[**Update** (September 2007): This paper was originally written for the
"Greenberg conference" – more formally the Conference on Language
and Prehistory in the Americas – that was held in Boulder, Colorado, in
1990. After many years of delays, the putative editor of the conference
volume announced that the book would never appear. Since the paper
was really suitable only for that particular venue, I decided not to revise
it for publication elsewhere. I also haven't changed my affiliation: the
final version of the paper was written before I moved from the Pitt to
Michigan in 1999.]

HYPOTHESIS GENERATION VS. HYPOTHESIS TESTING:
A COMPARISON BETWEEN GREENBERG'S CLASSIFICATIONS
IN AFRICA AND IN THE AMERICAS*

Sarah Grey Thomason

University of Pittsburgh

30 December 1994

Joseph Greenberg's *Language in the Americas* (henceforth *LIA* is about genetic classification—
that is, about the classification of the indigenous languages of the New World into families
(three of them). By 'families' Greenberg presumably means the same thing historical lin-
guists have meant since the beginnings of the discipline in the 19th century, and even further
back: all the languages in a given family are changed later forms of a single parent language
which was once spoken by some group of people in some particular place. Although most
historical linguists don't spell out what they mean by 'changed later form of a parent lan-
guage' in this definition, there would probably be quite general agreement that it means that
the parent language was passed down from generation to generation as a complex whole—an
interrelated set of phonological, morphosyntactic, semantic, and lexical structures. If this

definition is accepted, it carries some implications for any attempt to arrive at a comprehensive genetic classification of all the languages in a large area (such as Africa or the entire New World); and it causes serious problems, ultimately, for Greenberg's methodology. In this paper, which is primarily a methodological one, I'll elaborate on this point after laying some groundwork by comparing Greenberg's African classification to his American one.

According to Greenberg's own account, the method by which he arrived at his genetic groupings in the Americas is basically the same one he used forty years ago to classify African languages into four families: he calls it multilateral comparison or mass lexical comparison, and it involves searching word lists and paradigms for form/meaning similarities in both lexical and grammatical morphemes. But while most Africanists—especially Americans at first (Paul Newman, personal communication, 1988)—reacted favorably to his earlier work, most historical linguists who specialize in American Indian languages believe that the American classification is a deeply flawed piece of scholarship. This raises an interesting question: why should a method that seemed to produce good results in one part of the world inspire such a negative reaction when it is applied elsewhere?

I want to try to answer this question by addressing two major substantive issues. (I will ignore sociological issues, because they serve mainly to confuse the picture. And I will not discuss the charges of sloppy scholarship that have been made by specialists in various groups of Native American languages—among them Goddard for Algonquian [1987], Chafe for Iroquoian [1987], Campbell for Mayan [1988], Adelaar for Quechua and other South American languages [1989], Kimball for Muskogean, "Gulf", and Yukian [1992], Poser for Salinan and Yurumanguí [1992a], Berman for Yurok and Kalapuya [1992], and Rankin for Siouan [1992]—because, though the charges are serious, they seem to be a separate issue, unconnected with the basic problems with the methodology.) First, I think it's crucial to distinguish between hypothesis generation and hypothesis testing, and I'll argue that Greenberg's method is a perfectly good way of coming up with hypotheses but that it is not useful for testing them. The difference between Africa and the Americas in this respect is that new, or at least revived, genetic hypotheses were needed for African languages when Greenberg first applied his method there, but in North and Central America (though perhaps not South America) the main work to be done right now is hypothesis testing. Second,

Greenberg's innovative American groupings posit very distant genetic relationships in a context of great genetic diversity, whereas in Africa his most successful proposals grouped languages that were closely enough related to be established, and that eventually WERE established, by means of the standard Comparative Method. Moreover, the possibility of demonstrating a distant genetic relationship will vanish sooner for small families—by far the more common variety in the New World—than for very large families with many members, for instance Niger-Congo and Afro-Asiatic, the most widely accepted groupings in Africa (I owe this observation to Christopher Ehret, personal communication, 1988).

The main proposal I'll argue for is this: the problem with claims of distant genetic relationship is that beyond a certain time depth it is impossible to distinguish shared features that are due to accident and borrowing from those that are due to genetic inheritance. The basis for this proposal can be found in the standard assumptions that underlie all work in historical linguistics. One of these assumptions is the definition of a language family that I just gave. Since all the languages in a family have descended from the same parent language, they will—for a considerable period of time—display systematic correspondences in all their grammatical subsystems, which they inherited from their common parent. Another assumption is that all living languages are constantly changing in ways that are, over the long run, unpredictable; this means that, once a language has split into two or more daughter languages, the daughter languages will continue to diverge. Given enough time, therefore, the daughter-language structures will no longer correspond systematically in any grammatical subsystem, and at this point we lose the possibility of discovering their historical connection. Because we can't find out what their historical connection was, or even if there was any, we can only propose hypotheses about genetic links; we cannot test them.

An examination of the difference between Africa and the Americas before Greenberg's classifications, as far as the state of historical linguistic research is concerned, will set the stage for the more general methodological discussion. (I should emphasize that my comments about the African situation are not new; all of these points are mentioned in published reactions to Greenberg's African classification.) Before Greenberg presented his African classification, the most widely accepted groupings were based either on cultural or on typological similarities, or both. For instance, Guthrie's internal Bantu classification was explicitly ahis-

torical (1948:Ch. III), relying primarily on shared typology. And, notoriously, the "Hamitic" group comprised all East African languages which were not Bantu and which had masculine and feminine gender; in addition, the "Hamites" were cattle-raisers. These criteria would be recognized by any well-trained historical linguist as methodologically unsound. But the groupings were nevertheless maintained by most Africanists before Greenberg—partly or largely for sociological reasons having to do with the prominence of the people who advocated the old groupings and criteria—in spite of the obvious similarities that had already been noticed between some subgroups of putative "Hamitic" languages and various other languages that weren't classified as "Hamitic".

