "the Cushitic-derived amplificative of both nouns and verbs, -ša , is a very productive suffix" with "no direct parallel in neighboring Bantu languages." He also points to an apparently still productive adjective-forming suffix -'i of SC origin. Ma'a thus has a number of Cushitic derivational affixes beside Bantu affixes, but typologically there is no contrast between the two groups in this grammatical subsystem.

2.3 Syntax. As in the derivational morphology, the syntactic structures of Ma'a are divided between patterns of Cushitic origin and patterns of Bantu origin. But unlike the derivational patterns, two of the relevant syntactic differences between Cushitic and Bantu are typologically significant. First of all, Cushitic languages have dominant SOV word order. Most of them also have other word-order features often associated with SOV languages (cf. Greenberg [1966], e.g. postpositions and Adjective-Noun word order; but Noun-Adjective order is dominant in Iraqw and Dahalo, and in some Lowland

East Cushitic languages as well. Ma'a has as dominant word order patterns SVO and Noun-Adjective, and these are typically Bantu.[21] Ma'a also has prepositions rather than postpositions, e.g. locatives he 'to' and na 'from', both from Proto-SC verb roots [Ehret, p.c. 1982]. Prepositions are expected in an SVO language, but these particular ones cannot be attributed directly to Bantu influence, since, according to Ehret, in Asu and Shambaa "verbs of movement do not generally require the insertion of a directional marker" [p.c. 1982].

Bantu and Cushitic genitive constructions also differ, and Ma'a uses both types, but typologically the two are not far apart. Bantu uses an associative particle -a with appropriate noun-class prefixes. In Cushitic a possessor noun sometimes takes a case suffix or a subordinative particle, but sometimes the possessed and possessor nouns are simply juxtaposed without special marking of the construction. Iraqw and another SC language, Burunge, use the latter Cushitic construction, and so does Ma'a; but Ma'a also has the Bantu construction with prefixed -a between the two nouns. Ehret [p.c. 1982] remarks that the Cushitic usage seems almost as common as the Bantu pattern in Ma'a.

Another construction type in which Ma'a uses both Cushitic and Bantu patterns is the copula. The Bantu morpheme ni is used in Ma'a as the copula and to introduce the agent of a passivized verb [Goodman 1971:248], but, according to Ehret [p.c. 1982), forms of SC verbs for 'to be' are more common, and the copula relationship is obligatorily marked in Ma'a--unlike Asu and Shambaa, which frequently omit the marker. In this syntactic feature, however, as in the genitive construction, the typological difference between Bantu and Cushitic is not very great, so the mixture of construction types in Ma'a does not in any case seem likely to cause serious communication difficulties for speakers using different patterns.

The same point can be made for one of the two syntactic features that

---

[21]Ehret [p.c. 1982] believes that "earlier SOV order can be internally reconstructed for pre-Ma'a," and that the shift to SVO order is probably recent, dating from a period when Bantu influence had become significant.

have come to my attention in which Ma'a clearly matches Cushitic but not Bantu. Although most (but not all) Bantu languages have a class of words that can be identified as adjectives, the class is quite small in most of the languages [Welmers 1973:271]. But in Ma'a, as in other SC languages, this class is a large and important one [Ehret, p.c. 1982].

The other exclusively Cushitic feature of Ma'a syntax is the means of expressing possession, a "normal transitive verb -lo 'to have', fully conjugatable and occurring in the normal syntactically verbal contexts" [Ehret, p.c. 1982]. Ehret points out that this usage matches the pattern in SC languages but contrasts sharply with the patterns in all the nearby Bantu languages. These languages, he says, "combine person markers with a connective (a-na 'he-with') for expressing having in present action contexts and use a form 'to be with' for other tense/aspects." Like the SVO/SOV word order distinction, this difference in expressions of possession is typologically significant; but, while Ma'a agrees with Bantu in word order, it agrees with Cushitic in this feature.

