Reconstructing Proto-Muskogean Language and Prehistory: Preliminary results

George Aaron Broadwell University at Albany, State University of New York

Recent years have seen an upsurge in interest in Muskogean linguistics, and considerable progress has been made in understanding the prehistory of these languages and in reconstructing a vocabulary for proto-Muskogean.¹ This paper will argue that this reconstructed vocabulary provides us with information about the branching order of the languages within the family, tentative dates for language separation, and evidence about the environment of the Proto-Muskogeans.

1. The classification of the languages

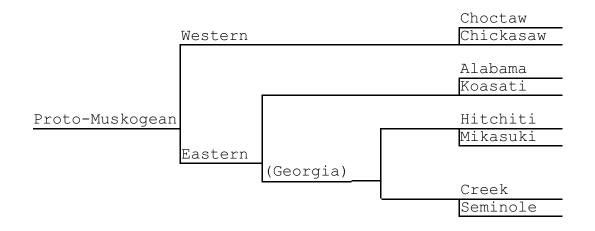
The Muskogean family contains four groups of closely related languages. Those spoken in this century are

- a.) Choctaw and Chickasaw
- b.) Alabama and Koasati
- c.) Hitchiti (now extinct) and Mikasuki
- d.) Creek and Seminole²

Classification above this level is controversial.

1.1 Theories of classification

The most generally known classification of the Muskogean languages is due to Haas (1941), who argued that the family contains two large groups -- Western Muskogean (consisting of Choctaw and Chickasaw) and Eastern Muskogean (composed of the other languages of the family). She was not explicit about the subgrouping of the Eastern Muskogean languages, but her remarks are generally interpreted as supporting the following tree:

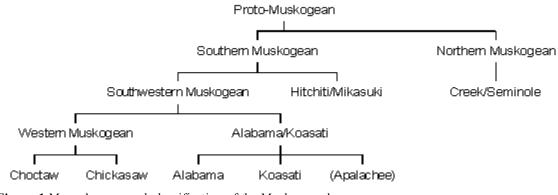


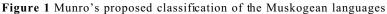
Note that one of the subgroups, Creek-Seminole and Hitchiti-Mikasuki, has no generally accepted name in the literature. For ease of discussion, I suggest that we call this suggested subgroup Georgia Muskogean, since all of the languages in it were spoken in the modern state of Georgia (along with adjacent areas of Alabama and S. Carolina).

Munro (1987, 1993) discusses evidence for another

2

classification, which is essentially the mirror image of Haas's classification. It is shown below:



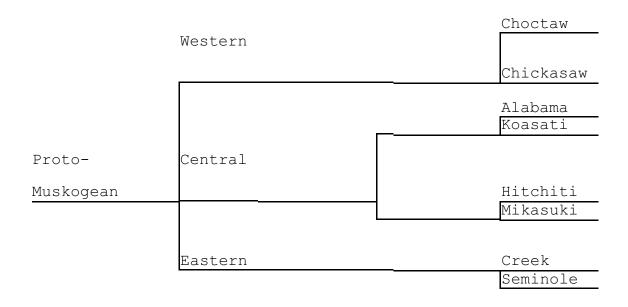


A disclaimer with respect to the labels is perhaps needed here. The terms Northern and Southern Muskogean are from Swanton (1922), and are used because they have a prior history in the literature. These geographical terms are inappropriate for some of the languages involved (e.g. Koasati, which was spoken to the north of the other languages). They are merely intended as convenient ways to discuss the groups. Note that nothing in this argument hinges on the particular names assigned to the groups.³

The two classifications differ most in their treatment of the `central groups', Alabama/Koasati and Hitchiti/Mikasuki, so arguments for one or the other of these classifications hinge on finding ways in which these groups resemble or fail to resemble the two most peripheral groups, Western Muskogean (Choctaw and Chickasaw) and Creek/Seminole.

A third suggestion for the classification of the family has been offered by Kimball (1989), who proposes that the Alabama-Koasati and Hichiti-Mikasuki groups form a subgroup which he calls Central Muskogean. Kimball suggests that Proto-Muskogean split into three branches: Western Muskogean, Central Muskogean, and Eastern Muskogean.⁴ I will argue below that this hypothesis is less successful than either the Haas or Munro proposal.

4



Nicklas (this volume) argues that the theory underlying such family tree models (or stammbäume) has been shown to be invalid, and thus argument about which of the classifications is correct is pointless.

However, it is a clear overstatement to claim that family tree models are discredited in historical linguistics. A glance at any textbook of historical linguistics shows that such models are widely accepted and used in the discipline (Bynon 1977, Anttila 1989).

The true historical relationships between languages are always complex; our diagrams can only selectively model certain aspects of this history. Family tree models of the sort shown above emphasize the relative chronology of change, and they are of course idealizations of linguistic history. When chronology is not the paramount consideration, other models of change may be more appropriate.

However, for the purposes of correlating linguistic and archeological prehistory, family tree models (in association with methods of estimating time depth) have clear advantages. They allow us to form hypotheses about the dates of prehistoric linguistic groups and to reconstruct elements of their culture and environment in ways which can be tested against the archaeological record.

1.2 Lexicostatistical data

Examining lexicostatistical data for the Muskogean languages provides us with information which is important both for deciding between alternate models of Muskogean prehistory and for correlating the results of historical linguistics with archaeology.

1.2.1 The concept of lexicostatistics.

Lexicostatistics (or glottochronology) is a method for estimating time depth in historical linguistics. It originated with Swadesh (1951, 1952, 1954, 1955) and is based on the analogy of carbon-14 dating. Swadesh

6

hypothesized that there is a core vocabulary in language which is resistant to borrowing, and that this vocabulary is retained at a constant rate, with about 14% lost per 1000 years. For two languages known to be related, the number of shared items is a function of the amount of time since they became distinct languages.

The method is imprecise and controversial (see criticisms in Hymes 1960, for example). However, lexicostatistics is the only quantitative method for estimating time depth in linguistics.⁵ Since correlations between reconstructed proto-languages and archeological cultures require some way of dating proto-languages, lexicostatistical estimates are a useful part of the discussion.

Two common critiques of lexicostatistical methods seem valid to this author. First, lexicostatistics should not be used to <u>establish</u> language families; it is instead a method for estimating their time depths. Second, the items compared in a lexicostatistical count ought to be true cognates, not chance resemblances.

Neither of these criticisms applies to the data in this paper. The Muskogean languages are indisputably related to each other, and the items compared are only true cognates, established independently through the comparative method.