Given this petrified and wrong-headed classificatory picture, Greenberg's method, which emphasized the search for form/meaning similarities in lexical items rather than typological or cultural similarities, was the proverbial breath of fresh air. The lexical similarities he presented cut across the old "Hamitic" boundaries, linking some "Hamitic" languages with Semitic and grouping others into new families. The outlines of the huge Niger-Congo family (itself a branch of Niger-Kordofanian in Greenberg's classification) also emerged from Greenberg's comparison. Once these hypotheses were in place, the work of testing them began. For most subgroups of Niger-Congo, attempts to establish genetic relationship by means of the Comparative Method have been successful, yielding ample evidence of systematic lexical and grammatical correspondences; the status of Mande as a Niger-Congo branch is still debated, but the other branches seem to be well established and universally accepted. Similar efforts have also been successful for Afroasiatic, though controversy still surrounds some corners of the proposed family (notably Omotic). Overall, the beneficial effects of Greenberg's proposed African groupings are clear; as one Africanist recently put it, 'G[reenberg] established order where there was prejudice and chaos, and a grateful set of Africanists adopted his labels, fully aware that they were problematic....He gave experts in different languages a basis for talking to each other' (Wald 1994).

As various scholars have observed, much of Greenberg's African classification was not new. Of the parts that were new, several are still controversial. The genetic unity of one of Greenberg's four families, Nilo-Saharan, is still in doubt. There are a number of examples in lower-level groupings too; for instance, Schadeberg 1981 argues that a grouping of the

Kadugli languages with Nilo-Saharan has at least as much support as Greenberg's grouping of Kadugli with Kordofanian (in the Niger-Kordofanian family). Another example concerns Greenberg's Khoisan family: Harold Fleming, who is certainly no foe of other proposals of distant genetic relationship, has recommended that Sandawe and Hadza should be 'de-classified from Khoisan, at least until a greater effort is made to classify them' (1983:555).[1]

In fact, it turns out that the parts of Greenberg's African classification that are now universally accepted by historical linguists who specialize in Africa are precisely those that have since yielded to the traditional Comparative Method, including evidence of systematic correspondences in both lexicon and grammar.[2] Some of his other proposed genetic groupings are still viewed as tentative, and still others have been rejected because such correspondences have not been found on closer examination of the data.

Now, if we compare the African situation—the dominant pre-Greenberg vs. the dominant post-Greenberg African classifications—to the New World situation, the first thing we notice is that in the Americas there was no pernicious or chaotic classificatory tradition to overcome. During the twentieth century, at least, the dominant approach to genetic classification in the Americas has been based on entirely appropriate linguistic criteria, not on typology or geography. Historical work on Native American languages has, for many decades now, used the same methods that proved so successful elsewhere, notably in Indo-European genetic linguistics but also in other areas. In fact, many twentieth-century Americanists were thoroughly trained in Indo-European linguistics before they turned their attention to Native American languages. In classifying the Indian languages, they and their students used the criterion of systematic sound/meaning correspondences in basic vocabulary and morphology—a criterion which emphatically was not employed by the most prominent pre-Greenberg classifiers of African languages. So there was no classificatory disaster area to clean up in the Americas. Even so, it is reasonable to ask whether or not Americanists are likely to learn from Greenberg's American classification, as Africanists learned from his African one.

In the last eighty years or so of Native American linguistic classification, an enormous amount of solid evidence has been presented for small and large genetic groupings, especially for the languages of North and Central America. In addition, many proposals have been

made that didn't get accepted in the end—most notably some of Sapir's six huge families in North America, but others too (see e.g. Campbell & Mithun 1979:38, which lists Sapir's Hokan-Siouan grouping as 'universally abandoned' and his Algonquian-Wakashan grouping as 'generally abandoned'; see also the comment by Krauss in the same volume [1979:838] to the effect that 'there is no detectible genetic relationship between Haida and [the other Na-Dene languages,] Tlingit and Athabaskan-Eyak'). Such proposals have been taken very seriously, even some that seemed quite speculative. One of Sapir's long shots did come through: his linking of Algonquian with Wiyot and Yurok in California is now universally acknowledged as a valid genetic grouping, because there is now enough solid evidence to satisfy everyone on that point. Some of Sapir's and other scholars' proposed groupings, especially less ambitious ones, are considered promising, and efforts to demonstrate their validity continue.[3] Still other proposed groupings have resisted all efforts to find systematic supporting evidence, and most people reject them today—not because the proposals seem bizarre, but because sustained efforts to support them by adducing systematic correspondences have failed; see, for instance, the references cited at the beginning of this paragraph. Although it should not be necessary to say so, I will add that rejecting such a proposal does not entail a belief, much less a claim to be argued for, that the languages in any proposed family are NOT related. Some of Greenberg's criticisms of people he calls "conservatives" seem to be based on a failure to understand the difference.[4]

This means that, when Greenberg published his new classification for the Americas, Americanists already stood in the same basic position with respect to his huge groupings that African linguists had reached on, say, Nilo-Saharan about thirty years after Greenberg first proposed that grouping: hypotheses for big Native American linguistic groupings were offered long ago; some have not yet been thoroughly tested, but some of the ones that have been tested have been rejected because they failed the tests. So it is not surprising that there is limited enthusiasm among Americanists for Greenberg's even bigger groupings in the Americas, most of which rest on evidence at least as shaky as the evidence Sapir used for his six-family classification.

