The Salinan and Yurumanguí Data in *Language in the Americas*

William J. Poser
Stanford University

1. Introduction

Joseph H. Greenberg's book *Language in the Americas* (Greenberg 1987 — hereafter *LIA*) has aroused much controversy, partly over its appeal to intuited resemblances rather than regular correspondances as evidence of genetic affiliation, and partly because of the questionable accuracy of the data cited together with the lack of detailed documentation (Adelaar 1989, Campbell 1988, Chafe 1987, Goddard 1987).[1] The question of documentation arises because the book lacks the scholarly apparatus expected in such studies. It does not, in general, give sources for the forms cited, justify morphological analyses, or give credit for equations to previous scholars. For further documentation the reader is referred to the notebooks upon which Greenberg relied in writing the book, which thus assume the status of an appendix to the book.

As a linguist not a party to the controversy over the classification of the indigenous languages of the Americas, I set out to satisfy my own curiosity as to the accuracy of the data in LIA by examining the data for Salinan and Yurumanguí, languages with small, closed corpora that can be carefully checked.

I here report the outcome of this investigation. By means of an exhaustive analysis of the Salinan and Yurumanguí data cited in LIA, I attempt to give an idea of the accuracy of the data in LIA, to evaluate the acceptability of the notebooks as documentation for LIA, and to gain insight into Greenberg's use of his sources and attention to the existing literature.

2. The Notebooks

LIA is based on a set of 23 handwritten notebooks, signed photocopies of which, made in 1981, are housed in the Green Library of Stanford University under the title *Regional Linguistic Notebooks: Amerindian*, Library of Congress catalog number

---

[1] I am grateful to Andrew Garrett, Joseph H. Greenberg, and Katherine Turner for their comments on a draft of this paper. I alone am responsible for its contents, with which they should not be assumed to agree.

P203.G7.f. The notebooks were microfilmed by the UC Berkeley library in the same year. The Hokan volume provides for a list of 528 words, the numbers from one to ten and the six pronouns of English, followed by 512 other words in alphabetical order.[2] Some slots are empty due to lack of data, while others contain more than one entry due to a multiplicity of sources or of relevant items. The notebooks do not justify morphological analyses, discuss the interpretation of variants or forms in related languages, or, in short, provide anything other than raw lexical material.

In addition to the publicly available notebooks, Greenberg notes (p. ix) that "... six grammatical notebooks have not been duplicated." as a result of which most of the morphological data cited in LIA is not to be found in the notebooks.[3]

Each page is divided into four columns, one per word, with languages and dialects assigned to the rows. The amount of space for each entry is therefore small and more-or-less fixed (approximately 6 mm high by 35 mm wide), with the result that entries are often crowded, cramped and difficult to read. Entries often spill over into neighboring rows and columns.

For this author to criticize another's handwriting would be for the pot to call the kettle black, but it is nonetheless true that between the crowding, Greenberg's handwriting, and the photocopying process, entries in the notebooks are often very difficult to read. Some, perhaps added in pencil or erased, are illegible or nearly so and give the notebooks the appearance of a palimpsest.

Two additional difficulties arise in using the notebooks as sources of data. The first is the lack of specific citations of sources. Sources used for each language are listed, in abbreviated form, at the beginning of each notebook. Individual entries contain only a brief annotation (e.g. "Pn"), without a page, section, or item number. When a form is cited from the body of the text of a source rather than from an alphabetized wordlist it can be quite difficult to track down. A note at the beginning of the notebook indicates that most of the sources are listed in the bibliography to volume 10 of *Current Trends in Linguistics* and that for central America one may

---

[2] A few words are slightly out of order, and the word *bear* is tacked on at the end of the notebook. My count of the headwords in the Hokan notebook is inconsistent with Greenberg's statement (p. ix) that there are 400 headwords, but the number of headwords seems to vary from notebook to notebook.

[3] An example is the Salinan pronoun *moʔ* "thou" (Mason 1918;22) cited on p.53, which is not listed in the notebook. Greenberg considers pronouns to be "grammatical" rather than "lexical".

also consult volume 5 of the *Handbook of Middle American Indians*, but some are nonetheless difficult to find.

A second difficulty is the fact that Greenberg generally cites forms in the original orthography of his source. While this is not unexpected in a scholar's working notes, it means that the data in the notebooks cannot be used without access to the original sources from which they are drawn. Otherwise one cannot know how to interpret the transcription.

For these reasons, quite aside from any questions of the accuracy of the data they contain, the notebooks fall short as an appendix to LIA or an archive of data. They do not contain any detailed discussion of etymologies or morphological analysis, they do not provide ready access to the original sources, and they cannot reliably be used without reference to the original sources.

3. The Sources

Like all too many extinct languages, the published descriptions of Salinan are less than adequate. No dictionary of Salinan exists, the closest thing to it being the vocabulary list in Mason (1918). Although in most respects inferior to the description in Turner (1987), the grammatical description in Mason (1918) is the only one in print, and many points remain obscure. Moreover, no systematic comparison of the two dialects exists, the literature on this topic consisting of a few passages in Mason (1918), and the passages in Sapir (1925) quoted below.[4]

The published sources on Salinan are:

Sitjar (1808)

Father Buenaventura Sitjar, a priest at the mission of San Antonio, died in 1808. His draft dictionary, the manuscript of which is in the Bancroft Library at the University of California at Berkeley, was edited and published by Shea as Sitjar (1861).

Coulter (1841)

---

[4] Turner (1987) gives a useful review of the available material for Salinan, including comparison of the transcriptions of Mason, Harrington, and Jacobsen, the most important sources.

A list of 61 Antoniaño words collected by Dr. Thomas Coulter appears in Scouler (1841; 247-251).[5]

Gallatin (1848)

Gallatin (1848;126) gives a 22 item list of Migueleño words that he collected himself during a brief visit to the mission.

Pinart (1878)

Alphonse Pinart collected vocabulary at the mission of San Antonio in 1878. This vocabulary, the manuscript of which is in the library of the University of California at Berkeley, was later published by Heizer (1952;73-82).

Kroeber (1904)

Kroeber (1904) contains a small amount of Migueleño data that he himself collected.

Mason (1918)

J. Alden Mason carried out a fairly extensive study of both dialects, the results of which were published in monograph form as Mason (1918). Although neither Mason's transcriptions nor his grammatical analysis are up to modern professional standards, this is by far the most extensive published study of Salinan, containing a grammatical description, texts, and a vocabulary of over 1000 words. Sapir (1917) cited data from Mason's field notes, prior to the publication of Mason's own monograph.

In addition to the published sources, there are a number of manuscript sources for Salinan. Most are merely word lists, including: Cabot & Dumetz (n.d.), de la Cuesta (1821), de la Cuesta (1833), Henshaw (1880, 1884), Merriam (1902, 1933), Taylor (1860), and Yates & Gould (1887).

There are also several sets of field notes: Kroeber (1901), Mason (1910), Harrington (1922, 1932-1933), and Jacobsen (1954-1958). The field notes of Harrington and Jacobsen are of considerable importance for any serious study of Salinan, as they contain large amounts of material collected by modern professional linguists.

---

[5] The list is reproduced by Gallatin (1848;129), with the omission of *kmopax* "strong", and the errors *painel* for Coulter's *paianel* "six", and *traam* for Coulter's *traan* "light", as well as the omission of the macron and breve diacritics.

Greenberg's Salinan material is in notebook 22, entitled *Hokan-Coahuiltecan with Keresan, Iroquoian, Timucuan.* According to the list of sources at the beginning of the notebook, Salinan data is cited from "Mason UCPAAE 14", "Pinart in Heizer Mission Vocabs", and "Jacobsen < Haas Hokan B + C ".

The first is Mason (1918), from which Greenberg indicates that Salinan data is drawn unless otherwise indicated. Mason's vocabulary list gives the Antoniaño form, the Migueleño form, the plural, where appropriate, for the Antoniaño dialect unless otherwise noted, and a gloss. Mason includes in his vocabulary material garnered from Sitjar (1861), all of which is for the Antoniaño dialect. Forms taken from Sitjar rather than from his own fieldwork are marked as such.

The second source, indicated in the notebook entries by the notation "Pn", is the Salinan vocabulary of Alphonse Pinart, cited from Heizer (1952).

The third source, Jacobsen, illustrates the difficulty of using the references in the notebook, for I have been unable to determine what is meant. William H. Jacobsen, Jr. has published a number of papers on topics in Hokan linguistics, but his work on Salinan remains unpublished, except for a few Salinan forms cited in Jacobsen (1958). To my knowledge, Mary Haas has not published any work with a title like "Hokan B and C". The only work with such a title known to me is McLendon (1964), which contains a few Salinan forms attributed to Jacobsen.

The Yurumanguí language once spoken in Colombia, now extinct, is known to us only through a short list of words and phrases recorded by Father Christoval Romero and given by him to Captain Sebastián Lanchas de Estrada, who included it in the account of his travels of 1768. The list, together with commentary, analysis, and comparisons to a variety of Hokan languages, was published by Rivet (1942), which is the source used by Greenberg.