7

1.2.2 Methods for this paper

The appendix to this paper contains a lexicostatistical list for five Muskogean languages: Choctaw, Chickasaw, Alabama, Mikasuki, and Creek.

Four words had to be omitted from the list: <u>bark</u> (of a tree), <u>feather</u>, <u>leaf</u>, and <u>moon</u>. In all the languages, <u>bark</u> is `tree' + `skin', <u>leaf</u> is `tree' + `hair', and <u>moon</u> is `sun' + `night'. In Choctaw and Alabama, <u>feather</u> is `bird' + `hair'. Since inclusion of these items would lead to the same lexical item being counted twice, they have been excluded, and the total sample is 96 words.

The Choctaw (Ct.) and Chickasaw (Cs.) data contain an additional complication. There are two words for hair in these languages: Ct. pāshi, Cs. pāshi' refers to hair on the head, while Ct. hishi, Cs. hishi' refers to body hair, animal hair, and feathers. The second of the two words is clearly cognate to the words in other Muskogean languages, but it is not clear that `body hair' is the primary sense of `hair'.⁶ Since deciding which Western Muskogean word to list affects the cognate count, I have counted the Western Muskogean as half cognate to the words in the other languages.

The Mikasuki data contain only 95 words because of the

lack of a Mikasuki word for `mountain'. Jack Martin (personal communication) tells me that it is not clear that there is a common Mikasuki word for `mountain'.⁷

In choosing the words for the list, I followed Swadesh's dictum that the most common conversational equivalent for the English word should be used. In several cases, this resulted in failure to list known cognates. For example, the usual Choctaw word for `mountain' is <u>habik</u>, which is not cognate with the other forms listed. There is a Choctaw word <u>bokko</u> which means `hill', and this is cognate to the Alabama form listed.

In determining whether the listed forms are cognate, I deferred in all cases to Munro et al (1991). I do not include words in that work which are cited as possible comparisons in the notes that follow the sets.

1.2.3 Discussion of the results

Assuming an 86% retention rate per thousand years (Swadesh 1954) for items on the 100 word list, the data yield the following dates of separation, rounded to the nearest decade:

Langua ge Pair	Percentage of cog- nates	Estimated separation date
Choctaw - Chickasaw	85% (82/96)	540 BP (AD 1450) ± 140 yrs.
Choctaw - Alabama	61% (58.5/96)	1640 BP (AD 350) ± 260 yrs.
Choctaw - Mikasuki	56% (53.5/95)	1920 BP (AD 70) ± 290 yrs.
Choctaw - Creek	41% (39. 5/96)	2950 BP (960 BC) ± 380 yrs.
Chickasaw - Alabama	66% (63.5/96)	1380 BP (AD 610) ± 230 yrs.
Chickasaw - Mikasuki	56% (53.5/95)	1920 BP (AD 70) ± 290 yrs.
Chickasaw - Creek	45% (43.5/96)	2650 BP (660 BC) ± 350 yrs.
Alabama - Mikasuki	63% (60/95)	1530 BP (AD 460) ± 250 yrs.
Alabama - Creek	51% (49/96)	2230 BP (240 BC) ± 320 yrs.
Mikasuki - Creek	55% (52/95)	1980 BP (10 BC) ± 290 yrs.

The range of error is computed at the 7/10 confidence level according to the procedures in Gudschinsky (1956).

1.2.3.1 Implications for subgrouping

Which of the three models of relationship do these figures most strongly support? We can decide by looking at

how similar the `central languages', Alabama and Mikasuki, are to Creek and Western Muskogean. Munro's model predicts that Alabama and Mikasuki should show the closest relationship to Western Muskogean, and further that Alabama should show the strongest similarities to Choctaw and Chickasaw; Kimball's model predicts that Alabama and Mikasuki should be equally close to Creek and Western Muskogean; and Haas's model predicts that Alabama and Mikasuki should be closest to Creek.

The figures above support the predictions of Munro's model. The situation is clearest with respect to Alabama, which shares 61% cognates with Choctaw and 66% with Chickasaw, but only 51% with Creek. Neither the Haas nor the Kimball model predict this degree of similarity.

Figures for Mikasuki are less conclusive. Mikasuki shares 56% cognates with Western Muskogean, but 55% with Creek. These results are unproblematic for the Kimball classification: his model predicts that the `central languages' are equally distant from Western Muskogean and Creek. The results are also unproblematic for Munro's model, since that classification claims that the relationship between Mikasuki and Western Muskogean is a more distant one than that between Alabama and Western Muskogean.

However, the failure of Mikasuki to show a particularly

close relationship to Creek <u>does</u> pose a problem for Haas's classification. Haas's model predicts that the closest relationship in the family ought to be that between Mikasuki and Creek, but the data do not bear that prediction out.

When the data for Alabama and Mikasuki are considered, the Munro model is the most successful model overall in predicting the degrees of similarity between the Muskogean languages. In the following sections I will discuss other evidence that also supports Munro's hypothesis about Muskogean linguistic pre-history.

1.2.3.2 Implications for dating

Accepting the most distant dates for each stage of the reconstructed proto-language, we arrive at the following figures:

1.)	Proto-Muskogean	2950 BP (960 BC) ± 380 yrs.
	Proto-Southern Muskogean	1920 BP (AD 70) ± 290 yrs.
	Proto-Southwestern Muskogean	1640 BP (AD 350) ± 260 yrs.

The dates above accord fairly well with the intuitive degrees of similarity between the languages.⁸

Although Alabama is somewhat more similar to Creek than Choctaw is, it is clear from these data that Alabama and Choctaw are most similar to each other. I suggest that the similarities between Alabama and Creek are due primarily to culture contact between these two peoples. As members of the Creek Confederacy, the Alabama had intense contact with Creek language and culture, and are likely to have borrowed words that increase the apparent similarity between the languages.⁹

1.2.3.3 Non-lexical evidence

The discussion so far has focussed on lexical evidence for the relationships within the family. There is fairly extensive non-lexical support for both the Haas and Munro classifications.

Haas's (1941) classification was supported by several proposed sound changes, most notably the development of proto-Muskogean *N and the sibilants.

However, as Munro (1987) notes, there are also several phonological and morphological isoglosses that support the Munro classification:

... the four Southwestern languages share a number of phonological and morphological traits which all appear to be innovations. These include a very unusual assimilation rule affecting the /-li/ auxiliary suffix (Munro 1985); the development of an /-l-/ passive infix; the use of plural *ha in first-person plural I affixes (Booker 1980:30; ...) and in the second-person plural II affixes (Munro [1993]); the spread of initial /a-/ through the complex II paradigm (Munro [1993]) and considerable similarity in the system of aspectual grade formation...