But if Greenberg's method turns out to produce some groupings that are eventually going to be rejected, both in Africa and in the Americas, then the method obviously isn't

a foolproof way of establishing genetic relationship. The question is, why? Why can't the unsystematic similarities that his method relies on be counted on to lead invariably to valid hypotheses of genetic relationship? The answer to this question brings us back to the issue of hypothesis testing: the method sometimes fails because, at great time depths and with great linguistic diversity, it provides no way of eliminating alternative explanations for the facts—of showing that alternative explanations are less well supported than a hypothesis of genetic relationship. As long as the similarities remain unsystematic, they could just as easily be due to accident or to language contact as to inheritance from a common parent. This is a controversial assertion, and Greenberg and his followers would surely deny it, so I'll try to back it up with some evidence.

First, consider accident (using the term loosely, to include similarities conditioned at least in part by universal markedness tendencies as well as genuine accident). Critics say that Greenberg's implicit criteria for phonetic and semantic similarity are so unconstrained that, except for closely-related languages whose relationships can be demonstrated by traditional means, his method would produce equally good results when applied to any randomly-selected collection of languages; Greenberg says it wouldn't. He does not accept the evidence of lists of comparable similarities that several critics have compiled from presumably unrelated languages, because, he says, these are merely two-language comparisons, whereas his method demands multilanguage comparison. The critics respond that, since he requires only a small number of similarities for a claim of relatedness, and since he does NOT require any correspondence to appear in all the languages of a proposed family at once, it is just as easy to compile comparable multilanguage lists as two-language lists. Various efforts have been made to develop statistical tests that will distinguish chance from historical links in this sort of comparison (cf. e.g. Justeson & Stephens 1980 and, in a more recent work with specific criticisms of Greenberg's own statistics, Ringe 1992). One problem with devising such tests is the difficulty of finding explicit criteria for deciding when a semantic or a phonetic correspondence is similar enough to count as a similarity for purposes of comparison.

Additional problems arise when we consider several other crucial points. One is that Greenberg does not take into account the possibility that markedness factors and sound symbolism might account for some of the widespread similarities he has noticed. A case

in point is the common set of first and second person singular morphemes consisting in part of nasal consonants, not just in the Americas but in many other parts of the world as well. Before concluding that first person $n$- and second person $m$- constitute evidence for Greenberg's Amerind family, one should keep the following points in mind. First, nasals and apical consonants are unusually common and stable in inflectional systems all over the world, in various functions; for the nasals, at least, this may be a result of phonetic salience and general historical stability (cf. Maddieson 1984:70, and see also Bailey 1970 on the subject of preference for apical consonants in word-final position). Second, it is common for a restricted subset of a language's consonant inventory to appear in its inflectional affixes (Floyd 1981). And third, affective and onomatopoetic uses of nasals, in particular, are much more widespread than mere accident could reasonably account for; probably the best-known examples are *mama* and *nana* as kin terms (see Jakobson 1962 and the related discussion in Rankin 1992:338-39). The point here is that too many of the inherited shared features in very distantly related languages are going to be unmarked, stable features that could just as easily be relics of parallel but historically unconnected attrition processes in unrelated languages.

A fourth item worth mentioning in this context is that, in languages with complex morphological systems of pronominal agreement, the general categories 'first person singular' (1sg.) and 'second person singular' (2sg.) may include diverse morpheme shapes. In such cases, the chances that one or two common segments will appear in one or more of these morphemes increase dramatically. Consider, for instance, the sets of 1sg. and 2sg. morphemes in Montana Salish (also called Flathead). In the 1sg. set we do indeed find some forms with $n$: -$n$ (transitive subject suffix), *in*- (possessive prefix), and *čn* (intransitive subject proclitic); but we also find $k^w u$ (object proclitic) and $q^w oy?é$ (independent pronoun). In the 2sg. set there is one form with an $m$: the object suffix -$m$ that occurs in verbs with a transitive-marking suffix -*st*. The other 2sg. forms are -$x^w$ (transitive subject suffix), *an*- (possessive prefix), $k^w$ (intransitive subject proclitic), *anwí* (independent pronoun), and -*sí* (object suffix in verbs with a transitive-marking suffix -*nt*). It's true that $m$ occurs only in the 2sg. in Montana Salish, not in the 1sg.; but $n$ occurs both in the 1sg. and in the 2sg. A considerable portion of the history of these morphemes has been revealed by comparative

reconstruction, but it is not possible to decide which of the forms are old and which are innovative simply by examining a list of the forms. Any method that sanctions plucking out a form here and a form there, and making historical claims about them without regard to the system in which they are embedded, is not likely to find favor with historical linguists (or any other historical scientists).

Another relevant methodological point is that phonetic correspondences that provide evidence for genetic relationship (when one applies the traditional Comparative Method rather than Greenberg's method) are by no means confined to "similar" sounds by any measure. For instance, correspondence sets like the Greek-Armenian forms in Table 1, though regular, would not be uncovered by Greenberg's method, because the forms are no longer similar: Greek has lost PIE *w, and the Armenian forms have undergone (among others) a change from PIE *d(u)w to *rk*.

Table 1

Some Greek-Armenian Correspondence Sets

| Ancient Greek | Armenian | |
|---|---|---|
| duo | erku | 'two' |
| de:r- | erkar | 'long' |
| dei- | erkiwl | 'fear/fright' |

Checking lists of words for similarities will fail to uncover real correspondences that are dissimilar, and will therefore make it difficult, if not impossible, to unravel layers of correspondences (e.g. older cognates vs. more recent borrowings, as in French and English correspondences). The important point here is that distantly related languages are quite likely to display 'odd' correspondences—correspondences which, though not generally reflecting changes as apparently peculiar as the Armenian ones, do not consist of obviously similar sounds. A rather typical example is found in the Salishan language family (whose time depth is about 4000 years), where $p$ in some of the languages corresponds regularly to $\check{c}$ in others. This is why most historical linguists are suspicious of obvious unpatterned

phonetic similarities in a scattering of lexical items as evidence for very distant genetic relationship: after five or ten thousand years, many correspondences resulting from changes in inherited vocabulary are likely to comprise sounds that are not very similar. It is also why recurring correspondences of DISSIMILAR sounds provide some of the most convincing evidence for genetic relationships, distant or otherwise.