4. Salinan

4.1. The Etymologies

4.1.1. Hokan

The first section of LIA in which Salinan data appears is the section of Chapter III "The Subgroups of Amerind", devoted to Hokan, comprising pp. 131-145.[6]

This section contains 37 entries with Salinan data, namely numbers: 8, 10, 14, 16, 18, 20, 25, 26, 33, 41, 46, 51, 54, 61, 62, 69, 76, 77, 82, 83, 87, 90, 92, 94, 111, 114, 121, 123, 125, 132, 139, 145, 151, 153, 160, 163, 166. Of these, the following 24 call for comment:

8 ASHES

LIA gives M. *ṭapai*. Mason (133) gives *ṭapai*ʰ, as does the notebook.

10 BACK

LIA gives M. *t-ič̆ʔoʔm*, A. *t-ič̆ʔoʔmoʔ*. The basis for the segmentation is unclear. If the initial /t/ were the definite article, we would expect *ṭ*, but both the notebook and Mason (128) show the plain alveolar.

16 BLOOD

LIA lists only A. *a:kat*ʰ, and fails to note M. *pa:kata* (Mason 1918: 128, recorded in the notebook), whose initial /p/ may well be original.

18 BODY

LIA gives *mat*ʰ for both M. and A. However, the notebook gives M. *maṭ*ʰ, as per Mason (127). The correct form appears in LIA entry A185.

20 BOIL₂

LIA gives *(k)-opototna*. The notebook gives *(k)opoṭoṭ*ʰ*naʔ*. Mason (145) gives *(k)opoṭoṭ*ʰ*n∧ʔ*. The inclusion of the prefix suggests that Mason considered the segmentation plausible but lacked direct evidence for it.

26 BURN

LIA mis-identifies *maltintak* as Migueleño, when in fact it is Antoniaño. This we know both from the fact that it is a plural (indicated in the notebook but

---

[6] Entries in this section will be referred to as H1, H2, etc. Similarly, numbers prefixed by *A* refer to the Amerind Dictionary, and those prefixed by G to the grammatical evidence in Chapter V.

not in LIA), and from the fact that Mason (145) attributes it to Sitjar. The A. singular form *maša:ł* given by Mason bears less resemblance to the putative cognates, such as Achomawi *malis* "fire", Northern Pomo *mali* etc. Even for Salinan, where the relationship between singulars and plurals is complex and often irregular, this is an odd-looking plural. Interestingly, the same plural form is given by Mason, again attributed to Sitjar, a few lines above, this time for *male:ntax* "remember, think", for which it is a more reasonable plural, derived by infixation of /t/. This latter is to be found in Sitjar (1861) at p. 40 under the headword *pensar*, but I have not found *maltintak* "burn, blaze" in Sitjar, suggesting that the form is spurious, due perhaps to a copying error by Mason.

46 EAT

The A. form *amma* listed in LIA appears neither in the notebook nor in Mason. It is to be found in Sapir (1917;11), who attributes it to Sitjar.

51 EXTINGUISH

LIA gives *šap* for both dialects, but the notebook gives A. *šap*ʰ, M. *šʌp*, as per Mason (139).

61 FIRE$_2$

LIA lists *ṭ-aʔauh*. Mason (133) gives *ṭaʔaûʰ*. The initial *ṭ* may well be the articular prefix, but the form is not segmented by Mason, and there appears to be no evidence for this analysis.

62 FLEA

LIA gives the segmentation *ṭ-a:jił*. The initial *ṭ* may well be the articular prefix, but the form is not segmented by Mason, and there appears to be no evidence for this analysis.

69 FULL$_2$

LIA identifies *ep*ʰ*ena:teł* as M. As a plural, however, it must be A. Mason (138) glosses it "fill" and lists it as a verb. A further question is why it is the plural form that is cited when the singulars, A. *ʌpeł* and *ep*ʰ*eł*, M. *ʌpel* and *epeł* are available. It is difficult to resist the suspicion that the reason is the greater similarity to the putative cognates, Clear Lake Pomo *minam*, Tequistlatec *imanna*, North Yana *baʔni* and Yurumanguí *pini-ta*, which have /n/ rather than /l/. Given the great variety and irregularity of Salinan plural

formation, there is a good chance that the /n/ of the plural is to be attributed to the plural morphology and not to the stem.

77 HAND$_1$

LIA identifies *itaʔl* as M., *etal* as A. (both meaning "shoulder"), whereas Mason (126) identifies these forms as A. and M. respectively. The notebook shows the correct dialect affiliations.

83 HEAVY

LIA gives A. *k-met'o*, M. *šmot*. The notebook gives A. *(k')meṭ'o*, M. *šmoṭ*, as per Mason (150). The inclusion of the prefix suggests that Mason considered the segmentation plausible but lacked direct evidence for it.

87 HOUSE

LIA makes use of the segmentation *ṭ-a:m* suggested by Mason (130). The inclusion of the prefix suggests that Mason considered the segmentation plausible but lacked direct evidence for it.

90 LARGE$_2$

LIA gives M. $k^{ʔw}aka$ "long, high, tall". The notebook records a long vowel $k^{ʔw}a:ka$ as per Mason (150), who omits the gloss "high". The gloss "long, high, tall" appears in Sapir (1917;10).

92 LAUGH

LIA gives A. *ilik'*. Mason (141) shows no glottalization. It appears that Greenberg misinterpreted the lower part of the semi-colon in his notebook after the entry above this one as the apostrophe that Mason used to indicate glottalization.

94 LEFT(SIDE)

LIA gives A. *oʔkel*. The notebook shows *oʔkelo*, as per Mason (150).

114 RABBIT

LIA gives A. *kol* "hare". The notebook shows *kol'* as per Mason (123).

121 ROPE

LIA lists A. *asol* under "rope" as it is listed in the notebook. Mason's gloss (130) is "cord, line, string".

123 SALT

LIA gives A. *ṭ-akai*. The notebook shows *ṭa:kaiʔ*, as does Mason (133). The segmentation is plausible but there appears to be no direct evidence for it.

132 SLEEP

LIA gives M. *p-apa* "copulate". The notebook shows *papa:*, while Mason (138) gives *papa:ⁱ*. The segregation of the initial /p/ is legitimate, if it is the active prefix, as plausibly suggested by Sapir (1925;416), but there appears to be no direct evidence for this segmentation.

145 TESTICLE

LIA gives A. *solo*, as does the notebook. Mason (126) shows *so:lo*.

160 WHITE

LIA gives A. *maṭaɫ*, as does the notebook. Mason (150) gives *maṭʰaɫ*.

166 WOMAN₃

LIA cites *hemuč*, which does not appear in Mason. *hemuč* is Chumash, and appears in the notebook in the row for Santa Cruzeño Chumash. The form is taken from the Chumash word lists in Powell (1877;561).

4.1.2. The Amerind Dictionary

The "Amerind Dictionary" in LIA contains 29 entries with Salinan data, namely numbers: 2, 7, 27, 43, 47, 53, 66, 79, 87, 102, 107, 128, 131, 137, 148, 174, 175, 181, 185, 198, 199, 217, 228, 238, 242, 244, 246, 248, 255. Of these, the following 18 call for comment:

2 ABOVE

LIA lists M. *o:pʰak*. According to Mason (126) this means "head", not "above". The notebook lists this form correctly under "head".

27 BEE₂

LIA cites *le-me'm* "bee, wasp" for both dialects. According to Mason (123) the forms are A. *lmém'*, M. *lemém'*, and both are plural. The notebook correctly

notes that both are plural but incorrectly inserts the extra *e* into the A. form. The segregation of the initial *le* is without known justification.

43 BODY₁

LIA lists M. *upi(-nit)* as meaning "fat". Mason (128) lists M. *upent*, A. *uprent*, which are the forms to be found in the notebook. The form *upi-nit*, which is not to be found in the notebook, is found in Sapir (1917;13), with the indicated segmentation, though without justification for it. Greenberg & Swadesh (1953;218) mis-cite the form as *upikit*. Sapir (1925;416) gives *upent*. Sapir very likely obtained the form *upinit* from the notes underlying Mason's citation in a set of morphological examples (14,19) of the form *ṭopinit-o* "fat", attributed to Sitjar (on p.19), where the initial *ṭ* is presumably the article. The associated abstract noun *ṭ'upint-éy'a* "corpulence" which shows the *u* that appears in the form given by Sapir, does not show the second *i*. Sitjar (p.26) lists several forms whose stems are plausibly regarded as *upinit* and *upint* under the headwords "gordo" (fat) and "gordura" (corpulence), e.g. *cúpinit* (Sitjar's orthography, not phonetic) "estoy gordo" (I am fat), though not the form *ṭopinit-o* cited by Mason. Since the form is from Sitjar, it must be Antoniaño, not Migueleño.[7]

47 BONE₂

LIA lists A. *axa:k*, M. *axak*. Mason (127) gives A. *axa:k*, M. *paxak*, which are the forms to be found in the notebook.