Other phonological and morphological innovations that support a subgrouping of Southern Muskogean are the development of Proto-Muskogean *k^w (Haas 1947), and the presence of subtractive morphology (Broadwell 1987a, 1993). Discussion of this evidence would take us too far afield, but I will note that while the phonological, morphological, and lexical evidence are all supportive of Munro's classification, the lexical evidence is perhaps the strongest.

2. Reconstructing proto-Muskogean environments

The lexicostatistical data provide estimates of the time depth for various stages of the historical development of the Muskogean languages. It is now possible to ask what vocabulary is reconstructable at these dates, and what sort of correlations there are between the linguistic data and the archaeological record.

Munro et al (1991) contains a particularly detailed set

of terms for flora, fauna, and agriculture. Examination of these items allows one to draw several conclusions about Muskogean environments and the ways in which they have changed through history.

2.1 Flora and fauna

The Proto-Muskogean vocabulary has 107 reconstructible terms for flora and fauna, which provide a rich view of their environment.¹⁰ The terms are given in (1) below. The names used are keyed to the names of the sets in Munro et al (1991), and additional information is provided in parentheses after some of the items.

2. <u>Reconstructable Proto-Muskogean terms</u>

apple (probably crab-apple or persimmon), bat, bee, beetle, bluejay, briar (blackberry), briar (smilax/arrowroot), buckeye, buffalo, chestnut, chicken snake, chickenhawk, chigger, chinquapin, chipmunk, civet cat (?), clam/spoon, copperhead, corn, cotton, crawfish, crane 1 (whooping crane), crane 2 (heron?), cricket, daddy-long-legs, deer, devil's shoestring, dove, duck 1 + 2, earthworm, falcon, flea, fly, frog 1, gizzard, gobble, goose, grape, grasshopper 1, grasshopper 2 (katydid),

hackberry, haw, hickory, hoe/plow, honey locust, hornet/wasp, horsefly, hummingbird, insect/worm, lamb's quarters, leech, lightning bug, lizard 1, locust/cicada, louse, martin, milkweed, mole, moss, mountain lion 2, muddauber, mulberry, muscadine, mushroom, oak 1 (postoak), oak 3 (white), onion, opossum, hoot owl, horned owl, screech owl, palmetto, perch, pigeon, plant sp. (cattail/beargrass), pokeweed, potato, prickly pear, pumpkin, quail, rabbit, redbud, ringworm, skunk, slippery elm, snake, spider, squirrel, stinging plant (poison ivy?), tadpole, thrush, trout, turtle 1, turtle 2 (soft-shelled), walnut, water lily, whippoorwill 1 + 2, wildcat, woodpecker 1, redheaded woodpecker, worm 1 + 2, wren 1, yellowhammer

In all the sets just listed, there is evidence that allows reconstruction of these items at the earliest point in Proto-Muskogean linguistic prehistory.

Of particular interest to this discussion, however, are terms which are not evenly distributed across the entire family, but only in some subset of the languages. Historical linguistics labels grammatical, phonological, or lexical features shared by some subgroup of a family as <u>isoglosses</u>. In general, the more isoglosses two languages share, the more closely related they should be. There are 81 terms which are reconstructable only at the level of some subgroup within Muskogean. They thus constitute lexical isoglosses, and they are listed below.

3. <u>Southern Muskogean</u> (19 items)

axe, bullfrog (2), butterfly, buzzard, cardinal, catfish 2, crow, cypress, fern, mosquito, mountain lion/panther, oak 2 (blackjack), persimmon, raccoon, sassafras, snail, tobacco 1, turkey, wolf

4. <u>Southwestern Muskogean</u> (35 items)

acorn 2, alligator 1, ant, ash (tree), bear 1, beech, bobcat, bottom land, bumblebee, cached food, catfish 1, cedar 1, cherry, corn silk, dogwood, eagle 2, elder, flying squirrel, frog 2, grass 1, huckleberry 1, lizard 2 (skink), lizard 3 (amphiuma), magnolia, meadowlark, mimosa, mint, oak (overcup), robin, strawberry, sycamore, tick 1, toad, winnow, wren 2

5. <u>Georgia Muskogean</u> (11 items)

acorn 1, bear 2, bullfrog 1, button snake root, cedar 2 (a possible loan from Cherokee), eagle 1, eel, gourd 1, tea (probably ilex vomitoria), thistle, tobacco 2.

- 6. <u>Eastern Muskogean</u> (6 items) alligator 2, cottonwood, grass 2, mouse, pelican, red bay tree
- 7. <u>Central Muskogean</u> (2 items) crane, pygmy rattlesnake
- 8. <u>Creek/Seminole and Alabama/Koasati</u> (8 terms) adze, buttonbush, catfish 3, doodlebug (pillbug), gourd 1, kingfisher, passion flower, woodtick

No sets show an isogloss for Western Muskogean plus Hitchiti/Mikasuki, the only other logical combination of languages.

2.2 Implications for classification

The genetic subgrouping implied by the the Munro classification is strongly supported by the terms for flora and fauna. Of the 81 isoglosses, those for Southern Muskogean and Southwestern Muskogean support the Munro classification, and constitute 66% (54/81) of the sample. Those terms that are reconstructed for Eastern Muskogean or Georgia Muskogean support the Haas classification, but constitute only 21% (17/81) of the sample. Even if one adds the sets with Creek/Seminole and Alabama/Koasati cognates, there are only 24 sets (30%).

Munro's classification is supported by more than twice as many isoglosses as Haas's classification is. The results here correlate well with the results reached in the previous section. There is thus very strong evidence for the existence of a Southwestern Muskogean group.

Kimball's Central Muskogean hypothesis fares far worse than either the Haas or the Munro classification. Only two items, constituting 2.4% of the data, support his position.

2.3 Directions of borrowing

Isoglosses are valid data for subgrouping only when we can be reasonably sure that the features which languages share are not the result of borrowing. Just as a historically oriented study of the English lexicon would attempt to discern French loan words, so too we must try to determine what portion of the isoglosses identified above are invalid due to borrowing.

I will use two interrelated criteria for identifying suspected borrowings:

a.) lack of a cognate in the most closely relatedlanguage. For example, if a cognate set consists only

of a Chickasaw and an Alabama word, then it is more likely to be a borrowing than a set that contains Chickasaw, Alabama, and Koasati cognates. b.) identical or nearly identical words in the languages. Of course, this is a necessary but not a sufficient criterion for identifying a loan.