2.4 Lexicon. Although the basic vocabulary of Ma'a is, as already mentioned, primarily of Cushitic origin, and although much cultural vocabulary is also Cushitic, the language has a very large number of Bantu words as well--at least 50% of the vocabulary, according to Ehret [p.c. 1982] (the other information in this section is from the same source). The earliest layer of Bantu words pre-dates the Bantu-influenced grammatical changes and shows features like Cushitic suffixes that are no longer productive in Ma'a. A later influx of Asu words constitutes the largest set of Bantu words, and more recently these have been supplemented, and sometimes supplanted, by Shambaa words. Included in the vocabulary of Bantu origin are many verbs and some body parts as well as cultural words.

Ehret notes, however, that in spite of the large proportion of Bantu words the lexical semantics of the SC portion of the vocabulary is still Cushitic, at least in some lexical fields. In particular, Ma'a uses "a five-part color division--black, white, red, yellow, green--with each expressed by a simple adjective," while Bantu languages typically "express only black,

white, and red with simplex terms."

Table 2 summarizes the features discussed in Section 2. Examination of this table shows that Ma'a agrees with Bantu rather than Cushitic in far more features in which the typological disagreement between Bantu and Cushitic is clearly significant.

3. How Did Ma'a Get Mixed?

The comparative discussion above substantiates the usual claim that Ma'a grammar--and most clearly the inflectional morphology, which is the grammatical subsystem usually assumed to be least susceptible to foreign interference--is mostly of Bantu origin. More to the point for the genetic question, it shows that Ma'a has few definite and productive Cushitic grammatical features, and some of the ones it does have (like the obligatory copula) do not differ much from Bantu typologically. If we set the paucity of systematic nonlexical Cushitic features against the presence of so many Bantu categories, grammatical morphemes, and even allomorphy, we must consider the possibility that Ma'a is not a changed later form of a single Cushitic parent language. Before we look at the available information about the history of the Ma'a people, we should ask what conclusions can be drawn from the linguistic evidence alone. That is, since genetic relationship of languages must in principle be established solely on the basis of the linguistic data, what are the possible routes of nongenetic development? And, does the structure of Ma'a fit any of these possibilities?

According to the model of language contact phenomena developed by Terrence Kaufman and me [1975 and Forthcoming], there are three basic lines of historical development that may culminate in a language native to a speech community whose basic vocabulary is demonstrably not from the same source as its grammar. First, speakers of a language A may shift to another language B under social conditions so pathological that only the vocabulary of B is successfully acquired. The languages most likely to have arisen through such a process are the Caribbean creoles, which emerged when enslaved Africans were forced (because they were put into linguistically diverse groups) to shift away from their various native languages, but without the opportunity

Table 2. Summary of typological agreements between Ma'a and Bantu *or* Cushitic

Cushitic		Ma'a		Bantu
/ɬ/	=	/ɬ/		-----
/ʔ x/	=	/ʔ x/		(/ʔ x/)
phonemic glottalics		allophonic implosion	=	allophonic implosion (Pare/Shambaa)
[ʃ ɣ]		/v ɟ ɣ/	=	/v ɟ ɣ/
-----		/^{m}b ^{n}d ^{n}ɟ ^{ŋ}g/	=	prenasalized voiced stop phonemes
/ʕ ɲ/		-----	=	-----
labialized dorsal phonemes		-----	=	-----
ejective stop phonemes		-----	=	-----
retroflex stop phonemes		-----	=	-----
phonemic vowel length		-----	$\stackrel{?}{=}$	-----(?)
C#		no C#	=	no C#
fem:masc; especially suffixes		Bantu N classes; prefixes	=	Bantu N classes; prefixes
optional sg:plu; singv.		oblig. sg:plu; no singv.	=	oblig. sg:plu; via prefix sets
postpositions	≠	LOC via prepositions	≠	LOC via -ni or particle -a
N-possessive pronoun	=	N possessive pronoun		N -a possessive pronoun
flexional; suffixes		agglutinative; mostly prefixes	=	agglutinative; mostly prefixes