53 BROTHER

LIA gives M. *pepeʔ* as "brother". According to Mason (134), it means "elder sister". The notebook lists the form under "sister" with the annotation "O.S.", presumably for "older sister".

79 DIE₁

LIA lists A. *ema-ṭʰ* "kill", explaining the -*ṭʰ*as a causative suffix. Mason (140) gives *(e)má:ṭʰ* with a long vowel. The notebook correctly reflects Mason. Whether this is a viable synchronic analysis is unclear, as the base form is presumably taken to be *me* "sleep". Whereas "cause to sleep" is a plausible euphemism for "kill", one wonders how likely it is to become the general word for "kill".

---

[7] This dubious form (without the unjustified hyphen) is also cited, presumably from Sapir, by Gursky (1974; 187), along with the two forms given by Mason. Like Greenberg, Gursky mistakenly identifies it as Migueleño.

Moreover, the analysis of the *ṭʰ* as a causative suffix is questionable. Mason lists a set of related causative suffixes *-at, -t, -te*, but none has a post-alveolar *ṭ*.

87 WATER

LIA lists A. *ṭšaʔ*. The forms in Mason (132) are *(t)šaʔ* and *tš̃ʰaʔ*. With regard to this form, Sapir (1917;8) noted parenthetically: "analysis into *t-šaʔ*, according to Dr. Mason, is probable, but bare stem *-šaʔ* is not found." A form with a post-alveolar, *ṭʰš̃ʌʔ*, is given by Mason for Migueleño.

137 HAND

LIA lists M. *maa* and *maʔa* "bring, carry". Neither appears in Mason (144), who gives the M. form as *maʔa:ʔuʔ*. A. has *ma:ʔa*. The forms in the notebook are correct. The forms in LIA appear to come from Sapir (1917;11), who gives *ma:a* and *maʔa*. It is surprising that Greenberg does not cite *me:n* "hand" (Mason 127), which does appear in the notebook, and resembles the putative cognates about as much as the forms cited.

148 HEAVEN$_2$

LIA cites A. *l-ema* "heaven". Mason (133) gives A. *léma*, M. *lem* as meaning "sky". The notebook lists the forms correctly but under the headword "heaven". There is no known basis for the segregation of the initial *l*. Indeed, Sapir (1917;4, fn.2a), was forced to note that: "According to Dr. Mason, however, there is 'no evidence whatsoever that *lem* "above, sky" can be analyzed into *l-em*.'"[8] As the comparanda lack *l*, this segmentation is important to the cognate set.

175 MAKE$_1$

LIA gives M. *ti:*. The notebook gives *ti:ʔ* as per Mason (146).

181 MANY$_2$

LIA cites only A. *k-i:sileʔ* "all", ignoring M. *k-i:silep'* (Mason 151, recorded in the notebook). As the final [p'] is likely original, this is the more appropriate comparandum. The putative cognates resemble the Salinan forms so little that this may not make much difference, but none of them have anything like a [p'] in them.

---

[8] Sapir quotes from a letter dated 26 July 1916, to be found in the Bancroft Library of the University of California at Berkeley.

185 MEAT$_3$

Salinan *maṭʰ* "meat" appears in a set of Hokan forms glossed "meat". However, Greenberg also lists *maṭʰ* in H18, whose gloss is "body". Moreover, the two cognate sets are inconsistent. Comecrudo *ewe* is given as a cognate here, while in H18 the putative Comecrudo cognate is *met*.

217 SAY

LIA gives M. *ṭeʔ* for "tell". Mason (147) shows aspiration, i.e. *ṭʰeʔ*, as does the notebook. Both Mason (147) and the notebook show no aspiration for Antoniaño.

228 SHOULDER

This etymology equates Antoniaño *tatal* and *itaʔl* "shoulder" with Achomawi *tala* "shoulder blade" and Northern Yana *dul* "neck", as representatives of Hokan "shoulder". This is inconsistent with the inclusion of *itaʔl* in H77 HAND$_1$. There is a good chance that *tatal* and *itaʔl* are one and the same form, with *tatal* an inaccurate transcription of *ṭ-itaʔl*, the initial *ṭ* being the articular prefix. *tatal* comes from Pinart, who might easily have missed the point of articulation of the *ṭ*, the glottalization, etc.

242 SOUR

LIA gives A. *t-erk*, M. *t-iek* for "animal's gall". The basis for segmenting out the /t/ is not given. If it were the article, we should expect *ṭ*. No mention is made of the different comparison proposed by Gursky (1974). Gursky's comparanda are Washo *ʔileg* and *ʔilek'* "liver", with headword "liver". Greenberg's headword (for the Hokan subhead) is "bitter", with the Washo comparandum *tˢ-iga-l* "kidney".

244 STAR$_2$

LIA cites A. *maša-lak* "morning star", without any internal evidence for this segmentation, which is justified by the statement (p. 259) "cf. Jicaque *lak-sak* 'sun', a compound in which *lak* = 'glow' and *sak* = 'sun'." Greenberg & Swadesh (1953;219-220) give a variant of this analysis, to wit: "The Jicaque is evidently a compound of two elements, perhaps *ala-k* 'glowing' and *sax* 'sun'." The fate of the initial *a* of *ala-k*, the segregation of the *k*, and the replacement of the *x* of *sax* by *k* are not explained. There is, moreover, a real question both as to the accuracy of the form *laksak* and as to whether this form may be

reconstructed for Proto-Jicaque. Jicaque has two dialects, the Western dialect of El Palmar, and the Eastern dialect spoken mainly in La Montaña de la Flor, but also in the Department of Yoro. The Western dialect is known only from a single word list recorded by Antonio S. Mordiaga in 1890, published by Alberto Membreño (1897), and reprinted by Lehmann (1920;656), one of the two sources cited by Greenberg & Swadesh (1953).[9] The Western dialect form is //luchoc//, probably reflecting [lučok]. The form *laksak* reflects two sources for the Eastern dialect, namely Conzemius (1920;167), the second source cited by Greenberg & Swadesh, who gives *laksak* and *loksak*, and Membreño (1897), as cited by Lehmann (1920), who gives //locsac//, presumably [loksak]. The other sources for the Eastern dialect give forms containing [ts] or [č] rather than [ks]. These include Lean and Mulia, the two archaic sources cited by Lehmann (1920;656), who give //lochac//, presumably [ločak], and all of the modern sources, e.g. Dennis & Royce (1983;14), who give *lots'ak*[h]. Campbell & Oltrogge (1980;219) reconstruct *lots'ak*. The [ks] reported by some sources is likely an erroneous transcription, due to lack of familiarity with ejectives.[10] Moreover, none of the sources contains any word of the form *sak* or *sax* that might be what Greenberg considers to be the second member of the form, or anything like *ala-k* "glowing" or *lak* "glow". Nor do these forms appear in the notebook. The notebook contains no headword "glow", and no entry for Jicaque under the headword "shine", nor has inspection of all of the Jicaque entries yielded any plausible sources for these forms.[11] The only entries under "sun" are *loksak*, *latsak*, and *loksaki*. It thus appears that there is no evidence whatsoever internal to Jicaque for analyzing *laksak* into *lak-sak*. [12] In other words, the comparison is between two words, each containing the sequence *lak* (if the Jicaque contains a [k] at all), in neither of which there is the slightest evidence for an analysis in which *lak* is a morpheme. In contrast, Greenberg fails to mention the etymology proposed by Mason (1918;19), namely that *mašalak* is derived from *mašal* "burn, blaze"

---

[9] Curiously, the notebook does not cite Lehman (1920) as a source for Jicaque, though it does cite it for other languages.

[10] Lyle Campbell, personal communication, 1991.

[11] The only forms that have approximately the right shape to be components of *laksak* are *sek* "chin" and *lats, laks* "snake".

[12] The problems mentioned here also invalidate etymology H142 SUN$_2$, in which Yurumanguí *siko-na* is compared to "Jicaque *lak-sak* 'sun, day' (lak = 'glowing')". Curiously, Greenberg omits the suggestion of Rivet (1942;29) that *siko-na* is cognate to Subtiaba *sigu* "year", and Antoniaño *tó:kena*, Migueleño *tó:kan* "day".

(cf. *mašale* "flames"). The weakness in Mason's etymology is that no suffix
-*ak* is known, but, there being evidence for one of the two morphemes, it is
nonetheless more plausible than Greenberg's.[13]

246 STONE₂

LIA cites M. *išak* and *išik* "knife" and *asak'a* "flint". The last is attested, but
is indicated by Mason (132) to be Antoniaño. Neither of the forms for "knife"
appears in the notebook, which gives M. *ča:k*, A. *čik*ʰ, *šik*ʰ, and *ṭišak*. Nor do
they appear in Mason (130), who gives M. *čak*, A. *(t)šik*. The form *išak* is cited
by Greenberg & Swadesh (1953;219). The forms in LIA as well as the form in
Greenberg & Swadesh (1953) appear to derive from Sapir (1917;8), who gives
the forms *(i)ša:k* and *(i)šik*. In the footnote to this entry, Sapir cites the form
as *šak*, suggesting that the forms with the initial *i* may be typographical errors,
with *i* in place of the articular prefix *ṭ*, the parentheses intended to segregate it
from the stem.