It is likely that the isoglosses most affected by borrowing are those for Southwest Muskogean and Georgia Muskogean, since in both cases there was close contact between the speakers of some of the languages within the subgroup. The isoglosses for larger groups such as Eastern Muskogean and Southern Muskogean are less likely to be contaminated by borrowing, since a lexical item which is so widely distributed through the family is unlikely to be borrowed.

2.3.1 Borrowing in Southwest Muskogean

By the criteria just mentioned, the following nine Southwest Muskogean isoglosses seem suspicious:

9.) ALLIGATOR (1). Al. haconcoba; Ct. hachõchobah, Cs. hachõ'choba' BEECH. Al. tomalaaha; Ct. hatõbalaaha

CHERRY. Al. ittotalikco, itocalikco; Ct. itti alikchi, Cs. itti' alikchi'. FROG (2). Al. cooto; Cs. chõ'ti'. HUCKLEBERRY (1). Al. osakohci, sosakohci; Cs. osakokchi'. MAGNOLIA. Al. kasaha, kaslhaha-hatka `white magnolia', kalhâaha; Ct. kolhaha. MINT. Al. sonok kilhâyli, sinoktilhâyli, snoktilhâyli; Ct. shinoktilhiili, MCt. anoktilhiili `medicinal plant used for fever', Cs. sholop tilhi'li', shoptilhi'li' `horsemint'. TICK (1). Al. satani; Ct, Cs shatanni, Ct. shitanni. WINNOW. Al. immaska; Ct. mashichih, Cs. mashka.

These sets are suspicious in that the Alabama word closely resembles the Western Muskogean, but there is no Koasati cognate. Additional problems in the etymologies of some sets also suggest borrowing.¹¹

As Munro et al. (1991) notes, ALLIGATOR appears to be composed of the words for `tail' (Al., K haci; Ct. has bish, Cs. hasimbish) plus the word for `big' (Al, K coba; Ct. chito, Cs. ishto). However, this etymology is only available for the Alabama form; neither `tail' nor `big' has the right form in W. Muskogean. This suggest that Western Muskogean borrowed the word from Alabama.

CHERRY and HUCKLEBERRY (1) have etymologies in the Western Muskogean languages, but either no etymology or an implausible one in Alabama. CHERRY is composed of the words for `tree' and `doctor' in the Western Muskogean languages, but this etmology will not work for the Alabama forms. The first form, <u>ittotalikco</u>, appears to be `tree' plus `tie', while the second has no etymology. It seems plausible that the Alabama words are borrowed from Western Muskogean, and subjected to folk etymological restructuring.

Similarly, HUCKLEBERRY (1) is analyzable as `hickory' + `juice' in Chickasaw, but has no etymology in Alabama.

Excluding these nine suspicious sets reduces the number of isoglosses supportive of Munro's classification to 45.

2.3.2 Borrowing in Georgia Muskogean

When we examine the reconstructable terms for Georgia Muskogean, there are also several sets that look suspiciously like borrowing. I think the following seven sets are possible borrowings: 10.) BEAR. Cr, S, OS nokosi; H, Mk nokos-i. BULLFROG (1) Cr. apatana; H. apatan-i. BUTTON SNAKE ROOT. Cr, OS paassa; H, Mk pas-i. CEDAR (2). Cr, S, OS acina; H, Mk acin-i. GOURD (1). Cr. ifipi; Mk. ifip-i. THISTLE. Cr. akaaca, akaaco; H. akac-i. TOBACCO (2). Cr. hici; H. hic-i.

A set is suspicious if the Creek word is identical or nearly identical to the Hitchiti or Mikasuki word, and more suspicious if it is attested only in Hitchiti or Mikasuki, but not both.

Further information makes some of the sets additionally suspicious. The ordinary word for tobacco in Hitchiti and Mikasuki is <u>akcomi</u>. The only instance of the Hitchiti word <u>hici</u> `tobacco' occurs in Swanton (1929) in a myth about the origin of tobacco. The myth is nearly identical to a Creek myth, in which the word <u>hici</u> is compared to the verb <u>hicita</u> `to see'. Since the story does not make sense unless the Creek word is used, we should be suspicious that <u>hici</u> is a true Hitchiti word.

CEDAR was identified by Haas (1941) as a loan from Cherokee, and therefore should also be discounted.

It is also clear that several of the isoglosses for Georgia Muskogean involve items important to Creek culture. 'Bear', 'eagle', and 'eel' are the names of Creek clans (Swanton 1946:658), and cedar, button snake root, ilex vomitoria (or `tea'), and tobacco all have ritual and medicinal purposes.¹² Hitchiti or Mikasuki speakers would have certainly been familiar with the Creek words for these items.

Excluding the seven suspicious sets mentioned above reduces the number of isoglosses supportive of the Haas classification to 17.

2.3.3 Revised figures for isoglosses

If we exclude the 16 suspicious sets identified for Southwest Muskogean and Georgia Muskogean, then we still have 65 isoglosses. 69% (45/65) of the isoglosses support the Munro model, 25% (18/65) support the Haas model, and 3% (2/65) support the Kimball model.

2.4 Implications for the environment of the Proto-Muskogeans

The reconstructable terms for flora and fauna give us a relatively rich view of the environment of the Muskogeans at different points in history.

Since cultigens are the plants whose chronologies are best understood, the appearance or lack of appearance of terms for particular cultigens gives us the opportunity to correlate the dates provided by the lexicostatistical estimates with archaeological findings.

2.4.1 Proto-Muskogean cultigens

While we cannot conclude from the presence of a word for a plant in the vocabulary of a proto-language that speakers of that language cultivated that plant, we <u>can</u> reasonably conclude that speakers of the langauge were familiar with the plant. Reconstructable for Proto-Muskogean are lamb's quarters (chenopodium), squash (cucurbitacea) and corn (zea mays).

Lamb's quarters is reconstructable on the basis of the following set:

11.) LAMB'S QUARTERS. Cr. taahwa; Cs. taani'.

In this case, we only have the word from two languages, but the correspondence is excellent. The linguistic results are fully in accord with generally accepted archeological estimates for the domestication of chenopodium (Muller 1978).

Squash/pumpkin is reconstructable on the basis of the following set: $^{\scriptscriptstyle 13}$

12. PUMPKIN. Cr/S/OS casi; H cosk-i; A/K coksi; Ct/Cs shokshi `watermelon', Ct. shokshobok `gourd'

Once again, the linguistic evidence is compatible with archaeological evidence for relatively early cultivation of squash in the Southeast (Muller 1978).