Table 2. (continued)

Cushitic		Ma'a		Bantu
SOV		SVO	=	SVO
GEN = N N(-gen.)	(=)	GEN = {N N / N -a N}	(=)	GEN = N -a N
obligatory copula	(=)	obligatory copula		optional copula
many adjectives	(=)	many adjectives		few adjectives
trans. verb 'have'	=	trans. verb 'have'		no trans. verb 'have'
vocabulary	=	± 50%; most basic vocabulary		vocabulary
5-part basic color system	=	5-part basic colors		3-part basic color term system

Note: (=) marks agreements in which the typological difference between Bantu and Cushitic is minor.

and/or the motivation to learn the slavemasters' language as a whole. Second, a well-established pidgin may get nativized (i.e. become a creole). An example is Tok Pisin (Neomelanesian), which has an English lexicon but markedly non-English grammar, and which is now learned increasingly as a first language by children in New Guinea. And third, speakers of a language A may maintain their own language for a long period of time in the face of interference from B so great that only the vocabulary of A is successfully maintained, in some cases only as a special alternative (secret) vocabulary. In such a case, we believe, all speakers of A will be bilingual, at least at one stage in the language's history. Anglo-Romani is an example: English gypsies all speak English, but they maintain Romani vocabulary as a secret code. That is, their version of the language is Romani only in its lexicon; the phonology, morphology, and syntax are all English, so that for them Romani is, in effect, a lexical substitution code. (The Romani spoken elsewhere, e.g. in Russia, is normally transmitted Romani--an Indic language--in grammar as well as in vocabulary.) All three of these processes result in languages that have arisen outside of normal transmission; their origins are therefore nongenetic.

If we oversimplify the Ma'a case for the moment, and suppose that all its vocabulary is Cushitic and all its grammar Bantu, we can imagine five possible origins for the language from these three lines of development, depending on whether language A is assumed to be Bantu or Cushitic. But of these five possibilities four can be eliminated as being in fact impossible or implausible.

If Ma'a were a case of unsuccessful shift from A to B, where A is Cushitic and B is Bantu, then we would expect Ma'a vocabulary to be basically Bantu: as Kaufman and I have argued, lexicon is acquired first in any process of language shift. But Bantu lexical morphemes in Ma'a are not common in the basic vocabulary--not, for instance, in a hundred-word Swadesh list--so that Ma'a could not reasonably be supposed to be the product of an imperfect shift from a Cushitic to a Bantu language.

Similarly, Ma'a could not be the product of massive interference from a Cushitic language B in a Bantu language A, because in this case we would

again expect Ma'a basic vocabulary to be primarily of Bantu origin. English is often said, for instance, to have a vocabulary that is mostly of French and/or Latin origin, but the basic vocabulary of English actually has only a scattering (ca. 7%) of Romance loanwords.

A more interesting argument might be made for interpreting Ma'a as a result of imperfect shift from a Bantu language A to a Cushitic language B. In this case we would expect, as we indeed find, Cushitic basic vocabulary. Problems arise, however, when we consider the grammar, in particular the morphology. For the most part, Ma'a inflectional morphology can be traced to a single Bantu language, Pare. The inflectional morphology is somewhat simpler than that of Pare, for instance in the optional use of some noun-class prefixes, but typologically it does not differ from ordinary Bantu. Its noun classification system is not significantly reduced, and both noun and verb inflections even include some Bantu allomorphy, as described above.

Now, this picture corresponds to no known shift situation. The closest analogue is perhaps the case of the Caribbean creoles, whose speakers originally acquired vocabulary from one of several B languages during a process of shift. In those instances the shift itself is explained by the desperate need of Africans from diverse linguistic backgrounds to communicate with one another. Without such a pathological social situation, it is hardly conceivable that any group would attempt a shift to a language that was available as a model only in the most superficial way.