255 THROAT

LIA lists A. *p-e:nik'a*. The notebook has *pe:nik'a:i*, which is the form given by
Mason (127). The basis for segregating the /p/ is unknown.

4.2. Grammatical Evidence

Chapter V "Grammatical Evidence for Amerind" cites Salinan data in eleven
places, sections 6, 13, 16, 19, 45, 74, 80, 84, 88, 90, 100.[14]

Of these, the following five call for comment.

In section 19 (p.288) LIA gives the Migueleño first person plural pronoun as *ka*.
The correct form, according to Mason (32) is *k*ʰ*aʔ*.

Section 80 (p. 311) reads in its entirety:

> The following etymology represents a past-tense marker in HOKAN:
> Yurumanguí *iba*, Coahuilteco *pa-*, Tequistlatec *-pa*, Salinan *be*,

---

[13] There is an agentive suffix -*mak* (Mason 1918;19), to which -*ak* might be related. Appeal to
a general phonological rule deleting /m/ after /l/ does not seem promising, as medial [lm]
clusters are permitted, but the small number of examples that I have seen do not rule out an
allomorphic relationship between -*ak* and -*mak*.

[14] G6 and G13 do not directly mention Salinan data, but G6 refers to the Salinan second person
pronoun discussed in Chapter 2, and G13 discusses the t-absolutive discussed in Chapter 2.

Pomo *(hi)ba*, Karok *-ʔipa ˜ pa-* (near past), and Shasta *p'-* (distant past, habitual past). The marker probably also occurs in Salinan *iwa-š*, which when suffixed to nouns means 'that which was formerly', e.g. *noqš* 'head', *noqš-iwa-š* 'skull'. The last element, *-š*, is a common noun formant in Chumash. The agreement in a form *ipa* among Yurumanguí in the extreme south, Coahuilteco in the middle, and Karok in the far north is striking.

This passage presents a number of difficulties. First, Salinan *be* is not a simple past tense morpheme as Greenberg glosses it. Mason (1918;35) glosses *be* as "when, definite past time", as in such examples as *be:-ya* "when I went". Sapir (1920;307) suggests that *be:ya* is really "an indicative *e:ya* 'I went' subordinated by the demonstrative stem *pe, pa* "the, that". Sapir's view is supported by Mason's statement that "Pure sonant *b* has been found only in the case of the demonstrative article *pe...*" (p.11). If Sapir is right, *be* is not a tense morpheme at all.

Second, the agreement among the forms is perhaps less striking than Greenberg suggests. Of the eight forms cited, only four (Karok, Pomo, Yurumanguí, Salinan *iwaš*) have the initial *i*, and as we shall see, the initial *i* of *iwaš* is probably not original. Moreover, some of these morphemes are prefixes while others are suffixes. Finally, as discussed below in section 5.3, it is unclear whether the Yurumanguí past tense morpheme is the infix *-iba-* or the suffix *-bai*.

Third, the discussion of *iwaš* is seriously flawed. To begin with, this suffix and the examples cited are not Salinan. The suffix *iwaš* is not mentioned in any source on Salinan that I have consulted, nor are the examples cited as illustrating the use of this suffix (*noqš* "head", *noqš-iwa-š* "skull"). Nor do they appear under the headings "head" and "skull" for Salinan in Greenberg's Hokan notebook. Indeed, Salinan has no [q] at all.

Rather, the suffix *iwaš* and the examples of its use are Chumash. *noqš* and *noqšiwaš* are among the examples of the use of the suffix *iwaš* in Barbareño Chumash given by Beeler (1976;259), one of the sources listed in the notebook. The notebook lists *noks* under "head" for Ineseño and Barbareño Chumash.

Nonetheless, the issue of the cognation of Chumash *iwaš* with the other forms cited still arises. Two facts militate against Greenberg's analysis. First, the initial *i* is very likely not original. There are two related suffixes in Chumash, the suffix *iwaš* which derives nouns meaning "dead, defunct, former", and the verbal past tense

waš.[15] The nominal suffix is invariant, as is the verbal suffix in Barbareño (Beeler 1976). In Ineseño, however, according to Applegate (1972;102-3), an epenthetic copy of the last vowel of the stem is inserted before the past tense marker when the stem ends in a sonorant. Although the history is not known, it seems very likely that the nominal suffix is historically derived from the verbal past tense, and that the initial *i* represents a frozen epenthetic vowel.

Second, there is little basis for segmenting out the final *š*. If *iwaš* is derived from the verbal suffix *waš*, it is unlikely that the final *š* is a noun-forming suffix. Noun-forming suffix *š* is not described in any source on Chumash that I have consulted, including the two most detailed grammatical descriptions Applegate (1972) for Ineseño and Beeler (1976) for Barbareño. There is a suffix -*Vš* described by both Beeler (1976;258) and Applegate (1972;213) as "resultative", and it is to this suffix that Greenberg refers.[16] The *V* means that the suffix takes the form of a vowel followed by *š*. According to Beeler (1976) this vowel is unpredictable, but Applegate (1972;93) gives rules for predicting the vowel. In any case, the derivation envisioned by Greenberg is not clear. If we start from a verbal past tense suffix *$^*ipa$, which in Chumash comes to be added to nouns as well as to verbs, it would seem unnecessary to add a resultative suffix, and one has to wonder why the basic past tense suffix would acquire the resultative suffix attached to it in nouns.

In section 84 (p. 311), LIA cites a suffix -*še* "desiderative" for Antoniaño. Mason discusses such a suffix, but with final glottal stop -*šeʔ*, on p.49, and although he glosses it "desiderative", he makes it clear that he is far from confident of this interpretation, a hesitation fully justified by the varied meanings of the examples adduced.

In section 88 (p. 313), LIA refers to a Salinan imperative morpheme -*i*-. This suffix is not described by Mason (1918) in his discussion of the verbal morphology (pp.34-54), nor in Turner (1987), so that it appears to be quite spurious, and given the lack of documentation, one despairs of tracking it down. The key turns out to be Rivet (1942;33), which presents the same equation as in LIA, minus the Karok form, which was evidently added by Greenberg. Rivet cites Sapir (1921;71), in which we find, as entry number 28, the following:

---

[15] In addition to the more detailed sources quoted below, the past tense *waš* is mentioned in such places as Kroeber (1904).

[16] Personal communication, July 1991.

Sal. -*i*-, imperative suffix with third person pronominal object
(e.g. *m-alel-i-k* ASK HIM!): Yana -*'i'*, imperative suffix.

This constitutes the entire discussion of imperative -*i*- in the literature. At the very least Greenberg is to be taken to task for using as evidence a morpheme for which the evidence is so skimpy, especially when it does not appear in Mason (1918), the only published grammatical description of the language, and the source of Sapir's data.

Whether Sapir's analysis should be accepted is unclear. Salinan has a number of third person singular objective suffixes, which according to Mason (1918;46-47) take the forms: -*o, -ko, -xo* and -*k*. The different suffixes are associated with different classes of verb. "... the *p*- prefix nearly invariably takes the suffix -*o* or -*ko* as its third personal objective form while the objective form in -*k* occurs exclusively with the *k*- prefix." (Mason 1918;39)[17] That is, the forms in -*k* are associated with what Mason considered to be the intransitive verbs, later argued by Sapir (1920;307-308, 1921;69-70) to be stative. In his discussion of the imperative, Mason (1918;41) states that: "The imperative takes its third person pronominal object in -*ik*, never in -*o* or -*ko*."

Sapir's reasoning appears to have been that the *i* of this suffix could be isolated as an imperative since the suffix appears only in imperatives and since other third person singular objective suffixes contain *k*. Against this we may consider the fact that imperative *i* occurs nowhere else and that it is not possible to analyze the third person singular objective morpheme simply as *k*; even if we could extract *k* from *ko*, we would be left with the forms in *o*. Moreover, judging from Mason's examples of the use of -*ik*, namely: *k-ámamp-ik* "take it out!", *ámes-ik* "shout to him!", and *m-alel-ik* "ask him!", it appears that -*ik* occurs even with active verbs, which according to Mason's generalization do not take the objective prefixes containing *k*. On balance, it seems that -*ik* must be treated as a unit, and that Sapir's analysis is overly aggressive.

In section 90 (p. 313) LIA cites *k*- as the imperative prefix. Actually, according to Mason (1918;41) this prefix appears only in the plural, and not in all cases.[18]

---

[17] A few pages later (p.47) Mason notes: "With a few inexplicable exceptions, forms involving the element *o* are found only in connection with the verbal prefix *p*-, those ending in -*k* only in connection with the verbal prefix *k*-."

