REPORT 8

SURVEY OF CALIFORNIA AND OTHER INDIAN LANGUAGES


both held at the 1993 Linguistic Institute at Ohio State University in Columbus, Ohio

Margaret Langdon, Volume Editor
Leanne Hinton, Series Editor
REPORT 8

SURVEY OF CALIFORNIA AND OTHER INDIAN LANGUAGES

Proceedings of the Meeting of the
Society for the Study of the Indigenous languages of the Americas
July 2-4, 1993

and the Hokan-Penutian Workshop
July 3, 1993

Both held at the 1993 Linguistic Institute at Ohio State University in Columbus, Ohio

Margaret Langdon
Volume Editor

Leanne Hinton
Series Editor
This volume is dedicated to

JAMES E. REDDEN

on the occasion of his retirement

for his enduring commitment to the publication

of the results of research on Yuman, Hokan, Penutian and

other American Indian languages

and also

for his contributions to the

documentation of the Hualapai language
INTRODUCTION

This volume includes a number of papers presented in conjunction with the 1993 Linguistic Institute at Ohio State University in Columbus, Ohio, at two conferences on American Indian Languages: the meeting of the Society for the Study of the Indigenous languages of the Americas, held July 2-4, 1993, and the meeting of the Hokan-Penutian Workshop, held on the morning of July 3, 1993.

This continues a tradition initiated during the Linguistic Institute at the University of Arizona in 1988, of offering conferences on American Indian languages during the summer Linguistic Institute of the Linguistic Society of America, which is held every two years on the campus of the host institution. The interaction thus afforded between students and faculty of the Institute and specialists in American Indian languages has proved mutually profitable.

We gratefully acknowledge the dedication of Catherine Callaghan in making these meetings thoroughly enjoyable, as well as the hospitality of Ohio State University.

The Hokan-Penutian Conference has a tradition of meetings dating as far back as 1970, when the first Hokan Conference was hosted by Margaret Langdon at UCSD. Since 1976, the Hokan (and later Hokan-Penutian) Conference proceedings were published most years by James Redden, as part of the series Occasional Papers on Linguistics, out of the department of Linguistics at Southern Illinois University at Carbondale. Beginning this year, with James Redden's retirement, the reports of these conferences are being published as part of the Survey Reports out of the Survey of California and Other Indian Languages at the University of California at Berkeley.

Margaret Langdon  
Volume Editor  

Leanne Hinton  
Series Editor
CONTENTS

GENERAL

Putting pronouns in proper perspective in proposals of remote relationships among Native American languages  
CAMPBELL, LYLE  
1

S- prefixation on Upper Chehalis (Salish) imperfective predicates  
KINKADE, M. DALE  
21

Comparative difficulties of the "Gulf" languages  
KIMBALL, GEOFFREY  
31

Like hair, or trees: Semantic analysis of the Coeur d'Alene prefix ne' 'amidst'  
OCCHI, DEBRA J., GARY B. PALMER, AND ROY H. OGAWA  
40

SOUTH AMERICA

Constituent order variation in Apurinã  
FACUNDES, SIDNEY  
59

Case, verb type, and ergativity in Trumai  
GUIRARDELLO, RAQUEL  
75

Nheengatu (Língua Geral Amazonica), its history, and the effects of language contact  
MOORE, DENNY, SIDNEY FACUNDES, AND NADIA PIRES  
93

Reconstruction of Proto-Tupari consonants and vowels  
MOORE, DENNY, AND VILACY GALUCIO  
119

Basic word order in Karitiana (Akirem family, Tupi stock)  
STORTO, LUCIANA R.  
138

HOKAN-PENUTIAN WORKSHOP

'How', and 'thus' in UA Cupan and Yuman: A case of areal influence  
ELLIOTT, ERIC  
145

Kroeber and Harrington on Mesa Grande Diegueño (lipay)  
LANGDON, MARGARET  
170

Conjunctions and reference tracking in Yuma  
MILLER, AMY  
183

The shifting status of initial glottal stop in Barbareño Chumash  
MITHUN, MARIANNE, TSUYOSHI ONO, AND SUZANNE WASH  
199

Final glottalization in Barbareño Chumash and its neighbors  
ONO, TSUYOSHI, SUZANNE WASH, AND MARIANNE MITHUN  
208
Putting Pronouns in Proper Perspective in Proposals

of Remote Relationships among Native American Languages

Lyle Campbell

Louisiana State University

1. Introduction. It is claimed (cf. Greenberg 1987) that there is a widespread pattern of \( n \) 'first person' and \( m \) 'second person' in pronouns which is characteristic of American Indian languages. Proponents of Greenberg's classification believe that evidence based on the distribution of these pronouns is totally convincing as an argument in favor of Greenberg's controversial "Amerind" proposal, which links most of the languages of the Americas — with the exception of Eskimo-Aleut and so-called Na-Dene — in a single large language family (cf. Fleming 1987:196, Ruhlen in press a:13, 1990:9), and they repeat this argument with great frequency. For example:

The enormously widespread \( n \) first person and \( m \) second person in the Americas as against \( m \) first person and \( n \) second person in Europe and Northern Asia is powerful evidence which requires explanation. (Greenberg 1990:19.)

One of the most salient traits of the Amerind family ... is the presence of 'first-person' \( n \) and 'second-person' \( m \) throughout the languages of North and South America. Furthermore, not only does this trait connect all eleven Amerind subgroups, it also serves to distinguish the Amerind family from the world's other language families. (Ruhlen 1989b:2.) [Emphasis mine, LC.]

Given the emphasis the pronoun argument has received and how misleading it has been, it is important to take it up in some detail, in order to set the record straight and to put the matter in clearer perspective. This is the intention of this paper.

2. Some history. Greenberg is by no means the first to have observed pronoun similarities among American Indian (and other) languages; it will be helpful to review some of this history in order to put the claims concerning these pronouns in context and to show the various sorts of explanations that have been offered to account for them. Though a number of earlier reports could be cited, let me begin with Brinton, who was fully aware of how common the \( n \) of first person forms was in American Indian languages and in languages generally around the world:

the \( N \) sound expresses the notion of the ego, of myself-ness, in a great many tongues, far apart geographically and linguistically. It is found at the basis of the personal pronoun of the first person and of the words for man in numerous dialects of North and South America. Again, the K sound is almost as widely associated with the ideas of otherness, and is at the base of the personal pronoun of the second person singular ... (Brinton 1890[1888]:396.)

Boas (1917:5) knew the American Indian pronoun facts, but thought they would submit to "psychological explanations" (and here Boas is cited also by Ruhlen [1987:222] and Haas [1966:102]). Boas cited Gatschet before him, who Boas says was already aware
of the widespread phonetic similarities among the pronouns in American Indian languages before him, but also did not attribute these necessarily to genetic inheritance (cf. Haas 1966:102).

Kroeber, too, was well aware of the American distribution of the n/m pronominal markers (stating that this was a well-known example), but realized that a genetic explanation was not necessarily required; he opted for diffusion/language contact as the probable explanation (on which more below):

Throughout the field of linguistic structure in the whole continent, there are abundant examples of the operation of the principle of territorial continuity of characteristics, and of the underlying one that even the most diverse languages affect each other, and tend to assimilate in form, if only contact between them is intimate and prolonged. Such are the exceedingly common occurrence [sic] of n and m to designate the first and second person pronouns; the geographical localization of families expressing sex gender; the prevailing tendency for pronominal elements, especially the possessive ones, and instrumental elements in verbs, to be prefixes rather than suffixes, as already mentioned for California. It is needless to multiply examples which are either familiar to the Americanist or readily compilable by him. (Kroeber 1913:399.)

The alleged n/m pronoun pattern in the Americas was cited by Trombetti (1905) and it became very well-known through the Sapir-Michelson debates (Michelson 1914, 1915, Sapir 1915a, 1915b). In 1920 Sapir listed "persistence of n- 'I' [and] m- 'thou'" as "Proto-American possibilities" (Sapir 1990[1920]:86; Golla 1984:452). Widespread n for 'first person' was talked about widely:

Getting down to brass tacks, how in the Hell are you going to explain general American n- "I" except genetically? It's disturbing to know but (more) non-committal conservatism is only dodging after all, isn't it? (Letter from Sapir to Speck in 1918, Darnell 1990:122; cf. also his letter to Kroeber in October of 1920, Golla 1984:316; also cited by Darnell and Sherzer 1971:27, Goddard 1986:201-2, and Ruhlen 1987:222.)

In fact, Greenberg's claim sounds very much like Swadesh's before him:

At least two short elements, n for the first person pronoun and m for the second ... are so numerous as to virtually eliminate the chance factor despite their brevity. In fact, even if one disregarded the cases which have one or the other and included only the languages which have both n and m for first and second person respectively, and if one holds to the restriction that both forms must belong to the same functional type [a restriction not imposed by Greenberg] -- whether independent pronoun or subject, object or possessive affixes -- the list of language groups would still be fairly impressive. It would include families of the Penutian and Hokan-Coahuiltecan phyla, Aztecan, Chibchan, and Mapuche. (Swadesh 1954:311-12.)

As an aside, it is interesting to note that here Greenberg and Ruhlen (see quotes above) conveniently forget the 'first-person' m which they find in several "Amerind" groups (Greenberg 1987:276; Ruhlen 1989b:1) and do not acknowledge the widespread *ta/tu 'thou' offered by Swadesh (1960:909) as Proto-American. That is, if the n/m
pattern distinguishes Amerind from Europe and Northern Asia with its alleged m/t pattern, then why do several Amerind groups exhibit pronoun forms (m/t) that Greenberg attributes to Europe and Northern Asia? This makes the pronoun claim less 'powerful', but no less in need of an 'explanation'.

Others, not even Americanists, had also noticed such similarities among the pronouns of diverse languages of the world, but denied the genetic explanation, e.g. Wundt (1900:33) and Trombetti (1905:44).

In short, these observations concerning pronouns are not new arguments, for example, for Greenberg's classification (as some mistakenly take them to be), but rather are old controversies (cf. Ruhlen 1987:222, Sapir 1917:184).

3. Possible explanations for the pronoun pattern. Since Greenberg and Ruhlen take this pronoun argument to be compelling, it is important to consider the matter in greater detail. Greenberg (1989:113) asserts of the alleged pronoun pattern "that a highly improbable event should have occurred more than a hundred times exceeds the bounds of credibility ... [it] cannot be explained plausibly except as the result of genetic inheritance." However, the assumption of genetic inheritance is by no means necessary nor is it the only explanation available (as the scholars cited above pointed out). There are strong reasons for believing that other factors are involved in explaining the sounds found to recur in these pronouns. Boas (1917:5, mentioned above) asked whether the pronoun pattern could be "due to obscure psychological causes rather than to genetic relationship." The following are a few factors which have been proposed which may make Boas "psychological causes" less "obscure".

(1) Certain sounds, especially nasals, are to be expected in grammatical morphemes, in particular in pronoun markers. As pointed out in Goddard and Campbell (in press:16-7; cf. also Campbell in press):

   The repeated appearance in different languages of the same consonants in grammatical functions is a real phenomenon of human language and as such requires an explanation. One contributing factor is the well known general linguistic trait that a single language typically uses only a fraction of its full complement of consonants to form its primary grammatical morphemes and hence must use the same consonants over and over in different functions (Floyd 1981). The consonants that are used tend to be the ones that are least marked ... specifically, the least marked consonants of the languages of the world include m, n, t, k, and s (cf. Ruhlen 1987:11). As a result of this economy and, so to speak, lack of originality in the use of consonants, there is a much greater than chance agreement among the languages of the world on what consonants are used in grammatical elements. It is thus to be expected a priori that these consonants will show up again and again in different languages and language groups marking, say, first or second person, and many languages will therefore come to have similar pronominal systems by this factor alone.

German inflectional endings are constrained such that the only vowel found is schwa, the only consonants /d, m, n, r, s t/. Of Latin's 15 or more consonants, only /b, d, m, n, r, s, t, w/ occur in inflectional endings; Hebrew permits only 8 of 22; and English has similar limits. Of Ancient Greek's 15 consonants, only /t, th, k, m, n, r, s/ occur in inflectional
morphemes (Floyd 1981). Even Trombetti (1905:89) had realized something of the limited sounds encountered in pronominal forms in the world’s languages:

In all these old pronominal forms only the vowels a, i, u, the stop consonants k, t and the nasals n, m are found. These are certainly primordial sounds ... [my translation, LC.]

(Cf. also Matisoff 1990:9.)

(2) Nasals in particular are found in grammatical morphological markers precisely because they are the most perceptually salient of all consonants (Maddieson 1984:70). "[T]he more distinctive speech sounds ... achieve the most successful transmission of a message." Nasals "are rarely subject to confusion with other types of consonants," and "there is value in incorporating such sounds [nasals] into any language (Maddieson 1984:70)." The dental/linguolateral nasal (n) is most common, with the bilabial (m) also common; most languages have both (Maddieson 1984:60, 69). These facts would seem to explain why nasals, especially n and m, show up in markers of pronouns so frequently in languages everywhere. This is borne out in, for example, Ruhlen’s (1989b) survey of first- and second-person pronouns in the world’s languages. Among the forms he lists for 40 different genetic units (some of which are very long-range and more controversial than others), for ‘I’ one finds 29 exhibiting a nasal consonant, and only 11 with no nasal; similarly for ‘thou’ (i.e. ‘you singular’) one finds 26 with n or m, and only 15 with no nasal. For ‘I’, 13 exhibit n. (Incidentally, the other consonants exhibited in the 11 non-nasal cases for ‘I’ are predominantly t, s, and k; for ‘thou’, among the 15 lacking nasals, the predominating consonants are t, s, c, and w. These recurrent sounds in the world’s pronoun systems are not accidental, but are predicted based on the perceptual saliency of the sounds employed (see (1) above).6

(3) Some consider the possibility of areal diffusion, including pronouns, among the various early groups which came into America, borrowing from one another either before they crossed the Bering Straits, or after, or both (cf. Bright 1984:15-6, 25; Milewski 1960; Kroeber 1913:399). Diffusion of pronouns in such a situation is not so unusual as some might prefer to think (for examples, see Everett in press, Matisoff 1990:113, Newman 1977:306-9, 1979a:218-23, 1979b:305-7, 1980:156, Rhodes 1977:9, Thomason and Kaufman 1988:219-20, 223-8, 235). Even Ruhlen (1989b:4) allows for the possibility that Nahali borrowed ‘thou’ from Dravidian, and several of the cases just cited involve Native American languages which have borrowed of pronouns. It is well-known that English they, their, them is borrowed from Scandinavian (replacing Old English hie, hiera, him, respectively; cf. for example Baugh 1957:120, 194). Surely we cannot deny the borrowing of pronouns elsewhere, when we have clear examples in our on linguistic backyard.7

Perhaps more to the point, there are a number of cases of borrowed pronouns documented in Native American languages. For example, Miskito borrowed its independent personal pronouns from Northern Sumu in relatively recent times: Miskito yan ‘I’ (cf. Sumu yan), man ‘you’ (cf. Sumu man).8 Another example of pronominal diffusion is particularly revealing, since it involves a Native American language, Mednyj Aleut, which has borrowed its verb morphology, including pronominal endings, from Russian. Some of the pronominal verbal paradigm is:
<table>
<thead>
<tr>
<th>Mednyj Aleut</th>
<th>Russian</th>
<th>cf. Bering Aleut</th>
</tr>
</thead>
<tbody>
<tr>
<td>uŋuči-ju</td>
<td>ja sižu</td>
<td>uŋuči-ku-q</td>
</tr>
<tr>
<td>uŋuči-iš</td>
<td>ty sidiš</td>
<td>uŋuči-ku-Xt</td>
</tr>
<tr>
<td>uŋuči-it</td>
<td>on sidit</td>
<td>uŋuči-ku-X</td>
</tr>
<tr>
<td>uŋuči-im</td>
<td>my sidim</td>
<td>uŋuči-ku-s</td>
</tr>
<tr>
<td>uŋuči-itı</td>
<td>vy sidite</td>
<td>uŋuči-ku-Xt-xichx</td>
</tr>
<tr>
<td>uŋuči-jat</td>
<td>oni sidjat</td>
<td>uŋuči-ku-s</td>
</tr>
<tr>
<td>ja uŋuči-ıł</td>
<td>ja sidel (masc.)</td>
<td>uŋuči-na-q</td>
</tr>
</tbody>
</table>

(Thomason and Kaufman 1988:234-5.)

It will be noticed that not only is the Russian pronominal system borrowed pretty much lock, stock, and barrel, but even these borrowed pronominal affixes afford parallels which match up with forms postulated by Greenberg as representative of American Indian languages. Thus Mednyj Aleut’s second person singular -iš can be compared with Greenberg’s (1987:278-9) -s ‘second-person marker’ (with various different shapes in a number of different languages). The -im of ‘first person plural’ apparently fits Greenberg’s (1987:276) -m ‘first person’, since he cites Miwokan -m, me ‘first-person plural subject of verbs’ and Takelma -am ‘first-person plural object marker’, among others, as evidence. The Mednyj Aleut ‘first person singular’ forms with -ju and ja appear to match Greenberg’s (1987:273) ‘first-person’ i which he sees widely in South American languages, where he cites as related such forms as Payagua ja- ‘my’, j-(am) ‘I’, Mataco ji- ‘my’, and Moseten je ‘I’. As for Mednyj Aleut’s -ıt ‘third person singular’, one easily sees parallels to Greenberg’s (1987:279) South American third-person elements where, for example, he says that i- and t- alternate in several of his groups of languages; here one sees both the i and the t in the Mednyj Aleut form. Since these morphemes are clearly borrowed into Mednyj Aleut from Russian (and therefore cannot have any direct historical connection to other New World languages), the fact that they parallel forms postulated by Greenberg as widespread among American Indian languages illustrates how fragile Greenberg’s pronominal argument is in general and also how weak his postulated grammatical evidence is on the whole. They show how easy it is to find accidentally similar forms that are as persuasive as those he listed. In particular, however, the Mednyj Aleut forms demonstrate that one cannot rule out the possibility of borrowing as an explanation for some of the similarities among pronouns that Greenberg asserts as genetic evidence for Amerind.9

Lest anyone suspect that cases of pronoun borrowing, as in Mednyj Aleut, might require European colonialism, I hasten to mention the existence of such examples as Alsea (Oregon). Kinkade (1978) found that while Alsea (often tentatively assigned to putative Penutian) has no discernible genetic relationship with Salishan (only a few loan words),10 it has undergone a remarkable convergence with Salish in the pronouns:

<table>
<thead>
<tr>
<th>Alsea</th>
<th>Proto-Salishan</th>
</tr>
</thead>
<tbody>
<tr>
<td>-an</td>
<td>-n</td>
</tr>
<tr>
<td>-ax</td>
<td>-xw</td>
</tr>
<tr>
<td>-o</td>
<td>-Ø</td>
</tr>
<tr>
<td>-a+</td>
<td>-+</td>
</tr>
<tr>
<td>-ap</td>
<td>-p</td>
</tr>
<tr>
<td>-a+x</td>
<td>(lx)</td>
</tr>
</tbody>
</table>

I point out here in conjunction with cases of pronominal borrowing that Haas
concluded from her comparison of languages in northern California that the n/m pattern is widely borrowed:

There are clear evidences of diffusion in pronominal forms in northern California ... belonging to a single diffusion area ... The most prominent feature is n- in the first person paired with m- in the second person ... But the total picture of diffusion of n- and m- in the first and second persons goes beyond the area being studied in this paper [Haas 1976] and so the problem really needs to be attacked on a larger scale. [Emphasis added, LC.]

While not all scholars today would agree that these are diffused in the way Haas sees it, from her report it is clear again that the pronoun pattern was well-known before it became so associated with Greenberg's claims and the explanation for it was not assumed automatically to be a genetic one.

(4) Another explanation that has been offered is child language; child-language expressions around the world abound in self-directed and other-directed words containing nasal consonants. The ultimate reason for this is the universal physical fact that a gesture equivalent to that used to articulate the sound n is the single most important voluntary muscular activity of a nursing infant. As Goddard (1986:202) pointed out, possibly this factor and the tendency for primary grammatical morphemes to consist of a single, unmarked (phonetically commonplace) segment account for the widespread appearance of n- in 'first-person' pronouns. Incidentally, in many societies, particularly among hunting and gathering groups, infants may continue to nurse until the age of five, sometimes longer, well into and beyond the age of language-acquisition. (Cf. Goddard and Campbell in press.)

4. Is the claim valid? More to the point, the claim for the ubiquity of 'first-person' n and particularly for 'second-person' m in "Amerind" is grossly overstated. To keep proper perspective, we need to ask also about all the American Indian languages which lack 'first-person' n or 'second-person' m, or both. And what of all the non-American Indian languages which have one or both of these? The second-person m is not nearly so common among American language groups as asserted by Greenberg and Ruhlen. To wit, in spite of the claimed generality for 'first-person' n and 'second-person' m (cf. Greenberg 1987:113), Greenberg finds South America typified by i 'first person', a 'second person', i 'third person' (Greenberg 1979, 1987:44-9, 273-5, 277-81; cf. Swadesh 1954:312) -- a totally distinct pattern, with second-person m particularly absent. If the i/a/i pattern is the hallmark of South America, then the n/m pattern is not as diagnostic for Amerind as a whole, as claimed (Greenberg 1979 notwithstanding). Moreover, Greenberg (1987:276) presents among his grammatical cases an Amerind 'first-person' m as characteristic of several Amerind groups (recall this is what he expects of Eurasian, claimed to be almost absent from America!), while several other groups have 'second-person' ka or s (Greenberg 1987:278; cf. Ruhlen 1989b). In brief, the distribution of pronouns in American Indian languages has been grossly overstated.

Furthermore, to round out the picture, it is important to keep in mind that in many languages outside the Americas 'first-person' n abounds, and it is not difficult to find non-American languages with both Greenberg's 'first-person' n and 'second-person' m pattern. I have undertaken no systematic search, but cite the following readily
available examples:
(1) Mbugu (Cushitic): ni 'first person sg', mu 'second person pl' (Goodman 1971:246).
(2) Proto-Munda *in 'I', *me 'you'. Some individual Munda languages have, for example: (a) Santali: in 'I' (absolute), -in (subjective), -in (direct object), -in (indirect object); am 'you' (absolute), -am me (subjective); (b) Mundari: (a)in 'I' (absolute), -in (subjective); am 'you' (absolute), -am me (subjective), -m, -me (direct object), am (indirect object); (c) Kurku in 'I' (absolute), -n (affixed); am 'you' (absolute), -mi, min (affixed); (d) Gutob niin 'I' (absolute, oblique); maai 'you' (absolute, oblique) (Pinnow 1966).
(3) Dravidian pronominal affixes (oblique):

<table>
<thead>
<tr>
<th>Language</th>
<th>Affix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kui</td>
<td>na-</td>
<td>you (pl)</td>
</tr>
<tr>
<td>Gonda</td>
<td>na-</td>
<td>mi-</td>
</tr>
<tr>
<td>Kurukh</td>
<td>en-</td>
<td>num-</td>
</tr>
<tr>
<td>Brahui</td>
<td>kan-</td>
<td>num-</td>
</tr>
<tr>
<td>Tamil</td>
<td>en-</td>
<td>num-</td>
</tr>
</tbody>
</table>
| Old Kanada | en-  | num- (Bloch 1954).