However, the date for corn correlates less well. The cognate set in question is given below:

13.) CORN. CR, S, OS aci; H, Mk asp-i; Al, K cassi; Ct. tãchi, Cs. tanchi'.

This is not a perfect set, by any means, but it seems likely that it is a valid one. I would tentatively reconstruct *(t/c)aci.

The Creek/Seminole and Hitchiti/Mikasuki forms are the most easily comparable, since the H/Mk forms appear to include a bound form of the noun <u>api</u> `body, stalk, cob'. The occurrence of /as/ rather than /ac/ can then be explained by a rule common to many Muskogean languages which turns / / to /s/ or /sh/ in syllable final position.

The Western Muskogean forms are also quite similar, the chief problem being the initial /t/.

I am inclined to regard the Alabama/Koasati word as non-cognate with the rest. However, if we accept the

Western Muskogean forms as cognate with the Creek and Hitchiti/Mikasuki forms, then clearly the set is reconstructable at the level of Proto-Muskogean.

How can we reconcile the presence of a word for corn with the generally accepted archaeological position that corn was not present in the southeast until considerably later, ca. A.D. 700?

Riley, Edging, and Rossen (1990) survey the evidence and find some evidence for positing corn as early as 200 B.C. At this early date, corn was not the staple crop it later became, but it might well have been present and familiar to the proto-Muskogeans. The lexicostatistical data support a date for Proto-Muskogean of 900 BC ± 380 years, so the linguistic evidence points to an earlier presence of corn in the southeast than the archaeological evidence does.

Two other reconstructable items strengthen the conclusion that the proto-Muskogeans knew corn. The first is the verb to SHELL CORN, and the second is a noun meaning corn riddle (BASKET 3).

A common approach in dismissing linguistic evidence that does not correlate with the archaeological results is to suggest that the reference of the words has changed through time (cf. Renfrew 1988). For example, the word for corn might have originally referred to some other grain. When corn was introduced to the southeast the word for the older grain might have been applied to the new-comer.

However, it seems unlikely that speakers of all the different languages in the family would have coincidentally decided to call the new grain the same thing. Once a language has split into two mutually unintelligable daughter languages, the speakers do not consult with each other about naming new phenomena.

The unlikeliness of this hypothesis increases when we realise that we must also assume that the words for shucking corn and corn riddle originally applied other actions and objects, and that once again widely separated people have coincidentally chosen the same words for actions and objects associated with the new grain.

I therefore conclude that presence of a word for corn in Proto-Muskogean constitutes a genuine conflict between the linguistic and archaeological data.

2.4.2 Tobacco as a Proto-Southern Muskogean cultigen

Tobacco is reconstructable at the level of Southern Muskogean, as shown in the following set:

14.) TOBACCO (1). H, Mk akcom-i; Al, K hakcomma, K hakcommi, Ap hakcoma; Ct. hakchoma, Cs. chomak. The Creek word for `tobacco' is <u>hici</u>, a word which is also found in Hitchiti. As argued in section 2.3.2, it likely that the Hitchiti word <u>hici</u> is a borrowing from Creek, and that <u>akcomi</u> represents the native Hitchiti word.

Fortunately, there is some additional information on the Creek word. Swanton (1929) recorded three myths of the origin of tobacco, and the first two of these give us a native explanation of the etymology of the Creek words. In these two versions, a couple have had sexual intercourse at a place where tobacco is later found to be growing. The people must decide what to call the plant.

The first version states ``the first name of the plant was `coeuns' (haisa). After they learned of it and came to value it, they made it a warrior (tasikaya) and gave it the name hitci (`finding') as a war name."

In the second version, the people discover the plant and say ``We shall call it hitci, and when we smoke we shall call it the same as quum coimus (haisa)."

Both versions of the story state that tobacco has two names: <u>haisa</u> and <u>hitci</u>, and both names have etymologies.

<u>Haisa</u> probably means `penis'. It is translated by `quum coimus' and `coeuns', Latin euphemisms for `penis', and the word Swanton cites, <u>haisa</u>, looks rather similar to the modern Creek word <u>haswa</u> `penis'. An etymology that

29

relates the word for tobacco to the word `penis' is also supported by the plot of the story.

<u>Hici</u> is derived from the verb <u>hic-ita</u> `to see', and in one version this is described as a war name for the plant. Since war names typically described attributes of the named person, it may be that tobacco is called `seer' because it causes visions. Alternately, the stories may mean that tobacco is the `seen' thing.

Speck (1909) records a similar Yuchi story of the origin of tobacco, in which the plant grows from sperm that falls on the ground (cited by Lankford 1987).

The folkloric connection between tobacco and sex found in Creek and Yuchi surprisingly corresponds to forms found in the Southern Muskogean languages. Consider the following forms for `penis', and compare them to the words for `tobacco' above:

15.) PENIS. H. akc-i; Al, K cici, ikci; Ct, Cs hakchin, Cs. inkilish.¹⁴

The additional syllable /om/ that appears in words for tobacco bears a similarity to an auxiliary meaning `like' in most of the Muskogean languages (Booker 1980). This suggests `penis-like' as an etymology for `tobacco'.

This rather surprising connection seems to be worth pursuing. Are phallic representations of tobacco to be

found in the southeastern archaeological record or in the folklore of other regions of the Americas?

Since there is not a reconstructable word for 'tobacco' in Proto-Muskogean, the family was probably split into mutually unintelligeable Northern and Southern languages when tobacco was introduced into the region.

Riley, Edging, and Rosen (1990) cite use of tobacco in the eastern U.S. as early as A.D. 100, and this correlates well with a date of about 1900 years for the split between Northern and Southern Muskogean.

Note that the term for `tobacco' sheds light on the difficulty we encountered with the word for `corn'. The linguistic evidence argues that the Muskogeans have been familiar with corn for a longer time than they have been familiar with tobacco, since `corn' is reconstructable at an earlier level.

The term for tobacco also demonstrates that the problem posed by the word for `corn' cannot be solved by simply recalibrating the time calculations for various stages of the proto-language. If we argue that Proto-Muskogean is significantly younger than the 960 BC (± 380) estimate, then we encounter the problem of explaining why `tobacco' is not reconstructable.

2.4.3 Non-reconstructable cultigens

In contrast to the other cultigens discussed, there is no reconstructable word for `bean'. Consider the following words from Ulrich (1986):

16.) BEAN. Cr. talaako; Al. castoki, K. palana; Ct. tobi (beans in general), bala (a particular variety), Cs. bala'.