[18] According to Turner (1987;161) it is not an imperative but is simply the second person plural subject prefix.

5. Yurumanguí

Yurumanguí forms occur in a total of 26 entries in LIA, to wit: H15, H32, H35, H44, H61, H69, H72, H87, H142, H158, H161, A9, A79, A86, A89, A102, A104, A191, A243, A269, G19, G61, G80, G88, G90, G102.

5.1 Segmentation

Morphological analysis is often difficult even in languages for which we have unlimited data; in a case like Yurumanguí where we have only a tiny corpus containing few related forms, we must proceed with the utmost caution and must expect to remain unsure on many points.

Rivet made a valiant effort at analysis on the basis of the data available to him, using both what little language-internal evidence there was, and what was suggested to him by similar forms in Hokan languages. Many of these proposals are interesting and would be worthy of pursuit if we were convinced of the affiliation of Yurumanguí with Hokan, but until such an affiliation be proven they remain the merest speculation.

The result of the dearth of Yurumanguí data combined with the fecundity of Rivet's etymological imagination is that most of the morphological analyses he proposed rest on hypothetical affiliations with Hokan, and so remain undemonstrated. Rather than exercising caution and utilizing only justifiable analyses, Greenberg simply accepts Rivet's analyses and presents them as if they were clearly justified internal to Yurumanguí, as we will now see in detail.

Rivet (pp. 28-29) posits a prefix *a-*, appearing on both verbs and nouns. Other than the fact that a considerable number of words begin with *a*, he cites no language internal evidence for the existence of such a prefix. In many cases he gives no evidence of morphological complexity of any kind. Where he does give evidence, it consists of comparisons with other languages. For example, he compares *aikan* "wings" with Cochimi *ixquan*. Rivet admits (p.28) that he is unable to assign any meaning to this prefix, either in Yurumanguí or in Hokan. In other words, there is no evidence whatever that Yurumanguí has a prefix *a-*.

This putative prefix appears in four forms. In *a-ikan* "wing" in entry A269, and in *a-umi-ssa* "sit down", and *a-uma-sa* "chair" in entry H87, Greenberg transmits Rivet's unjustified segmentation. In A79 Greenberg cites *ima-sa* "kill". The attested form is actually *aimasa*. If the cited form is not simply a mistake, it would appear

that Greenberg has stripped it of the initial *a* that Rivet considered to be a prefix. There is no evidence in any of these cases that the *a* is not part of the stem.

Rivet (pp. 31-32) posits a suffix -*a*, appearing on verbs, nouns, and adjectives, on the grounds that many words end in *a* and that he finds correspondences in other languages. There is no evidence that *a* is a distinct morpheme in any of the words cited. Here again he is unable to assign the suffix any meaning in Yurumanguí. This putative suffix appears in one form in LIA, namely *sia-a* "heaven" in entry A243. There is no evidence that this form is morphologically complex.

Rivet (p. 29) posits a suffix -*na* appearing on nouns, his only evidence being comparison with other languages. There is no evidence that any of the nouns cited is morphologically complex. As for meaning, he says: "As in the Hokan languages, the meaning of this suffix remains vague."

This putative suffix is cited in entry G61, and is segmented out of three other forms in LIA. In two, *sipa-na* "hat" in H32 and *siko-na* "sun" in H142, Greenberg merely follows Rivet's segmentation. In the third, *(ko)-u-(na)* "eye" in A104, the segregation of *na* is due to Greenberg, not Rivet. In none of the three is there any evidence of morphological complexity.

Rivet (p. 37) posits a negative suffix -*ta* on the basis of three forms. In two of the three cases, *pini-ta* "empty thing" and *na-ibi-ta* "we are not [angry]", neither is the form clearly negative nor is there evidence of morphological complexity. That "empty thing" actually means "not full" is an assumption. The claim that *na-ibi-ta* contains a negative morpheme rests on Rivet's reinterpretation of an utterance glossed in the original phrase list (p. 10) as "Yes, we are friends." This reinterpretation is conceivable, but there is no evidence whatever in support of it. The third case is *aupita* "I do not hear.", which is compared with *aupeitaia* "I have heard." That it is *ta* that is the negative morpheme is hardly clear from this comparison. Indeed, if we consider the alternative form glossed "I have heard" in the phrase list *austage* (p.14), we might propose that it is *ta*, which is common to all three forms, that is the stem of "hear".

This putative suffix appears in H69 where it is segmented out of *pini-ta*, which Greenberg glosses incorrectly as "empty" rather than "empty thing".

Rivet (pp. 30-31) posits a suffix with variants -*s, -sa, -isa, -za, -iza* that appears on nouns and in the infinitives of verbs. The suggestion that such a suffix forms infinitives is plausible given the large percentage of forms glossed as infinitives that end in this way together with the fact that there is one alternation, that between

saisa "to die" and *saibai* "he died", in support of it. There is, however, no evidence that any of the other forms claimed to contain this suffix is morphologically complex, and hence, no evidence that it attaches to nouns.

This putative suffix appears in two forms: *mai-sa* "night" in entry H15 and *ba(-isa)* "day" in entry H35. There is no evidence that either form is morphologically complex.

joima "saliva" is analyzed into *jo* "mouth" and *ima* "water", following Rivet's proposal (p.40), and the two components cited under A191 MOUTH and A89 DRINK respectively. There is no evidence whatsoever that *joima* is morphologically complex. Neither of the putative component morphemes is independently attested. The closest to independent evidence that we have is Rivet's suggestion (p.40) that *ima* is related to the *uma* of the verb *čuma* "to drink", which he segments as *č-uma*. Here again, there is no language-internal evidence that *č* is a prefix. The two examples given by Rivet (p.28) are *čuma* "to drink" and *čuma-é* "drink!", which yield no evidence that *č* is a prefix. The remainder of his evidence for this analysis is the existence of allegedly cognate prefixes in other languages and alleged cognates of *uma*, including most of the Hokan examples cited by Greenberg in A89.

In G90 Greenberg cites an imperative prefix *k-*. Rivet (1942;34) posits such a prefix, but his evidence is inadequate. Two forms are cited as containing this prefix: *k-aska-ti* "place it", and *ku-koko-na* "may one cook it quickly". There is not a shred of language-internal evidence for Rivet's segmentation of these forms; no other form of either verb is attested, and most of the attested imperatives do not begin with *k* but rather contain the imperative suffix *-i* cited by Greenberg in G88. Rivet's only evidence for the existence of an imperative prefix *k* other than the fact that two imperatives happen to begin with *k*, is the existence of allegedly cognate imperative prefixes in other languages. In sum, there is no language-internal evidence that Yurumanguí possessed an imperative prefix *k-*.

5.2. Spurious Forms

Two forms appear to be spurious. The first is *ita(-asa)* "wife" in H161. No such form is to be found; the only item with this gloss is *ki-tina*. It is possible that this is an error for *ataisa* "sister".

The other spurious form is *-fa* "excrement", in A102, where Greenberg analyzes *anga-fa* "ashes" as "fire-excrement". *anga* is independently attested, with the meaning "firewood", as are two other semantically related words beginning with *anga*:

anga-iaka "burning tobacco" and *anga-isa* "light". It is thus quite plausible that this form is a compound, but the identity of *-fa* is unknown. There is no other attestation of *fa*, either as an independent word or in another compound. Nor does the assumption that "ashes" is of the form "firewood + X" force the conclusion that X = "excrement". Other possibilities that come to mind include: "remainder", "residue", "dead", "charred", "blackened", "powder", "result", and "cool". That it might mean "excrement" is suggested only by Rivet's comparisons (p.40) with Pomo, Salinan, and Subtiaba. In short, there is no language-internal evidence that *fa* means "excrement".

5.3. Miscellaneous

In H61 FIRE$_2$ Greenberg cites *anga(-fa)* "ashes" with the implication that *anga* means "fire". As discussed above, *anga* appears in two other apparent compounds as well as in isolation, where it means not "fire" but "firewood". The word for "fire" is *angua*, which fits the other members of Greenberg's cognate set about as well as *anga*.

In G80 Greenberg cites the past tense suffix *-iba*. The evidence for such a suffix consists of the single pair of words: *saisa* "to die", *saibai* "he died" (Rivet 1942;32). Rivet analyses these as *sa-isa* and *sa-iba-i*, with *iba* an infix into the stem *sai*. Given this analysis, it is unclear why the infinitive should not be *sai-isa*. Since *isa* is but one of a whole group of variants of what appear to be the same suffix, including *sa* (p. 30), it seems less problematic to analyze these forms as *sai-sa* and *sai-bai*, with the past tense morpheme thus taking the form *bai*. The form of the suffix is thus unclear.