Still another methodological observation is that the experimenter effect is liable to produce very serious skewing in any application of Greenberg's method where the researcher knows—as Greenberg did in his "Amerind" comparisons—the approximate geographical locations of the languages.

A final methodological issue that is crucial to the question of accidental similarity is the line of argumentation that Greenberg uses in *LIA* against Levine's 1979 claim that there is insufficient evidence to support the grouping of Haida with Athabaskan-Eyak and Tlingit in the Na-Dene family.[5] Instead of citing Greenberg himself, I will quote J. David Sapir's comment on this point (1987:664), because it shows that Greenberg is not alone in his view of the methodological issue:

> [Levine] removed Haida [from Na-Dene] by using stringent criteria to dismiss a large portion of [Edward] Sapir's lexical correspondences. Greenberg's brilliant counter is to take the very same criteria and apply them to Celtic [standing for Athabaskan] and Albanian and Armenian [standing for Haida and Tlingit]. On the basis of Levine's analysis [Albanian and Armenian] could not be considered related to Celtic, an absurdity given what we know about Indo-European as a whole.

Now, in spite of Greenberg's detailed discussion of this issue in Ch. 6 of *LIA* (pp. 321-30), it is not at all clear that he is right in thinking that it would be impossible, using Levine's fairly standard criteria, to show that Celtic, Albanian, and Armenian are related to each other, if they were the only attested IE languages.[6] But if in fact they could not be proved to be related by those criteria, then it is certain that that result would NOT be absurd, and what we know about Indo-European as a whole on the basis of other information would be completely irrelevant.

It is entirely possible that we might be unable to demonstrate the genetic relationship of some group of distantly related languages with data from those languages alone. A simple thought experiment shows why: all we need to do is assume that many centuries and many changes have left the languages without systematic correspondences in their structures. The existence of factual inaccuracies in the resulting historical picture is not a relevant criterion to use in evaluating the historical methodology; many or most of our historical hypotheses are inaccurate in some respects, due to the inevitable loss of information over several millennia (see Thomason 1993 for a more detailed discussion of this point). The crucial question is this: to what extent are our hypotheses supported by the available data? If, in presenting a hypothetical example, we restrict the available data in a way that excludes lots of information, then we cannot turn around and use the excluded information to prove that the methodology is flawed. The methodology has to be evaluated in terms of what it enables us to do with the data at hand, not on the basis of additional information.

For all these reasons, although Greenberg asserts that accident cannot account for the comparative data he claims as evidence, probably most historical linguists would agree that he has failed to rule out accident as the source of his similarities.

The second reason for rejecting the results of Greenberg's method as evidence for the establishment of distant genetic relationships has to do with borrowing. His position on borrowing is that it 'is accepted when plausible', but that 'it can never be an overall explanation of a mass of resemblances in basic items, lexical and grammatical...over an extended area' (1986:496). Similarly, in his presentation at the Boulder conference (March 1990) Greenberg asserted that borrowing will never be widespread enough to interfere with genetic classification (and see his discussion of the Altaic problem in this volume).[7]

One problem with this position is that, if it is hard to distinguish inherited resemblances from accidental ones after several thousand years, it will be even harder to distinguish inherited resemblances from borrowings after such a long time, because borrowings often show systematic correspondences with source-language morphemes. But a more dramatic problem with Greenberg's position on borrowing is that it is demonstrably false. Consider the case outlined in Table 2—the famous mixed language Ma'a (Mbugu) of Tanzania (data from Thomason 1983). Using Greenberg's methodology, it would be easy, and perhaps even neces-

sary, to argue for two distinct genetic groupings for Ma'a: Bantu and Cushitic. The evidence would be compelling on both sides, as the two columns of features in Table 2 indicate. But probably the strongest evidence is on the Bantu side, by Greenberg's criteria, because of the language's Bantu inflection. Note especially the irregular allomorphy in the 1sg. negative prefixes and the Class 1 markers, which Ma'a shares with Bantu languages; and compare Greenberg's remark in *LIA* that shared grammatical irregularities, in particular, have 'enormous probative value...their presence tells us that there is a relationship, but not at what level' (1987:30).[8]

Table 2

Evidence linking Ma'a to Bantu and Cushitic

| Features shared by Ma'a (Mbugu) and Bantu | Features shared by Ma'a (Mbugu) and Cushitic |
| --- | --- |
| 1) Allomorphy: *si-* 1.sg. NEG vs. *te-* other NEG; *a-* Class 1 subj. vs. *m-* Class 1 obj. | 1) none |
| 2) Inflection: sg./plu. class prefixes on nouns & verbs; tense/aspect prefixes; person agreement on verbs. | 2) Inflection: only the pronominal possessive suffixes. |
| 3) Derivational suffixes: causative *-ija*, passive -*wa*, stative -*Vka*, vb. extensions. | 3) Derivational suffixes: amplificative suffix *-sa*, vb. extension *-ti*, etc. |
| 4) Lexicon: at least 50% of the lexicon, including some basic items – e.g. some body parts (C. Ehret, p.c. 1982). | 4) Lexicon: much of the rest of the vocabulary, including most of the basic items. |

It is not really very hard to determine (at least) the more recent stages in the development of Ma'a, but figuring them out requires close study of the history of the language and of its speakers; a simple search for form/meaning resemblances won't help to solve the puzzle. Ma'a used to be a non-Bantu language with much Cushitic lexicon, though it may never have been an actual Cushitic language.[9] Its speakers have all been bilingual in the Bantu language Pare for several hundred years, and they have Pare-speaking kinfolk; more recently they have become trilingual, adding Shambala to Ma'a and Pare (see Thomason 1983 and the references cited there for discussion). The linguistic result of the contact situation has

been a wholesale replacement of earlier Cushitic grammatical structures (and maybe some structures from other non-Bantu languages as well) and much Cushitic vocabulary by Bantu structures and lexicon.