(5) Kusaiean: nga 'I'; kom 'you', vom 'below you', iyom 'of you' (Lee 1976).
(6) Gilbertese: na 'I' (subjective); -m 'you' (sg), -mi (pl)
   'possessive' (Cowell 1951).
(7) Sonsorol (SW Micronesia): na'n 'I'; -m 'you' (possessive)
   (Capell 1969).
(8) Tanga (Melanesia): -ng 'I' (possessive), -m 'you' (possessive) (Bell 1977).
(9) Wagay (Australian): an 'I' (bound), imh 'you' (Breen 1976).13

(For other cases and discussion, see Trombetti 1905:80-90.)

Perhaps most convincing of all, Dryer (in as yet unpublished work, personal communication) found in his sample of 333 languages that 7% of the languages (excluding Amerind; i.e., 17 out of 252 languages) had both an n in first person and an m in second person, with both either singular or both plural. These include (in addition to some cited above) Enka, Chuave (Papuan); Chrau (Mon-Khmer); Akan (Niger-Congo); seven Bantu languages; Tamazight (Berber); and Hebrew, Arabic, and Tigrinya (Semitic). In Dryer’s research, only 17% of the languages from Greenberg’s Amerind (14 out of 81 languages) had this pattern. He found further from pronoun data on 289 languages, that 118 had more nasals in first person singular than in third person singular, 128 had the same number of nasals in both, and in only 47 were there more nasals in third person singular than in first person singular. At the same time, 34 of these languages had more nasals in second person singular than in third person singular, while 48 had more nasals in third person singular than in second person singular.14

For other examples and discussion of non-American pronoun systems similar to those claimed by Greenberg for Amerind, see also Benjamin (1976), and Matisoff (1990).

The repetition of consonants from the set n/m/t/k/s in pronoun forms around the world, as in the cases just cited, reflects the likelihood of these sounds occurring in grammatical morphemes, specifically pronouns. These sounds are selected from the set of least marked, most perceptually salient consonants.15 In this context it is sobering to
recall Callaghan's (1980:337) presentation of the accidental coincidences in the Miwokan and Indo-European pronominal affixes:

<table>
<thead>
<tr>
<th>Proto-Eastern Miwok</th>
<th>Late Common Indo-European</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative Suffixes</td>
<td>Secondary Affixes (Active)</td>
</tr>
<tr>
<td>1 sg *-m</td>
<td>-*m</td>
</tr>
<tr>
<td>2 sg *-s</td>
<td>-*s</td>
</tr>
<tr>
<td>3 sg *-Ø</td>
<td>-*t &lt;**Ø</td>
</tr>
<tr>
<td>1 pl *-mas</td>
<td>-*me(s)/-mo(s)</td>
</tr>
<tr>
<td>2 pl *-to-k</td>
<td>-*te</td>
</tr>
</tbody>
</table>

If it is conceded that Proto-Eastern Miwok and IE can accidentally share so much in the paradigm of pronominal affixes, where the sounds involved are selected from a set of highly unmarked consonants, then what is to prevent such a coincidence from arising in other languages elsewhere, including in different American Indian languages?

In light of the examples listed here, it is interesting to recall the claims about the purported Amerind n/m pattern:

In a recent survey of pronominal patterns around the world [Ruhlen 1989b]
I found that the Amerind pattern, so pervasive in the Americas, was virtually absent elsewhere in the world. (Ruhlen 1989a:5.)

I collected the first- and second-person pronouns for all the world's major linguistic families ... I did not find a single family anywhere else in the world that shared the Amerind pattern, which turns out not only to define the Amerind family, but at the same time to differentiate it from the world's other language families. (Ruhlen 1990:9)

Perhaps this illustrates how easy it is to see only what one is looking for; there are abundant cases of the n/m pattern outside the Americas.

Greenberg's case is also not helped by languages which behave in ways opposed to his claims, e.g. such Amerind cases as those with n 'second person', e.g. Cayuse n, Cherokee nihi, Atakapa na, Tonkawa na,16 Siouan -nx, Cheyenne ne,17 Proto-Guahibo *ni-hi, Guambiano ni, Tupinamba enê, Proto-Tupi-Guarani *ne, etc. (see also additional cases with m 'first person': Tunica li'ma, Yana wa?ma, Alsea -hm-t 'first person object', Tacana ema, Catio mi, etc.), and cases with the reverse of expectations, i.e. n 'second person / m 'first person' as in Lakota niye 'first person singular', niye 'second person singular' (cf. iye 'third person singular') (Matthew Dryer, Personal communication, Matteson 1972:65, 89).18 Greenberg's case is also not helped by Amerind languages which have neither n nor m in first or second person pronoun forms (e.g. Zuni, Kuikuro, etc.). Proto-Muskogean, with *a 'I, *ç 'you' (and *p or *l 'we'), is such a case, in spite of Greenberg's misanalysis of specific Muskogean languages.19

5. Other criticisms. Finally, Nichols (in press: 37-40) criticizes the Amerind pronoun argument on the basis that it lacks the "paradigmaticity" found in "stock after stock and language after language", when 'first' and 'second person' 'singular' and 'plural' are compared in the Nostratic hypothesis. That is, Nostraticists find in the various component families that they believe make up Nostratic a recurring pronoun system with different but paradigmatically related forms for first- and second-person, singular and plural, subject and object pronouns. This is exemplified in Shevoroshkin's (1989:3-4) Nostratic reconstructions of *mi 'I' / *minV 'oblique form of first person singular', *ti 'thou' / *tínV 'oblique form of second person singular'. There is no such paradigmatic
pattern to the n/m of Greenberg’s pronominal argument. 

6. Conclusion. Whatever the explanation for the frequency of ‘first-person’ n, and for the recurrence of ‘second-person’ m, it will not do to look only at American languages which contain them, ignoring at the same time the dual fact that many American tongues lack them while their presence in non-American languages is amply attested. The n ‘first person’ / m ‘second person’ is by no means unique to, diagnostic of, or ubiquitous in American Indian languages. Moreover, several explanations in addition to genetic inheritance have been proposed by various scholars, with some new possibilities also presented here. 

In short, then, the pronoun evidence has been misleadingly simplified and vastly overstated.

Footnotes

1. For example, Gilij already was aware of widely shared pronominal similarities: La semejanza de las partículas de persona antepuestas, y a veces también postpuestas a los nombres y a los verbos, es conocidísima de todos. Pero estas partículas en algunos lenguajes muy diferentes entre sí son algunas veces semejantes, otras veces son distintas, pero poco. De la primera clase son por lo común las de los maipures y de los mojos, como uno puede ver claramente poniéndolas en comparación entre sí: De la segunda clase son las de los mejicanos y de los maipures en las partículas de primera persona. Lo que puede verse en la palabra mejicana nócal [no-kal] casa, puesta en comparación con la maipure nupaná. Las de los chiquitos convienen con las tamanacas en los signos de segunda persona. Mata campo, amatări tu campo, son voces tamanacas. Poos casa, apo tu casa, son chiquitas. La lengua tamanaca conviene con la mejicana en los signos de tercera persona: ical su casa es voz mejicana, itéuti, del mismo significado, es tamanaca. (Gilij 1965[1782]:274).

The similarity of the person particles, preposed, and sometimes also postposed to the nouns and verbs is well known by all. But these particles in some languages which are very different from one another are sometimes similar, other times distinct, but only slightly so. In the first class are commonly those of the Maipures and the Mojos, as one can see clearly putting them in comparison with each other. In the second class are those of the Mexicans [Nahual] and the Maipures in the first-person particles. That can be seen in the Mexican word nócal [no-kal] ‘house’ placed in comparison with Maipure nupana. Those of the Chiquitos correspond to the Tamanacos in the second-person signs. Mata ‘field, country’ [and] amatări [a-mata-rí] ‘your field, country’ are Tamanaco forms. Poos ‘house’ [and] apo po ‘your house’ are Chiquito. The Tamanaco language corresponds with Mexican in the third-person signs: ical [i-kal] ‘his/her/its house’ is a Mexican form; itéuti, with the same meaning, is Tamanaco.

Edwards (1783) observed similarities between Mohegan and Hebrew first and second person pronouns reminiscent of forms Greenberg (1987) claims as strong
evidence for his Amerind hypothesis (Edwards 1823[1787]:18).

Gallatin (1836:161) reported:

The pronouns of the first and second person belong also in the Indian
languages to the class of primitive words ["that class, which has generally
been considered as so absolutely necessary in any state of society, that the
words of which it consists must have been in use everywhere in its earliest
stages, and could not have been borrowed by any nation from any other"].

By 1874, the widespread n of 'first person' in various American Indian languages
was well known, as we see in Sayce's (1874:216) matter-of-fact discussion of
'grammaticalization' (not his term):

[There] is the phenomenon which meets us in several of the North
American dialects, where the pronoun na, ni, or nu, "my," has become an
inseparable and meaningless affix of numberless words, just as in the
Continental "milord".

2. Note that Greenberg (1987:278) also presents a second-person singular pronoun akin
to Brinton's K forms in a wide variety of American languages (with forms ka, ikia, aki,
ka-, -ke-, -ga, etc. in Greenberg's list).

3. Fleming (1987:196) both asserts the strength of Greenberg's pronoun argument and
declares a misleading position regarding pronouns in linguistic change that is, alas,
harbored by a good number of linguists:

While pronouns are not the only good evidence in the world [of genetic
relationship], I agree with the Semiticists and Nostraticists that pronouns
really do not get borrowed very much, nor do they change easily. This is
an empirical matter for me, not a matter of faith in one kind of evidence.
When pronouns have changed, by replacement, or seriously disguised
phonetic change, as in the Chadic and Omotic sectors of AA [Afro-Asiatic],
then everything becomes more difficult. (Fleming 1987:196.)

Indeed, it is an empirical matter, and the facts are quite plain. As shown later in the
text of this paper, the borrowing of pronouns, while not frequent, is not at all
uncommon, and several examples are mentioned. Those who feel that pronouns "do not
change easily" ought to take heed of the well-known shifts, for example, involving second
person pronouns throughout Europe (as in a raw comparison of you/Se/usted/vou,
etc.), or the changes among forms for first person pronouns across languages of South
and East Asia in honorific language. The empirical reality is that pronouns can and do
change, and as discussed later in this paper, there are good reasons why pronouns often
share similarities.

4. This could explain such observations as, "Thus, word-initial nasal consonants such as
m- and n- often remain intact for millennia" (Ruhlen 1989a:2), and "the old n of proto-
Indo-European [was] retained in English practically intact" (Swadesh 1960:898).

5. Collinder had observed that -m for 'first person' is common in many different and
unrelated languages (Callaghan 1980:39).

I suspect that the perceptual salience of nasals and the importance in
communication of being able successfully to distinguish negative utterances from affirmative ones combine to help explain why negative markers in language around the world so typically have n or m, why ma and nV or something very similar are so frequent.

6. Similar facts have not gone unnoticed by other linguists. Callaghan (1991:53), reporting David Stampe’s observation, points out:
   the frequency of dentals and sonorants [including nasals] in the verbal paradigms of unrelated languages, possibly because of the ease with which these types of consonants enter into clusters with other consonants. Such a tendency would greatly increase the possibilities of chance resemblance. It is noteworthy that Greenberg places heavy emphasis on pronouns and pronominal paradigms in "Grammatical Evidence for Amerind" (1987, Chapter 5).
   (Cf. also Callaghan 1980:39.)

7. These findings are, of course, in sharp contrast to the assertion that "pronominal affixes are among the most stable elements in language: they are almost never borrowed" (Greenberg and Ruhlen 1992:97). Moreover, it is well established that certain aspects of pronominal systems are quite easily influenced by contact from other languages, e.g. the widespread diffusion of the inclusive/exclusive pronominal category in a number of languages of Western North America (Jacobsen 1980), and the shift from independent plural pronominal affixes to ones structurally composed of the singular pronominal marker plus a plural affix (cf. Robertson 1992). In particular, pronominal systems seem to be subject to analogical reformations, and many believe that they are also dominated by tendencies towards iconic symbolization, as other deictic markers are.

8. Miskito ba ‘third person, that (one)’ appears to be in origin a demonstrative pronoun. Southern Sumu (Ulwa) pronouns differ, and in general Miskito owes much to diffusion from and convergence with Northern Sumu. I owe these Miskito and Sumu observations to Kenneth Hale (personal communication).

9. It could also be mentioned that since Mednyj Aleut pronominal affixes fit "Amerind" forms, then by definition Russian pronominal affixes fit equally as well, since the Mednyj Aleut forms were taken over directly from Russian.

10. Kinkade does not rule out entirely a possible genetic relationship between Alsea and Salish, but notes that the conclusions in his investigation are negative (Kinkade 1978:5), and he adds further that at present it is not possible to answer the question, "how much of this Alsea pronominal system is originally Alsea and how much borrowed?" (Kinkade 1978:6).

11. Greenberg’s (1990:11) only response to the battery of negative evidence and alternative explanations which have been offered and which severely challenge his Amerind pronoun claims, presented at the Boulder conference and mostly represented in this section, was to single out of this one concerning child language acquisition (–a minor case, certainly not the most serious–) for ridicule:
In a remark at the Boulder Conference Campbell attributed such a preponderance of nasals [in the pronouns] to the phonetic nature of infant sucking reflexes!

Greenberg has repeated this misrepresentation in a variety of interviews and papers; he implies this is an unreasonable hypothesis without stating why. However, his 'infant sucking reflexes' is not an accurate account of what was said. What he alludes to appears in a list of explanations (some my own, many from other scholars) which have been offered to explain the putative n/m pronoun pattern -- this one comes originally from Ives Goddard (1986:202), where the matter of nursing is only a part of the story, where Greenberg's is a garbled caricature of it. The text he is citing actually reads (Campbell in press):

Another explanation [not my own in this case, rather originally from Ives Goddard] that has been offered is child language; child-language expressions around the world abound in self-directed and other-directed words containing nasal consonants ... [It] is a universal physical fact that a gesture equivalent to that used to articulate the sound n is the single most important voluntary muscular activity of a nursing infant. As Goddard (1986:202) points out, possibly this factor and the tendency for primary grammatical morphemes to consist of a single, unmarked (phonetically commonplace) segment account for the widespread appearance of n- in 'first-person' pronouns. Incidentally, in many societies, particularly among hunting and gathering groups, infants may continue to nurse until the age of five, sometimes longer, well into and beyond the age of language-acquisition.

Greenberg failed to mention the other more relevant and damaging facts, repeated in this paper.

12. Actually, Swadesh (1954:312) goes even further than Greenberg in that he derives the South American pattern from what he assumes to be the more general American n/m pattern:

In parts of Aztecan and Chibchan, and in Arawakan, second person m gives way to p or b, and it is at least possible that the bilabial stop may be somehow derived from the nasal. In much of South America, first person n is replaced by y; in certain areas one finds ñ. The palatalized b may be a transition form, which could have easily arisen as the result of a preceding front vowel. That is, ñina could give ñina, and the latter could have developed into Tiya. In fact, forms approximating all stages of this transition can be attested, thus Sahapin ñin, Esselen ñene, Yuman ña, Chontal ñiya. If we can thus derive first person y from n, then the n-m pronominal set extends from Chinook (naika 'I', maika 'thou') in northwestern United States to ona (ya 'I', ma 'thou') on the Straits of Magellan.

13. Some very clear evidence of the non-uniqueness of n 'first person' and m 'second person' in 'Amerind' comes from controversial sources, from the remote comparisons proposing very, very far-flung genetic relationships involving large segments of the
world's languages which exhibit these. Moreover, the m 'first person' and t 'second person' pattern, which Greenberg and Ruhlen assert to be diagnostic of their Eurasiatic grouping, is documented also in "Amerind" by several of these comparisons:

(1) Swadesh (1960:907-8) reconstructed *(te)ne, often *(t)ni 'I' for Proto Ancient American, noting such "interhemisphere linguistic connections" as, e.g. Malayan *-na(w) 'I', Polynesian *na-ku 'I', Basque nick 'I', Hebrew ni 'T', Somali-Galla 2ani 'I'. For 'thou, thy' he reconstructed *ma/mu and related it to such Old World forms as: Melanesian *mu, Malay-Bugis mu, and Dayak ma. Swadesh (1960:909) also presented a number of American Indian languages to illustrate his *ta/tu 'thou/thy', together with Old World comparisons (e.g. IE *te(w) 'thou', Hebrew 7ataa, Uralic *-t, Avar dun, Kvarshi do, etc.)

(2) Shevoroshkin (1989:19) presents as cognates of his Proto/proto-language **ni 'I': Nostratic *ni 'I', Nivkh i 'I', Khoisan *ni 'I' (also *an 'I': cf. Dravidian *vān 'I', Austronesian *NV 'I', Indo-Pacific *n[i] 'I', Amerind *ni 'I'. These he relates to Nostratic *nAH 'we exclusive', Dene-(Sino-)Caucasian *nV 'we', Amerind *nAH 'we'. He presents similar evidence for a Western group of this Proto/proto-language (including Amerind, Austric, and Indo-Pacific) for *n[i] 'thou' (also *kV 'thou') (Shevoroshkin 1989:20). He also presents, however, evidence for Proto/proto-world **mi 'I', apparently *I and you', in which he includes "Amerind": ' cf. Nostratic *mi 'I', Dravidian *ma-/m- 'we' (< Nostratic *mā 'we inclusive'), Khoisan *mi 'I' (also *me 'my'), Dene-(Sino)-Caucasian *mV 'I', Amerind *m 'I' (preserved as archaism in different Amerind languages: see Greenberg 1987, Chapter 5) (Shevoroshkin 1989:19). Shevoroshkin also gives evidence of a widespread ** t 'you' (see Shevoroshkin 1989:4 for supporting forms).

(3) It is worth pointing out that while both the m and t versions for second person are widely attested in American Indian languages (and elsewhere), a widespread pattern with k has also been proposed. For example, early in American Indian linguistic studies Michelson (1914) reported it in in Wiyot k-, Yurok qe- 'thy', and in Molala k[t]. Greenberg (1987:277-8) cites for 'you' among his grammatical sets Kalina ka(-be), Auake kai(-kite), Proto-Ge *ka, Erkibatsa ikia, Bororo aki, Coroado ga, Allentiac and Millcayac ka, Xinca ka, etc. 'second person singular pronoun', and, under grammatical set 8, Quechua kam, Gennaken kem, Aymara huma, and Kahuapan kem, huma, k̃ma 'second-person singular independent pronoun'. One could easily associate these with Shevoroshkin's (1989:19) forms for Western Proto/proto-world *kV 'thou': Nostratic *k/gV 'thee', Dene-(Sino)-Caucasian *kV 'thou', Amerind *kV 'thou', Austronesian *kev/*keH 'thou', Indo-Pacific *kV 'thou'.

14. Dryer (personal communication) cautions that, while the overall numbers support his test hypotheses about nasals in pronouns, the nature of the sample precludes applying statistics to test for significance. He also is inclined to believe that the n/m pattern in American Indian languages may suggest genetic relationship, but points out that if Hokan and Munda can share n/m due to chance, then why can't Hokan and Penutian?

15. Also, cases should not go unnoticed which have n in both 'first' and 'second' person forms (e.g. Proto-Salish *n- 'first person possessive', *ʔen- 'second person possessive; Newman 1979a:211), or a case such as Chimane which has nasals in the pronouns for all three persons, singular and plural (Martín and Pérez Diez 1990:576). Greenberg's case is also not helped by Amerind languages which have neither n nor m in
16. Manaster Ramer (1992:2) points out that in this regard Tonkawa, with its saː- 'I' and naː- 'you' is more similar to putative Na-Dene (cf. Navajo shi/ni, Chipewyan si/nen) than putative Amerind.


18. Proto-Siouan had *w- and *r- respectively for these, reflected in Sioux as wa-/ma- ‘I’ and ya-/ni- ‘you’. Note how Lakota and Sioux are dead opposites of what Greenberg expects to find. (Goddard 1988.)

19. That is, Greenberg (1987:53) gave Creek une ‘I’, a misreading of the orthography for what should be ane, and "Apalachee" ani, a non-existing form erroneously copied from Creek sources. However, Muskogean independent pronouns attach prefixes to -ni/-no bases, cf. ani ‘I, hasno ‘you’, pilno ‘we’; it is a- that is T, not the n as supposed by Greenberg (Kimball 1993:448-9).

20. DeLancey (in press) argues that paradigmatically related matching pronominal forms support the hypothesis of a genetic relationship between Klamath and Sahaptian. He goes on to show similar patterns in some other languages from both North and South America. However, when compared among themselves, DeLancey’s cases illustrate basically a recurring pattern of n ‘first person’ and m ‘second person’, but no paradigmatic among distinct first person forms or differing second person markers. DeLancey argues that such a recurrent matching pattern in different American Indian language groups is unlikely to be due to diffusion or chance. True. However, as pointed out here, a number of other possible explanations still remain in addition to the possibility of genetic inheritance. Given the extreme frequency of first person pronouns with n in the world’s languages, as explained above, we cannot at present determine whether the frequent first person n in American languages has genetic status. When we set aside n, we are left with recurrent m alone (which is not as general in American languages as Greenberg claimed), and m by itself is not a paradigm (or a pattern). DeLancey notwithstanding, the possibility of diffusion also cannot be ruled out either (as argued earlier in this paper).

21. Even in American cases where Greenberg finds the highly touted n/m pronoun pattern, it is clear that its presence is not always due to inheritance from an ancestor which had it. That is, even in cases where the documentable history demonstrates that some currently existing ‘first-person’ n- or ‘second-person’ m- actually derives from some other sound as the result of recent change, such cases are, nevertheless, taken by Greenberg as positive evidence for his claims (see Campbell 1988 for examples, see Goddard and Campbell in press for additional discussion).
Bibliography


Edwards, Jonathan, Jr. 1788[1787]. Observations on the language of the Muhhekaneeew Indians; In which the extent of that language in North America is shewn; its genius is grammatically traced; some of its peculiarities, and some instances of analogy between that and the Hebrew are pointed out. (Communicated to the Connecticut Society of Arts and Sciences, and published at the request of the Society.) New Haven: Josiah Meigs. (Reprinted 1788, London: W. Justins, Shoemaker-Row, Blackfriars; reprinted with notes by John Pickering, 1823 in the Massachusetts Historical Society Collection, second series, 10:81-160, and Boston: Phelps and Farnham.)