Moreover, the most famous characteristic of Caribbean creole grammar--the near-total absence of inflectional morphology--is explained by the diversity of the languages that entered into the formation of the creoles. African *pidgins* which have arisen among speakers of exclusively Bantu languages show morphological reduction far more sweeping than the minor sorts of simplification observed in Ma'a (see e.g. Polomé [1971:58], on Katanga Swahili, and Nida & Fehderau [1970] on Kituba). The inevitable conclusion, even if we did not know that Pare existed, is that the elaborate Ma'a morphology must come from a single source, or perhaps from one primary source with later influence from a second of the same group--but not from several sources that were all influential at the same time. So, if Ma'a represented a case of unsuccessful

shift, then we must assume that a group of speakers who shared a single Bantu language (Pare) shifted to a Cushitic language whose availability as a model was as severely restricted as that of the European languages in Caribbean creole formation. The linguistic facts of Ma'a would fit such a hypothesis, but the necessary social assumption is so improbable that it constitutes strong evidence against the hypothesis.

The same basic argument applies to the hypothesis that Ma'a represents a nativized pidgin. In all pidgins, the morphological structures of the participating languages are significantly simplified, even when--as with the Bantu-based African pidgins--all the languages involved in their formation share general principles of morphological organization to a very large extent. Given this salient characteristic of pidgins, an assumption that Ma'a had its origin in a pidgin which arose from contact between one Bantu language and one Cushitic language is untenable. (Goodman [1971:253] gives a similar argument.) Another difficulty with such a hypothesis, though maybe not an insurmountable one in itself, is the relative rarity of pidgins in two-language contact situations. Usually more than two languages are involved.

The fifth and last possible line of nongenetic development is the one that most closely fits Ma'a: Ma'a arose as a product of massive interference from a Bantu language B (Pare) in a Cushitic language A. The circumstances that must be assumed to support this claim are by no means unique so far as the linguistic outcome is concerned, and the hypothesized social situation, though rare, is neither unknown in general nor dubious in this case, given the recorded history and cultural traits of the Ma'a people.

The external history of Ma'a has been described from oral traditions by Kimambo [1969] and Feierman [1974] in their studies of the Pare and the Shambaa people, respectively. According to these traditions, the Ma'a people came to the South Pare mountains about three hundred years ago, settling in Vudee in the northwestern region of South Pare. Some time after that, apparently in "an attempt to resist encroachment on their ways of living" [Kimambo 1969:62], a large group of Ma'a moved southward to settle in the Usambara Mountains. The Ma'a clans that remained in South Pare shifted to Pare, but the main Ma'a group in Usambara did not shift either to Pare or to

Shambaa. One small offshoot of the Ma'a had moved to Usambara before the main migration; these people, the Nango, shifted to the Shambaa language. As Feierman [1974:77] puts it, citing several oral traditions, "the Nango chose to live the Shambaa way," but the Mbugu (Ma'a) did not.

This picture of the Ma'a as resisters of cultural assimilation is supported by comments by Copland and Green. Copland emphasizes that "the Mbugu were not disposed to sink their individuality" as some clans had done [1933-4:242], and Green remarks that "the Wambugu are a reserved and uncommunicative people regarding their past history and present customs... They keep themselves apart from the indigenous Wasambaa [Shambaa]" [1963: 175]. Green also refers to their "present independent attitude," to their "isolation from the Wasambaa," and to their "extreme conservatism" as seen in their resistance to cultural influence from Europeans as well as from the Shambaa (p. 177). This desire for isolation was no doubt reinforced by the efforts of some Shambaa chiefs, late in the nineteenth century, to enrich themselves by selling Ma'a (among other people) into slavery--efforts which forced the Ma'a to barricade themselves behind palisades "where they could farm in peace and not be enslaved" [Feierman 1974:172].