Now, Ma'a itself will not be an immediate problem for classification. Although Greenberg has classified it as a Cushitic language, he could perhaps claim that it's an exceptional case that should not be classified into either family. In fact, he will have to do that if he does not wish to abandon the standard assumption that a daughter language is a changed later form of its single parent language, because Ma'a is clearly neither a normally-transmitted Cushitic language nor a normally-transmitted Bantu language. (But his methodology has no provision for making such a determination, so some modification of his approach would be necessary before he could claim Ma'a as an exception.)

The trouble is that such a strategy would only work for recent cases of this type—cases in which the changes in the language are so recent that the linguistic history can be recovered from the data. If, after a few millennia, Ma'a has diversified into numerous daughter languages spoken over an extended area, just as (for instance) Bantu spread over a large part of Africa and diversified in former times, its daughter languages will present serious problems for Greenberg's method: by that time, the resemblances to both Bantu and Cushitic may still be detectable, but they won't be as obvious, and it is unlikely that the historical puzzle will be solvable by then (assuming that all the current documentation has vanished). Quite likely, since most of the morphology (including virtually all of the inflection) is already Bantu, only Bantu morphology will be left; in that case, the borrowed bits of Ma'a would steer a future Greenbergian classifier in the wrong direction.

And even though Ma'a itself is an unusual type of case, as far as we know, it is not unique. Another example is the North American language Michif, with almost all nouns, noun modifiers, and nominal syntax from French and all verbs with their morphology and syntax from Cree. A third such case is Mednyj Aleut, in which the entire finite verb morphology has been borrowed from Russian, including—in the past tense only, where Russian verbs do not inflect for person—Russian pronouns. There are a few other well-documented cases of such dramatic mixtures, too. Less spectacular but still very substantial cases of structural and lexical linguistic interference, including all sorts of lexical and grammatical features, are

rather common. The point, of course, is that we have every reason to suppose that past contact situations also produced such mixtures, so that a methodology that is based solely on resemblances in one or two grammatical subsystems can't ever provide adequate support for a proposed genetic relationship in the standard meaning of that term. Only a careful search for systematic correspondences—not absolute regularity, which has never been insisted upon for classificatory purposes, but recurring correspondences—between all of a language's grammatical subsystems and those of its proposed sister languages will reveal a discrepancy in the sources of its structures, as Boas suggested long ago (1917:4; see Thomason & Kaufman 1988 for detailed arguments on this point).

Where does that leave us? I have argued that accident can't be ruled out as the source of Greenbergian similarities in cases where the languages being compared are related either very distantly or not at all. I have also argued that in those same cases language contact can't be ruled out as the source of such similarities. These issues will be especially problematic when we have to choose between claiming that two or more languages are distantly related and admitting ignorance—that is, saying that we can't tell whether or not they are related. The traditional proof of genetic relationship, which Greenberg finds unnecessarily restrictive, is the discovery of recurring sound/meaning correspondences that permit phonological reconstruction, together with the establishment of systematic correspondences in other grammatical subsystems that permit grammatical reconstruction. (This last requirement is not often emphasized in the literature, but it is needed to weed out the products of intensive language contact, which typically, as Boas predicted, show a discrepancy in correspondences between the lexicon and the rest of the language.)

If we want to say, with Greenberg, that demonstrating genetic relationship does not require showing that reconstruction is possible,[10] then I think it is appropriate to ask what the purpose of our classification is. If it is merely a way of bringing some order into a long list of languages so that people whose research requires reference to those languages can label them as members of particular groups (recall, in this connection, the comment cited earlier from Wald 1994), then historical linguists will have no quarrel with the enterprise as long as it is not called genetic classification. But if we use the term 'genetic relationship' and thereby commit ourselves to a historical claim, namely that the languages in a given

group are descended from a single parent language, then we must surely be required to show that descent with modification is better supported than either of the other two possibilities. Unless we can adduce evidence in the form of systematic correspondences, we won't be able to show this.

At the risk of repeating myself, I will re-emphasize the crucial point here. Given enough time—10,000 years, based on the estimated time depths of language families that have been convincingly established to date, is a standard estimate of the outer limits of the applicability of the Comparative Method—related languages are bound to diverge so much in all their grammatical subsystems that the few remaining shared inherited features cannot be distinguished either from accidental correspondences or from ancient foreign interference. No reconstruction is possible, because there are (for instance) so few recurring sound correspondences in words of comparable meaning that we couldn't be confident about any posited protolanguage phonemes or etymologies. In such a case, hypothesizing that the languages are related is an empty exercise, because the hypothesis cannot be tested. Hypothesizing that the languages are unrelated, or that one of them borrowed extensively from another, will have exactly the same result: nothing. So why bother? It's better, it seems to me, to admit, however reluctantly, that we are ignorant about the nature of their historical connection, if any.

I am not claiming that all of Greenberg's hypotheses are empty. Some of them, including some of the ones that are original with him, may turn out to be valid and fruitful, and may lead scholars to discover historical connections that they would not otherwise have discovered. I do believe that Greenberg's method is entirely appropriate as a means of constructing hypotheses, where these are needed—as they were in Africa and still are in other areas, perhaps including South America. His method is very likely to achieve good results in instances where genetic relationships are recent enough to yield eventually to demonstration by applying the Comparative Method. But his method is not reliable all by itself as a DEMONSTRATION of genetic relationship at any level, and a language family proposed on the basis of such data cannot be accepted as a valid genetic grouping until and unless the Comparative Method has been successfully applied to the proposed member languages.