6. Discussion

6.1 Segmentation

Determining the morphological analysis of comparanda is important in order to know what may legitimately be compared, since only comparisons of whole morphemes are meaningful. Where languages are already known to be related one language may of course shed light on the etymology of words in another, but where languages are not known to be related comparisons in which the morphological analysis itself depends on the relationship carry considerably less weight than those in

which the segmentation is clearly established, for the simple reason that the additional degrees of freedom increase the probability with which similarities may be due to chance. Greenberg acknowledges this when he says (p.xvi):

> A further difficulty attends the citation of forms that were accompanied by prefixes or suffixes in the original source. I have sought to be scrupulous in excluding as part of the stem, in etymological entries, those elements for which there was good internal evidence concerning their affixal status. But I have not been entirely consistent in including these in citations. The following general convention has been adopted for morphologically complex forms. Individual morphemes are separated by hyphens ...

This passage indicates that hyphens are to be interpreted as indicating morpheme boundaries clearly justified internal to the language cited, as is the usual practice. However, analysis of the morphological analyses in LIA reveals a number of cases of questionable or unjustifiable segmentation.

We have already seen, in section 5.1, that most of the morphological analyses in the Yurumanguí data lack language internal justification, for a total of eleven unjustifiable segmentations (H15, H35, H69, H87, H142, A79, A89, A104, A191, A243, A269) and two unjustifed affixes (G61 and G90).

Cases of possible but undemonstrable segmentation arise with some frequency in Salinan because of the relatively poor quality of the available material. Mason was often unsure of the analysis of forms, and listed as stems forms that subsequent scholars have suggested to be complex, or listed forms with a putative prefix in parentheses, presumably because it is a plausible analysis but one for which he lacked direct evidence. When an analysis is plausible, data showing the alternations that would clearly demonstrate the analysis are frequently unavailable.[19]

One example of a plausible segmentation that is not directly justified is H132, where it is likely that the p- is the active prefix, but no other form directly establishes this. Similar are H20 and H83, where Mason seems to have believed the initial k to be the stative prefix but to have lacked evidence for this belief. In A255, the initial p may be the poorly documented articular p, but there is no direct evidence of this.

---

[19] One might hope to find examples justifying segmentations in the texts in Mason (1918). However, my reading of these texts produced no examples that justify the segmentations that I have marked as questionable.

In H61, H62, H87 and H123 the initial *ṭ* may well be the articular prefix, but again there is no evidence that these forms are morphologically complex.

In two cases, H10 and A242, the initial *t* may be the articular prefix,[20] as there are some clear cases recorded with *t*, but the fact that the forms begin with an alveolar whereas the articular prefix is regularly post-alveolar, casts some doubt on this analysis.

Finally, A79 is one of the rare cases in the Salinan data in which Greenberg attempts to justify his analysis. As indicated above, this analysis is questionable.

In other cases, A27, A43, A148, and A244, there is no language-internal basis whatever for the segmentation in LIA.

6.2 Lower-Level Reconstruction

The status of *p* and *p'* that appear in Migueleño but not in Antoniaño presents a number of difficulties. The existence of such pairs was noted by Mason (1918;16), who cited the following as a subset of many examples:

Antoniaño	Gloss	Migueleño
aš	elk	páš
ʔakata	blood	pakata
axák	bone	paxák
atʰ	oak	p'atʰ
as	son	p'as
ṭšxaʔ	stone	šxap
ka'	acorn	kʰap'
titsʰéʔwu	tail (his)	ṭ-itsʰéʔp
tʰɪšeléʔ	fingernail	išilip
ṭišxéʔwu	foot (his)	ṭišxép
ṭaáʔ	deer	ṭaáʔp
sánat'	hide	spanat
leáṭ'	duck	elpáṭ'
ṭʰáʔak	head	ṭópaka
ṭáaiʔ	ashes	ṭop'ai

---

[20] For A242 this is suggested by Gursky (1974;196 — my translation), who says: "t- might be the nominal prefix."

Mason also pointed out (p. 15) that:

> ... the final *p* of a stem in the Migueleño dialect, which is normally
> lost or replaced by a glottal stop in the Antoniaño form, frequently
> reappears in the latter dialect when the stem is expanded.

citing the examples:

ṭišxéxeʔ	feet	*ṭišxe:pléto*	their feet
tšxaʔ	stone	*tšxápaneł*	stones

If Mason is right that Migueleño preserves instances of *p* that have been lost in Antoniaño, the *p* of the Migueleño forms must be taken into account in comparisons outside of Salinan.

This issue arises in three entries. In H16, Greenberg cites only A. *a:kat*ʰ, and fails to note M. *pa:kata*.[21] Similarly, in A181, he cites A. *k-i:sileʔ*, ignoring M. *k-i:silep'*. Finally, in A47 he cites M. *axak*, when in fact the correct form is *paxak*. Unless this is simply an error, it reflects an unexplained analysis into *p-axak*.[22]

Sapir (1925; 500-501 — transcription modernized) regarded the initial *p* of the Migueleño forms as a prefix:

> A nominal *p*-prefix can be pretty clearly made out for Salinan
> (e.g. M. *p-a:kata*, A. *a:kat*ʰ "blood": Hokan *ₐxwati*, no.3; M. *p-akenai*, A. *akainai* "animal's womb"; M. *p-aktainaʔ*, A. *akatšanai*
> "thumb"; M. *p-axak*, A. *axa:k*ʰ "bone": Hokan *ihyaka*, no.4; M.
> *p-*ʰ*a:t*ʰ, A. *at*ʰ < *hat*ʰ [?] "white oak"; M. *p-axakił*, A. *askle-t*
> "live oak"; M. *p-axuwe* "bow": Chum. *t-axa, ax*, no. 49); less
> safely in Chumash (e.g. *p-ako-waš, p-akü-was* "old man": Chon.
> *akwe*, Moh. *kwo-ra-*, Cochimi *aku-so*, Ton. *ku-ša* "old"; *p-awa-yiš*
> "house": Hokan *awa*, no .50); and in Pomo (e.g. N. Po. *b-isi-l*
> "rabbit-skin robe": *iši, itši, ši, ši-ts*; N. Po. *b-atsiya* "yellowham-
> mer": *k-atsiya, k-otšiyo, k-otšiya*).

---

[21] In contrast, Gursky (1974;180 — my translation) cites the M. form, with the note: "only if p- in M. *pá:kata* is a prefix, as Sapir 1925; 500 Fn. 14 suggested."

[22] Gursky (1974;181 — my translation) felt compelled to note: "only if M. p- is a prefix, as Sapir 1925:500 Fn.14 proposed".

In a endnote to this sentence he went on to point out that there are difficulties with the proposal that M. *p* is uniformly lost or reduced to glottal stop in A.:

> It is not altogether clear whether in cases like this the Antoniaño dialect has actually lost a *p-*, as Mason states, or has merely not used the nominal prefix. An examination of his material suggests that original p^h and *p* (intermediate) remain in both dialects (e.g. M. $p^h xat$ "excrement": A. $p^h xat^h$, no.8; M. *penan* "milk": A. *penano*; M. *pala:kak'* "California woodpecker": A. *pela:k:aʔ*; M. *spoket* "fur, hair": A. *spoket*; M. *pasił* "chia": A. *pasił*; M. $pat^h ak$ "manzanita": A. $pat^h ax$; M. *pe'łⁱ* "pil": A. *pili*; M. *t'ənəpleʔ* "firesticks": A. *tapleya*; M. *tipintša* "whiskey": A. *tepenša*. Such examples are far too numerous to allow one to say that original *p* disappears in Antoniaño. It is probably nearer correct to say that *p'* becomes *ʔ* in Antoniaño (e.g. M. $k^h ap'$ "acorn": A. *kaʔ*; M. *p'as* "son": A. *as*, read *ʔas*; M. *išilip'*, read *išilip*, "fingernail": A. t^h-ⁱ*šeleʔ*) but that in cases of type M. *p-*: A. zero we are really dealing with parallel forms with and without *p*-prefix. As Mason does not always write *p'*, it seems that cases like M. *spanat* "hide": A. *sanat'* should really be understood as *sp'anat* : *s'anat* (contrast M. and A. *spoket* above). Our interpretation is supported by the fact that the active verbal *p*-prefix does not disappear in Antoniaño and by the further fact that in derivatives of nouns with *p-* this consonant is replaced by other elements (e.g. *k-akat-e* "be bloody", *k-exako-p* "bony").

Sapir's appeal to Hokan correspondences in order to establish the existence of the prefix is of course putting the cart before the horse if the goal is to establish genetic affiliation. Moreover, such an argument cannot demonstrate the synchronic analyzability of the form, so that the unexplained omission of the initial *p* of *paxak* cannot be justified as appropriate presentation of data, even if it is legitimately to be ignored in comparisons.

Sapir's other two arguments regarding initial *p* are potentially valid, although alternatives must be explored before they can be accepted as definitive.[23]

---

[23] To my knowledge, Sapir's paper is the most recent treatment of this issue in the literature.