Nevertheless, the wish to preserve their cultural autonomy clearly did not keep the Ma'a from having regular contacts with their Bantu-speaking neighbors. Kimambo and Feierman report such contacts with the Pare and the Shambaa, and Ehret [p.c. 1982] points out that even now the Ma'a are neighbors of Asu groups who live on the edges of the Usambaras. The intimacy of these contacts is evident from the fact that all the Ma'a today are apparently fluent in both Pare and Shambaa [Ehret, p.c. 1982]. (But it is not clear that this was true fifteen years ago, because Feierman had to use a translator when he collected Ma'a oral traditions in 1968, though he was fluent in Shambaa [1974:8).

But, though continuing regular and close contact with the Asu and the Shambaa is certain, this is not the most striking sociolinguistic feature of Ma'a history. Probably the most important Bantuizing influence in the development of Ma'a, and the one that accounts for the unusual nature and degree of linguistic mixture in the modern language, is the ethnic link

between the Ma'a speakers in the Usambaras and the Pare-speaking Ma'a clans who remained in Vudee in the South Pares. Long after the main body of the tribe had moved to the Usambaras, tradition demanded that they return to Vudee for the annual initiation rites. Copland, Green, Kimambo, and Feierman all report independently that, as Kimambo puts it [1969:62], "the original residence in Vudee remained the shrine of all the Mbugu," and that the Ma'a of Pare and Shambaai were regularly reunited there to initiate their young men. Feierman believes that this practice may have helped to create the impression among the Shambaa that the Ma'a of the Usambaras were not fully cooperative, and therefore not trustworthy, neighbors [1974:81]. Copland was told that "the last pilgrimage from Usambara took place in 1921" [1933-4: 242], which suggests that the practice was on the decline by 1933. This is reinforced by Green, who comments that the Ma'a of the Usambaras used always to go to Vudee for the rites, but that "lately the rite has been performed at Shume in the Usambaras. They still occasionally go to Vudei, four or five hundred strong when the mood takes them" [1963:175-6].

In other words, recent use by the Ma'a of Bantu languages has never been confined to trade and other limited communicative functions with Bantu neighbors, from whom the Ma'a have deliberately kept their cultural distance. The remaining Ma'a speakers[22] have maintained contact with their Bantu-speaking kinsfolk in the South Pares, and in earlier years, at least, this contact necessitated regular unrestricted communication during the annual visits to Vudee. The clans that had remained in Pare--by implication, those Ma'a who did not resist cultural assimilation so strongly--had shifted to the Pare language and forgotten the Ma'a language, so that the Ma'a of the Usambaras would have to know Pare in order to talk to them. Meanwhile, however, the Ma'a of Pare may have kept the memory of at least some Ma'a vocabulary for use in the rituals, much as the Copts of Egypt still use Coptic in their religious ceremonies even though they all speak Arabic natively. That is, in

[22]Ehret [1980:11] says there are several thousand of them today. Whiteley says that there were about 11,000 Ma'a in the Usambaras at the time of the 1948 census [1960:96].

the context of ethnic reunion for ritual purposes, the most salient part of the ethnic-heritage language--the original Cushitic vocabulary--may well have played an important role, even though the language for ordinary communication between the two Ma'a groups must have been Pare. In any case, as far as the pilgrim Ma'a were concerned, this same salient part of their language was eventually the major relic of their Cushitic heritage, stubbornly maintained in spite of the pressure to shift to Bantu in their Bantu-surrounded homeland and the more intimate pressure to adopt Pare so that they could talk to their kin in Pare.

The degree of resistance to total cultural assimilation that is well attested in Ma'a is probably unusual, but other situations can be found that are comparable, though not identical, both in their cultural and in their linguistic results. One is the case of the English gypsies, who speak English but also maintain Romani vocabulary, the Indic lexicon of their original ethnic language, for use (with English grammar) as a secret code. But, while most English gypsies speak English most of the time, Ma'a is still (as far as I can tell from sources that are largely silent on this point) the ordinary every-day language of its speakers. Another difference between Ma'a and Anglo-Romani is that, as we have seen, Ma'a still has a few productive Cushitic grammatical features, while Anglo-Romani has no Romani grammar at all. So Anglo-Romani is an even more extreme example of grammatical replacement than Ma'a.