The history of some prominent hypotheses developed from multilateral comparison underscores this point. As noted above, major portions of Greenberg's influential African classification have been verified through the Comparative Method, and are now universally accepted; other portions of that classification still await verification, and are not universally accepted by Africanists.[11] Greenberg's 1971 Indo-Pacific hypothesis—which links all the Papuan languages of New Guinea with each other and with Tasmanian and the languages of the Andaman Islands—has not been accepted by specialists in those languages. For instance, the hypothesis is not even mentioned in Foley 1986, the standard general work on Papuan languages. Foley says that, on the basis of present evidence, Papuan languages 'belong to at least sixty different language families' (1986:3), with no genetic links among the various families; and he comments in his final chapter ('Papuan languages and New Guinea prehistory') that, aside from 'the question of genetic links to Australian languages [which he considers dubious], no Papuan language family has been demonstrated to have any genetic affiliation outside the immediate New Guinea area' (1986:275). It is worth noting that Foley cites massive interference among Papuan languages as a major hindrance to any efforts to establish genetic relationships in the region. In particular, he makes the methodological observation that, in order to distinguish cognates from borrowings, one must assume the existence of 'a core of the vocabulary of a language that is resistant to borrowing'. But, he goes on to say, 'it does seem that this assumption is of questionable value in Papuan languages' (1986:210). He cites as evidence cases of borrowed basic vocabulary items, including entire pronominal systems.

Finally, as Greenberg has often pointed out, a method of multilateral comparison similar or identical to his own has been applied in the past. Sir William Jones' 1798 proposal linking Sanskrit, Greek, Latin, Gothic, Celtic, and Persian is the most famous example, and is often cited in introductory linguistics textbooks—justifiably—as an important milestone in the development of the field of Indo-European linguistics. Jones' place in history is not threatened by the fact that other groupings he proposed on the basis of the same kinds of evidence did not turn out to be correct, among them his classification of Malay and several Iranian languages other than Persian as Semitic (Jones 1799a:52, 1799b:7-8, 10) and his classification of Tibetan and Austronesian languages other than Malay as Indo-European

(Jones 1799b:12-13).[12] These hypotheses did pick out groups of languages to which the Comparative Method might be applied, but they did not—as the results make clear—provide adequate evidence for the genetic relationship of any language groups.

*This essay is a revised version of a paper presented at the 19th African Linguistics Conference at Boston University on April 15, 1988, and at the conference on Language and Prehistory in the Americas in Boulder, Colorado, March 24, 1990. I'm grateful to members of the 1988 audience, especially Christopher Ehret, for helpful comments on the first draft, and to Terry Kaufman, Dell Hymes, and Matthew Dryer for equally helpful comments on a slightly later written version. I also thank Paul Newman for helping me find relevant literature on Greenberg's African classification. Any remaining errors of fact or interpretation are of course my responsibility alone. Inevitably, given the long delay between the Boulder conference and the publication of this proceedings volume, some points in this paper are no longer as novel as they were in 1990; I have not attempted to mention all later publications that discuss points similar to some of the ones covered here, though I have updated references and several substantive points as well.

[1] Fleming has changed his mind on this point since 1983; he now believes that Hadza and Sandawe should be classified as Khoisan languages after all (personal communication, 1991), though he does not, as far as I know, cite new evidence to support his new position.

[2] Take as an example the case of Chadic. Although, according to Newman (1980:5-6), this group is still not accepted by all Africanists as a branch of Afroasiatic, the solid evidence that he provides in his 1980 article of systematic correspondences between Chadic and the rest of Afroasiatic should settle the question.

[3] At least one of these proposals has received new attention due to Greenberg's reformulation of an earlier suggestion: see Munro 1994, an evaluation of his proposed Yuki-Gulf grouping. As Munro points out (1994:130), Mary Haas was the first to propose a Gulf group (1951), and Haas 1954 suggests a grouping of Yukian and Gulf languages. Munro rejects many of Greenberg's specific word comparisons, remarking that 'this paper owes little but its inspiration to Greenberg' (135), but concludes that the data she has collected makes the proposed group promising and worthy of further investigation.

[4] One example is in his address to the 18th Annual North American Congress of Afroasiatic Linguistics, published as Greenberg 1990. Commenting on the controversy over *LIA*, he says that many American specialists in American Indian languages 'have reacted in a

Guthrie-like manner, fighting for the uniqueness of the languages they study' (1990:9). This is a very serious distortion of his critics' views. Neither here nor anywhere else does he cite a single reference to support this repeated charge, for the very good reason that his critics do not hold any such view. Probably all Americanists who find Greenberg's proposals unconvincing would agree with Janhunen, a specialist in North Asian languages whose response to a question about his opinion of comparable proposals in that area is (in part) as follows (1989:28, 30): 'I do have serious doubts concerning the validity of the long-range comparisons carried out *so far* using North Asian material....Personally I am ready to believe in any distant relationship, if only the facts can convince me. So far no sufficiently convincing facts have been presented.'

[5] Levine is not alone in this assessment; see the quote from Krauss 1979 above, and see also Thompson 1979:752. The following discussion of Greenberg vs. Levine is taken from Thomason 1993; it is repeated here because of its importance for the methodological issues stressed in this paper. There is also some overlap elsewhere between this paper and Thomason 1993, especially in the concluding paragraphs, because the two papers concern related (though distinct) methodological issues.

[6] Calvert Watkins (personal communication, 1989) tells me that it would actually be quite easy to demonstrate the relationship by the standard methodology—though not, probably, by Greenberg's methodology.

[7] At the 1990 conference, Greenberg cited as evidence the example of the Balkan languages, where, in spite of extensive borrowing, it is quite easy to determine that Rumanian is a Romance language, that Serbo-Croatian and Bulgarian are Slavic languages, and that Albanian and Greek are more distantly related to each other as well as to Rumanian and the two Slavic members of the Sprachbund. But the Balkan languages do not show the kinds of sweeping interference that one finds in more extreme cases, and the Indo-European family, with a time depth of 5000-6000 years, is well within the limits for the application of the Comparative Method. Sydney Lamb, supporting Greenberg in his paper at the Boulder conference, made a similar statement about contact-induced language change; but there are many counterexamples (e.g. in Thomason & Kaufman 1988) to Lamb's claim that the 'degree to which transfer is possible is a function of the degree of closeness [of the languages]

and durability [of the features]'.