Fleming, Harold C. 1897. Review article: Towards a definitive classification of the world’s languages (review of A guide to the world’s languages, by Merritt Ruhlen). Diachronica 4.159-223.


——. 1988. Pre-Cheyenne *y. In honor of Mary Haas: From the Haas festival conference on Native America linguistics, ed. by William Shipley, 345-60. Berlin:
Mouton de Gruyter.


____. 1979b. The Salish object forms. IJAL 45.299-308.


S- PREFIXATION
ON UPPER CHEHALIS (SALISH) IMPERFECTIVE PREDICATES

M. Dale Kinkade

University of British Columbia

A prefixed s- is common throughout Salish. Sometimes more than one s- prefix is identified for a language, and one of these is nearly always glossed 'nominalizer'. Upper Chehalis also has this prefix, although it is difficult to say whether there is more than one s-there, or if s- simply has a range of functions including nominalization. Whichever might be the better analysis, I will give primary consideration here to the use of a prefixed s- as an aspect marker. Since this is probably ultimately a derived usage, I will also make limited comments about some of the other functions of s-.

It has been claimed for Upper Chehalis that "the formal sign of a continuative aspect form is the prefix /s-/l, although it can be determined from many suffixes as well" (Kinkade 1964:33-34). This is true for all elicited sentences and phrases in this language, both in my field notes and in those collected by Boas in 1927. Several suffixes are, however, better taken as diagnostics for this aspect (which I now call 'imperfective', rather than 'continuative'), notably subject and passive suffixes. This is because it turns out that texts do not usually have a prefixed s- on imperfectives. This seemingly contradictory occurrence—imperfectives marked by s- in elicited material but not in texts—has, in fact, a systematic explanation, largely determined by discourse structure.

Boas does not directly address the issue; he notes that "all verbs have two forms, completive or momentary, and continuative" (1934:105, note 12), but he says nothing about the s- prefix in this context (he otherwise identifies it only as a nominalizer). It does indeed follow the patterns discussed below in sometimes being present and sometimes not in the text fragment he presents. However, the various paradigms he presents all show continuative (i.e. imperfective) forms beginning with s-.

Examples of imperfective predicates with s- from elicited material are given in (1).1

---
1 Abbreviations used are: AUT 'autonomous', CAUS 'causative', COP 'copula', DEF 'definite', DESCR 'descriptive', DETR 'detransitive', DIMIN 'diminutive', EXT 'extender', (f) 'feminine gender', FUT 'future', HAB 'habitat', IMPF 'imperfective', INDEF 'indefinite', INDIR 'indirective', INTR 'intransitive', MDL 'middle voice', MOD 'modal', OBJ 'object', OBL 'oblique', PASS 'passive', PERF 'perfective', PL,pl 'plural', POSS 'possessive', Q 'question', QUOT 'quotative', REFL 'reflexive', sg 'singular', SUBJ 'subject', TRANS 'transitive', UNR 'unrealized, future'.

21
(1) *šáw’ a-w-n čsa tit manó·mš.*
  [s-play-INTR-3SUBJ(IMPF) again DEF children]
The children are playing again.

*pit šáw’uy’i ō*

*s-máx”a-t-n tit x’miyità-sqš.*
  [s-drive-TRANS-3SUBJ(IMPF) DEF automobile]
He’s driving the car.

*pit máx”-n ō*

*s-t’ayóc’-tu-stš pət tit táwn.*
  [s-parade-CAUS-PASS(IMPF) in DEF town]
There is a parade in town.

*pit t’ayóc’-t-m ō*

*s-šá”x-ən-mit-anš + s-šá”xaláš.*
  [s-hunt-MDL-1SGSUBJ(IMPF) OBL deer]
I’m hunting deer.

The final suffix in each of these predicates has an imperfective form different from the perfective aspect equivalent (i.e. ō for 3SUBJ,-tm for PASS, and čn for 1SGSUBJ); these endings would themselves be sufficient to mark which aspect is represented (and in turn affect the specific shape of what precedes them). The perfective forms of each of these predicates is given to the right of the example. The s- is thus redundant for marking aspect.

It was well after I had written my initial description of Upper Chehalis, in which I claimed that s- marked imperfectives, that I realized this prefix is most often missing from imperfectives in texts. Although I have been aware of this discrepancy for some time, I had never bothered to investigate it further until I recently read the draft of a paper by Paul Kroober, in which he also noted, and was mystified by, the absence of s- in texts. A sample text fragment is given in (2), where there is a series of imperfectives, not one of which has s- prefixed to the main predicate of the line, which in each case is unmistakably imperfective.²

(2) *šáwas tawé·la-t-n.*
  [first sit-AUT-3SUBJ(IMPF)]
First he sits down.

---
² The first line is somewhat ambiguous in this regard, since sequences of two or more s’s tend to collapse into one in allegro speech, and an s- prefixed onto the form meaning ‘sit’ might either be omitted or missed by the transcriber.
qʷaʔqʷiƛʷ-t-n tóʔ-a-t-tí tit s-mákʷm-umš.
[cut-trans-3subj(imppf) short-pl-intr-pl def s-prairie-people]
He cuts the grass up short.

yél-t-n s-át lámé·tn.
[thread-trans-3subj(imppf) to-in string]
He threads it on a string.

tám=éʔus-tw-n.
[tie-middle-caus-3subj(imppf)]
He ties them together.

ʔáqa n yánq=anus-n.
[now and put.around.neck=ʔ-3subj(imppf)]
He puts it around his neck.

This s- prefix may be replaced by a proclitic t ‘unrealized, future’ (which can also occur with perfective forms); t and s- do not co-occur. This means that constructions with this t do not provide information on the actual occurrence of s-. It also needs to be noted that s- is used regularly to mark subordinate predicates (which may be either participial-like, gerundial, or possessed), and this usage is consistent both in elicited material and in texts. Examples of such subordinated material are given as (3) through (7).

(3) ʔtu tó·ʔ-stw-ìtti t s-ʔtlan-n.
[then hear-caus-3plsubj(imppf) indef s-sing-3subj(imppf)]
Then they hear singing.

(4) xáwas ʔiʔkʷ-ml-n t qaʔ tuʔat t s-cétx-mit-n tuʔat t s-máníča.
at.first fetch-detr-3subj(imppf) obl indef water from at indef s-dribble-mdl-3subj(imppf)
from at indef s-mountain]
First he fetches water from a spring dripping from a mountain.

(5) mé·tla t s-kʷáw=əq-s tit pósa.
[not indef s-join=voice-3poss def monster]
The monster doesn’t answer a thing.

(6) X'áł-stw-n t čá· X'a s-q'al-óm-s.
[look.for-caus-3subj(imppf) indef where fut s-camp-mdl-3poss]  
He looks around to where he will camp.

(7) ?at t s-nápm-itn-s, n m'ús-mit-ìtti.
[when indef s-finish-eat-3poss and sleep-mdl-3plsubj(imppf)]
When they have finished eating, and they sleep.
In each of these sentences, s- nominalizes what follows; the nominalization is further indicated by the use of an article (τ 'indefinite' in 3, 4, 5, 7, and x'a 'future' in 6), and -s 'third person possessive' in (5) through (7). The first two examples are very much like English gerunds (i.e. nominal and based on imperfective forms); the other three require possessive inflection to indicate the subject of the subordinated form. In (3) the gerund is direct object, in (4) it is the object of a (compound) preposition. In (5) the initial negative requires the predicate following to be subordinate; in (6) a question-word does the same. In (7) the subordinate predicate occurs in a clause beginning with a subordinating conjunction.

In order to try to determine what, if any, pattern there is to the presence or absence of s-, I extracted 15 pages (containing about 540 clauses) of a long text, then marked all the imperfective predicates. This identified roughly 275 examples—imperfective being the predicate of choice in narration; only 49 had a prefixed s-. These results were surprising enough, given the uniformity of the presence of s- in elicited material, although more startling was the distribution of forms with and without s-. It turned out that directly quoted speech consistently used s- on imperfectives, while it was consistently absent from the rest of the narrative text. There are a very few exceptions both ways which I have not yet figured out, although most of the s-prefixes forms in the non-quoted narrative text turned out to be subordinate clauses, where their occurrence is regular. Some of the forms in quoted speech are indeterminate as to the use of s- because either the stem itself begins with s- or there is a particle or word preceding that ends in s, and, as noted earlier, sequences of more than one s tend to collapse into one segment. Other imperfective forms in quoted speech are marked with τ 'unrealized, future', and thus cannot include s-. Thus (8) through (10), which include quoted speech, can be contrasted with (2) above. Unlike the usage without s- in the narrative text, quoted speech does not contain strings of imperfectives. Rather, one finds them interspersed with various subordinate constructions, and with the quoted speech moving along in short sequences interspersed with narrative text.

(8)  cút-nax-n t x"enéx"wane,
    [say-def-3subj(impf) indef x"enéx"wane]
    x"enéx"wane says,
    "táta ta s-k"aná-t-s,
    [when past s-get/take-trans-3poss]
    when he gets it,
    "ró- nax-wi-q"ulati tit ?a-s-?üm-c.
    [oh true-? def 2sgposs-s-give.food-1sgobj(perf)]
    "Oh, thank you for feeding me.
cilačs 1-s-qáx-čt;
[five indef s-many-1plposs]
We are five in number;
múš n-s-nésči-im.
[four 1sgposs-PL-younger.brother-PL]
I have four younger brothers.
χʷágʷ u čt q’ic’-t-áliwan-x.
[all yet 1plsubj(perf) thus-?=appearance-def]
All of us look alike.
ʔám u ʔa[l]aṭ ʔůqʷ-n
[when yet when[2sgposs] find-3obj(perf)]
So when you find him
čá. ʔaṭ tii s-mánichi.
[where on def s-mountain]
anywhere on the mountain.
ʔám u q’ic’-t-áliwan-x uʔ t’ónca,
[when yet thus-?=appearance-def yet to I]
If he should look like me,
wi táx wi tóč’s tu ʔaṭ n-s-nésči-im.
[and that one cop one from in 1sgposs-PL-younger.brother-PL]
and that is one of my younger brothers.
wi ʔáqa s-tawá-mi-n-anš.
[and now s-leave-2sgobj(impf)-n-1sgsubj(impf)]
And now I am leaving you.
tan s-wákʷs-anš ʔaṭ 1-s-xá’x.
[now s-go-1sgsubj(impf) to-in indef s-bush]
Now I am going into the bush.
tu šán’-x
[from there-def]
From there
n t’l’a-yá’xʷ-á-w-anš."
[and unr again-go.home-intr-1sgsubj(impf)]
and I will go back home."}

cún-t-nax-n cic póśaʔ,
[say-?=def-3subj(impf) def(f) monster]
The monster says,
"hýy č.
[goodbye 2sgsubj(perf)]
"Goodbye.

?áqa s-wák"w-s-anš.
[now s-go-1sgsubj(impf)]
Now I am going."

wá' k"w-s-n t xʷənēxʷone s-at t s-xá̱l.
[go-3subj(impf) indef Xʷənēxʷone to-in indef s-bush]
Xʷənēxʷone go·es into the bush.

(9) kʷáw-aq-n c man-s c malé,
[join-voice-3subj(perf) indef(f) child-3poss indef(f) Malé]
The daughter of Malé answers,

"x-ławá-mi-n-anš,
[s-leave-2sgobi(impf)-n-1sgsubj(impf)]
"I will leave you,

?áqa s-?íkk"a-t-anš t s-šam'=álax".
[now s-fetch-trans-1sgsubj(impf) indef s-?=people]
now I will fetch the people."

k"w-á-w-n s-at x"áq u t pé:p-s-ayu.
[get.to-intr-3subj(impf) to-in all yet indef ?=animal]
She gets to all the animals/birds.

yáw-š-ni-t-n
[tell-indir-indir-trans-3subj(impf)]
She tells them

?it ?íkʷtaqí-t-m t ta máñ-s.
[perf steal-trans-pass by past child-3poss]
that her child was stolen. ["My child was stolen."]

(10) cáw-nax-n t ūk"w-at s?-at tit s-šam'=álax",
[say-def-3subj(impf) indef Moon to-in def s-?=people]
Moon says to the people,

"?ó· s-?ini-n-ap s-at tit cá[·]pš."
[oh s-do-n-2plsubj(impf) in def stream[dimin]]
"Oh, what are you doing in the stream?"

"s-yúś-tawt s-at t s-q"áq".
[s-work-1plsubj(impf) to-in indef Raven]
"We work for Raven.
s-x'a?=úl-itn-stawt."
[s-look.for-EXT-food/fish-1pSUBJ(IMPF)]
We are salmon fishing."
cút-nax-n,
[say-DEF-3SUBJ(IMPF)]
He says,
"t čús na nk"e wínwín-nax-ap."
[indefalwaysQHABdo-DEF-2pSUBJ(IMPF)]
"Do you always do that?"

These passages, then, show how direct quotations may be identified by the use of a prefixed s-

Other s- prefixes occur where expected, but what is striking is the contrast between quoted speech and narrative text. We are familiar with the common requirement in various European languages to flag indirect speech, often by switching verbal mood to subjunctive or by using modal auxiliaries, as well as shifting pronominal referents. In Upper Chehalis, it is direct speech that is flagged, and by using the usual signal for subordinate predicates. (Mode may be marked, although not in any way that could be said to distinguish direct from indirect speech.) It does not seem necessary, however, to claim that quoted speech is in fact subordinate. It is only imperfective forms that are so marked; perfective forms show no difference whatever from the regular narrative text usage, and it seems unlikely that one aspect would be categorized as subordinate in direct speech while others would not. The use of s- to mark these imperfectives is, however, most likely derived from its use as a sign of subordination.

To clarify the difference between direct and indirect speech, it is necessary to turn now to examples of the latter. It is possible to identify some indirect speech in Upper Chehalis, although it is used much less frequently in the texts available to me than is direct speech. Traditional stories are far more likely to be dramatized by quoting the actual utterances of characters than to refer to their speech indirectly. Examples of indirect speech, however, show nothing out of the ordinary. I have identified only ten or twelve instances of indirect speech in the entire Adventures of Xwənəxwəne text, which is over 4000 lines (or clauses) long, and replete with directly quoted speech. Examples are given in (11) through (16); those in (11) and (12) are instances of indirect speech from the narrative portion of the story, those in (13) through (16) are instances within direct quotations. For each example of indirect speech I have added (in English) what the equivalent would be as direct speech; this shows in particular the pronominal displacements.
(11)  $k^{"enk"}^{"en}^{-\cdot}nus-mit-n,$
      [pay.attention=inside-MDL-3SUBJ(IMPF)]
He wonders,
$tam-anin$
[what-now]
What is it
$q\acute{a}t\ s\acute{a}\-t\-\an\ c\acute{\i}tm.$
[MOD make-3OBJ(PERF) food]
how he can make food?  ['"How can I make food?"

(12)  $k^{"enk"}^{"en}^{-\cdot}nus-mi\-n$
      [pay.attention=inside-MDL-3SUBJ(IMPF)]
He wonders
$?e\-nm\ t\ q\acute{a}t\ s-q\acute{a}t-s\-\in\-s\-n\-s$
[how INDEF MOD s-MOD-s-do-?-3OBJ(PERF)-3POSS]
how he can do it  ['"How can I do it?"
   $cu\ q\acute{a}t\ k^{"ena\-x\"}\ t\ t\ c\acute{\acute{a}}\uw.$
[so.that MOD get/take-3OBJ def spring.salmon]
so that he can get the spring salmon.  ['"(How) can I get the spring salmon?"

There are only a few examples in the narrative portion like those in (11) and (12) (only four have been noted), and three of them are introduced with 'he wonders'. These three also have the modal particle $q\acute{a}t$ after an interrogative word, and in all three the main predicate of the indirect speech is a perfective transitive form with third person arguments. The example in (12) has a subordinated form of 'do' with a third person possessive suffix as well. A third instance is in the last two lines of (9) above; the indirect speech there is an ordinary perfective passive form.

Examples (13) through (16) illustrate indirect speech within quoted speech. In (13) a primary second person becomes first person in indirect speech. In (14) first and second person arguments reverse their roles for subject and object. In (15) a first person possessor is referred to in indirect speech as third person. In (16) a third party is cited as referring to the person addressed, resulting in a shift from third to second person.
(13a) "ʔó· sʔínwat-n t p’ayók"—
[oh s-say.what-3SUBJ(IMPF) INDEF Bluejay]
"Oh, what is Bluejay saying—

n-k’uwy ači."
[1sgpess-mother quot]
my mother, he says." ["Your mother."]

(13b) "wi ?it cúń-c
[and PERF say/tell-1sgobi(PERF)]
"And he told me

ʔ’áq" ači s-yəc’á-w-anš.
[better! quot s-turn.back-INTR-1sgobi(IMPF)]
I’d better turn around and go back, he says.

ʔónca."
[I]
I (should)." ["You’d better turn around & go back."]

(14a) "q’at c’áp’-nt tit ?a-s-cún-c-x
[MOD disagreeable-DESCR DEF 2sgposs-s-tell-1sgobi(PERF)-DEF]
"What you told me was disagreeable

q’at ʔ’ómx-c č tač tit ?a-ʔ’éʃ ʔ’xš’š."
[MOD stab-1sgobi(PERF) 2sgobi(PERF) with DEF 2sgposs-stick[dimin]]
that you would stab me with your little stick." ["I will stab you."]

(14b) "sʔínwat-š.
[s-say.what-2sgobi(IMPF)]
"What are you saying?

ʔí-cút-x č na
[ʔí-say-DEF 2sgobi(PERF) Q]
Didn’t you say

ʔ t yucá-mš č."
[UNR INDEF kill-1sgobi(PERF) 2sgobi(PERF)]
you will kill me?" ["I will kill you."]

(15) "wi ?it cúť
[and PERF say]
"And he said
None of these shows any notably unusual morphology or syntax in the indirect speech. The instance in (13a) is merely a possessed form. The predicates in (14) through (16) are all ordinary perfectives; both (14a) and (14b) are simple transitives, (15) is reflexive, and (16) is intransitive with a quasi-auxiliary (x'áq"’). The indirect speech in (13b) looks superficially as if it is an imperfective predicate with a prefixed s- that might be explained as simply being the result of its occurrence within quoted speech. However, it is more likely that this s- is a subordinate marker; x'áq"’ is a quasi-auxiliary that can be followed by either perfective or imperfective predicates, although imperfectives are the more common. In all such cases, this imperfective has the s- prefix, suggesting that it is not the quoted speech that requires it.

Returning to the earlier point that within texts, imperfectives in directly quoted speech have an s- prefix, and those outside quoted speech do not, we have a simple explanation for why it is, as was observed at the outset, that all separately elicited imperfectives have s-: they are quoted speech. The usual way of eliciting is to ask "How do you say ‘XYZ’?" That which is asked for is given as a quotation. The response, with or without an introductory "you would say" is likewise quoted speech, hence may require a prefixed s-.

REFERENCES

Boas, Franz. [1927] [Chehalis Field Notes.] (Manuscript No. [30(S2c.1)] [Freeman No. 589] in American Philosophical Society Library, Philadelphia.)


Comparative Difficulties of the “Gulf” Languages

Geoffrey Kimball

Tulane University

0. Introduction. In the late 1940s Mary R. Haas postulated that there was a “Gulf” language stock in the Southeastern United States. She stated that the stock was composed of the Muskogean family of languages (Alabama, Apalachee, Choctaw, Chickasaw, Creek, Hitchiti, Koasati and Mikasuki) and the language isolates Atakapa, Chitimacha, Natchez, and Tunica. Haas suggested that the Natchez language and the Muskogean family were more closely related to each other than to the other three isolates (1956), and accepted Swadesh’s association of Atakapa and Chitimacha (1946; 1947). However, after a burst of early work, Haas set this theory aside, and in recent years has expressed skepticism about the unity of the “Gulf” stock. Research on “Gulf” languished for many years until revived by Greenberg (1987) in his Language in the Americas, and recently, inspired by Greenberg, Pamela Munro (1992) has studied it.

1. Grouping of the Languages. The geographical relations of the Gulf languages have never been adequately discussed, since contact or lack thereof can affect whether two languages can be perceived as related. The spatial relationships of the languages can be diagrammed as follows:

Note that there are three Atakapan languages (Akokisa, Western Atakapa, and Eastern Atakapa), a fact that is not well-recognized. The only languages without geographic contact are the Atakapan languages and Tunica, and the Atakapan languages and the Muskogean family. By proximity, one would expect to have more similarities between Eastern Atakapa, Chitimacha, and Natchez, Natchez and Muskogean, Tunica and Muskogean, and Chitimacha and Muskogean. If one accepts the proposed relationships between Atakapan and Chitimacha and between Natchez and Proto-Muskogean, one gets the following tree:
"Gulf"

The question arises, are these three separate branches, or are two of the branches closer to each other than to the third?

2. Comparative Morphology. The grammatical similarities between the five groups are those which have suggested that the languages are ultimately related, and are much more impressive than the lexical evidence. However, see Appendix II, which shows that phonological similarities among the grammatical morphemes are almost absent.

Active Verbs vs. Stative Verbs
(inflected by patient affixes)
- All languages
- Constituent order SOV
- All languages
Postpositions
- All languages
Locative cases
- All languages
Stative Verbs inflected by patient affixes and stative verbs inflected by dative affixes.
- Natchez
- Muskogean
Nominal Cases (Ergative-Absolutive/Nominative-Accusative)
- Natchez (ergative)
- Chitimacha (ergative)
- Muskogean (nominative)

Independent inflected verbs vs. auxiliary inflected verbs
- Tunica
- Muskogean
- Natchez
- Chitimacha
Reference tracking
- Switch-reference (-t/-k/-n) and focus (-o-) Muskogean
- Reference tracking (-k) and focus (-o-k) Natchez
- Reference tracking (-man) Tunica
- Focus (-\$) Chitimacha
- Focus (-\$) Atakapa
Possessive prefixes (alienable and inalienable possession)
- Tunica
- Muskogean
Possessive suffixes (no distinction)
- Natchez
Independent possessive words (no distinction)
- Chitimacha
- Atakapa

These selected features reinforce the tripartite division of the family, with Tunica appearing to lean ever so slightly more closely to the Natchez-Proto-Muskogean branch. However, since it is possible that many of these features are areal in nature, rather than genetic, they may just mark the Southeast as a linguistic area, rather than indicate anything about possible genetic relationships.