Our inability to reconstruct a word for bean is compatible with the view that beans were introduced into the southeast about 1000 years ago (Muller 1978). Note that the terms are different even in Alabama and Koasati, which are closely connected languages. We might therefore view the split between these two groups as predating the introduction of beans, perhaps not long after the split of Southwestern Muskogean into its two constituent groups, Western Muskogean and Alabama/Koasati.

2.4.4 Correlating linguistic and archaeological time depths

The dates for various stages of the break-up of the Muskogean family can be correlated with the archaeological dates as follows:

Estimated date	Linguistic events	Other events
960 BC ± 380	Proto-Muskogean	(early) corn, chenopodium, squash
AD 70 ± 290	Northern and Southern separate	introduction of tobacco
AD 350 ± 260	Southwestern and Hitchiti/Mikasuki separate	
AD 700?	Alabama and Koasati separate	
AD 1000		introduction of beans
AD 1450 ± 140	Choctaw and Chickasaw separate	

There are two crucial anchor points here. First, Northern Muskogean and Southern Muskogean must have become distinct prior to the introduction of tobacco, since they have different reconstructable words for the plant. Second, since all of the languages except Choctaw and Chickasaw have different words for beans, they were probably already been separate and mutually unintelligeable before the introduction of beans about 1000 years ago.

3. Conclusion

The work represented here is a preliminary effort towards applying the data from ongoing efforts in Muskogean historical linguistics to problems of history in the Southeast. I hope that the dialogue now beginning between linguists, archaeologists, ethnologists, and ethnohistorians of the Southeast will continue to explore these connections. Appendix: The 100 Word list¹⁵

(Items identified as cognate are co-boldfaced or co-underlined)

English	Chickasaw	Choctaw	Alabama	Mikasuki	Creek
all	mõma	mõma	óyha	maamos-	omalka
ashes	hottok	hitokchobi	histo	tolhambi	iisso
belly	ittakoba'	iffoka	ikfi	lampi	nalhki
big	ishto	chito	<u>coba</u>	<u>coob-</u>	lhakkii
bird	foshi'	hoshi	foosi	foosi	foswa
bite	<u>kisili</u>	kopooli	<u>kachalhlhi</u>	kabalikci	akkita
black	losa	losa	loca	looci	lasti
blood	issish	issish	lhakhani	picikci	caati
bone	foni'	foni	cokfoni	-fooni	iffoni
breast	ip shik	ip shik	pisi	owaaci	hokpi
burn	lowa	lowah	libatli	yill-	noklhita
claw	iyyakchosh	iyyakchosh	iyyaksi	iiyakoosi	ilinkososwa
cloud	hoshonti	hoshõti	onoolici	hosoti	aholocii
cold	kapassa	kapassa	kasatka	kapaali	kasappi
come	minti	m ti	ila	ont-	atita

English	Chickasaw	Choctaw	Alabama	Mikasuki	Creek
die	illi	illi	illi	il-	ilita
dog	ofi'	ofi	ifa	iifi	ifa
drink	ishko	ishko	isko	isk-	iskita
dry	shila	shila	solotka	sokook-	kalhpii
ear	haksibis	haksobish	hakco	hacoobi	hakco
earth	yakni'	yakni	ihaani	yakni	iikana
eat	impa	pa	ipa	imp-	hompita
egg	akankoshi′	akãkoshi	akaakocóòsi	onaasi	costaki
eye	<u>ishkin</u>	<u>nishkin</u>	ittilhi	iti	tolhwa
fat (grease)	niha	bila	nitokci	niihi	nihaa
fire	lowak	lowak	tikba	iiti	tootka
fish	nani'	nani	lhalho	lhaalhi	lhalho
fly, to	wakaa	hika	wakayka	yakaal-	tamkita
foot	iyyi'	iyyi	iyyi	iyi	ili
full	kayya	kayya	kayya	labakni	fackita
give	ima	ima	inka	iik-	imita
good	chokma	achokma	kano	<u>hiilhi</u>	<u>h lhi</u>
green	okchamali	okchamaali	okcakko	honotbitalakci	laani
hair	pãshi '/hishi'	pãshi/ hishi	hissi	tokisi	issi
hand	ilbak	ibbak	ilbi	ilbi	inki

English	Chickasaw	Choctaw	Alabama	Mikasuki	Creek
head	ishkobo'	noshkobo	isbakko	yoosi	ika
hear	hánglo	haklo	haalo	hakl-	pohita
heart	chõkash	chõkash	conoska	conosbi	fiiki
horn	lapish	lapish	lapihci	lap-i	yapi
I	ano'	ano	ana	aani	ani
kill	abi	abi	ibi	<u>ill c</u>	<u>iliicita</u>
knee	iyyinto'lhka'	iyyi kalaaha	ittôlhpa	tolhpi	tolhkowa
know	ith <u>á</u> na	ikhana	sobayli	ataalh	kilhlhita
lie down, to	tí'wa	talaaya	baláàli	talaal	wakkita
liver	salakha	<u>salakha</u>	illopi	lopi	lopi
long	falaa	falaaya	baski	backi	capki
louse	issap	issap	icha	hicahci	icka
man	hattak nakni'	hattak nakni	naani	nakni	honanwa
many	lawa	lawa	lawa	aconki	solkii
meat (flesh)	nipi'	nipi	nipo	akni	apiswa
mountain	onchaba	habik	bokkoscaaha		iikanhalwii
mouth	iti	iti <u>albi</u>	<u>icokhalbi</u>	ici	<u>cok</u> wa
name	holhchifo	hohchifo	holcifa	hocilki	hocifka
neck	nokhistap	ikkõla	nokbi	nokbi	nokwa
new	himitta	himmona	hahpa	himaci	mocasi

English	Chickasaw	Choctaw	Alabama	Mikasuki	Creek
night	oklhili'	ninak	tanka	niilhaki	nilhii
nose	ibichchala'	ibishakni	ibisaani	ibi	уороо
not	<u>ki'yo</u>	<u>kiiyo</u>	mánko	maati	monks
one	chaffa	achaffa	caffaaka	lhaamin	hamkin
person (hu-	hattak	hattak	aati	yaati	isti
man)					
rain	omba	õ ba	oyba	okoob-	oskita
red	homma	homma	homma	kitisci	caati
road (path)	hina'	hina	hini	hini	nini
root	haksish	hakshish	assikci	aski	yalomka
round	lhibokta	kalaaha	bonotka	polocki	polooki
say	aachi	<u>aachi</u>	manka	kaac	maakita
sand	shinok	shinok	sanco	samooci	oktaaha
see	p sa	p sa	hicha	hica	hicita
seed	nihi'	nihi	hilhikci	yiilhi	nilhka
sit	bínni'li	biniili	<u>cokóòli</u>	<u>cokool-</u>	leykita
skin	hakshop	hakshop	affakci	<u>halbi</u>	<u>halhpi</u>
sleep	nosi	nosi	noci	nooc-	nocita
small	iskanno'si	osi	cinoofa	wink-	cotki