The fact that *p* is not always lost was noted already by Mason (1918), who quite correctly did not view it as necessarily arguing against his proposal. It is quite possible that the loss may have been subject to phonological and/or morphological conditions of which we are not aware. Such conditions sometimes remain obscure even after extensive study — in the case of Salinan, there has been virtually no work on the comparison of the two dialects.

Sapir's observation that in certain cases the *p* is absent even in Migueleño when the stem appears in an adjective is interesting and may indeed show that it is not an integral part of the stem. However, one would like to rule out other hypotheses. For example, what happens with stems that clearly do begin with a *p*? Could the loss of the *p* in these cases be due to phonological causes, for example, reduction of a [kp] cluster?[24]

In sum, the status of word-initial correspondences between M. *p* and A. zero remains unsettled, and although Sapir casts significant doubt on the necessity of reconstructing *p* for Proto-Salinan in these cases, it seems irresponsible to ignore the M. forms. In the case of non-initial *p'*, as in A181, the evidence is that *p'* ought to be reconstructed, whence Greenberg's citation only of the *p*-less A. form is inappropriate. Greenberg's treatment appears not to reflect a considered analysis of these correspondances. While Migueleño *p* is ignored where convenient, in H139, where the putative cognates all have a labial, he cites both A. *šxaʔ* and M. *šxap*.[25]

6.3. Errors

By far the largest number of errors in LIA are incorrect forms, which are found in 24 Salinan entries: H8, H18, H20, H51, H83[bis], H90, H92, H94, H114, H123, H132, H145, H160, A27, A47, A79, A87, A137[bis], A175, A217, A246[bis], A255, G19, G84, and one Yurumanguí entry: H161. The errors vary in nature. Some involve the omission of phonetic detail of dubious importance, such as the distinction made by Mason between [a] and [ʌ], or the presence of aspiration, neither of which appears to have been phonemic. Other omissions of phonetic detail are less trivial, including failure to note the alveolar/post-alveolar distinction, vowel length, and

---

[24] This does not seem promising as, according to Turner (1987), [kp] clusters are permissible, even in initial position.

[25] Katherine Turner (p.c. April 1991) indicates, on the basis of her study of all of the available Salinan material, that the correspondances for *p* are quite messy, with *p* sometimes present in Migueleño and absent in Antoniaño, with *p* sometimes present in Antoniaño and absent in Migueleño, sometimes present in both dialects, and sometimes absent in both dialects.

glottalization, all of which appear to have been be phonemic. Stress, which appears to have been phonemic, is never marked in LIA, though it is generally indicated in the notebook. [26] In still other cases, one or more entire segments are missing.[27]

A second type of error consists of incorrect glosses, which are found in one Yurumangui entry, H69, and seven Salinan entries: H90, H121, A2, A53, A148, G80, G90. Some are probably trivial (e.g. the addition of "high" to "long, tall" in H90), while others may undermine the validity of the cognate set.

A third type of error is incorrect identification of the dialect, found in four Salinan entries: H26, H69, H77, A43.

Questions of segmentation arise in the fifteen Salinan cases and twelve Yuru-manguí cases (ten forms and two affixes) discussed above. Of these, four Salinan analyses and all twelve Yurumanguí analyses lack any language-internal justification.

The most serious errors are spurious forms, of which there are eight clear cases in Salinan entries, in the four entries: H166, A244, G80, G88, and two in the Yurumanguí entries: H161 and A102. Of these, four forms claimed to be Salinan are actually Chumash. Two Jicaque forms, one Chumash form, and one Salinan form are not attested, at least in the sources cited by Greenberg. One Yurumanguí form is simply wrong; another exists but is of unknown meaning. Such spurious forms not only make LIA unreliable as a source of data, but it goes without saying that they are of no comparative value, no matter what methodology one may favor.

Such a large number of errors raises the question of whether they are random or tend to improve the equations. Given that LIA makes no attempt to demonstrate regular phonological relations between members of the comparison sets and the looseness of the criteria for semantic similarity, in many cases it is difficult to say whether the errors have any impact on the equation. Nonetheless, in some cases it is possible to form an opinion.

---

[26] Turner (1987;39-42) gives a phonemicization together with minimal pairs supporting it. A similar phonemicization, based on Jacobsen's notes, is given by Hester (1978;500). While Greenberg could not have made use of Turner (1987) and might have missed Hester (1978), in the absence of a phonemicization the conservative route is to include information rather than to omit it. In any case, the fact that aspiration, vowel length, glottalization, and the alveolar/post-alveolar distinction are sometimes marked and sometimes omitted makes it unlikely that the errors result from an implicit phonemicization.

[27] Greenberg (personal communication, July 1991) informs me that many errors are typographical errors arising from difficulties in the publishing process.

The two inconsistent equations (H18 with A185 and H77 with A228) clearly increase the number of equations — of each pair, only one can be correct.

In every case of questionable segmentation (Salinan: H10, H20, H61, H62, H83, H87, H123, H132, A27, A43, A79, A148, A242, A244, A255; Yurumanguí: H15, H35, H69, H87, H142, A79, A89, A104, A191, A243, A269, G61, G90) the segmentation in LIA makes the cited form more closely resemble the comparanda. Other equations that seem to be improved by errors are: H16, H166, A2, A47, A53, G80, G88 for Salinan and A102 and G80 for Yurumanguí.

On the other hand, one equation, A246, seems to be the worse for Greenberg's error, as the forms with which the Salinan forms are compared have no initial vowel corresponding to Greenberg's erroneous *i*.

In sum, of the 62 questionable entries, the equation is improved by the error or questionable segmentation in 39 cases and worsened in 1, with a neutral effect in the remaining 22 cases.

6.4. The Relationship between LIA and the Notebooks

Although LIA is putatively based entirely on the material in the notebooks, the above analysis of the Salinan data reveals a number of problematic examples. Four entries, H46, A43, A137, and A246, contain forms not to be found in the notebook, which appear to be taken from Sapir (1917). In H90, while the form itself appears in the notebook, the gloss does not and appears to be taken from Sapir (1917). Finally, the Jicaque form *laksak* cited in H142 and A244 does not appear in the notebook. These examples appear to establish that LIA derives in part from sources other than the notebooks. However, Greenberg (personal communication, July 1991) informs me that subsequent to the deposit of photocopies of his notebooks in the library in 1981, he continued to add entries to the originals in his possession until the book went to press in 1984.

6.5. Use of Sources

Greenberg's selection from the published sources for Salinan is appropriate. The sources ignored are generally of low quality and, with the exception of Sitjar, contain only small amounts of data. Moreover, Mason (1918) incorporated most of the previously published data into his own work. Use of the unpublished field notes of Harrington and Jacobsen might, have provided more and higher quality data, had

Greenberg been able to make use of them. The source used for Yurumanguí is the only one available.

The forms in H46, A43, A137, and A246 that appear to derive from Sapir (1917) are attributed by Sapir to Mason's unpublished notes. Since Mason's monograph was published the following year and was used by Greenberg as his main source of data, it is difficult to understand why he cites these forms rather than the presumably more reliable forms in Mason's published work.

7. Conclusion

Of the 81 forms cited in LIA entries for Salinan, 51 or 63.0%, are questionable in some way, ranging from inaccurate forms and glosses and misidentifications of dialect to unjustified segmentations and spurious forms.[28] If we count only clear errors, omitting marginal errors like the omission of non-phonemic detail, matters of analysis, such as morphological segmentation, and inconsistent equations, there are 37 errors, or 45.7%.

In the case of Yurumanguí, there is only one outright error out of 26 entries (3.8%), but when we add in the dubious analyses and other entries calling for comment, the number rises to 17 (65%).

Not all of the errors and questionable points bear clearly on the validity of the equations, but insofar as LIA is intended also as a source for further work, such as attempts to extend or modify comparison sets, construction of phonological correspondences, reconstruction, or detection and elimination of loanwords, even minor errors are potentially problematic.

Judging from the Salinan entries, the notebook is considerably more accurate than LIA; it is incorrect only in nine relatively minor cases: H26 (dialect), H69 (dialect), H121 (gloss), H132 (vowel length), H145 (vowel length), H160 (aspiration), A27 (spurious vowel), A148 (gloss), and A246 (vowel length). Most errors arise between the notebook and LIA. Nonetheless, that 12% of the notebook entries should be in some way erroneous, together with the difficulty of using it, and the

---

[28] The number of forms exceeds the number of entries because section 80 of Chapter V, discussed above, refers to no less than five forms (*be-, noqš, noqšiwaš, -iwaš,* and *-š*), all of which are erroneous or otherwise call for comment. Inconsistent equations have been counted as single errors.

need to refer in any case to the original sources, make the notebooks less useful as documentation for LIA and as a source of data than they are represented as being.