3. Comparative phonology. This is the point at which grave difficulties arise. Part of the problem is that the historical phonologies of the isolates are essentially unknown and unstudied, so that conditioning of reflexes is unclear. But even allowing some slack to account for this problem, still, repeated and regular correspondences, such as occur among
the Muskogean languages and the Atakapan languages, and which are illustrated below, are almost impossible to find.

REGULAR CORRESPONDENCES

<table>
<thead>
<tr>
<th>Muskogean</th>
<th>*coki ‘pumpkin’</th>
<th>*θaθixo ‘fish’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creek</td>
<td>časi</td>
<td>ɁaɁo</td>
</tr>
<tr>
<td>Mikasuki</td>
<td>čoks-i</td>
<td>ɁaɁ:i-i</td>
</tr>
<tr>
<td>Alabama</td>
<td>čoksi</td>
<td>ɁaɁo</td>
</tr>
<tr>
<td>Koasati</td>
<td>čoksi</td>
<td>ɁaɁo</td>
</tr>
<tr>
<td>Choctaw</td>
<td>šokši</td>
<td>nani</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Atakapan</th>
<th>*wọj ‘eye’</th>
<th>*yil ‘light; day’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akokisa</td>
<td>oɁ</td>
<td>iy</td>
</tr>
<tr>
<td>Western</td>
<td>wọ:1</td>
<td>yil</td>
</tr>
<tr>
<td>Eastern</td>
<td>wil</td>
<td>iy</td>
</tr>
</tbody>
</table>

Compare the examples above, with the following two Tunica-Atakapan correspondences:

- ‘black’ ‘to call’
  - Tunica mélő wáli
  - Atakapan *méj wan (W)

Taken separately, these two examples look plausible; however, together they illustrate one of the great problems of comparative “Gulf”: multiple correspondences. Here Tunica l can be compared with Atakapan l and Atakapan n in the same sort of root shape. It is highly unlikely that there are conditioning factors that could explain this variation, so that the only explanations available are that 1) one of the correspondences is correct, and the other is chance, 2) one or both are loans, or 3) both correspondences are by chance. Because of the lack of geographical contact between these two languages, 2) can be ruled out. However, for adjoining languages, loans have to be strongly considered. Loanwords are a likely possibility when two words in adjoining languages are almost identical phonologically, as in the following.

- ‘hundred’ ‘two’ ‘vulture’ ‘to hear’
  - Natchez pu·p -ahp ?o-ši ?e·p-le-hal?išt
  - Chitimacha pu·p ?upa ?o-š wopi-

The absolute identity of the terms for ‘hundred’ in Chitimacha and Natchez, in comparison with the more likely historical relationship between the obsolete Natchez bound form for ‘two’ and the Chitimacha form, indicates that it was borrowed from one language into the other. The same is true of the term ‘vulture,’ which begins with a glottal stop in both languages, as compared to ‘to hear’ which shows a more frequent correspondence of Natchez initial glottal stop to Chitimacha initial w.

A number of comparative sets consist of bird names, and with these borrowing and onomatopoeia must first be discounted. A good example is ‘robin,’ with forms attested in
Tunica, Proto-Muskogean, and Natchez:

Tunica
PM
Natchez
wišk?ohku
*ciskoko/*k'iskoko
miškoko

These look plausible, until the fact that there is a resembling form in the neighboring Siouan language Quapaw: šiikkoköke (Rankin, 1986). Here widespread borrowing seems to be the likely source of these words. Onomatopoeia, on the other hand, is the likely origin for the terms ‘pileated woodpecker,’ and ‘red-headed woodpecker’ which are found in the same three contiguous languages:

Tunica
PM
Natchez
pâhpahka-na
*kwahkwa-ka
pakpakä-ši-L
čuhčuhi-na
*čaxča-ka
cawcah

In this case the pileated woodpecker is ‘the one who goes ‘pakpak’ and the red-headed woodpecker ‘the one who goes ‘čahčah.”

4. Comparative Orphans. Another typical problem is the comparative orphan, that is, a comparison that looks excellent, but is without any further examples. The following is a typical example:

Tunica
Chitimacha
to die
lúpí
nu·p-

Since Chitimacha has no l phoneme, it is natural that it would be replaced by n, and the vowel correspondence illustrates a postulated feature about Tunica historical phonology, that intervocalic stops that were preceded by a historical long vowel were not preaspirated. However, no other example of a Tunica/Chitimacha l/n correspondence can be found.

Conclusion. As one can see from the appended vocabulary comparisons, good apparent cognate sets are not common, and when one has to apply the possibilities of borrowing, onomatopoeia and chance, the number of sets shrinks further. This is the real frustration of comparative “Gulf:” there is just enough to suggest the languages are related, but there is not enough to provide clear and unequivocal proof. To put this problem into a form that can be more widely appreciated, what would the difficulties have been in working on Proto-Indo-European, if the only languages available to reconstruct it had been the Romance languages, Wend and Sorb, Albanian, Armenian, and Manx?

Footnote
¹Tunica material is from Haas (1940; 1953), Natchez is from Haas (1934-36), Chitimacha is from Swadesh (nd.), Atakapa has been rephonicized from Gatschet and Swanton (1932), and all Proto-Muskogean reconstructions are my own.
Bibliography

Gatschet Albert S. and John R. Swanton

Greenberg, Joseph H.

Haas, Mary R.
1934-1936 Natchez fieldnotes. Mss.

Munro, Pamela

Rankin, Robert L.

Swadesh, Morris
Appendix I: Lexical Comparisons

TWO LANGUAGE COMPARISONS

Tunica-Muskogean

<table>
<thead>
<tr>
<th>Tunica</th>
<th>Gloss</th>
<th>PM</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>-áhaya</td>
<td>sibling of opp sex</td>
<td>*xaya</td>
<td>sister-in-law</td>
</tr>
<tr>
<td>-wána</td>
<td>to want</td>
<td><em>k</em>anna</td>
<td>to want</td>
</tr>
<tr>
<td>wátoru-hki</td>
<td>big gray heron</td>
<td>*watola</td>
<td>whooping crane</td>
</tr>
<tr>
<td>?élů</td>
<td>fruit; to bear fruit</td>
<td>*aȟi</td>
<td>fruit, bear fruit</td>
</tr>
<tr>
<td>?țšą</td>
<td>willow</td>
<td>osí (Koasati)</td>
<td>willow</td>
</tr>
</tbody>
</table>

Tunica-Natchez

<table>
<thead>
<tr>
<th>Tunica</th>
<th>Natchez</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>čáhka</td>
<td>cak-</td>
<td>to stick in</td>
</tr>
<tr>
<td>čólu</td>
<td>col-</td>
<td>to drip</td>
</tr>
<tr>
<td>háhka</td>
<td>haku</td>
<td>corn</td>
</tr>
<tr>
<td>kára</td>
<td>kolkol-</td>
<td>to gulp</td>
</tr>
<tr>
<td>káhpun</td>
<td>hi-kap-</td>
<td>to put in the mouth</td>
</tr>
<tr>
<td>kápaši</td>
<td>kapa·ht(į)</td>
<td>chicken</td>
</tr>
<tr>
<td>lálahki</td>
<td>lá·lak</td>
<td>wild goose</td>
</tr>
<tr>
<td>pihču</td>
<td>pic-</td>
<td>to fart</td>
</tr>
<tr>
<td>réma</td>
<td>leM-</td>
<td>to shine</td>
</tr>
<tr>
<td>róhku</td>
<td>lo·k-</td>
<td>to snore</td>
</tr>
<tr>
<td>rswasi</td>
<td>wa·</td>
<td>sassafras</td>
</tr>
<tr>
<td>šímu</td>
<td>ši·M-</td>
<td>to blow the nose</td>
</tr>
<tr>
<td>wiwu</td>
<td>*wi·W-</td>
<td>to blow (of wind)</td>
</tr>
<tr>
<td>?šška-</td>
<td>?ac</td>
<td>wild potato</td>
</tr>
<tr>
<td>?úhu</td>
<td>?oho·-</td>
<td>to cough</td>
</tr>
<tr>
<td>-nahku</td>
<td>-neke</td>
<td>like, resembling</td>
</tr>
<tr>
<td>-yóni</td>
<td>?u·nuh</td>
<td>intestines</td>
</tr>
</tbody>
</table>

Tunica-Chitimacha

<table>
<thead>
<tr>
<th>Tunica</th>
<th>Chitimacha</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ší</td>
<td>?asi/?ayš</td>
<td>man/men</td>
</tr>
<tr>
<td>-ála-wéča</td>
<td>wa?aš</td>
<td>ear</td>
</tr>
<tr>
<td>čárina</td>
<td>čana</td>
<td>kingfisher</td>
</tr>
<tr>
<td>háhku</td>
<td>?ak-šuš</td>
<td>cypress</td>
</tr>
<tr>
<td>lúpi</td>
<td>nu·p-</td>
<td>to die</td>
</tr>
<tr>
<td>náka</td>
<td>nakš</td>
<td>war</td>
</tr>
</tbody>
</table>

Tunica-Atakapa

<table>
<thead>
<tr>
<th>Tunica</th>
<th>Atakapa</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>méli</td>
<td>me:l</td>
<td>black</td>
</tr>
<tr>
<td>wáli</td>
<td>wan</td>
<td>to call</td>
</tr>
</tbody>
</table>
Natchez-Muskogean

<table>
<thead>
<tr>
<th>Natchez</th>
<th>Muskogejan</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ci-p-hakiʔiš</td>
<td>*čoʔpa</td>
<td>to buy</td>
</tr>
<tr>
<td>col</td>
<td>*čolyi</td>
<td>pine tree</td>
</tr>
<tr>
<td>co·Y</td>
<td>*tahayo</td>
<td>squash</td>
</tr>
<tr>
<td>kaWkp</td>
<td>*kaxʷ-ka</td>
<td>fox/bark or yelp like a fox</td>
</tr>
<tr>
<td>lahanaW</td>
<td>*xana-li</td>
<td>six</td>
</tr>
<tr>
<td>ša·š(i)</td>
<td>*θaθi-xo</td>
<td>perch/fish</td>
</tr>
<tr>
<td>?a·L</td>
<td>*xalki</td>
<td>wife</td>
</tr>
<tr>
<td>?eNt</td>
<td>*innoti</td>
<td>tooth</td>
</tr>
<tr>
<td>?o·ko</td>
<td>*poko-li</td>
<td>ten</td>
</tr>
<tr>
<td>-a·nah</td>
<td>(K.)-nána</td>
<td>nothing but</td>
</tr>
<tr>
<td>-c</td>
<td>-*t</td>
<td>ergative/nominative</td>
</tr>
<tr>
<td>-n</td>
<td>-*n</td>
<td>absolutive/accusative</td>
</tr>
</tbody>
</table>

Natchez-Chitimacha

<table>
<thead>
<tr>
<th>Natchez</th>
<th>Chitimacha</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>kuN</td>
<td>ku·</td>
<td>water/liquid</td>
</tr>
<tr>
<td>pu·p</td>
<td>pu·p</td>
<td>hundred</td>
</tr>
<tr>
<td>wasṭa·N</td>
<td>waštik</td>
<td>cow</td>
</tr>
<tr>
<td>weykoL</td>
<td>way'</td>
<td>spiderweb</td>
</tr>
<tr>
<td>?e·p-le-halʔiš</td>
<td>wopi-</td>
<td>to hear</td>
</tr>
<tr>
<td>?o·ši</td>
<td>?o·š</td>
<td>vulture</td>
</tr>
<tr>
<td>?o·k-ahp</td>
<td>?upa</td>
<td>twenty / two</td>
</tr>
</tbody>
</table>

Natchez-Atakapa

<table>
<thead>
<tr>
<th>Natchez</th>
<th>Atakapa</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>?inu</td>
<td>eːŋ</td>
<td>name</td>
</tr>
<tr>
<td>?o·</td>
<td>oːl</td>
<td>persimmon</td>
</tr>
</tbody>
</table>

Chitimacha-Atakapa

<table>
<thead>
<tr>
<th>Chitimacha</th>
<th>Atakapa</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>heyči</td>
<td>hišinj</td>
<td>ten</td>
</tr>
<tr>
<td>kesi</td>
<td>keːc(k)</td>
<td>liver</td>
</tr>
<tr>
<td>kiča</td>
<td>kiš</td>
<td>woman</td>
</tr>
<tr>
<td>kupu</td>
<td>kipaco</td>
<td>gourd</td>
</tr>
<tr>
<td>ney</td>
<td>neː</td>
<td>earth</td>
</tr>
<tr>
<td>si·c</td>
<td>šiːt</td>
<td>Spanish moss</td>
</tr>
<tr>
<td>sisč'up</td>
<td>ses</td>
<td>mulberry</td>
</tr>
<tr>
<td>?a·ci</td>
<td>oc</td>
<td>beaver</td>
</tr>
</tbody>
</table>

Atakapa-Muskogean

<table>
<thead>
<tr>
<th>Atakapa</th>
<th>PM</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>iwal</td>
<td>xʷolo</td>
<td>shell</td>
</tr>
</tbody>
</table>
THREE LANGUAGE COMPARISONS

Tunica-Muskogean-Natchez

Tunica  Muskogean  Natchez
čuhčuhina  *čaxča:ka  cawcah
páhpahkana  *kʷahkʷa-ka  pakkakú-ši-L
wišk?ohku  *č/kʷiskoko  miškokʷ
?úču  *piči  suckle; breasts  šu-

redheaded woodpecker  pileated woodpecker  robin [Quapaw šiikkokóke]  breast

Tunica-Chitimacha-Natchez

Tunica  Chitimacha  Natchez  hackberry
kó-  kamu  konŋ

Tunica-Natchez-Atakapa

Tunica  Natchez  Atakapa
-ʔaha  -ha·t  -hah  negative

Tunica-Chitimacha-Atakapa

Tunica  Chitimacha  Atakapa  howi  hi  wind
húri

Natchez-Chitimacha-Atakapa

Natchez  Chitimacha  Atakapa
?aweh  waʔa / wa-  wahš  uncle
?i-š  waši  wo:š / wi:š  hand

Natchez-Chitimacha-Muskogean

Natchez  Chitimacha  Muskogean
ha-ku-ši-ʔiš  ?a·-  *im-aka  to give

FOUR LANGUAGE COMPARISONS

Tunica-Natchez-Chitimacha-Atakapa

Tunica  Natchez  Chitimacha  Atakapa  skunk
šíki  šic  kišʔeʔe  šikiši
## Appendix II: Morphological Comparisons

<table>
<thead>
<tr>
<th>English</th>
<th>PM</th>
<th>Tunica</th>
<th>Natchez</th>
<th>Chitimacha</th>
<th>Atakapa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INDEPENDENT PRONOUNS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>*ano</td>
<td>?ima</td>
<td>take·ha</td>
<td>?iš</td>
<td>wiš</td>
</tr>
<tr>
<td>you</td>
<td>*ično</td>
<td>má(M); hćma(F)</td>
<td>?akahni</td>
<td>him?</td>
<td>naš</td>
</tr>
<tr>
<td>s/he</td>
<td>–</td>
<td>?ou(M); tihci(F)</td>
<td>?išna</td>
<td>hus</td>
<td>haš</td>
</tr>
<tr>
<td>we</td>
<td>*posno</td>
<td>?inima</td>
<td>takahni·</td>
<td>?us</td>
<td>yukiš</td>
</tr>
<tr>
<td>you (pl)</td>
<td>*hačno</td>
<td>winima(M); hinima(F)</td>
<td>?aNkahni·</td>
<td>was</td>
<td>nakt</td>
</tr>
<tr>
<td>they</td>
<td>–</td>
<td>sćma(M); sinima(F)</td>
<td>?išna·ni·</td>
<td>hunks</td>
<td>hakitiš</td>
</tr>
<tr>
<td><strong>POSSESSIVE PRONOUNS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>my</td>
<td>*ca-/*am-</td>
<td>?i-</td>
<td>-niš</td>
<td>?iš</td>
<td>wi</td>
</tr>
<tr>
<td>your</td>
<td>*či-/*čim-</td>
<td>wi-(M); hi- he-(F)</td>
<td>-piš</td>
<td>him?</td>
<td>na</td>
</tr>
<tr>
<td>her/his</td>
<td>*i-/*im-</td>
<td>?u-(M); ti-(F)</td>
<td>-?iš</td>
<td>hus</td>
<td>ha</td>
</tr>
<tr>
<td>our</td>
<td>*po-/pom-</td>
<td>?i-n</td>
<td>–</td>
<td>?us</td>
<td>yuki</td>
</tr>
<tr>
<td>you (pl)</td>
<td>*hači-/*hačim-</td>
<td>wi-n-(M); hi-n-(F)</td>
<td>–</td>
<td>was</td>
<td>nakt</td>
</tr>
<tr>
<td>their</td>
<td>–</td>
<td>si-(M); si-n-(F)</td>
<td>–</td>
<td>hunks</td>
<td>hakit</td>
</tr>
<tr>
<td><strong>AGENTIVE PRONOUNS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>*-iš</td>
<td>-ni</td>
<td>ta/-ya/-?a-(ka-)</td>
<td>-ki</td>
<td>-o</td>
</tr>
<tr>
<td>you</td>
<td>*ič-/*či-</td>
<td>wi-(M); hć-(F)</td>
<td>pan/-pi/-pa·</td>
<td>-i?i</td>
<td>naš</td>
</tr>
<tr>
<td>s/he</td>
<td>*θ-</td>
<td>?u-(M); ?a-(F)</td>
<td>na/-?i/-?a-·</td>
<td>-?i</td>
<td>haš</td>
</tr>
<tr>
<td>we</td>
<td>*il-/*li</td>
<td>?ina-</td>
<td>–</td>
<td>-naka</td>
<td>-cel</td>
</tr>
<tr>
<td>you (pl)</td>
<td>*hač-/*hači-</td>
<td>wina-(M); hēna-(F)</td>
<td>–</td>
<td>-na?a</td>
<td>-tem</td>
</tr>
<tr>
<td>they</td>
<td>*θ</td>
<td>?ūna-(M); sīna-(F)</td>
<td>–</td>
<td>-na?a</td>
<td>-o+</td>
</tr>
<tr>
<td><strong>PATIENT PRONOUNS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>*ca-</td>
<td>?ihk-</td>
<td>-t-</td>
<td>-ki-</td>
<td>hi</td>
</tr>
<tr>
<td>you</td>
<td>*či-</td>
<td>wihk-(M); hihk-(F)</td>
<td>-p-</td>
<td>-θ-</td>
<td>n</td>
</tr>
<tr>
<td>s/he</td>
<td>*θ-</td>
<td>?uhk-(M); thik-(F)</td>
<td>-θ-</td>
<td>-θ-</td>
<td>ha</td>
</tr>
<tr>
<td>we</td>
<td>*po-</td>
<td>?iṅk-</td>
<td>–</td>
<td>-kuy-</td>
<td>iš</td>
</tr>
<tr>
<td>you (pl)</td>
<td>*hači-</td>
<td>wink-(M); hink-(F)</td>
<td>–</td>
<td>-θ-</td>
<td>nak-</td>
</tr>
<tr>
<td>they</td>
<td>–</td>
<td>sīnk-(M); sink-(F)</td>
<td>–</td>
<td>-θ-</td>
<td>šak-</td>
</tr>
<tr>
<td>reflexive</td>
<td>*ili-</td>
<td>–</td>
<td>-hši-</td>
<td>–</td>
<td>hat-</td>
</tr>
<tr>
<td>reciprocal</td>
<td>*ixti-</td>
<td>–</td>
<td>-tahn-</td>
<td>–</td>
<td>bok-</td>
</tr>
<tr>
<td><strong>STATIVE VERB SUBJECT PRONOUNS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>*ca-</td>
<td>?i-</td>
<td>-t-</td>
<td>-ki-</td>
<td>hi</td>
</tr>
<tr>
<td>you</td>
<td>*či-</td>
<td>wi-(M); hi-(F)</td>
<td>-p-</td>
<td>-θ-</td>
<td>n</td>
</tr>
<tr>
<td>s/he</td>
<td>*θ-</td>
<td>?u-(M); ti-(F)</td>
<td>-θ-</td>
<td>-θ-</td>
<td>θ-</td>
</tr>
<tr>
<td>we</td>
<td>*po-</td>
<td>?i-n</td>
<td>–</td>
<td>-kuy-</td>
<td>ic-</td>
</tr>
<tr>
<td>you (pl)</td>
<td>*hači-</td>
<td>wi-n-(M); hi-n-(F)</td>
<td>–</td>
<td>-θ-</td>
<td>–</td>
</tr>
<tr>
<td>they</td>
<td>–</td>
<td>si-(M); si-n-(F)</td>
<td>–</td>
<td>-θ-</td>
<td>θ-</td>
</tr>
</tbody>
</table>
Like Hair, or Trees: Semantic Analysis of the Coeur d’Alene Prefix ne’ ‘amidst’

Debra J. Occhi, Gary B. Palmer, and Roy H. Ogawa

University of Nevada, Las Vegas, NV

1. Introduction

This paper investigates the semantic structure of a spatial prefix of Coeur d’Alene, a Salishan language spoken in northern Idaho and eastern Washington state. The prefix under study is ne’, which is realized through phonological conditioning as allomorphs [niʔ], [neʔ], [nəʔ], and [naʔ]. It contrasts in usage with other spatial prefixes, including n ‘in’2, čit ‘on something broad’, č ‘on, narrow point of contact’, t ‘on, against’ and cən ‘under’. The prefix ne’ has been written and glossed as nəʔ ‘among’ by Johnson (1975: 34) and ni’ ‘amongst’ by Reichard, who also observed that “the combination ni’-….i’qs refers to hair of the nostril, but is generalized for nose” (1938: 596). Reichard identified a second prefix with the same shape that she glossed as indicating “superlative degree” (596). Where a brief gloss is appropriate, we use ‘amidst’ because it seems to subsume the more abstract usages for which ‘among’ or ‘amongst’ seems sometimes inappropriate. Near synonyms in English, amidst connotes a more spatially restricted subject surrounded by its object, while among connotes comingling.3 Both senses seem to occur in our data.