English	Chickasaw	Choctaw	Alabama	Mikasuki	Creek
smoke	shobohli	shobohli	sobotli	<u>ockoci</u>	<u>ikkoci</u>
stand	híkki'ya	hikiiya	<u>lokóòli</u>	lokooka	hoylhita
star	foshik	fichik	hociilhi	owaaciki	kocacampa
stone	tali'	tali	tali	tali	cato
sun	hashi'	hashi	hasi	haasi	hasi
swim	yopi	okshiniili	oohapka	opahk-	omeyyita
tail	hasimbish	has bis	haci	haaci	haci
that	yamma	ma	akki	ma	ma
this	yappa	pa	уа	уа	ya
thou	ishno'	chishno	isna	cihn-	ciimi
tongue	isõlash	ittõlas	icoolaksi	cokolaasi	tolaaswa
tooth	noti'	noti	innati	-nooti	noti
tree	itti'	itti	itto	ahi	ito
two	toklo	toklo	tôklo	toklan	hokkoolin
walk	nõwa	nowa	<u>ciyahli</u>	<u>cayahl</u>	yakapita
5warm (hot)	lashpa	lashpa	ikba	hãyyi	<u>hayyita</u>
water	oka'	oka	oki	ooki	оуwа
We	poshno'	pishno	posna	pohni	poomi
what	nanta	natah	náàsi	naaki	naaki
white		<u>tohbi</u>	hatka	hatki	hatki

yellow	lakna	lakna	laana	lakni	laanii
woman	<u>ihoo</u>	<u>ohooyo</u>	tayyi	tayki	hoktii
who	kata	katah	náksi	noolh-	isteyma
English	Chickasaw	Choctaw	Alabama	Mikasuki	Creek

1. I thank Emmanuel Drechsel, Penelope Drooker, Heather Hardy, John Justeson, Geoff Kimball, Pat Kwachka, Jack Martin, Pamela Munro, Dale Nicklas, Dean Snow, and the audiences at the Southern Anthropological Society and University at Albany for their comments and suggestions on this paper. All mistakes are my own.

NOTES

The following abbrevations are used: : Al = Alabama, Ap = Apalachee, Cr = Creek, Cs = Chickasaw, Ct = Choctaw, H = Hitchiti, K = Koasati, MCt = Mississippi Choctaw, Mk = Mikasuki, OS = Oklahoma Seminole, S = Seminole. The following orthographic conventions are used in the citation of data from Muskogean languages: nasalized vowels are indicated with a tilde; <ch> (in Western Muskogean) and <c> (in other languages) represent []; <lh> represents [] (a voiceless lateral fricative); and <sh> (in Western Muskogean) represents [š].

The financial support of the Departments of Anthropology and Linguistics and Cognitive Science at the University at Albany, State University of New York is gratefully acknowledged.

2. In general this discussion omits the extinct Muskogean languages Apalachee (Kimball 1987, 1988), Guale and Yamasee (Broadwell 1991), since the available data are fragmentary. Haas (1949) argues that Apalachee is most closely related to

41

Alabama and Koasati. Broadwell (1991) suggests that Guale and Yamasee are most closely related to Creek.

3. Thus Kimball's (1989) objection to Munro's classification on the grounds of the names assigned to the branches is misguided.

4. Kimball is explicit about his hypothesis of a ternary split in the text of his paper, but the accompanying diagram shows all three possible branchings for Central Muskogean with question marks (i.e. as a separate branch, joined with Western Muskogean, and joined with Eastern Muskogean).

5. As Diamond (1992) notes, critics of lexicostatistics are often nevertheless willing to make estimates of time depth based on intuition.

6. In fact, there is some possibility that pāshi/pāshi' may also be cognate with the other words for `hair'. The etymology of these forms is something of a mystery, but one derivation might be PM *ik^wa `head' + *hisi `hair'. While PM *k^w generally develops into Western Muskogean /b/, there are some cases in Choctaw where initial /p/ and /b/ alternate, e.g. <u>pichilli/bichilli</u> `to ooze out', <u>bishlichi/wishlichi</u> `to milk' and <u>pishi</u> `to suck'.

7. Presumably this is due to the environment where Mikasuki is currently spoken.

8. Muskogean seems comparable in diversity with the

Romance family, which has a time depth of about 2500 years. This may suggest that the date for Proto-Muskogean is somewhat more recent than 960 BC figure given. A date of ca. 600 BC is within the range of error of the lexicostatistical calculations and consistent with the comparison to Romance.

9. Nicklas (this volume) suggests that the fact that the figures for an Alabama-Creek split and a Choctaw-Creek split differ from each other shows that lexicostatistical methods are invalid. This is a mistaken interpretation of the data.

The figures in the table above attempt to give us some quantitative measure of the degree of similarity between the Muskogean languages. That similarity is then used to estimate the degree of separation between the languages. However, some portion of that similarity is due to a distant common history, while another portion of the similarity is due to more recent influence and borrowing. It is entirely consistent with the results above to claim that the Choctaw and Alabama languages are equally distant from Creek, but that Alabama shows a greater modern similarity due to more recent Creek influence. In deciding which of the estimates of the separation date is more likely to be correct, we should thus prefer the more distant figure suggested by the Choctaw-Creek comparison.

43

10. The list below includes all the terms for flora and fauna listed in Munro et al. (1991), with the following exceptions: I did not include BIRD, BIRD SP., FISH SP, FISH (because they are not specific enough to give us useful information) or obvious borrowings from European languages (BACON, CAT, COFFEE, COW, GOAT, OKRA, RICE, TOMATO, WHEAT).

I reject the proposed cognate sets for BEAN, BEAVER, HUCKLEBERRY 2, PEANUT as improbable.

The cognate set for CHICKEN is a special problem. There is a similar word for `chicken' in many languages, but it must have originally applied to some other sort of bird, since the chicken is a European introduction.

Some sets contain lexical material duplicated in other sets, and I have attempted to include only one set in such cases. I include OPOSSUM, but not HOG, since they are from the same root. I exclude OWL 4, since it is the same root as HOOT OWL. I do not include the compounds ROADRUNNER (`fast bird') or WHALE (`water blow').