On the basis of the preceding analysis of the Salinan and Yurumanguí data in LIA, we may draw the following conclusions, the generality of which depends upon the extent to which the treatment of these two languages is representative of the work as a whole:

- LIA contains material not to be found in the notebooks on which it is putatively based. Not only are the notebooks containing grammatical material not publicly available, but LIA contains lexical data not to be found in the copies of the lexical notebooks housed in the Green library.

- In spite of the representation that segmentations indicated by hyphens are independently justifiable internal to the language in question, it is frequently the case that the indicated segmentation is questionable, in some cases wholly without language-internal justification, in other cases, while perhaps plausible, not demonstrable on the basis of the available data.

- Although works like Sapir (1917) are used, without acknowledgment, as sources of forms and analyses, little attention appears to have been given to any other aspect of the literature, such as the discussion of p and p' by Mason (1918) and Sapir (1925), the caveat regarding segmentation in Sapir (1917), or the sometimes conflicting cognate sets in Gursky (1974).

- No distinction is made between data from amply documented languages with well-understood morphology, and data from poorly documented languages the analysis of which is quite speculative. It is possible that Yurumanguí did have a past tense infix *iba* (G80) and that Salinan did have an imperative suffix *-i-* (G88), but one can hardly put much confidence in these forms. The reader is entitled to know the provenance and reliability of the data.

- The data in LIA contain numerous errors, including incorrect forms, spurious forms, incorrect glosses, incorrect dialect affiliations, and attribution of forms to the wrong language.

- The data in the notebooks, though imperfect, are not nearly as error-ridden as those in LIA. The errors are both fewer and in general less serious. Nonetheless, since the notebooks contain nothing but raw lexical material, in the orthography of the sources, and without detailed source references, they fail to provide the detailed documentation that is missing from LIA.

References

Adelaar, Willem F. H. 1989. Review of *Language in the Americas*. *Lingua* 78.249-255.

Applegate, Richard B. 1972. *Ineseño Chumash Grammar*. Unpublished Ph.D. dissertation, Department of Linguistics, University of California at Berkeley.

Beeler, Madison S. 1976. "Barbareño Chumash Grammar: A Farrago," in Margaret Langdon & Shirley Silver (eds.) *Hokan Studies*. The Hague: Mouton. pp. 251-269.

Cabot, Padre Pedro & Pedro Dumetz (n.d.) *A Vocabulary of the (illegible) words from Mission San Antonio taken in the handwriting of (illegible) A.S. Taylor, Nov. 1853.* Given to the Boston Athenaeum by Mrs. Schoolcraft, March 6, 1871. $2B7 C11

Campbell, Lyle. 1988. Review of *Language in the Americas*. *Language* 64.591-615.

Campbell, Lyle & David Oltrogge. 1980. "Proto-Tol (Jicaque)," *International Journal of American Linguistics* 46.205-223.

Chafe, Wallace. 1987. Review of *Language in the Americas*. *Current Anthropology* 28.652-653.

Conzemius, Edward. 1921. "The Jicaques of Honduras," *International Journal of American Linguistics* 2.163-170.

de la Cuesta, Fr. Felipe Arroyo. 1821. *Vocabulary of Salinan*. Transcription by Albert S. Gatschet, National Anthropological Archives, Smithsonian Institution, Washington, D.C. Bureau of American Ethnology Ms. 850.

de la Cuesta, Fr. Felipe Arroyo. 1833. *Lecciones de Yndios*. H.H. Bancroft Collection, University of California, Berkeley. Ms. C-C 63b.

Dennis, Ronald K. & Margaret Royce de Dennis. 1983. *Diccionario Tol (Jícaque) - Español y Español - Tol (Jícaque)* Tegucigalpa, Honduras: Instituto Lingüistico de Verano.

Dixon, Roland B. & Alfred L. Kroeber. 1919. "Linguistic Families of California," *University of California Publications in American Archaeology and Ethnography* 16.47-118.

Gallatin, Albert S. 1848. "Hale's Indians of Northwest America and Vocabularies of North America," *Transactions of the American Ethnological Society* 2.1-134.

Goddard, Ives. 1987. Review of *Language in the Americas*. *Current Anthropology* 28.656-657.

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

Greenberg, Joseph H. & Morris Swadesh. 1953. "Jicaque as a Hokan Language," *International Journal of American Linguistics* 19.216-222.

Gursky, Karl-Heinz. 1974. "Der Hoka-Sprachstamm. Eine Bestandsaufnahme des lexikalischen Beweismaterials," *Orbis* 23.170-215.

Haas, Mary. 1964. "California Hokan," in William Bright (ed.) *Studies in Californian Linguistics* University of California Publications in Linguistics, vol. 34. Berkeley and Los Angeles. pp. 73-87.

Harrington, John Peabody. 1922 and 1932-33. *Field Notes on Salinan* National Anthropological Archives, Smithsonian Institution, Washington, D.C.

Heizer, R.F. 1952. *The Mission Indian Vocabularies of Alphonse Pinart*. Anthropological Records 15.1. Berkeley: University of California Press.

Henshaw, Henry H. 1880 and 1884. *Vocabulary collected at San Miguel and San Antonio*. National Anthropological Archives. Smithsonian Institution, Washington D.C. BAE Ms. 3077a and 3077b.

Hester, Thomas Roy. 1978. "Salinan," in Robert F. Heizer (ed.) *Handbook of North American Indians* Washington, DC: Smithsonian Institution. Vol.8: California. pp. 500-504.

Jacobsen, William H. Jr. 1954-1958. *Salinan Field Notes*. Unpublished manuscript.

Jacobsen, William H. Jr. 1958. "Washo and Karok: An Approach to Comparative Hokan," *International Journal of American Linguistics* 24.195-212.

Kroeber, A. L. 1901. *Salinan Field Notes*. Unpublished manuscript on file in the Survey of California and Other American Indian Languages, Department of Linguistics, University of California, Berkeley.

Kroeber, A. L. 1904. "Languages of the Coast of California South of San Francisco," *University of California Publications in American Archaeology and Ethnology* 2.29-80.

Lehmann, Walter. 1920. *Zentral-Amerika.* Berlin: Dietrich Reimer.

McLendon, Sally. 1964. "Northern Hokan (B) and (C): A Comparison of Eastern Pomo and Yana," in William Bright (ed.) *Studies in Californian Linguistics* (University of California Publications in Linguistics 34) pp. 126-144.

Mason, J. Alden. 1910. *Salinan File Slips.* American Philosophical Society, Philadelphia.

Mason, J. Alden. 1918. "The Language of the Salinan Indians," *University of California Publications in American Archaeology and Ethnology* 14.1-154.

Membreño, Alberto. 1897. *Hondureñismos: Vocabulario de los provincialismos de Honduras.* Tegucigalpa: Tipografia Nacional. 2nd ed.

Merriam, C. Hart. 1902 and 1933. *Vocabularies of North American Indians* and *Field Check Lists.* The Manuscript Division, Library of Congress, Washington, D.C.

Powell, John Wesley. 1877. "Linguistics," *Contributions to North American Ethnology* (Washington, D.C.: U.S. Department of the Interior) 3.439-613.

Rivet, Paul. 1942. "Un Dialecte Hoka Colombien: Le Yurumangí," *Journal de la Société des Américanistes de Paris* 34.1-59.

Sapir, Edward. 1917. "The Position of Yana in the Hokan Stock," *University of California Publications in American Archaeology and Ethnology* 13.1-34.

Sapir, Edward. 1920. "Review of 'The Language of the Salinan Indians' by J. Alden Mason." *International Journal of American Linguistics* 1.305-309.

Sapir, Edward. 1921. "A supplementary note on Salinan and Washo," *International Journal of American Linguistics* 2.68-72.

Sapir, Edward. 1925. "The Hokan Affinity of Subtiaba in Nicaragua," *American Anthropologist* 27.402-435, 491-527. [This paper is cited in LIA (p. 401) with the incorrect title "The Hokan Affinities of the Subtiaba in Nicaragua".]

Scouler, John. 1841. "Observations on the Indigenous Tribes of the North West Coast of America," *Journal of the Royal Geographic Society (London)* 11.215-251.

Sitjar, Fr. Buenaventura. 1861. *Vocabulario de la Lengua de los Naturales de la Mision de San Antonio, Alta California.* Shea's Library of American Linguistics, no. 7. New York: Cramoisy Press. Reprinted 1970 at New York by AMS Press.

Taylor, Alexander S. 1860-1863. "California Notes: The Indianology of California," *California Farmer*, Vols. 13-20. Salinan vocabulary 2-22-60 and Vol. 13, no. 10, 4-27-60. H.H. Bancroft Collection, University of California, Berkeley.

Turner, Katherine. 1987. *Aspects of Salinan Grammar*. Unpublished Ph.D. dissertation, Department of Linguistics, University of California at Berkeley.

Yates, L.G. and G. H. Gould. 1887. *Salinan and Chumashan Vocabularies*. National Anthropological Archives, Smithsonian Institution, Washington D.C. Bureau of American Ethnology Ms. 4239.