While we believe that all of these glosses are correct as far as they go, when compared to actual usages they are incomplete, and we are not convinced that Reichard was correct in proposing the homonymy of ni’ ‘amongst’ and ni’ ‘superlative degree’. Furthermore, it is not immediately obvious exactly what any of these interpretations contribute to the understanding of such abstract words as s-ni’-k’wi’n ‘choice’ and na’-qhit ‘maybe’. Our analysis will show how such usages are motivated by reasonable abstractions and extensions from a single prototypical sense of ne’. We treat ne’ as a complex category as defined by Langacker to be a “network of related units centering on a prototype” (1991: 119). This resolves the homonyms identified by Reichard into a single category. While the connection may seem unintuitive, it appears that, in Coeur d’Alene, the superlative degree may have something in common with the hairs of the nostrils.

1.2 Related Work: Cora u and a (‘inside’ and ‘outside’)

The approach of studying affixes as complex categories has been used previously by Casad and Langacker in the analysis of two prefixes of Cora, a Uto-Aztecan language spoken in the state of Nayarit, Mexico (Casad and Langacker 1985; Langacker 1991). The prefixes u and a symbolize a basic ‘inside’/‘outside’ contrast in Cora, but in some instances words that contrasted only in their usage of these two prefixes were assigned the same translation and could be “employed to described precisely the same objective situation” (Langacker 1991: 34). Such a situation might lead one to conclude that the choice
of prefixes was arbitrary or determined by grammatical rules rather than semantic considerations, but Langacker argued that "a clear semantic rationale for the choice can almost always be found" (1991: 34). Casad and Langacker demonstrated that the puzzling usages could be explained if speakers were construing the same objective situation in terms of contrasting imagery. A dog's tail might be described as 'inside' from a perspective behind the dog, because it is construed as inside the viewing area in the line of sight presented by the rump. But it can be described as 'outside' when imagined as viewed from the dog's side, because it is outside the viewing area in the line of sight. The base notion of a viewing area within a line of sight is itself an extension of the more prototypical concepts pertaining to 'containment in physical space'.

In addition to the contrast 'in line of sight'/outside line of sight', Casad and Langacker also found other variants on the prototypes of /u/ and /a/, such as the specializations 'on inner surface'/on outer surface' and the extensions 'deep penetration to interior'/shallow penetration to interior', 'accessible'/inaccessible', and 'on back side'/on face/front'. Such polysemous complexes naturally raise the question of whether any meaning can be discovered that is schematic for all senses of a prefix and that would therefore suffice to define it. Langacker concluded that "it is most improbable that a single abstract meaning can be found that would be schematic for all the specific values attested for /u/ and /a/" (1991: 55). He argued further, that even if a fully schematic sense were found it would not suffice as a definition, because "it would also be schematic for indefinitely many values that /u/ and /a/ happen not to have" and "it would fail to provide an explicit account of the facts of the language, in particular the range of conventionally established senses and usages characteristic of these morphemes." (1991: 55).

Understanding the Coeur d'Alene prefix ne' also requires an understanding of extensions from a prototypical sense, but our data require no changes in construal or perspective comparable to those adduced by Casad and Langacker in their study of Cora /u/ and /a/. Unlike those authors, we have found a very abstract sense of AMIDST COLLECTION OF ENTITIES that we believe is fully schematic for all of the attested senses of ne', but we agree with them that defining the schema is insufficient to provide an adequate account of conventionally established senses and usages. Furthermore, the schema subsumes no usages that are not also subsumed by one of its subschemas. In this respect, ne' contrasts with other spatial prefixes in Coeur d'Alene for which defining the schema of a complex category sometimes enables one to understand usages that are highly abstract or metaphorical in psychological or social domains, usages that are not subsumed within subschemas.

1.2 ne'

The prefix ne' appears to have a clear prototype with an anatomical component of meaning in its sense of AMIDST HAIR (OF). The prefix also appears in terms referring directly or indirectly to the intestines, pubic hair, and the clitoris. This centrality of anatomical usage is found in other Coeur d'Alene spatial prefixes, such as cen 'under', but not in all, as it is not a prominent
characteristic of any of the senses of n ‘in’ or č ‘on’ (Palmer 1990). However, Coeur d’Alene and other Salish languages have a set of suffixes whose primary meanings are clearly anatomical (e.g. head, face, back, hand, leg, etc.) (Palmer and Nicodemus 1985; Palmer 1993). In this respect Coeur d’Alene grammar might be said to have an “anthropocentric perspective” (Wierzbicka 1985).

A more schematic conceptualization of the meaning of ne’ is AMIDST LONG THIN VERTICAL THINGS, a sense that underlies the usage in (8) ni’tgelish ‘it crept in the grass’ and (9) the place name ni’lokhwawq ‘Cut in the Woods’, an important camping site in former times. An extension from this is AMIDST THIN FLAT THINGS, which subsumes the term for ‘bookmark’ and the notion of setting something in the corner (in some sense, between the walls).

An even more schematic characterization of ne’ is AMIDST OBJECTS, which underlies the term for ‘blizzard’ and the probable neologism for ‘orgy’. An extension of this schema seems to be involved in the equally abstract IN THE CENTER, which we believe to underlie such terms as those meaning ‘approximately’ and the term for center itself: ni’mit’wes. A further specification of AMIDST OBJECTS is the notion of ONE SELECTED FROM MANY (hereafter abbreviated ONE FROM MANY), which subsumes terms for ‘choice’, ‘appoint’, and ‘election’. The notions of ‘farthest point’, ‘highest point’, and ‘best in the crowd’, which all seem to pertain to the sense of EPITOME, may be derived by further specification of ONE FROM MANY. An extension of the notion of AMIDST OBJECTS is the notion of MIXED, which involves a cluster of unspecified objects construed reflexively. This sense subsumes terms for ‘assorted’, ‘clutter’, and ‘blend’.

Similar abstractions from body-part prototypes and similar extensions and metaphorical usages that we see in such terms as those meaning ‘approximately’, ‘choice’ and ‘best in the crowd’ have been observed in other Native American languages. For example, Friedrich’s report of extensions of body-part suffixes of locative space in Tarascan reveals a similar phenomenon to that emerging from our analysis of Coeur d’Alene prefixes (1979b; 1979a). The processes of abstraction and extension from terms for body-parts have also been reported by Brugman (1983) and summarized in Palmer (1990: 267-268):

In Chalcatongo Mixtec, Brugman (1983) observed a hierarchy of extensions of body part terms in which literal uses are extended to metaphorical partitive uses, thence to spatial relational uses, and finally to abstract relational uses. The term nuu face, for example, also refers partitively to the face of an object, or relationally to the space in front of an object. The term čīi belly can be extended beyond spatial senses to the abstract relation "because." Brugman also found that body part terms differ in their susceptibility to abstraction and extension, with the term ndaʔa hand/arm being most restricted, followed in rough order by čīi belly, siki back, and the least restricted, nuu face.
Thus, this Coeur d'Alene prefix appears to behave much like other locative particles in other Native American languages. They are readily described as complex categories having both prototypical and schematic senses. Where obvious prototypical senses occur, these are often based on body parts. This work also parallels that of Brugman (1988) and Lakoff (1987) who proposed that the English preposition over be analyzed as a complex category made up of "instance links", "similarity links", and "transformational links". Since transformational links are based not on shared subschemas but on related subschemas, then, in the case of over, there can be no single schema that subsumes all of its usages. Thus, like Casad and Langacker, they would necessarily reject a single schema definition. But this conclusion of Brugman and Lakoff has been questioned by Dewell, who has proposed a central schema from which all of the spatial variants of over "can be derived either directly or indirectly using nothing but natural, independently-motivated image-schema transformations" (1993). Dewell's finding, then, appears to lend support to our findings concerning the Coeur d'Alene prefix ne'. This is not to argue that all such grammatical morphemes must always have a central schema, but it suggests that one should not overlook the possibility. In the following sections, we outline our methods and the actual analysis of terms with ne' in an attempt to provide a more precise characterization of ne' as a complex category.

2. Methods

The first step was to identify candidate words which might contain the prefix. We examined several reference works and grammars in Coeur d'Alene/English translation, producing a list of 59 words containing the prefix ne' (Reichard 1938; Johnson 1975; Nicodemus 1975a; Nicodemus 1975b; Palmer and Nicodemus 1985; Palmer, Nicodemus and Connolly 1987; Palmer, Nicodemus and Felsman 1987). Each word was parsed, using the general morphological analyses and glosses provided in Reichard (1938), Johnson (1975) and Nicodemus (1975a; 1975b). We used the morphological analyses to interpret the constructional meaning of each word in terms of the meaning of its constituent morphemes. Our interpretations were constrained to be consistent with the full-word-glosses provided by native speakers. For example, in (1) our constructed interpretation would be 'on the surface in the hair at the top of the head', which is consistent with the published gloss of 'in the hair at the top of the head'. Our constructed glosses were based in part on a schema of grammatical relations to be discussed in more detail below. Only the morpheme glosses and published native-speaker glosses are provided here.

Each word evokes an image based on its constituent morphemes. Commonalities in the imagery of several words were drawn as image-schemas (Figure 1). Commonalities among schemas were diagrammed as superschemas (Figure 2). We used the concepts of extension and elaboration developed by Langacker (1987, 1991), discussed below, to describe how schemas relate to one another (Figure 2). Finally, we proposed a single schema that subsumes all the others.
Perhaps the weakest link in this process lies in the visualization of the spatial image-schemas diagrammed in Figure 1. Our visualizations of image-schemas are based entirely on our interpretations of whole word-glosses and of morpheme glosses in context, as provided to researchers by native speakers, so they are largely inferential. However, the spatial schemas that govern the use of *ne‘* and other Coeur d’Alene affixes are not necessarily available to conscious inspection by native speakers either. Therefore, while recognizing this limitation on our methodology, we nevertheless proceeded in the belief that new and valid results might be possible and that future research in Coeur d’Alene and other Salish languages may clarify the issues. At the very least, our results have provided us with a sharper focus for future elicitation and interpretation of texts.

In discussions of an earlier draft of this paper, it has been argued that the lines of Figure 2 could be drawn otherwise. It is certainly possible that Figure 2 is not a valid interpretation of the semantic relations among this set of terms. Furthermore, it is not exhaustive because one can always posit more schemas and lines of extension. But the organization in Figure 2 is the one that makes the most sense to us given the criteria that we will present in this section for identifying elaborations and extensions of schemas. As additional data become available from Coeur d’Alene and comparative data are adduced from other Interior Salish languages it will become more clear whether Figure 2 is an adequate representation of *ne‘* as a complex category.

The validity of our analysis depends in part on our ability to discover and define schemas for a group of similar usages of the prefix. In cognitive linguistics schematization has been given a simple definition that gives great latitude to the researcher. Talmy defined it as “a process that involves the systematic selection of certain aspects of a referent scene to represent the whole, while disregarding the remaining aspects” (Talmy 1983: 225). According to Langacker, “a schema is an abstract template representing the commonality of the structures it categorizes, which thus elaborate or instantiate it” (Langacker 1991: 59-60). Langacker also asserts that “a schema differs from a list of critical attributes in being an integrated concept in its own right”.

Concepts may be linked by elaboration or by extension. Elaboration of a schema, also called instantiation, is symbolized with a solid arrow as [A] → [B], where [B] is a more detailed instance that nevertheless conforms to all specifications of the schema [A] (1991: 267). Extension of a prototype category “implies some conflict in specifications between the basic and extended values; hence [A] ----> [B] indicates that [B] is incompatible with [A] in some respect, but is nevertheless categorized by [A]” (Langacker 1991: 266). Langacker also introduced a third kind of relationship [A] ↔ [B], which symbolizes “a perception of mutual similarity”. Following these conventions, a complex category can be described as a network of concepts, including schemas, prototypes, and variants.

This simple distinction between elaboration and extension is not always easy to discern in the data. Langacker regards elaboration as simply a limiting case of extension in which the incompatibilities in the specifications of the two concepts being compared are reduced to zero. Therefore, it is usually sufficient
to simply determine whether they are plausibly related by commonalities. Schemas are identified in our data wherever they seem clearly warranted.

The meanings of the spatial prefixes of Coeur d’Alene can all be characterized as instantiations of a special, but more general, kind of schema in that they are relational, as opposed to nominal. In Langacker’s framework, a relation connects two entities that are called trajector and landmark. Langacker defined a trajector as “the figure within a relational profile” (1987: 217). By analogy, a landmark would be the ground within a relational profile. The terms trajector and landmark are abbreviated as tr and lm. Every relation connects a trajector and a landmark, either of which may or may not be further elaborated. We have adopted the hypothesis that, in Coeur d’Alene, trajectors of spatial prefixes are always supplied by the word root or stem. These may have meanings that pertain to things (nominal), processes, or states. Though it may seem more intuitive to limit the concept of trajector to figures in motion, Langacker’s definition applies to static, as well as dynamic, figures. Landmarks in Coeur d’Alene predications are often instantiated by lexical suffixes, and these often specify body parts. Some common ones are qən ‘head’, us ‘face, eye’, qən ‘mouth’, ixən ‘arm’, ɨxt ‘hand, arm’, ɨḵən ‘back’, and ɨxən ‘leg, foot’.

These concepts of relation, trajector, and landmark will enable us to describe the various schemas, elaborations, and extensions of ne’ with greater precision. Thus, to provide an example, the term č-qən ‘in the hair at the top of the head’ can be analyzed as follows: ne’ (realized as [ni?]) has the sense AMIDST HAIR OF. The trajector for ne’ is an abstract entity instantiated by the root morpheme čən ‘surface’. The landmark, characterized by ne’ as a hairy entity, is instantiated or elaborated by the suffix complex awəs-qən ‘the top of the head’.

The prefix č ‘on’ requires its own separate analysis.

In the following section we will show that most elaborations and extensions of ne’ operate on the landmark. Thus, our definitions in small caps often describe the landmark (AMIDST OBJECTS, AMIDST THIN FLAT THINGS, AMIDST HAIR OF, AMIDST INTESTINES, ONE SELECTED FROM MANY), but they are neutral as to the nature of the trajector. They could have been written more completely as TRAJECTOR AMIDST OBJECTS, etc. Other definitions locate the trajector with respect to more abstract or tacitly understood landmarks (IN THE CENTER, EPITOME). The notion of MIXED seems to conflate trajector with landmark, so that trajectors and landmarks are defined reflexively.

2. Data

In this section the data are presented in schematic groupings. In all cases, the landmark of ne’ is a plural spatial domain; the prototypical landmark is the hair atop the human head. Sources are coded by initial and page as follows:
Terms from Nicodemus’s dictionaries and from Palmer, Nicodemus, & Felsman (1987) and Palmer, Nicodemus, and Connolly are written using Nicodemus’s practical orthography, which uses underlining for stress and does not write schwas. The first example of this orthography is term (2), below. Terms from Johnson (1975) and from Palmer and Nicodemus (1985) are written in Americanist phonemic orthography, as in term (1). Terms from Reichard are converted to this orthography.

2.01 AMIDST HAIR (OF)

ne' describes the location of something (the trajector), as amidst hair on something (the landmark).

(1) čn'įʔemáwesqen, č-neʔ-ʃem-iwes-qen, on-amidst-surface-middle-head, top, 'in the hair at the top of the head' (P&N:355)
(2) sniʔch'maq'wasqen, s-ni'ʔ-ch'm-a'was-qen, NOM⁴-amidst-surface-waist, middle, between-top, 'crest (lit. the top of the head)' (Nb:103)⁵
(3) niʔráqqaq, neʔ-ʃráq-qen, amidst-wide-head, top, 'Wide Forehead, Wide Surface Under Hair' (place name, probably referring to a wide ridge or mountain) (PFN:60)
(4) eniʔkʷúselsčen, ʔic-neʔ-ʃkʰus-ulscčen, CUST-amidst-curly-forehead, 'hair curls back from forehead' (J:236)
(5) sniʔch'ami'qs, s-ni'ʔ-ch'am-i'qs, NOM-amidst-surface-nose, 'nose (lit. surface of the ...)' (Na:218)
(6) sniʔch'amch'am'q's, s-ni'ʔ-(REDUP)ʔch'am-i'qs, NOM-amidst-surface-nose, 'nostrils (lit. surfaces amidst the nose)' (P & N:356)

2.02 AMIDST LONG THIN VERTICAL THINGS

The abstract conceptualization for hair would be LONG THIN VERTICAL THINGS. The conceptualization arises from the qualities of human hair, being of greater length than width and usually observed in a vertical orientation relative to the scalp and the ground due either to its pattern of vertical growth or to the fact that gravity causes it to hang down. Native American hair is characteristic in a straight. Therefore, this schema subsumes the schema for AMIDST HAIR (OF). This schema can be instantiated by such diverse things such as grass, brush, logs, houses, and humans. ne' locates the trajector amidst the plural elements of the landmark. In fact the frequent occurrence of terms with the frame [niʔ__-ilqʷ/ amidst__log, tree] 'in the trees/forest' suggests a
prototype that may rival AMIDST HAIR in its salience. That this is a highly
conventionalized combination is further suggested by the contraction of ilq” in
(11).

(7) ni’tgwkw, ni’-√tekw, amidst-stuffy, choke, ‘thicket, brushwood, brake (lit.
a woods whose interior is, suffocating)” (Nb:72)
(8) ni’tgdish, ni’-√ted-ish, amidst-squirm-act of, ‘it crept in the grass’, ‘He
acted foolishly’ (Na:160)
(9) ni’lukhwalqw, ni’-√lukhw-alqw, amidst-cut-tree, log, ‘Cut in the Woods’
(place name) (PNF:25)
(10) ni’ch’ch’a’ra’lqw, ni’-(REDUP)√ch’ar-alqw (DIM GLOT), amidst-band lies-
tree, log, ‘Small Cut in the Woods’ (place name) (PNF:124)
(11) ni’nsi’, ni’-nes-i’[lqw], amidst-damp-tree, log, ‘Damp in the Woods’
(place name) (NPF:29)
(12) ni’gwalpalqw, ni’-√gwal-p-alqw, amidst-burn-INCHOT-tree, log,
‘burned (lit. the forest was ...’ (Nb:82), (contrasts with t-gwal-p-alqw
‘the log burned.’)
(13) ni’gwept, ni’-√gwept, amidst-hairy-inherent, ‘bushiness (lit. the forest
is bushy with)’ (Nb:83)
(14) sni’sharus, s-ni’-√shar-us, NOM-amidst-hang-fire, ‘boiled beef’ (hanging
over fire) (Nb:65)
(15) sni’tata’ri’tkhw, s-ni’-(REDUP)√te’ri-t-khw, NOM-amidst-covered with
trails-house, ‘alley’ (Nb:13)

In (15), we suggest that the landmark is houses, or house walls and corners,
construed as long thin things (as in the mission town of DeSmet at the turn of the
century), and that the trajector (√te’ri ‘covered with trails’) is placed amidst them.
The possibility has been raised by discussants that the landmark of this term
should be regarded as the trails. In this interpretation, the trajector is amidst the
trails rather than amidst the houses. We prefer our interpretation because the
great preponderance of terms supports an argument based on morphosyntax:
trajectors are always supplied by the root, and landmarks, when given, are
supplied by a lexical suffix. To draw a parallel in English, given the phrase,
“covered with trails among the houses”, we would argue that the trajector of the
preposition ‘among’ is ‘covered with trails’ and the landmark is ‘the houses’.
Although the order of occurrence of these elements is different in Coeur
d’Alene—[rel tr lm] as opposed to English [tr rel lm]—it must still be considered
in assigning an interpretation. Note that (15) is parallel to (16), which clearly
supports our assignment of trajectors to roots and landmarks to lexical suffixes.

(16) tk”e’niʔcélul”x”, tk”e’-7ic-neʔ-√cél-ul-hx”, distributed-CUST-amidst-one
stands-position-house, ‘he was standing around among the houses’
(J:234)
(17) na’qwa’qwe’elstkhw, na’-(REDUP)√qwe’l-stkhw, amidst-one speaks-IMP,
‘you (sg.) are to speak to h/h’ (Na:154)
(18) sni?gʷ’épšen, s-neʔ-ʔgʷ’ep-šen, NOM-amidst-hair-legs, 'pubic hair' (P & N:85)
(19) sni’ch’mysshn, s-ni’-χ’m-us-shn, NOM-amidst-surface-face, eye-leg, 'clitoris (lit. surface of 'small hill' at upper end of vulva)' (Nb:114)
(20) ni’y’ilxw-us-shn, ni’-vly’lxw-us-shn, amidst-cover with fabric-face, eye-leg, 'apron (lit. a covering for the lap)' (Nb:25)
(21) ni’bmpa’wes, ni’-vbem-p-a’wes, amidst-whir-INCHOAT-waist, middle, between, 'orgy (lit. there is speeding or intoxication among them)' (Na:159)

2.03 AMIDST THE INTESTINES

Another extension occurs with the intestines, which are conceived of as long thin things with the trajector located somewhere within them. The trajector is either diffuse or indeterminate as to its discreetness. Note that the suffix in’č ‘belly, stomach’ is restricted by the glosses in this data to the intestines and bowels.

(22) sni’tgsstspench, s-ni’-(REDUP)štes-ts-p-ench, NOM-amidst-bulge-VB?-INCHOAT-stomach, 'bulkage (lit. swellings among the intestines)' (Nb:80)
(23) sni’tš’q’tsa’rench, s-ni’-(REDUP)šts’ar-ench (GLOT), NOM-amidst-hurt-stomach, 'intestinal pains' (Na:219)
(24) sni’ch’gsch’ssmenč, s-ni’-(REDUP)šch’es(REDUP?)-m-ench, NOM-amidst-bad-MDL-stomach, 'constipation (lit. something going wrong with the bowels)' (Na:218)

2.04 AMIDST THIN FLAT THINGS

This group shares the idea of landmark length being greater than thickness, like human hair, and it extends the dimension of width. In this way the trajector can be located amidst thin flat things such as walls or the pages of a book.