11. The following discussion relies heavily on the discussion of the cognate sets in Munro et al (1991).
12. As Martin (1987:117) notes, the phonology of the Mk. *nokosi* `bear' is unusual for a Mikasuki word, since we would expect lengthening of the initial syllable in words of this shape. This strengthens the case for treating this word as

44

a loan.

13. In many of the other languages, `watermelon' appears to be a compound based on `pumpkin' + a root /tal(ak)/ which may mean `lie down'.

14. Al, K *cici* probably originates as a children's word for penis. The origin of Cs. inkilish is obscure, and it is not clear that it is cognate to the other items.

15. Data for this appendix were provided by Heather Hardy (Alabama), Jack Martin (Creek and Mikasuki), and Pamela Munro (Chickasaw). Choctaw data comes from Byington (1915) and Broadwell (1987b).

REFERENCES

Anttila, Raimo. 1989. <u>Historical and comparative</u>

linguistics. Amsterdam: John Benjamins.

- Booker, Karen M. 1980. <u>Comparative Muskogean: Aspects of</u> <u>Proto-Muskogean verb morphology</u>. Ph.D. dissertation. University of Kansas.
- Broadwell, George A. 1987a. Subtractive morphology in Southwestern Muskogean. Paper read at the Kentucky Foreign Language Conference, Lexington, KY.
- Broadwell, George A. 1987b. <u>A Mississippi Choctaw -- English</u> <u>dictionary with an English -- Choctaw index</u>. Mississippi Band of Choctaw Indians, Philadelphia, MS. Photocopy.
- Broadwell, George A. 1991. The Muskogean connection of the Guale and Yamasee. <u>International Journal of American Linguistics</u> 57:267-270.
- Broadwell, George A. 1993. Subtractive morphology in Southern Muskogean. <u>International Journal of American</u> <u>Linguistics</u> v. 59, no. 4.
- Byington, Cyrus. 1915. <u>A dictionary of the Choctaw language</u>. Edited by J.R. Swanton and H.S. Halbert. Bureau of American Ethnology Bulletin 46.
- Bynon, Theodora. 1977. <u>Historical linguistics</u>. Cambridge: Cambridge University Press.

- Diamond, Jared. 1992. <u>The third chimpanzee</u>. New York: Harper Collins.
- Gudschinsky, Sarah C. 1956. The ABC's of lexicostatistics (glottochronology). <u>Word</u> 12: 175-210.
- Haas, Mary R. 1941. The classification of the Muskogean languages. In <u>Language, culture and personality: Essays</u> <u>in memory of Edward Sapir</u>, ed. L. Spier, et al. 41-56. Menasha, WI:Banta Publishing.
- Haas, Mary R. 1947. The development of Proto-Muskogean *k^w. <u>International Journal of American Linguistics</u> 13:135-37.
- Haas, Mary R. 1949. The position of Apalachee in the Muskogean family. <u>International Journal of American</u> <u>Linguistics</u> 15:121-27.
- Hymes, Dell. 1960. Lexicostatistics so far. <u>Current</u> <u>Anthropology</u> 1:3-44.
- Kimball, Geoffrey. 1987. A grammatical sketch of Apalachee. International Journal of American Linguistics 53:136-74.
- Kimball, Geoffrey. 1988. An Apalachee vocabulary. <u>International Journal of American Linguistics</u>. 54: 387-98.
- Kimball, Geoffrey. 1989. Another proposal for subgrouping the Muskogean languages. Paper read at the Society for the Study of the Indigenous Languages of the Americas,

July 1989.

Lankford, George E. 1987. <u>Native American legends</u>. Little Rock, Ak: August House.

- Martin, Jack B. 1987. Vowel lengthening in Hitchiti and Mikasuki and its relation to Proto-Muskogean metrical structure, in <u>Muskogean linguistics</u>, ed. by Pamela Munro. UCLA Occasional papers in linguistics, no. 6.
- Muller, Jon D. 1978. The southeast, in <u>Ancient Native</u> <u>Americans</u>, ed. by Jesse D. Jennings. San Francisco: Freeman.
- Munro, Pamela. 1985. Proto-Muskogean LI and <u>li</u> deletion. Paper read at the Conference on American Indian Languages, American Anthropological Association, Washington, DC.

Munro, Pamela, ed. 1987. <u>Muskogean linguistics</u>. UCLA Occasional papers in linguistics, no. 6. Los Angeles.

- Munro, Pamela. 1993. The Muskogean II prefixes and their significance. <u>International Journal of American</u> <u>Linguistics</u> v. 59, no. 4.
- Munro, Pamela, George A. Broadwell, David M. Cline, Abigail C. Cohn, Emanuel J. Drechsel, Heather K. Hardy, Geoffrey D. Kimball, and Jack Martin. 1991. <u>Muskogean</u> <u>cognate sets</u>. unpublished ms.
- Renfrew, Colin. 1988. <u>Archaeology and language: The puzzle</u> <u>of Indo-European origins</u>. Cambridge: Cambridge

University Press.

Riley, Thomas J., Richard Edging, and Jack Rossen. 1990. Cultigens in prehistoric eastern North America: Changing paradigms. <u>Current Anthropology</u> 31:525-541.

Speck, Frank G. 1909. Ethnology of the Yuchi indians. <u>Anthropology Publication of the University of</u> Pennsylvania Museum, no. 1.

- Swadesh, Morris. 1951. Diffusional cumulation and archaic residue as historical explanations. <u>Southwestern</u> <u>Journal of Anthropology</u> 7:1-21.
- Swadesh, Morris. 1952. Lexicostatistic dating of prehistoric ethnic contacts. <u>Proceedings of the American</u> <u>Philosophical Society</u> 96:452-63.
- Swadesh, Morris. 1954. Time depths of American linguistic groupings. <u>American Anthropologist</u> 56:361-377.
- Swadesh, Morris. 1955. Towards greater accuracy in lexicostatistic dating. <u>International Journal of</u> <u>American Linguistics</u> 29:283-88.
- Swanton, John R. 1922. <u>Early history of the Creek Indians</u> <u>and their neighbors</u>. Bureau of American Ethnology Bulletin 73. Washington: U.S. Government Printing Office.
- Swanton, John R. 1946. The indians of the Southeastern United States. Bureau of American Ethnology Bulletin 137. Washington: U.S. Government Printing Office.

Swanton, John R. 1929. Myths and tales of the Southeastern Indians. Bureau of American Ethnology Bulletin 88. Washington: U.S. Government Printing Office.

Ulrich, Charles. 1986. Muskogean legumenonymy. Unpublished manuscript.