Another language that can usefully be compared with Ma'a is the one spoken by members of the scattered ethnic Greek communities in Asia Minor. When they were studied by Dawkins early in this century, these communities had been under constant cultural pressure from surrounding Turkish speakers for hundreds of years. Many Greeks had adopted the Turkish language and culture (including, for instance, the Moslem religion), but others had retained both their language and their religion, along with other cultural traditions. However, through centuries of bilingualism the Greek spoken in Asia Minor became heavily Turkicized, though not to the extent that Ma'a has become Bantuized: numerous Greek grammatical features, including the bulk of the inflectional systems, remained in Asia Minor Greek [Dawkins 1916].

A third example is the spectacular case reported by Menovščikov [1969], in which the Aleut spoken on Mednyj, one of the Commander Islands, has had its entire elaborate finite verb morphology replaced by that of Russian, though the original elaborate noun morphology and non-finite verb morphology are retained. The historical interpretation of this case is difficult, partly because both Aleut and Russian speakers on Mednyj were apparently bilingual to some extent in each other's language during the relevant period (see Thomason [1981] for discussion), but it is clear that Aleut was partially maintained in the face of very strong pressure from Russian.

Of these three cases, the one that most closely resembles Ma'a in its developmental characteristics is Asia Minor Greek. To the extent that Anglo-Romani is spoken natively at all, it has become a native language (in some gypsy families) through expansion in function of the secret-code jargon consisting of Romani words and English grammar. This has taken place since the shift to English by English gypsies, so that it is probably best viewed as a re-emergence of first-language learning of Romani vocabulary--that is, after a break in transmission, and after a period in which the first-learned language in the community was English. The case of Mednyj Aleut is more complicated, but here too it seems fairly likely that the transmission process was abnormal, because second-language learners, mainly Russian husbands of Aleut wives, must have participated in the transmission of the language to children born into the culturally and linguistically diverse community.

But Asia Minor Greek, like Ma'a, developed in communities which, though largely bilingual, were relatively homogeneous culturally. I see no room for doubt in either case about the existence of cultural and ethnic continuity from a period pre-dating Turkish and Bantu influence, respectively. This continuity includes normal transmission of the community's language, in that (unlike Anglo-Romani and Mednyj Aleut) there is no evidence that would force, or permit, us to infer a break in the transmission of an entire language at any period in the language's history. As far as we can tell, the ancestors of the current Ma'a speakers did not shift to Bantu while keeping their original Cushitic lexicon, and they did not experience the disruption of mixed Bantu-Ma'a households that might have prevented children from learning Ma'a as a

whole language. So to this extent, at least, Welmers is right in claiming that the development of the language is "within the familiar framework of continuous language history" [1973:8].

Nevertheless, in one respect Ma'a resembles Mednyj Aleut and especially Anglo-Romani rather than Asia Minor Greek, and this is the crucial point for the question of genetic relationship: it is not possible to show for Ma'a, any more than for Anglo-Romani or Mednyj Aleut, systematic form/function correspondences in all grammatical subsystems. The history of Ma'a is similar in kind to the history of Asia Minor Greek, but the amount of foreign interference is far greater in the Ma'a case. Therefore, although the Bantu interference features presumably accumulated gradually in Ma'a, so that the difference in the language of any two adjacent generations was minor, the net effect is a language whose grammatical morphemes are almost entirely of Bantu origin (if we assume, as many linguists do, that derivational affixes belong to the lexicon), and whose definitely Cushitic phonological and syntactic features are also outweighed by the Bantu features.