(25) ni’q’e’i’wesn a q’e’y’minn, ni’-vq’e’i’wes-n a q’e’y’minn, amidst-stick-middle,waist-nom ART book, 'bookmark (lit. placed between the pages of a book)' (Nb:67)
(26) ni’cqáqwasqen, nɛʔ-vcqq-is-wis-qen-en, amidst-upright-between-head-NOM, 'I set a hollow object in the corner' (J:233)

2.05 AMIDST OBJECTS

Here, a highly abstract use of nɛ’ describes a trajector amidst various kinds of objects that are not specified as to whether they are long, thin, or vertical. In some words of this group, the landmark need not even be specified. The trajector may be diffuse or indeterminate as to its discreetness.
(27) kupsniʔáq̲áq̲wesšutemš, kup-s-niʔ-ʔl̩q̲-i̲w̲es-ʃet-əmš, 2p. sg.-NOM-amidst-search for-waist, middle-BEN-MOD, ‘she is to look amongst for something as a favor to you’ (R:596)
(28) sniʔtmtmihysmsh, s-niʔ-REDUP\t̩am-iʔ-us-m-ush, NOM-amidst-sorch-ʔ-fire-MDL-act of, ‘blizzard’ (Nb:61)
(29) niʔpuxʷi̲w̲es, niʔ-\puxʷ-ii̲w̲es, amidst-blow-waist, middle, between, ‘he blew among’ (J:237)
(30) naʔq̲h̲esmicht mnakisgwən, naʔ-\q̲hes-m-icht-mn-tkhw, amidst-good-MDL-hand-INST-2p. sg, ‘elaborate (lit. you manipulate it carefully)’ (Na:155)

2.06 IN THE CENTER

The landmark is a diffuse space. The schema locates the trajector at the very center of the landmark.

(31) neʔshi,t, neʔ-\shi,t, amidst-exact, ‘approximately’ (Nb:25), ‘circa’ (Nb:108)
(32) sniʔtsl̩gʷ wasq̲n̩, s-niʔ-\tsl-aʔ-was-q̲n̩, NOM-amidst-five[poss. mistranscription of ts̩el ‘one stands’]-waist, middle, between-head, ‘porcupine quill over head’ (Na:219)
(33) niʔsát̩q̲en̩, niʔ-\sát̩q̲en̩, amidst-twist-nose, beak-NOM, ‘crank’ (J:238)
(34) niʔm̩iʔteʔwes, niʔ-\m̩iʔ-teʔwes, amidst-centered-waist, middle, between, ‘among, center, midpoint, middle, axis’ (Nb:16), ‘amid’ (Na:99)

Nicodemus has provided two interesting variants on this term for ‘center’ in (35) and (36):

(35) ep̩ sniʔm̩iʔt̩ewes, ep̩ s-niʔ-\m̩iʔ-t̩ewes, s/h/it has NOM-amidst-centered-waist, middle, between, ‘concentric (lit. It has a common center)’ (Na:73)
(36) teʔl niʔm̩iʔteʔwes, teʔl niʔ-\m̩iʔ-teʔwes, from amidst-centered-waist, middle, between, ‘centrifugal (lit. moving or directed from the center)’ (Nb:99)

2.07 MIXED

The tacit landmark for this group is a cluster of unspecified objects. The trajector describes turning motion.

(37) eniʔsglm, e-niʔ-\sl-m, CUST-amidst-turn-MDL, ‘assorted (lit. It is a mixture)’ (Nb:31)
(38) niʔsglm, niʔ-\sl(REDUP)-m, amidst-turn around-cause, ‘clutter’ (lit. it was piled in a disordered state) (Nb:116)
(39) niʔsglm stm, niʔ-\sl-m-stm, amidst-turn-MDL-PASS, ‘blend (lit. It was mixed, i.e. with other entities)’ (Nb:60)
2.08 ONE SELECTED FROM MANY

The landmark is a set of objects. The trajector is one of those objects. All of our examples are based on the same root k’w’i’n ‘choose’.

(40) s’ni’k’w’i’n, s-ni’-vlk’w’i’n, NOM-amidst-choose, ‘choice (lit. something chosen from among several)’ (Na:219)
(41) s’ni’k’w’i’nm, s-ni’-vlk’w’i’n-m, NOM-amidst-choose-MLD, ‘election (lit. the act of choosing someone)’ (Na:219)
(42) ni’k’w’i’nnts, ni’-vlk’w’i’n-nts, amidst-choose-TRNS, ‘appoint’ (Nb:25)

2.09 EPITOME

Terms in this group single out something (or someone) that possesses the greatest value of a quality. Therefore, it is a special case of the ONE FROM MANY schema. These are the terms that fit Reichard’s observation that “The prefix [ni?] is used to indicate the superlative degree” (Reichard 1938: 39, 596). The trajector is the quality that characterizes the person or thing and is designated by the stem. It is the stem that introduces a scale of value. In (45) - (48) it also seems to connote a range of possibilities that serve as the tacit plural landmark. In (49) - (51) the landmark is specified. Note that (51) has the [ni?-, ___-ilq”] frame, but here it belongs to the EPITOME schema rather than the AMIDST THE TREES subschema of LONG THIN VERTICAL THINGS. However, it seems quite likely that the available metaphor of trees is a motivating influence on this construction.

(43) s’ni’lgkut, s-ni’-vlék-ut, NOM-amidst-distant-be in position, ‘apogee (lit. the farthest point)’ (Na:219)
(44) s’ni’ngwst, s-ni’-vn-gwes-t, NOM-amidst-in-high, ascend-inherent, ‘acme (lit. the highest point)’ (Nb:6)
(45) s’ni’t’ik’wt’ik’ut, s-ni’-(REDUP)’vt’ik’w-ut, NOM-amidst-old-be in position, ‘eldest (lit. h/s is oldest person in group)’ (Na:219)
(46) s’ni’qhest, s-ni’-vnqhes-t, NOM-amidst-good-inherent, ‘h/s is the best in crowd, best, elite’ (Na:219)
(47) s’ni’ch’ch’qne, s-ni’-vc’ch’q’e-ne, NOM-amidst-small, ‘atom (lit. the very smallest entity)’ (Na:218)
(48) s’ni’siysiyyu’, s-ni’-(REDUP)vn-siy-us, NOM-amidst-able-face, eyes, ‘champ (lit. one who is most capable)’ (Nb:100)
(49) s’ni’cécw’temš, s-ne’-(REDUP)vn-ciw-t-emš, NOM-amidst-youngest adult-inherent-people, ‘the little one who was youngest’ (J:233)
(50) s’ni’cécw’umš, s-ne’-(REDUP)vn-ciw-t-umš[Dim Glot], NOM-amidst-youngest adult-inherent-people, ‘the youngest of the small ones’ (R:596)
(51) s’ni’céšalq”, s-ni’-vcéš-alq”, NOM-amidst-be long, tall-long stiff object, ‘the tallest’ (R:596)
2.10 POSSIBILITY

*ne'* has also acquired an abstract sense of POSSIBILITY, perhaps as a
metaphorical extension of ONE SELECTED FROM MANY. Thus, Nicodemus (1975b: 259) gives the following terms:

maybe, adv. na'qhît
maybe, adv. ne' (stem)
maybe, adv. ne'gwnîkhw
maybe so, adv. ne''nìs

The complex terms in this group can be analyzed as follows:

(52) *na'qhît, na'-√qhît*, amidst-to leave, desert, abandon, 'maybe, possibly' (Na:155)
(53) *ne'gwnîkhw, ne'-√gwnîkhw*, amidst-true, 'maybe'
(54) *ne''nìs, ne'-√mine'-us?* (GLOT), amidst-to be apt, likely-eye, face, 'maybe so' (Na:259,151)

To these can be added the imperative form in (55)

(55) *ne''wi'intkhw, ne'-√wi'in-tkhw*, amidst-call-IMP3, 'call (lit. you shall ... h/h aloud!)' (Nb:87)

All of the above terms have been attested and published. Other terms that
apparently include the prefix *ne'*, but which were unanalyzable (59) or whose
attestations or transcriptions seem less reliable (56-58) include the following:

(56) *na'nsțq'ělkhw* 'Place for Hooking On' (place name; NPF:52)
(57) *na'ulståq'ělkhw* 'Maybe Belonging to Place for Hooking On' (place name; NPF: 55)
(58) *ne'atsq̈hagšt'm* 'Where Crows Call' (place name; NPF:43)
(59) *ni'țukhkwelch'î* (personal name; well attested but meaning not known; √ţukhw 'pull on' NPC:51)

3. Conclusions

Various senses of the Coeur d'Alene prefix *ne'* are depicted as schemas that
belong to a complex category. Both schematic relations and prototype-to-variant
relations figure prominently in the semantic structure of *ne'*. Given that the
prefix predicates a relation that connects a trajector to a landmark, we have
found that it is changes in the landmarks that account most frequently for
variations in meanings. Prototypical landmarks refer to hair and trees. The
reference to hair is not surprising given the importance of body part symbolism
in spatial and locative terms in other Native American languages, e. g. Mixtec
and Tarascan.
Our findings support the work of Casad and Langacker that shows that spatial prefixes can be described as complex categories. Unlike these authors, and contrary to expectations generated by the work of Brugman and Lakoff, we find that a central schema can be reasonably posited for the prefix *ne*. However, we concur with Langacker that a central schema needs to be supplemented with other conventional schemas to provide an adequate account of a complex category.

Few clearly metaphorical usages emerge, though some usages seem vaguely metaphorical, as in terms that refer to epitomes, centers, and possibility. Perhaps the lack of clearly metaphorical usages reflects the high level of abstraction of spatial prefixes.

**Endnotes**

1We wish to thank Dale Kinkade, Margaret Langdon, and George Urioste for useful comments on this paper. They are not responsible for the mistakes that remain.


3*Websters Third New International Dictionary of the English Language Unabridged* (Springfield: G. & C. Merriam Co., 1976, pp. 70, 72) has *among* (or *amongst*) "surrounded by: in the midst of: intermingled with" (p. 72) and *amid* (or *amidst*) "in or into the middle of: surrounded or encompassed by: AMONG" (p. 70).

4Many of the terms in this paper are taken from the two volume dictionary by Nicodemus (1975a,b). The dictionary was apparently produced by presenting Nicodemus with a list of terms in English in order to elicit Coeur d’Alene equivalents. The result was undoubtedly a number of invented words never before uttered in Coeur d’Alene, but nevertheless revealing of the semantic principles underlying their construction. Unfortunately, it is often not possible to say with certainty which are traditional usages and which are neologisms.

5Adding a new wrinkle to the search for the function of the puzzling *s* prefix in Salish, Kinkade (1993) has found that *s* in Upper Chehalis functions to mark imperfectives in directly quoted speech.

6This was glossed as ‘chest’ in the source, an apparent typographical error.
References Cited

Brugman, Claudia  

——  

Casad, Eugene and Ronald Langacker  

Dewell, Robert B.  

Friedrich, Paul  

——  

Johnson, Robert  
1975 The Role of Phonetic Detail in Coeur d'Alene Phonology. [Ph.D, Washington State University, Pullman.]

Kinkade, M. Dale  
1993 S- prefixation on Upper Chehalis (Salish) imperfective predicates. Paper presented to the SSILA Summer Meeting, Columbus, Ohio, July 2, 1993.

Lakoff, George  

Langacker, Ronald W  

——  

Nicodemus, Lawrence G.  


Palmer, Gary, Lawrence Nicodemus and Thomas Connolly  1987  Khwi' Khwe Gu+ Schitsu'umsh 'These Are the Coeur d'Alene People. Plummer: Coeur d'Alene Tribe of Idaho.

Palmer, Gary B., Lawrence G. Nicodemus and Lavinia Felsman  1987  Khwi' Khwe Hntmikh'umukhw: This is My Land. Plummer: Coeur d'Alene Tribe of Idaho.


Figure 1: Image-Schemas for the Coeur d’Alene Prefix *ne’*
Figure 1, continued

* AMIDST THE INTESTINES
  * *snii'tsa'ts'a'rench*
  `intestinal pains`

* IN THE CENTER
  * *ni'mii't'wes*
  `among, center, midpoint, middle axis`

* MIXED
  * *eni'sel'm*
  `assorted`

* ONE FROM MANY
  * *snii'k'wi'j'n*
  `choice`

* EPITOME
  * *snii'lekut*
  `apogee' (the farthest point)
Figure 2: Prefix *ne* as a Complex Category
Table 1: Consonant Phonemes of Coeur d'Alene

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Stops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>p</td>
<td>t</td>
<td>k*</td>
<td>q</td>
<td>q*</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ejective</td>
<td>p'</td>
<td>t'</td>
<td>k*</td>
<td>q'</td>
<td>q'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>b</td>
<td>d</td>
<td>g*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>c</td>
<td>ç</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ejective</td>
<td>c'</td>
<td>ç'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>medial</td>
<td>s</td>
<td>ʃ</td>
<td>x*</td>
<td>x</td>
<td>x*</td>
<td></td>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td>ɹ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonorants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-ejective</td>
<td>l</td>
<td></td>
<td>r</td>
<td>ç</td>
<td>ç*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ejective</td>
<td>l'</td>
<td></td>
<td></td>
<td>ç'</td>
<td>ç'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-ejective</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ejective</td>
<td>m'</td>
<td>n'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semivowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-ejective</td>
<td>y</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ejective</td>
<td>y</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[\text{a}\] Based on Table 1 in Johnson, Robert E., The Role of Phonetic Detail in Coeur d'Alene Phonology. (Ph.D. dissertation in Language and Linguistics, Washington State University, Pullman, WA, 1975). Johnson and other Coeur d'Alene scholars have used /R/ to represent the pharyngeal phoneme /ʃ/ and its labial and ejective counterparts.

Table 2: Vowel Phonemes of Coeur d'Alene

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>e</td>
<td>æ</td>
<td>œ</td>
</tr>
<tr>
<td>Low</td>
<td>æ</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

\[\text{a}\] Coeur d'Alene vowels undergo both progressive harmony and, under the influence of low-back consonants, regressive lowering.
CONSTITUENT ORDER VARIATION IN APURINÁ

(Arawakan)

Sidney Facundes

University of Oregon, Eugene, OR, USA

Museu Paraense Emílio Goeldi, Belém, Pará, Brazil

0. Introduction*

The present paper is an attempt to present some descriptive facts about clitics and constituent order variation in Apuriná,1 with a discussion about whether Apuriná has a configurational or non-configurational constituent structure organization. This leads to discussion of issues related to the properties of configurationality, following mainly the works of Jelinek (1984) and Hale (1990).

The discussion about configurationality focuses on one main property of configurational languages, namely the fixed order of clausal constituents. Other properties generally associated with the configurationality issue (e.g. hierarchical vs. flat clausal constituent structure organization, continuous vs. discontinuous expressions, and so on -- cf. Jelinek (1984) and Hale (1990)) will be left aside in this paper.

Section 1 presents the description of grammatical relations

* I wish to thank Doris Payne, Spike Gildea and Denny Moore for initial discussions and suggestions on preliminary versions of this paper. I wish also to thank the Inter-American Foundation, the Conselho Nacional de Pesquisas (CNPq/Brazil) and the Museu Paraense Emílio Goeldi for financial support of this research. As part of an ongoing language research project, the syntactic description of the language in focus here is rather preliminary; much of the grammar needs to be known in more detail.

1 Apuriná belongs to the Maipuran branch of the Arawakan linguistic family (David Payne 1991). It is spoken mainly along the tributaries of the Purus River in the Western Amazonian region of Brazil. There are more than 2,000 Apuriná, and at least 50% still speak the native language (Facundes 1990); however, the Apuriná language has been increasingly replaced by Portuguese. In most of the villages Portuguese is being learned by children as their first language. Dialectal variation can be found in some of the nearly 20 Apuriná villages. The present analysis is intended to cover only the dialect spoken in the Japiim village, along the Paciá River, near Lábrea city, in the state of Amazonas.
and all the possible variations of constituent order; section 2 focuses on the syntactic status of the nominal clausal constituents in the various orders, as well as on the syntactic status of the verbal person markers; section 3 suggests a syntactic analysis of the given facts based on the notion of configurationality, and offers a brief discussion of basic constituent order; finally, section 4 is a brief conclusion.

1. Constituent Order

Previous analyses of Apurinā constituent order (Pickering 1974, Aberdoor 1985, Derbyshire & Pullum 1981, and Facundes 1992b), have considered all nominal clausal constituents (preverbal and postverbal nominals in OSV, SVO, OVS and VOS orders) as expressing arguments of the verb. Moreover, the person markers which are attached to the verb were considered as verb agreement markers.

Pickering, Derbyshire & Pullum suggest an analysis based on structural facts of the Apurinā syntax to argue for OSV as the basic constituent order of the language. Aberdoor presents a study of frequency counting in which she shows that OSV is very rare in text. A reanalysis of the data may show that the role played by clausal constituent order is correlated with the role played by the person markers on the verb; that such person markers are in complementary distribution with preverbal nominals, but not with postverbal ones; and that clitics can also express verb arguments, while nominals split in argumentative and adjunctive functions depending on whether they are pre- or postverbal.

1.1 Grammatical relations

Before getting into the description of constituent order variation, one would like to understand how the verb argument structure is syntactically marked; that is, how core grammatical relations are marked in Apurinā. In a rather simplistic way, it is possible to distinguish at least the grammatical relations subject and object in this language. The distinction between subject and object here is primarily based on morphosyntactic evidence, namely the person markers on the verb (cf. Table 1). As will be seen in

2. The term 'argument' is used here to refer to the core grammatical relations. 'Core' is defined to mean the grammatical relations which are structurally required by the verb as part of its subcategorization frame. At this point only subject and direct object are clearly core grammatical relations in Apurinā (see also fn.3).

3. The discussion of grammatical relations and constituent order in this paper does not consider clauses with possible trivalent verbs. There are interesting phenomena in Apurinā related to the verbs which commonly behave as trivalent verbs across many languages; although they may be important in arguing about grammatical relations and constituent order, their description will be delayed until the results of further research are available.
(1.2), such person markers may be coreferential with certain overt NOMINALS.  

<table>
<thead>
<tr>
<th>PERSON</th>
<th>SUBJECT</th>
<th>OBJECT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>PL</td>
</tr>
<tr>
<td>1</td>
<td>nī-/ñí-/n-</td>
<td>a-</td>
</tr>
<tr>
<td>2</td>
<td>pī-/pí-/p-</td>
<td>hi-/h-</td>
</tr>
<tr>
<td>3M</td>
<td>i-/i-/φ-</td>
<td>...-na</td>
</tr>
<tr>
<td>3F</td>
<td>u-/û-</td>
<td>...-na</td>
</tr>
</tbody>
</table>

Note that the pronominal marker system above allows the identification of person, number and gender of both subject and object. As the same set of subject pronominal markers is used to refer to both intransitive and transitive subjects while the set of object markers is only used to refer to objects, the system of grammatical relations follows the nominative-accusative pattern.

1.2 Constituent Order Variation

There are six logical possibilities for the relative order of subject, verb and object: SOV, SVO, OSV, VSO, OVS and VOS. Of the six logical possibilities only VSO has not been found. Pickering (1974) has stated that the SO sequence is simply ungrammatical; however, at least in elicitation it is possible to collect cases of SOV. An analysis in detail of the discourse-pragmatic functions of the several constituent orders given in this paper is still required and, therefore, is not discussed here.

The sentences in (1-3) show one of the nominals referring to subject or object in preverbal and another one in postverbal position, i.e. OVS and SVO orders. As the ungrammaticality of the sentences in (4-5) indicates, while the postverbal nominals ARE coreferential with the pronominal marker on the verb, preverbal

---

4 For the time being, until their status is discussed further, pre- and postverbal nominal constituents will be simply referred to as NOMINALS.

5 It appears that all allomorphs of person markers are morphophonologically conditioned. Not all possible allomorphs are listed here; some forms also have a nasal counterpart whose conditioning has not yet been clearly determined.

6 The plural for the 3rd person masculine or feminine is formed by the prefixation of the subject marker of the 3rd person singular masculine or feminine plus the plural suffix -na in the end of the verb.
nominals are NOT.  

(1) O  V  S
   iwa  u  -mapuruka  uwa
   it  3SG.FEM.SUBJ-root.up  she
   'She rooted it up'

(2) S  V  O
   uwa  mapuruka-ro  iwa
   she  root.up  -3SG.MASC.OBJ  it.MASC
   'She rooted it up'

(3) S  V  O
   iwa  mapuruka-ro  uwa
   he  root.up  -3SG.FEM.OBJ  it.FEM
   'He rooted it up'

(4) S  V  O
    *iwa  i  -mapuruka-ro  uwa
    he  3SG.MASC.SUBJ-root.up  -3SG.FEM.OBJ  it.FEM
    (he rooted it up)  

(5) O  V  S
    *uwa  i  -mapuruka-ro  iwa
   it.FEM  3SG.MASC.SUBJ-root.up  -3SG.FEM.OBJ  he
   (he rooted it up)

Moreover, the ungrammaticality of the sentence in (6), in contrast with the one in (7), indicates that once both nominals are postverbal the nominal referring to the object must precede the subject; thus, VOS is allowed, but VSO is not.

(6) V  S  O
    *a  -mapuruka-ro  ata  iwa
  1PL.SUBJ  -root.up  -3SG.MASC.OBJ  we  it.MASC
  'We rooted it up'

---

7 Abbreviations and special symbols used:

O  Free Object  S  Free Subject
O- Bound Object  s- Bound Subject
MASC Masculine  FEM Feminine
SG Singular  PL Plural
# High Central Unround Vowel  c Flosive Alveo-Palatal
ã Nasal Palato-Alveolar  s Fricative  

8 Aberdoor (1985) does mention cases of preverbal nominals expressing grammatical relations which are coreferenced by person markers on the verb. Nevertheless, such examples appear to be extremely rare in her texts. This fact offers some clues about the development of pronominal clitics which, however, are beyond the scope of this paper.
The additional examples below confirm that preverbal nominals referring either to subject or object are not cross-referenced by person markers on the verb. In (8,11) both (OSV, SOV) nominals precede the verb, which bears no person marker. In (9-10 and 12-13) the sentence is not accepted when either of the nominals is cross-referenced on the verb.

(8) O S V
    iwa  ata  mapuruka
    it  we  root.up
    'We rooted it up'

(9) O S V
    *iwa  ata  mapuruka-ri
    it.MASC  we  root.up -3SG.MASC.OBJ
    (we rooted it up)

(10) O S V
    *iwa  ata  a  -mapuruka
    it.MASC  we  1PL.SUBJ-root.up
    (we rooted it up)

(11) S O V
    ata  iwa  mapuruka
    we  it  root.up
    'We rooted it up'

(12) S O V
    *ata  iwa  mapuruka-ri
    we  it.MASC  root.up -3SG.MASC.OBJ
    (we rooted it up)

(13) S O V
    *ata  iwa  a  -mapuruka
    we  it  1PL.SUBJ-root.up
    (we rooted it up)

As might be expected, nominals referring to the subjects of intransitive verbs follow the same behavior as those referring to transitive subjects. In (14) the subject precedes the verb and no pronominal marker is attached to the verb; in (15) the presence of the pronominal marker on the verb when the nominal is preverbal leads to an ungrammatical sentence. In (16) the nominal is postverbal and the pronominal marker occurs, whereas in (17) the lack of the pronominal marker with a postverbal nominal causes the sentence to be ungrammatical.
Based on the examples given up to now, it is possible to summarize the constituent orders in Apuriná as follows:

### Table 2: Apuriná Core Constituent Orders

<table>
<thead>
<tr>
<th>Transitive</th>
<th>Intransitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>O S V</td>
<td>O V S S S V O</td>
</tr>
<tr>
<td>S O V</td>
<td>*V S O S V S</td>
</tr>
</tbody>
</table>

Looking at the table above, one would tend to postulate that Apuriná is almost completely "free" constituent order language. In the next section, some arguments against a "free" constituent order language are discussed.