⁸ Nevertheless, at the Boulder conference Greenberg dismissed the Ma'a grammatical evidence as irrelevant, since Ma'a is in his view obviously related to Cushitic. He suggested during the discussion of this paper that the Bantu morphology is, or was, not used regularly in Ma'a. But recent fieldwork by Matthias Brenzinger (personal communication, 1990; see also Brenzinger 1987) and Maarten Mous (1993) has shown that in fact Ma'a now consists of an entirely Bantu grammatical structure, with only some Cushitic lexicon remaining. Brenzinger's and Mous's data thus represent a stage later than the one reflected in the sources I examined for Thomason 1983. I would argue that Greenberg can't have it both ways: either his methodology as a whole is claimed as a valid way of detecting and establishing genetic relationship (in which case Ma'a is an absolute counterexample to the claims made for the methodology in Greenberg 1987 and elsewhere), or one must study a language's history in order to determine which lexical and grammatical features are inherited and which are borrowed (in which case Greenberg's methodology is not being employed at all).

⁹ Maarten Mous, a specialist in Southern Cushitic (the branch to which Ma'a would belong if it were a Cushitic language), has conducted a comparative study that suggests a more complex origin for the language than bantuization of an earlier Cushitic language (as proposed in Thomason 1983): he argues persuasively that, although Ma'a has indeed been heavily bantuized in the past 100-300 years, it may have originated as a lexical and grammatical mixture rather than as a Cushitic language (Mous 1993).

¹⁰ Note that it is not necessary to carry out a full-scale reconstruction in order to show that reconstruction is possible. In particular, absolute regularity in sound correspondences is not needed; all that's needed is a large enough body of recurring correspondences to rule out chance and—because the correspondences will be in all grammatical subsystems—borrowing.

¹¹ See, for instance, the following comment by Heine (1992, cited in Poser 1992b): 'Although Greenberg's work represents considerable progress over that of previous writers, it leaves a number of questions open. His approach is largely inadequate for the PROOF of genetic relationship; it can do little more than offer initial hypotheses, to be substantiated by more reliable techniques like the comparative method....The Nilo-Saharan family, in particular, must be regarded as a tentative grouping, the genetic unity of which remains

to be established.' Compare, in this connection, a recent comment by Dimmendaal in a review of a book on the genetic affiliation of the West African language Songhay (Nicolaï 1990): 'I hold the…view that Greenberg arrived at the best hypothesis regarding Songhay, namely, that it is a Nilo-Saharan language. I am confident that we will find more grammatical (as well as lexical) support in favour of this hypothesis.…After all, the initial hypothesis about the affiliation of, for example, Phrygian to Indo-European also was based on a few diagnostic features, mainly of a morphological nature' (1992:612). But if the evidence is at present insufficient to support Greenberg's hypothesis about the affiliation of Songhay, then believing in his hypothesis is a matter of faith, not science. Moreover, Dimmendaal's point about Phrygian is irrelevant: at the time Phrygian was grouped with Indo-European, specialists already knew what Indo-European was like, so they could tell whether, and how, the few Phrygian facts fit into the overall picture. But no one knows much about what Proto-Nilo-Saharan (if there ever was such a language) was like, so the Songhay situation can't legitimately be compared with the Phrygian situation. The same is true for all comparisons between new hypotheses of genetic groupings and Indo-European: isolated facts can readily be fit into the painstakingly developed matrix of Indo-European lexicon and grammar, but many decades were required to develop the picture we now have of Indo-European. No such solid matrix is available in any entirely new grouping to help interpret isolated facts.

[12] These citations from Jones 1799a and 1799b are from Poser & Campbell 1992; see Poser & Campbell for a discussion of their implications for the history of Indo-European studies and for the methodology of linguistic comparison.

## REFERENCES

Adelaar, Willem F.H. 1989. Review of Greenberg 1987. Lingua 78:249-55.

Bailey, Charles-James N. 1970. Toward specifying constraints on phonological metathesis. Linguistic Inquiry 1:347-49.

Berman, Howard. 1992. A comment on the Yurok and Kalapuya data in Greenberg's *Language in the Americas*. IJAL 58:230-33.

Boas, Franz. 1917. Introductory. International Journal of American Linguistics 1:1-8.

Brenzinger, Matthias. 1987. Die sprachliche und kulturelle Stellung der Mbugu (Ma?a). Cologne: University of Cologne M.A. thesis.

Campbell, Lyle. 1988. Review article on Greenberg 1987. Language 64:591-615.

Campbell, Lyle, and Marianne Mithun. 1979. Introduction: North American Indian historical linguistics in current perspective. The languages of Native America: Historical and comparative assessment, ed. by Lyle Campbell and Marianne Mithun, 3-69. Austin: University of Texas Press.

Chafe, Wallace. 1987. Review of Greenberg 1987. Current Anthropology 28:652-53.

Dimmendaal, Gerrit. 1992. Review of Nicolaï 1990. Bulletin of the School of Oriental and African Studies 55.610-12.

Fleming, Harold. 1983. Surma etymologies. Nilotic studies: Proceedings of the International Symposium on Languages and History of the Nilotic Peoples, Cologne, January 4-6, 1982, ed. by Rainer Vossen and Marianne Bechhaus-Gerst, 524-555. Berlin: Dietrich Reimer.

Floyd, Edwin D. 1981. Levels of phonological restriction in Greek affixes. Bono homini donum: Essays in historical linguistics in memory of J. Alexander Kerns, ed. by Yoel L. Arbeitman and Allan R. Bomhard, 87-106. Amsterdam: John Benjamins.