If we consider only the linguistic structures of modern Ma'a, I do not believe that a convincing case can be made for treating the language as a changed later form of Proto-Southern Cushitic: there are too many Bantu substitutions in the grammar, and too few remaining systematic Cushitic grammatical features. If this is true, then Ma'a cannot be said to be genetically related to Cushitic languages, unless we reformulate the notion of genetic relationship to fix on the vocabulary as the sole criterion for establishing relationship. But then genetic relationship would cease to be a historical concept and become merely a synchronic taxonomic one. Probably the most significant consequence of such an approach would be that comparative reconstruction could no longer be claimed to be a guess at the structure of a real language that was spoken in the past, and that consequence would surely be unacceptable to most historical linguists. Ma'a morphological data certainly cannot be treated, for purposes of reconstruction, as gradually modified Proto-SC, and Ma'a phonology is also of limited value for reconstructing Proto-SC phonology, given the highly irregular effects of contact-induced

changes in phonological shapes of words.[23]

Just as we could infer from the structures of the least decreolized Caribbean creoles that they arose through imperfect shift or as nativized pidgins, we can infer from the structure of Ma'a, as I argued at the beginning of this section, that it arose in a long-term situation of language (and culture) maintenance under conditions of intense cultural pressure from Bantu. We can therefore establish that Ma'a arose from a Cushitic language. But since Ma'a as a whole is no longer an appropriate object for comparative reconstruction with Cushitic, the integrity of genetic relationship as a claim about gradual linguistic divergence over time is best preserved by putting Ma'a, along with languages like Anglo-Romani, Taki-Taki, and Tok Pisin, outside the genetic model.

---

[23]This does not, of course, mean that Ma'a is irrelevant for the reconstruction of Proto-SC. But its usefulness is similar to the usefulness of borrowed elements in neighboring languages, e.g. Finnish borrowings from early Germanic: the Ma'a data must, in most instances (/ɬ/ is an exception), be considered in light of Bantu phonology as well as Cushitic in order to arrive at a reasonable interpretation of the relevance to Proto-SC reconstruction.

REFERENCES

Bender, M. L. (ed.). 1976. *The Non-Semitic Languages of Ethiopia*. East Lansing, MI: African Studies Centre.

Bender, M. L., E. Mulugeta and D. L. Stinson. 1976. "Two Cushitic languages." In M. L. Bender, J. D. Bowen, R. L. Cooper and C. A. Ferguson (eds.), *Language in Ethiopia*, pp. 130-54. London: Oxford University Press.

Black, P. 1976. "Werizoid." In Bender [1976], pp. 222-31.

Bliese, L. 1976. "Afar." In Bender [1976], pp. 133-65.

Copland, B. D. 1933-34. "A note on the origin of the Mbugu with a text." *Zeitschrift für Eingeborenen-Sprachen* 24:241-5.

Dawkins, R. M. 1916. *Modern Greek in Asia Minor*. Cambridge: University Press.

Dolgopolskij, A. B. 1973. *Sravnitel'no-istoričeskaja Fonetika Kušitskix Jazykov*. Moscow: Nauka.

Ehret, C. 1974. *Ethiopians and East Africans: The Problem of Contacts*. Nairobi: East African Publishing House.

Ehret, C. 1976. "Cushitic prehistory." In Bender [1976], pp. 85-96.

Ehret, C. 1980. *The Historical Reconstruction of Southern Cushitic Phonology and Vocabulary* (Kölner Beiträge zur Afrikanistik, 5). Berlin: Reimer.

Ehret, C. and D. Nurse. 1981. "The Taita Cushites." *Sprache and Geschichte in Afrika* 3:125-68.

Elderkin, E. D. 1976. "Southern Cushitic." In Bender [1976], pp. 278-97.

Feierman, S. 1974. *The Shambaa Kingdom: A History*. Madison, WI: University of Wisconson Press.

Gleason, H. A., Jr. 1955. *Workbook in Descriptive Linguistics*. New York: Holt, Rinehart & Winston.

Goodman, M. 1971. "The strange case of Mbugu." In D. Hymes (ed.), *Pidginization and Creolization of Languages*, pp. 243-54. Cambridge: University Press.

Gragg, G. 1976. "Oromo of Wellega." In Bender [1976], pp. 166-95.

Green, E. C. 1963. "The Wambugu of Usambara (with notes on Kimbugu)." *Tanganyika Notes and Records* 60-61:175-89.