2. The Syntactic Status of Nominals and Person Markers

There is a problem for the interpretation of Table 2 above as presenting indications of a "free" constituent order language. Such an interpretation only works under the assumption that pre- and postverbal nominals have identical syntactic status. This section presents arguments against that assumption.

2.1 Postverbal Nominals as Adjuncts

In all the examples given above which show at least one postverbal nominal, such a nominal can be missing. For instance, the examples in (18-19), in contrast with those in (1-2), show that postverbal nominals occur optionally in a clause when there is a coreferential person marker on the verb.
(18) O  V
    ivate    u    -mapuruka
it.MASC  3SG.FEM.OBJ  -root.up
'She rooted it up'

(19) S  V
    uwa  mapuruka-rí
she  root.up -3SG.OBJ.MASC
'She rooted it up'

Thus, Table 2 above can be revised as follows:

<table>
<thead>
<tr>
<th>Transitive</th>
<th>Intransitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>O S V</td>
<td>V (O) (S)</td>
</tr>
<tr>
<td>S O V</td>
<td>*V (S) (O)</td>
</tr>
<tr>
<td></td>
<td>S V</td>
</tr>
<tr>
<td></td>
<td>V (S)</td>
</tr>
</tbody>
</table>

In addition, in (20-21) the sentences show that neither the postverbal nominal referring to object nor the postverbal nominal referring to subject can occur without the coreferential verb person marker. This fact should falsify the claim that the postverbal nominal triggers an 'agreement' marker on the verb. Also, if a clause can 'lose' postverbal nominals without affecting its propositional content, it seems strange to argue that such nominals express arguments of the verb

(20) *uwa  mapuruka  ivate
she  root.up  it.MASC
(she rooted it up)

(21) *ivate  mapuruka  uwa
it  root.up  she
(she rooted it up)

The label OPTIONAL is used here to refer to the quality that a nominal may have of OPTIONALLY occurring in a clause when it corefers to a person marker on the verb. Since only postverbal nominals can (and have to) be coreferential with a verbal person marker, a distinction can be drawn between pre- and postverbal nominals: The latter are optional whenever they can occur whereas the former are not. This distinction can be used as an evidence that these pre- and postverbal elements have a different syntactic status. The hypothesis is that while preverbal nominals express core grammatical relations, those which are postverbal are adjunctive elements. Therefore, since postverbal nominals are optional whenever they occur, it is possible to claim that they behave syntactically as oblique elements. Such oblique elements
will be labeled **ADJUNCTS**.\(^9\)

### 2.2 Preverbal Nominals as Core Grammatical Relations

The evidence used here for categorizing postverbal nominals as adjuncts does not apply to the preverbal ones. Preverbal nominals are NOT optional whenever they occur. The optionality of postverbal nominals can be seen in the examples given in (2) and (19) (which are repeated for convenience in (22) and (23)). The non-optionality of preverbal nominals can be seen by contrasting the examples (22) and (23) with the one in (24): (24) is ungrammatical because the preverbal nominal referring to the subject is missing. Since preverbal nominals are not cross-referenced on the verb, once they are missing there has to be another element to express the argument of the verb (cf. the subject marker in (25)). The fact that the verb person marker shows up in the absence of a preverbal nominal (or vice-versa) is a clue to the role played by the former in a clause, as will be seen in the next subsection.

\[
\begin{align*}
(22) & \quad S \quad V \quad O \\
& \quad \text{uwa} \quad \text{mapuruka-rì} \quad \text{ìwa} \\
& \quad \text{she} \quad \text{root.up} \quad -3SG.MASC.OBJ \quad \text{it.MASC} \\
& \quad '\text{She rooted it up}'
\end{align*}
\]

\[
\begin{align*}
(23) & \quad S \quad V \\
& \quad \text{uwa} \quad \text{mapuruka-rì} \\
& \quad \text{she} \quad \text{root.up} \quad -3SG.MASC.OBJ \\
& \quad '\text{She rooted it up}'
\end{align*}
\]

\[
\begin{align*}
(24) & \quad V \\
& \quad \text{*mapuruka-rì} \\
& \quad \text{root.up} \quad -3SG.MASC.OBJ \\
& \quad (\text{she rooted it up})
\end{align*}
\]

\[
\begin{align*}
(25) & \quad V \\
& \quad u \quad -\text{mapuruka-rì} \\
& \quad 3SG.FEM.SUBJ\text{-root.up} \quad -3SG.MASC.OBJ \\
& \quad '\text{She rooted it up}'
\end{align*}
\]

Therefore, in a certain way it is possible to argue that free-standing nominals may or may not express verb arguments in Apurinā depending on whether they precede or follow the verb in a clause.

### 2.3 Person Markers as Clitic Arguments

Given the evidence that, due to their optionality, postverbal

---

\(^9\) The syntactic categorization suggested here for postverbal nominals in Apurinā has already been suggested elsewhere (Jelinek (1984) and Hale (1990)) for nominals with similar behavior in other languages.
nominals cannot be the verb arguments, one, then, needs to ask what syntactically represents the verb arguments in sentences with no preverbal nominal. As shown above, preverbal nominals cannot be optional whenever they can occur because they are not cross-referenced by verb person markers. Thus, preverbal nominals and verb person markers are in complementary distribution. The reason to be in complementary distribution is that both accomplish the same syntactic function; that is, they play the same syntactic role in a clause. This role is to express the argument of the verb.

As a syntactic element, rather than simply a morphological affix attached to the verb, the person markers behave as PRONOMINAL CLITICS. To consider person markers as pronominal clitics does not mean simply to find another label for an atypical affix; more than that, it is to try to describe more precisely the syntactic behavior and function of such an element.

As the syntactic elements which express verb arguments, the clitics behave as pronouns which are attached to the verb; that is, clitics seem to function as normal pronouns, except that they are phonologically bound morphemes. Furthermore, the difference, for instance, between a free-standing pronominal and a clitic is that while the former has the typical syntactic distribution of a nominal which can function as subject or object grammatical relation, the latter has a distribution which is morphologically determined. The distribution of the clitics is morphologically determined in that they fill up a fixed slot in a verbal construction formed by morphological rather than syntactic operations.

Nevertheless, although free-standing nominals expressing arguments find themselves identified by means of syntactic rather than morphological rules, their syntactic function are basically the same: Both free-standing nominal arguments and clitic arguments function as the syntactic bearing elements of verb arguments.

3. The Syntactic Analysis

It has been said above that the arguments of the verb can be expressed either as clitics or as free-standing nominals; however, nominals which co-occur with (and are coreferential with) clitics are NOT arguments of the verb but, rather adjuncts. In this section, the discussion is focused on some possible syntactic implications of the stated analysis. Configurationality and constituent order are issues related to the clitic and preverbal arguments plus postverbal nominal adjunction claim.

3.1 The Theoretical Notion of Non-configurationality

As Jelinek (1984) has pointed out, the recent interest on the
non-configurational property of some languages has been motivated principally by Ken Hale's work on Australian and Native American languages. The initial discussion by Hale (1980, 1981, 1982, 1983) aimed to account for some of the characteristics usually found in non-configurational languages. Some of these supposedly common non-configurational properties would be free clausal constituent order variation, syntactically discontinuous expressions and null anaphora. In these works Hale suggested some parameters of configurationality, which have been revised by Jelinek (1984).

For a better understanding of the notion of non-configurationality, it may be helpful to look at one Warlpiri example. "In the following Warlpiri sentence, any word order is possible, with the provision that the AUX clitic sequence occur in the second position.

(26) Ngarrka-ngku ka waviri panti-rni.
    man-ERG AUX kangaroo spear-NONPAST
    'The man is spearing the kangaroo'"
    (Jelinek 1984:39-40)

Jelinek has claimed that the arguments of verbs in Warlpiri are expressed by the clitics;\(^\text{10}\) the nominals which are coreferential with the clitics would be optional and, thus, non-argumental features. Finally, based on these ideas, she has proposed an extended configurationality parameter for languages which share some of the Warlpiri grammar features. This configurationality parameter would be as follows:

(27) "a. In a configurational language, object nominals are properly governed by the verb.
    b. In a W-[Warlpiri]type non-configurational language, nominals are not verbal arguments, but are optional adjuncts to the clitic pronouns that serve as verbal arguments."
    (Jelinek 1984:73)

This new analysis of the non-configurational properties in Warlpiri has been endorsed in Hale's 1990 Core Structures and Adjunctions in Warlpiri Syntax:

"there might exist languages... whose free word order results simply from the fact that (certain or all) overt phrasal expressions are adjuncts..."  
(Hale 1990:36-7).

Considering such a theoretical approach to constituent order variation, it may be interesting to draw some attention again to the Apurinã constituent order variation described above.

\(^{10}\) For details and examples, see Jelinek 1984.
3.2 Configurationality in Apurinā

The summary of constituent orders given in Table 2 above suggested a system of partially free constituent order variation in Apurinā. Revising Table 2, Table 3 represents the occurrence of postverbal nominals as adjuncts. Under the analysis of pronominal clitics as arguments of the verb and postverbal nominals as adjuncts, Table 3 should be further revised as follows:

Table 4: Apurinā Core Constituent Orders

<table>
<thead>
<tr>
<th>Transitive</th>
<th>Intransitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>O S V</td>
<td>O s-V</td>
</tr>
<tr>
<td>S O V</td>
<td>s-V-o</td>
</tr>
<tr>
<td>s-V-o</td>
<td>S V</td>
</tr>
<tr>
<td>s-V</td>
<td>S-V-o</td>
</tr>
</tbody>
</table>

Such a revision is due to the fact that if postverbal nominals are adjuncts, they cannot also be core grammatical relations at the same time. With respect to the analysis of the transitive verbs, the result of this revision is a three-way system of syntactic expressions of the verb argument structure: (i) Both the verb arguments can be syntactically realized as free-standing nominals, or (ii) both the arguments can be simultaneously realized as clitics, or yet (iii) one verb argument can be realized as a free-standing nominal while the other is realized as a clitic.

Analogously, in relation to the intransitive verbs, the verb argument can be syntactically expressed as either (i) a free-standing nominal or as (ii) a clitic.

Transitive and Intransitive verb sentences can be grouped into three order types based on the syntactic realization of their arguments in a sentence, as seen in Table 5 below. Type I groups the sentences with only phonologically free-standing syntactic elements expressing arguments; Type II groups only phonologically bound syntactic elements expressing arguments; and Type III groups sentences which mix Types I and II.

Table 5: Grammatical Relations Organization

<table>
<thead>
<tr>
<th>Type I</th>
<th>Type II</th>
<th>Type III</th>
</tr>
</thead>
<tbody>
<tr>
<td>O S V</td>
<td>s-V-o</td>
<td>O s-V</td>
</tr>
<tr>
<td>S O V</td>
<td></td>
<td>S V-o</td>
</tr>
<tr>
<td>S V</td>
<td></td>
<td>s-V</td>
</tr>
</tbody>
</table>

For the purpose of analyzing configurationality in the distribution of the syntactic elements expressing arguments in Type I, only sentences with transitive verbs are relevant here. The indications are that OSV and SOV can be used interchangeably as long as their interpretation is not ambiguous. However, when
ambiguity exists, the object is required to precede the subject. For instance, the variation seen in the unambiguous examples in (28-31) cannot occur in ambiguous examples like (32-33). The second interpretation of the sentences in (32-33) is not possible when the context does not disambiguate; when this happens, the position of the syntactic elements is fixed and, therefore, the order of the constituents is configurational.

(28)  O  S  V
      yuwata  nota  etama
      knife  I  see
      'I see the knife'

(29)  O  S  V
      nota  yuwata  etama
      knife  I  see
      'I see the knife'

(30)  O  S  V
      hākiti  kīkī  keta
      jaguar  man  shoot
      'The man shoots the jaguar'

(31)  S  O  V
      kīkī  hākiti  keta
      man  jaguar  shoot
      'The man shoots the jaguar'

(32)  O  S  V
      anāpa  kīkī  etama
      dog  man  see
      'The man sees the dog'
      *'The dog sees the man'

(33)  O  S  V
      Pedro  Paulo  keta
      shoot
      'Paulo shoots Pedro'
      *'Pedro shoots Paulo'

Contrasting Type I with Types II and III, and following Jelinek's configurationality parameter given in (25), the tendency would be to argue for a partial configurationality in Apurinã. This tendency follows from the configurationality parameter because preverbal nominals (as in OSV, Os-V and SV-o) ARE arguments of the verb, whereas postverbal nominals (as in SV-o(O), Os-V(S) and s-Vo(O)(S)) are NOT verbal arguments, but rather optional adjuncts to the clitic pronouns which are the verbal arguments. Therefore, while Type I would be configurational, since the clausal constituent order is syntactically relevant, and Type II would be non-configurational, since postverbal nominals are adjuncts and clitics are arguments, Type III would be both.

However, if we adopt a notion of grammatical relations which
is not necessarily defined only in terms of hierarchical structures of clausal constituents (e.g. NPs, VPs or N"s, V"s), but mainly in terms of how a nominal element can syntactically interact with the verb element in a clause, there may be an alternative way to analyse the Apurinā system summarized in Table 5.

The syntactic realization of verb arguments in Type II sentences, which consists only of clitics attached to the verb, has a fixed order. The order of the clitics is morphologically determined by the position class they occupy in the verbal construction. However, syntactically these clitics are in complementary distribution with free-standing grammatical relations; and lexically, as portmanteau morphemes, they bear grammatical information as case roles (nominative-accusative, cf. Table 1), person, and gender (feminine and masculine). Therefore, considering the syntactic behavior of the clitics, the Type II does not necessarily poses a problem to a notion of configurationality based on the functional features of the clausal constituents. That is, the only additional feature is that the arguments are phonologically realized as bound morphemes.

Finally, Type III which is constituted of a mixture of types falls out from the description of Types I and II. In a language that allows the syntactic expression of verb arguments by means of free-standing as well as bound morphemes, a hybrid kind of argument expression, including both the Types I and II should be expected to occur. Rather than posing a problem, Type III reinforces the analysis suggested for the first two types.

3.3 Basic Constituent Order

The criteria usually used to determine the basic constituent order of languages can be grouped into three sets: Descriptive simplicity, statistical frequency, and pragmatic neutrality (following Mithum 1992). It is beyond the scope of this paper to present a study on pragmatic values of clausal constituent orders. What will be mostly considered here is the descriptive simplicity criterion and a very preliminary study on statistical frequency.

As has been mentioned earlier in this paper, in previous analyses of Apurinā constituent order (Pickering 1974, Derbyshire & Pullum 1985, and Facundes 1992b), the suggested basic constituent order was based on the descriptive simplicity criterion: Having assumed that pronominal markers on the verb were agreement triggered by the postverbal subject and object, all the constituent order patterns found could be derived from an unmarked one, namely OSV.11

11 Aberdoor (1985) also did some work on Apurinā which included statistical frequency and discourse-pragmatic functions; however, her analysis is based on the assumption that pronominal markers are verb agreement markers and that missing nominals are result of zero anaphora.
The descriptive simplicity criterion permits one to say that OSV is more basic than SOV, since the occurrence of the latter is predictable. SOV can occur only when there is no ambiguity in a clause. Another indication in favor of OSV would be that, at least in texts, the adjuncts which are coreferential with the clitics can only occur in the sequence OS.

The decision of whether or not to consider order types like Os-V and s-V-o in the analysis of constituent order is usually related to pre-established theoretical assumptions. For instance, one assumption could be to consider as pertinent for the analysis of constituent order only the occurrence of the free-standing grammatical relations; another one could be to postulate that free-standing elements tend to be more neutral than bound morphemes.

Based on the description of the system of grammatical relations suggested here, one would tend to consider bound clitics as relevant in analysing Apurinã constituent order. If clitics behave as normal subjects and objects, except that they are phonologically bound, their role on constituent order might be as important as that of any other subject or object.

On the other hand, one of the possible consequences of the phonological attachment of clitics to verbs is that the clitics, then, follow the rules of the morphology and no longer of the syntax. What this might mean is that additional syntactic or morphological tests are required to establish the relevance of argument clitics for an analysis of basic constituent order which is based only on the descriptive simplicity criterion.

The statistical frequency criterion would lead one to choose OSV as more neutral than SOV, since SOV does not seem to occur in text but only in elicited data (cf. Facundes 1993) or in quotative clauses (cf. Aberdoor 1985). However, OSV is extremely rare in frequency (0.9%).

Os-V and s-V-o occur with equal frequency (33%) and are the most frequent order types, which makes either of them good candidates for basic order and reinforces the idea that bound grammatical relations are (statistically) relevant in defining basic constituent order. SV-o, however, occurs with low frequency (2.5%).

Based on a structural analysis including formal rules, an alternative syntactic analysis of constituent order would postulate a right-dislocation to generate postverbal nominals followed by the attachment of the person markers on the verb. If clitics are not considered, the result of this analysis would be that OSV would be the basic order of the language. This approach to the data, however, would require further details (i.e. deletion of postverbal nominals) in order to account for the optionality of such nominals. No motivation for right-dislocation has been found up to now, and, besides, the evidence that postverbal nominals are adjuncts rather
than core grammatical roles also poses problems to such an analysis.

Also based on formal rules, another alternative syntactic analysis of constituent order would be to postulate a left-deslocation of postverbal nominal(s) and person marker(s) deletion. By this analysis the basic constituent order would be VOS. Such an approach, however, only would account for the occurrence of preverbal nominal in a formal description, but would not say anything about their syntactic function, or their syntactic status in contrast with the status of the postverbal nominals.

There is no definitive hypothesis to be presented at this point in relation to a basic constituent order. Additional information about the grammar of this language might provide better clues. Nevertheless, it seems pretty clear that a synchronic description of aspects of the Apurinã syntax must involve a certain degree of complexity.

4. Conclusion

The ideas about the grammar of Apurinã described here were intended to show how the verb argument structure is syntactically expressed and how it corelates with constituent order. Such ideas are bound into the synchronic internal evidence of the language. Historical considerations and discourse-pragmatic functions were delayed until the results of further research are available.

The initial appearance of "free" constituent order is not actually confirmed when the grammar of the language is studied more carefully. Constituent order is relevant for Apurinã syntax, which motivates the attempt to describe its unmarked, most frequent or most neutral realization. However, this last task requires additional research; any hypothesis about basic constituent order would be more theoretically dependent than motivated by the internal language structure. Nevertheless, the preliminary description of Apurinã shows a corelation which may exist between pronominal clitics, argument roles and configurationality.

5. References


Facundes, Sidney da S. 1990. Levantamento Linguístico do Rio Purus. Relatório de pesquisa para o Conselho Nacional de
Pesquisas, Brazil.

-------. 1992a. *Split Case-Marking in Apuriná (Arawakan, Maipuran)* (ms.)


-------. 1993. *Apuriná Word Formation* (ms.)


CASE, VERB TYPE AND ERGATIVITY IN TRUMAI

Raquel Guirardello
Museu Paraense Emílio Goeldi (Pará - Brazil)

1. Introduction

This paper will present a short study of the Trumai verbal case-marking system.

In a first approach to the data, it would be possible to say that Trumai has four types of verbs: intransitive, transitive, ditransitive, and a fourth type which is our main interest in this paper, for although it seems to be transitive (it basically has two participants), it is morphosyntactically distinct from prototype transitive verbs. We will show that (a) this fourth type is in reality intransitive, with one participant the subject and the other an oblique NP; we will label these verbs extended intransitives; (b) these four verbal types collapse into only two basic types, intransitive and transitive.

The lexical category of extended intransitive verb is motivated by both semantic features of the verbs themselves - i.e. the actions which these verbs express take a location rather than a patient for a second participant - and by pragmatic features of the second participant for those verbs which do take patients as second participants - i.e. they are stereotypical or otherwise of little importance. Such pragmatic factors are also important in the choice of when to use an extended intransitive verb versus a transitive verb when either is available, e.g. for pairs of the type *fa* 'kill, hit' and *disi* 'kill, hit', one of which is transitive and the other extended intransitive. In essence, the speaker chooses which verb to use depending upon the topicality of the patient/locative second participant.
2. Description: the verb types and the case-marking frames in Trumai

First, we will illustrate the four verb types of Trumai, showing how the morphosyntactic system is organized.

The object of a transitive verb (O) receives the same treatment as the subject of an intransitive verb (S), that is: the same case-marking (unmarked); the same position in the clause (in adjacency with the verb); occurrence of the third person clitic -n when the overt nominal does not occur; and a certain control of the position of these functions in the case of lexical item deletion, through the i (or ii) morpheme. In contrast, the subject of a transitive verb (A) receives a distinct case-marker (the suffix -k), can vary in position in the clause and, when omitted, leaves behind neither a marker on the verb nor another morpheme anaphorically (such as the i/ii of absolutive). Since S and O pattern together morphosyntactically, and since A receives its own unique morphosyntax, this language is clearly ergative-absolutive. There are other ways in which ergativity manifests itself, as for instance the imperative construction, which we will treat later. We now offer illustrative examples of the four verb types in Trumai.

Type 1 - Intransitive verbs: Absolutive S (unmarked); verbal cross-reference to Abs (V-n '3 Abs')

The intransitive subject can be expressed by a lexical NP (examples 1,2) or, if that is deleted, by means of the third person enclitic -n (example 3). In the case of first and second persons, the use of pronouns is obligatory, that is, these pronouns cannot be deleted. The third person enclitic also presents the allomorph -e for verbs which end with a consonant.