Foley, William A. 1986. The Papuan languages of New Guinea. Cambridge: Cambridge University Press.

Goddard, Ives. 1987. Review of Greenberg 1987. Current Anthropology 28:656-57.

Greenberg, Joseph H. 1971. The Indo-Pacific hypothesis. Current trends in linguistics, vol.8: Linguistics in Oceania, ed. by Thomas F. Sebeok, 807-71. The Hague: Mouton.

Greenberg, Joseph H. 1987. Language in the Americas. Stanford: Stanford University Press.

Greenberg, Joseph H. 1990. Address to NACAL 18, Emory University, March 1990. California Linguistic Notes 20.3:8-10.

Greenberg, Joseph H.; Christy G. Turner II; and Stephen L. Zegura. 1986. The settlement of the Americas: A comparison of the linguistic, dental, and genetic evidence. Current Anthropology 27:477-97.

Guthrie, Malcolm. 1948. The classification of the Bantu languages. London: Oxford University Press.

Haas, Mary R. 1951. The Proto-Gulf word for *WATER*. IJAL 17:71-79.

Haas, Mary R. 1954. The Hokan-Coahuiltecan word for 'water'. Papers from the Symposium on American Indian Linguistics, ed. by C. Douglas Chrétien et al., 57-62. (University of California Publications in Linguistics 10.) Berkeley & Los Angeles: University of California Press.

Heine, Bernd. 1992. African languages. International encyclopedia of linguistics, ed. by William Bright, vol. 1:31-36. Oxford: Oxford University Press.

Jakobson, Roman. 1962. Why "mama" and "papa"? Selected writings, vol.1:538-45. The Hague: Mouton.

Janhunen, Juha. 1989. Mother tongue: Newsletter of the Long Range Comparison Club (January), 28-30.

Jones, Sir William. 1798. Third anniversary discourse: On the Hindus. [Delivered 2 February 1786.] Asiatick Researches 1:415-31.

Jones, Sir William. 1799a. Sixth anniversary discourse: On the Persians. [Delivered 19 February 1789.] Asiatick Researches 2:43-66.

Jones, Sir William. 1799b. Eighth anniversary discourse: On the borderers, mountaineers, and islanders of Asia. [Delivered 24 February 1791.] Asiatick Researches 3:1-20.

Justeson, John S., and Laurence D. Stephens. 1980. Chance cognation: A probabilistic model and decision procedure for historical inference. Papers from the Fourth International conference on historical linguistics, ed. by Elizabeth Traugott et al., 37-45. Amsterdam: John Benjamins.

Kimball, Geoffrey. 1992. A critique of Muskogean, "Gulf", and Yukian material in *Language in the Americas*. IJAL 58:447-501.

Krauss, Michael E. 1979. Na-Dene and Eskimo-Aleut. The languages of Native America: Historical and comparative assessment, ed. by Lyle Campbell and Marianne Mithun, 803-901. Austin: University of Texas Press.

Levine, Robert D. 1979. Haida and Na-Dene: A new look at the evidence. IJAL 45.157-70.

Maddieson, Ian. 1984. Patterns of sounds. Cambridge: Cambridge University Press.

Mous, Maarten. 1993. The making of Ma'a. Paper presented at the Colloquium on Synchronic and Diachronic Sociolinguistic Methods and Interpretations, University of Bayreuth, December 1993.

Munro, Pamela. 1994. Gulf and Yuki-Gulf. Anthropological Linguistics 36:125-222.

Newman, Paul. 1980. The classification of Chadic within Afroasiatic. Leiden: Universitaire pers Leiden.

Nicolaï, Robert. 1990. Parentés linguistiques (à propos du songhay). Paris: CNRS.

Penn, William. 1683. Letter to the Free Society of Traders.

Poser, William J. 1992a. The Salinan and Yurumanguí data in *Language in the Americas*. IJAL 58: 202-29.

Poser, William J. 1992b. Message posted on the LINGUIST list, moderated by Anthony Rodrigues Aristar and Helen Dry, 3-327 (8 April 1992).

Poser, William J., and Lyle Campbell. 1992. Indo-European practice and historical methodology. Paper presented at the Berkeley Linguistics Society, 15 February.

Rankin, Robert L. 1992. Review of Greenberg 1987. IJAL 58:324-51.

Ringe, Donald A., Jr. 1992. On calculating the factor of chance in language comparison. (Transactions of the American Philosophical Society, vol. 82, part 1.) Philadelphia: The American Philosophical Society.

Sapir, J. David. 1987. Review of Greenberg 1987. Current Anthropology 28:663-64.

Schadeberg, Thilo C. 1981. The classification of the Kadugli language group. Nilo-Saharan: Proceedings of the First Nilo-Saharan Linguistics Colloquium, Leiden, September 8-10, 1980, ed. by Thilo C. Schadeberg and M. Lionel Bender, 291-305. Dordrecht: Foris.

Thomason, Sarah Grey. 1983. Genetic relationship and the case of Ma'a (Mbugu). Studies in African Linguistics 14.195-231.

Thomason, Sarah Grey. 1993. Coping with partial information in historical linguistics. Historical linguistics 1989: Papers from the 19th International Conference on Historical Linguistics, ed. by Henk Aertsen and Robert Jeffers, 485-96. Amsterdam: John Benjamins.

Thomason, Sarah Grey, and Terrence Kaufman. 1988. Language contact, creolization, and genetic linguistics. Berkeley & Los Angeles: University of California Press.

Thompson, Laurence C. 1979. Salishan and the Northwest. The languages of Native America: Historical and comparative assessment, ed. by Lyle Campbell and Marianne Mithun, 692-765. Austin: University of Texas Press.

Wald, Benji. 1994. Comparative method in linguistics: Karomojong/Jie. LINGUIST list, moderated by Anthony Rodrigues Aristar and Helen Dry, 5-1404 (7 December 1994).