Greenberg, J. H. 1955. *Studies in African Linguistic Classification*. New Haven.

Greenberg, J. H. 1963. *The Languages of Africa*. Bloomington: Indiana University Press.

Greenberg, J. H. 1966. "Some universals of grammar with particular reference to the order of meaningful elements." In J. H. Greenberg (ed.), *Universals of Language*, pp. 73-113. Cambridge, MA: MIT Press.

Hall, R. A., Jr. 1958. "Creolized languages and genetic relationships." *Word* 14:367-73.

Hudson, G. 1976. "Highland East Cushitic." In Bender [1976], pp. 232-77.

Hudson, R. A. 1974. "A structural sketch of Beja." *African Language Studies* 15:111-42.

Hudson, R. A. 1976. "Beja." In Bender [1976], pp. 97-132.

Kimambo, I. N. 1969. *A Political History of the Pare of Tanzania, c. 1500-1900*. Nairobi: East African Publishing House.

Kimenyi, A. 1978. "Grammatical tone neutralization in Kinyarwanda." *Studies in African Linguistics* 9:303-17.

Loogman, A. 1965. *Swahili Grammar and Syntax*. Pittsburgh: Duquesne University Press.

Meillet, A. 1921 (1958). "Le problème de la parenté des langues." Reprinted in his *Linguistique Historique et Linguistique Générale*, pp. 76-101. Paris: Champion.

Meinhof, C. 1906. "Linguistische Studien in Ostafrika. X. Mbugu." *Mittheilungen des Seminars für Orientalischen Sprachen* 9(3):293-323.

Menovščikov, G. A. 1969. "O nekotoryx social'nyx aspektax èvoljucii jazyka." In *Voprosy Social'noj Lingvistiki*, pp. 110-34. Leningrad: Nauka.

Nida, E. and H. W. Fehderau. 1970. "Indigenous pidgins and koinés." *IJAL* 36:146-55.

Polomé, E. 1971. "The Katanga (Lubumbashi Swahili Creole." In D. Hymes (ed.), *Pidginization and Creolization of Languages*, pp. 57-59. Cambridge: University Press.

Sasse, H.-J. 1976. "Dasenech." In Bender [1976], pp. 196-221.

Thomason, S. G. 1980. "Continuity of transmission and genetic relationship." In E. C. Traugott, et al. (eds.), *Papers from the Fourth International Conference on Historical Linguistics*, pp. 27-35. Amsterdam: Benjamins.

Thomason, S. G. 1981. "Are there linguistic prerequisites for contact-induced language change?" Paper presented at the Tenth Annual University of Wisconsin-Milwaukee Linguistics Symposium: Language Contact.

Thomason, S. G. and T. Kaufman. 1975. "Toward an adequate definition of creolization." Paper presented at the 1975 International Conference on Pidgins and Creoles, Honolulu.

Thomason, S. G. and T. Kaufman. Forthcoming. *Language Contact, Creolization, and Genetic Linguistics.*

Tucker, A. N. and M. A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa. Handbook of African Languages.* London: Oxford University Press.

Tucker, A. N. and M. A. Bryan. 1974. "The 'Mbugu' anomoly." *BSOAS* 37:188-207.

Welmers, W. E. 1952. "Notes on the structure of Saho." *Word* 8:145-63; 236-51.

Welmers, W. E. 1973. *African Language Structures.* Berkeley: University of California Press.

Whinnom, K. 1971. "Linguistic hybridization and the 'special case' of pidgins and creoles." In D. Hymes (ed.), *Pidginization and creolization of Languages*, pp. 91-115. Cambridge: University Press.

Whiteley, W. H. 1960a. "Linguistic hybrids." *African Studies* 19(2): 95-7.

Whiteley, W. H. 1960b. "The verbal radical in Iraqw." *African Language Studies* 1:79-95.

Zaborski, A. 1976. "Cushitic overview." In Bender [1976], pp. 67-84.