Examples:
S  V
(1)  ha-Ø pita
    1-Abs  go.out
    'I go out'

S  V
(2)  hine-Ø pita
    3-Abs  go.out
    'He goes out'

V-n
(3)  pita-n
    go.out-3Abs
    'He goes out'

(4) Dative marking goal of action

GOAL    S  V
ole-g  ka in  ha-Ø kawa
manioc-dat  ?  ?  1-Abs  go
'I'm going to get manioc (right now)'
(lit. 'I'm going for manioc')

(5) Dative as mark of locative

S  V  LOC
ha-Ø  ara'tsi  ka in  tehene-ki
1-Abs  sit.down  ?  ?  floor-dat
'I sat down in the floor'

Type 2 - Transitive verbs: Ergative A (-k/-ts),
Absolutive 0 (-Ø; V-n)

The subject of a transitive verb is marked with the suffix
-k (cf. examples 6,8) which is preceded by epenthetic vowel e or a
when attached to words which end in a consonant; the first person
singular allomorph of the ergative marker is -ts (cf. examples 7
a-b). The third person object (absolutive) nominal (or free
pronoun) can be deleted; in this case, the clitic -n/-e occurs
(cf. 7 a-b). It is possible to find some variation in the word
order (AOV, OVA - cf. 8 a-b) but the order AVO is not allowed (cf. 8c) and the orders OAV, VAO and VOA are not attested. We see then that the sequence OV is not broken.

(6) hine-k atlat-Ø mapa
3-Erg pan-Abs break
'He broke the pan'

(7) a. hai-ts kasoro mud-Ø husa
1-Erg dog neck-Abs chain
'I chained the dog (by the neck)'
(lit. 'I chained the dog's neck')

b. hai-ts ka in husa-Ø
1-Erg ? ? chain-3Abs
'I am chaining it'

(8) a. wirix ma'may ka in Yaka-k
manioc-pap mix ? ? proper name-erg
'Yaka is mixing the manioc pap'

b. Yaka-k ka in wirix ma'may

c. *Yaka-k ka in ma'may wirix

Type 3 - Ditransitive verbs: Erg A (-k/-ts), Abs O (-Ø;V-n), Postverbal Dat IO (-t1/ -ki/ -s)

In the few verbs of this type, we can also see that the order OV is preserved. The indirect object (IO) is not obligatory; unlike O (but like the A), the IO nominal (or free pronoun) can be omitted with no resultant marker on the verb. For instance:
Type 4 - Extended Intransitive: Abs Agent (-∅; V-n), postverbal Dat Patient (-tl/ -ki/ -s)

The extended intransitive verb class is interesting because semantically such verbs seem to be transitives, for they can present two participants, with one of them the agent and the other one the presumed patient of the action. The problem is that, unlike what we found for other transitive verbs, the agent here is treated as the absolutive and the 'patient' as an indirect object. See the examples below:

AGT | V | PAT
(13) kiki-∅ fa hine-tl
man-Abs hit/kill 3 - Dat
'The man hit/killed him'

AGT | V | PAT
(14) kiki-∅ fa kodesiš-eg
man-Abs hit/kill snake - Dat
'The man hit/killed the snake'
AGT  V  PAT
(15) ha-∅  some  cafe-s
    l-Abs  drink coffee-Dat
    'I drank coffee (a lot)'

AGT  V  PAT
(16) ha-∅  some  cafe-ki
    l-Abs  drink coffee-Dat
    'I drank coffee (a little)'

V-3n  PAT
(17) ma-n  hUMAN-ki
    eat-3Abs  bean-Dat
    'He ate beans'

The dative marker varies according to the kind of NP which occurs in this position: singular pronouns and human nouns receive the suffix -tl (with an epenthetic vowel added to the forms which end in a consonant); plural pronouns and human nouns receive the suffix -ki. Non-human nouns can receive two kinds of marking, -tl or -s, according to the verb: for instance, fa 'hit/kill' requires the marker -s, while make 'bite' requires -tl. Thus, extended intransitive verbs can be subcategorized into two classes on the basis of which dative marker they choose for non-human patients. These classes are lexically determined and are mutually exclusive, that is, verbs which require one suffix do not accept the other. Either -tl or -s, however, can alternate with -ki, when the patient consists of a small quantity of something (see examples 15 and 16 above). The following table summarizes the allomorphy of the dative marker:

<table>
<thead>
<tr>
<th>PRON - HUMAN NOUN</th>
<th>NON-HUMAN NOUN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>Verb Class I</td>
</tr>
<tr>
<td></td>
<td>-tl</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>Verb Class II</td>
</tr>
<tr>
<td></td>
<td>-ki</td>
</tr>
<tr>
<td></td>
<td>-s</td>
</tr>
</tbody>
</table>
3. **Analysis: Verb type 4 as morphosyntactically intransitive**

When we look more carefully at the entire system of the Trumai language and its morphosyntactic elements, we are forced to conclude that verb type 4 is actually intransitive. Only one of the participants is essential, the subject; the second participant is not obligatory and can be omitted without troubles. This is prima facie evidence that we are dealing with a syntactically intransitive verb type, albeit one which can be extended by taking the 'Patient' NP as a syntactically IO. Thus the label extended intransitive, which emphasizes that the verb in question is basically intransitive, rather than transitive.

The arguments in support of this analysis are: constituent order, verbal person marking, case-marking, and the morphosyntax of the imperative construction. Observe the following table:

**a) Basic Word Order**

<table>
<thead>
<tr>
<th>Type</th>
<th>Word Order</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRANSITIVE</td>
<td>S V (LOC)</td>
<td>or S V</td>
</tr>
<tr>
<td>TRANSITIVE</td>
<td>A O V</td>
<td></td>
</tr>
<tr>
<td>DITRANSITIVE</td>
<td>A O V (IO)</td>
<td></td>
</tr>
<tr>
<td>EXTENDED TRANSITIVE</td>
<td>S V (IO)</td>
<td></td>
</tr>
</tbody>
</table>

We see from clause types 2 and 3 that the position of O is preverbal, whereas in type 3, the IO (and the oblique LOC in type 1) comes after the verb and is optional (it can vary its position, occurring before the subject, but not before the verb - cf. note 7). In type 4, the erstwhile patient occurs postverbaly and is optional, hence it patterns with IO and LOC rather than O.

**b) Verbal person marking**

<table>
<thead>
<tr>
<th>Type</th>
<th>Person</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRANSITIVE</td>
<td>[ ]</td>
<td>V-n</td>
</tr>
<tr>
<td>TRANSITIVE</td>
<td>A [ ]</td>
<td>V-n</td>
</tr>
<tr>
<td>DITRANSITIVE</td>
<td>A [ ]</td>
<td>V-n IO</td>
</tr>
<tr>
<td>EXTENDED TRANSITIVE</td>
<td>[ ]</td>
<td>V-n IO</td>
</tr>
</tbody>
</table>

As we said before, when the third person S or O nominal (or free pronoun) is omitted, the enclitic -n/-e occurs. If verb type
4 were transitive, the enclitic should refer to the O, as it does in type 2; but on the contrary, it refers to the S.

c) Case System

<table>
<thead>
<tr>
<th>Case Type</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRANSITIVE</td>
<td>S-∅ V Goal-s/-ki</td>
</tr>
<tr>
<td>TRANSITIVE</td>
<td>A-k O-∅ V</td>
</tr>
<tr>
<td>DITRANSITIVE</td>
<td>A-k O-∅ V IO-s/-t1/-ki</td>
</tr>
<tr>
<td>EXTENDED INTRANSITIVE</td>
<td>S-∅ V IO-s/-t1/-ki</td>
</tr>
</tbody>
</table>

Looking at the case system, it is clear that the second participant of the extended intransitive is not an O, for it presents exactly the same case-markers as the indirect object of a ditransitive verb.

d) Imperative

The final argument, which confirms that type 4 verbs are intransitive, is the morphosyntax of the imperative mood in Trumai. The imperative particle wana is employed to mark imperatives for intransitive verbs, while transitive and ditransitives verbs use the particle waki. The extended intransitive verbs pattern with intransitive verbs, taking the particle wana.

<table>
<thead>
<tr>
<th>Case Type</th>
<th>Imperative Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRANSITIVE</td>
<td>wana</td>
</tr>
<tr>
<td>TRANSITIVE</td>
<td>waki</td>
</tr>
<tr>
<td>DITRANSITIVE</td>
<td>waki</td>
</tr>
<tr>
<td>EXTENDED INTRANSITIVE</td>
<td>wana</td>
</tr>
</tbody>
</table>
(18) *wana pita*
   Imp go.out
   'Go out!'

(19) *wana pes hen*
   Imp run then
   'Run, then!'

(20) *waki husa hen*
   Imp chain then
   'Chain (it), then!'

(21) *hine-tl waki kiti*
    3 - Dat Imp give
    'Give (it) to him!'

(22) *wana sone*
    Imp drink
    'Drink (it)!

(23) *wirix-ki wana sone*
    manioc.pap-Dat Imp drink
    'Drink the manioc pap!'  

As this final test indicates most clearly, Trumai basically has not four different morphosyntactic verb types, but only two, TRANSITIVE and INTRANSITIVE, with some transitives (type 3) and some intransitives (type 4) extended by means of the adjunction of an optional indirect object. Although the morphosyntax is clear, we are left with a puzzle: why would semantically transitive (i.e. two-participant) verbs be obligatory codified in morphosyntax as intransitive?
4. A possible explanation: semantically and pragmatically reduced transitivity

The fact that an action involves two participants doesn't necessarily mean it has high transitivity. As shown by Hopper and Thompson (1980), transitivity involves not only the number of participants, but an entire set of components, such as volitionality of the agent, punctuality and telicity of the action, the degree of affectedness of the patient, individualization of the patient, etc. The relevance of these components can vary from language to language in determining how a given action will be codified in morphosyntax.

The Trumai language codifies as intransitive verbs those two-participant verbs which have inherently reduced transitivity, due to either semantic features of the verb (section 4.1) or pragmatic features of the second participant, the erstwhile patient (section 4.2).

4.1. The semantic factor: locative object verbs (eg. 'bite', 'hit', etc)

In some actions, the second participant (i.e. the non-agent) is not a true patient (which is completely affected by the action), but rather is a kind of location. That is, while contact is made with the second participant, the effect of the contact may or may not be transferred to it (cf. Hopper and Thompson's 1980 conclusion that 'transferral' of the action is the most basic component of transitivity). It is thus more basically a location where the action occurs than a patient affected by the action. Some languages, such as Trumai, mark the difference between patient-objects and locative-objects in surface morphosyntax. Other languages, such as English, mark this difference only in syntactic variation (cf. Fillmore 1970). Observe the following examples:

(24) (a) I hit him (he may be affected)
(b) I hit at him (he is not affected)

c) I killed him (he is affected)
(d) * I killed at him
(25) (a) I gave him a hit
(b) * I gave him a kill

From such examples, Fillmore (1970) argues that the erstwhile patient of hit is a location (in example 25, in reality, it is a recipient, which is a kind of metaphorical location). In (24a), the blow arrives at the second participant, but the second participant may or may not be affected. In (24b) the blow is aimed at the second participant, but either the blow does not connect or it has no effect; the use of a locative preposition to mark the second participant argues for a semantic case role of location rather than patient.

The second participant of kill, in contrast, is necessarily affected; therefore it is a patient to whom the action is transferred (and, for this reason, the use of a locative preposition here is impossible). In the case of hit the actual patient is the blow which is created through the action, and the erstwhile patient him is a kind of location where the 'real' patient is created. Example (25) shows this more clearly: it is possible to give a hit (the patient which is invisible in 24a-b) to the second participant (who is recipient), but is impossible to give a kill to the second participant, precisely because he IS the patient.

Different languages treat the lexical category of locative object verbs differently. Where English uses syntactic variation, Lhasa Tibetan allows only the syntax of (25) above for such verbs (DeLancey - p.c. with Gildea), and in Trumai, type 4 verbs allow only the syntax of (24b) above. A brief list of the semantically conditioned type 4 verbs includes:

The Trumai Extended Intransitive Locative Object Verbs

make 'bite' ; fa 'hit/kill'
xom 'suck' ; laxod 'smell (action/perception)'
xu'tsa 'look/see' ; fa'tsa 'listen/hear'
lax 'hunt'

Verbs like 'hunt' also have a locative object, for the object need not be necessarily present to conduct the action (we can hunt.
all day and at the end of hunting have never even encountered a likely patient).

4.2. The pragmatic factor: the non-topical patient

With some verbs the patients are habitual, very predictable, often indefinite and unindividuated. Most of the remaining extended intransitive verbs in Trumai are of this type. While the agent is topical, the patient has little importance or individuation and is often stereotypical. The Agents of 'eat' and 'drink', for example, are affected by these actions in a way that is likely to be more salient to a speaker than the effect on the thing eaten. Hopper and Thompson (1980) note that lowered topicality in a patient is likely to lead to less transitive morphosyntax, such as antipassives, etc. A subclass of type 4 verbs in Trumai are of this type, where stereotypically non-topical patients obligatorily occur as indirect, rather than direct, objects:

The Trumai Extended Intransitive Verbs with non-topical patients

\[\text{xoxan 'wash' ; sone 'drink' ; ma 'eat' ; maska 'sew'}\]

At this point, we should point out some potential problems for our analysis: most troubling is the verb \textit{suda} 'make (something)' which in Trumai is an extended intransitive. This verb is a problem for the explanation given above because the object of the action falls into neither of the categories above: it is not predictable nor it is semantically a locative, but rather a patient. It is hard to understand why the language would codify this apparently fully transitive action with intransitive morphosyntax. In another way as well, the Trumai system is not totally coherent. Like \textit{lax} 'hunt', \textit{padi} 'wait' is neither a telic action nor must its object be necessarily present. But while 'hunt' in Trumai is extended intransitive, 'to wait' receives fully transitive morphosyntax.
5. The pragmatic uses of transitive/extended intransitive pairs

In Trumai, some pairs of verbs are semantically equivalent (that is, they express the same action), but belong to different morphosyntactic categories. For each of these verbs it is possible to obtain a paradigm with all persons of subject and object (1st, 2nd, 3rd) and with nouns. These verbs are the followings:

<table>
<thead>
<tr>
<th>EXTENDED INTRANSITIVE</th>
<th>TRANSITIVE</th>
<th>GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>fa</td>
<td>disi</td>
<td>'hit/kill'</td>
</tr>
<tr>
<td>make</td>
<td>tako</td>
<td>'bite'</td>
</tr>
<tr>
<td>dama</td>
<td>tuxa'tsi</td>
<td>'pull'</td>
</tr>
</tbody>
</table>

These verbs alternate with one another in actual use: it is possible to use either the extended intransitive or the transitive forms any time the basic concept needs expression. What is interesting to see is that speakers seem to select which category of verb to use based on the persons of agent and patient. When first person acts on third, speakers prefer the extended intransitive form; when third acts on first, the preferred verb used is the transitive.

This preference was observed in two kinds of data:

1. In elicitation of paradigms from various consultants; e.g., when the paradigm was requested for 'hit', the consultants systematically selected the verb type depending on the person of A and O: 1A -> 30, 1A -> 20 and 2A -> 30 were most commonly expressed with extended intransitive verbs, whereas 3A -> 10, and 2A -> 10 were most commonly expressed with transitive verbs. When only third person was involved, the changes were not so systematic;

2. Texts. For example, Monod-Becquelin (1976) observes that in a text where a Trumai person tells about the killing of an uncle by the Kayabi indians, the transitive verb ('form ergatif' in her terminology) occurs very frequently when the consultant speaks about the actions of the aggressors (Kayabi) on the victims (Trumai). While telling of the revenge (Trumai on Kayabi - the Trumai indian speaks of the Kayabi people) the intransitive
('construction Sujet-Objet' in her analysis) occurs. Here we have again the difference between NPs:

attack: Kayabi (3rd pl) on Trumai (1st pl) = transitive
revenge: Trumai (1st pl) on Kayabi (2nd pl) = intransitive

But sometimes the selection does not occur in the expected way (that is, it is a tendency rather than a rigid rule).

This new pattern can be explained by the difference in topicality in objects: first person is inherently more topical than third person, for the speaker will always consider himself to be the center of speech.\(^9\) Morphosyntactically, first person is given a topical position, as the object. As third person has inherently lower topicality, it tends to hold the less topical position of indirect object. From this follows the selection of verbal forms:

(26) 3Agt --> lPat = inherently topical object uses the transitive verb.

\[
\begin{array}{ccc}
\text{O} & \text{V} & \text{A} \\
\text{ha-ta} & \text{disi-tke} & \text{ka in hinak wan-ek} \\
\text{l-Abs} & \text{hit/kill-Desid} & ? \ ? \ 3 \ \text{pl} \ \text{-Erg} \\
\text{They want to kill/hit me}' \\
\end{array}
\]

(27) lAgt --> 3Pat = inherently less topical object uses the extended intransitive verb

\[
\begin{array}{ccc}
\text{A} & \text{V} & \text{IO} \\
\text{ha-ta} & \text{fa-tke} & \text{ka in hine-t1} \\
\text{l-Abs} & \text{hit/kill-Desid} & ? \ ? \ 3 \ \text{-Dat} \\
\text{'I want to hit/kill him'} \\
\end{array}
\]

It is interesting also to observe that, in the case of \textit{fa} and \textit{disi}, there are two possible meanings for these verbs: 'hit' and 'kill'. Since 'hit' takes a locative object and 'kill' does not, one might ask how it is that (a) a single verb can mean both and (b) one verb with both meanings ends up as transitive and another as extended intransitive. We can make the hypothesis that in old Trumai there might have been one verb to one meaning, i.e. probably the extended intransitive verb \textit{fa} meant only 'hit' and
the transitive verb *disi* meant only 'kill'. We would them assume that the semantics of these historically distinct verbs evolved closer to one another (i.e. *fa* developed the meaning 'kill' and *disi* the meaning 'hit'), while the morphosyntactic category of each remained constant.

6. Conclusion

These are the results obtained by our analysis. There are other facts we intend to investigate in the future, such as the following:

* the selection when, in the use of *fa/disí* and other pairs, the involved NPs are third person nouns (i.e. pronoun versus common noun; human noun versus non-human noun, etc.)

* If the choice of dative markers (-*s/-tl/-ki*) is partially conditioned by other characteristics of the NP-Indirect object, e.g. degree of individuation; degree of affectedness (total/partial); animacy; human versus non-human in relation to pronouns (3rd pronoun referring to human versus 3rd pronoun non-human), etc.

* It was suggested to us that, although the choice of verb from pairs (like those shown in section 5 above) is essentially lexical rather than morphological, the syntactic effect greatly resembles that of an antipassive: for one verb, the second participant is an O (cf. the active in a language with a morphological anti-passive); for the other verb, the second participant is an oblique, the IO (cf. antipassive construction which demotes the O to oblique). The difference is that in Trumai this is a nonproductive lexical pair rather than a morphosyntactic process (reminiscent of the distinction between so-called 'lexical causatives' like the pair *die/kill* and truly productive morphological causatives).

That is an interesting idea to be discussed; the question is if such pairs do function as anti-passives, and if so, how productive this system would be. As we need more information about these pairs and the entire system of the Trumai language (i.e. about the possibility of a morphological anti-passive
construction in Trumai), we prefer to reserve this discussion until another paper.

NOTES

1. My research on Trumai has been conducted since 1989 in the Xingu Reserve in Mato Grosso, Brazil. The data used here were given by the consultants Kumaru, Amati and other persons from Terra Preta Village. Past research was funded by Brazilian foundations: CNPq (Conselho Nacional de Pesquisa), FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), FAEP (Fundação de Apoio à Pesquisa), CAPES (Coordenadoria de Aperfeiçoamento ao Pessoal de Ensino Superior); current research is funded by CNPq and FINEP (Fundação Nacional de Ensino e Pesquisa). A previous analysis of the Trumai verbal case-marking system benefitted from comments by Dr. R. Dixon. An earlier version of this paper was presented at the 1993 Summer Meeting of the Society for the Study of Indigenous Languages of the Americas (SSILA); financial support for participation in this meeting was given by CNPq (Brazil), the Linguistic Society of America and the AID SUNY/Training Program. Thanks to Denny Moore for encouragement and Spike Gildea for detailed comments and discussion; some ideas presented here were suggested by Scott DeLancey. Thanks also for comments received during the SSILA meeting presentation. Remaining mistakes are my own responsibility.

2. Trumai is an isolated indigenous language spoken in the Xingu Reserve, in the central region of Brazil. The Trumai people live distributed in four different places (three villages and near the P.I. Pavuru, an administrative post). While there are between 100 and 109 inhabitants in Trumai communities, due to intermarriage with indians of other Xingu tribes, the actual number of speakers is between fifty five and sixty.

3. We adopt in this paper the abbreviations used in Dixon (1979), adding also the term IO:

S    intransitive subject
A    transitive subject
O    transitive direct object
IO   indirect object
4. As the analysis of the *i/i* morpheme is complex, we will not treat it in this paper. In brief, this morpheme optionally marks overt NPs. When A or IO NPs are deleted, the *i/i* morpheme is also deleted; in contrast, when the S or O NP is deleted, the *i/i* morpheme can remain.

5. The following abbreviations are used in examples:

AbS    absolutive case
Dat    dative case
Erg    ergative case
AGT    agent
PAT    patient
SG     singular
PL     plural
Imp    imperative

6. The words *kasoro*, *aeros* and *cafe* are borrowings from Portuguese ('cachorro, arroz and café', respectively).

7. In reality, as LOC, IO can present variations in position:
   Extended Intransitive : (IO) S V
   Ditransitive :        (IO) A O V

8. Perhaps there would be other pairs. At the moment (at this stage of our knowledge about the language) these are the pairs attested by us.

9. This is reminiscent of Dixon's (1979) proposed hierarchy for person-based split ergativity, and of Gildea's (to appear) discussion of a similar hierarchy for inverse systems. In both, the 1 > 3 is a basic part of the hierarchy. Other kinds of NPs are considered. Dixon's hierarchy is the following:

   1 > 2 > 3 pronouns  > 3 proper nouns  >  nouns

   /    |    \
   /    |    
   /    |    
   /    |    
      human > animate > inanimate
While in some languages it is obligatory to follow this hierarchy, in Trumai is not, although as yet we have no evidence for the rest of hierarchy (human > animate > inanimate).

REFERENCES


NHEENGATU (LÍNGUA GERAL AMAZÔNICA), ITS HISTORY, AND THE EFFECTS OF LANGUAGE CONTACT

Denny Moore, Sidney Facundes and Nádia Pires
Museu Paraense Emílio Goeldi, Brazil

INTRODUCTION

One of the most fascinating cases of a language altered by contact with other languages has remained largely unknown to English-speaking linguists—the case of Nheengatu, also called Língua Geral Amazônica. This language was once dominant throughout the settled Brazilian Amazon region and is still spoken in its modern form in some areas, especially in the region of the Upper Rio Negro.

The indigenous language which was the source of Nheengatu, Tupinambá, is known through descriptions written by Jesuit missionaries (for example, Anchieta 1595 and Figueira 1621), sources which provided the basis for the modern analysis of this now extinct language by Rodrigues (1958, 1990). Old documents in Nheengatu survive from each successive century. There are collections of texts and amateur grammatical descriptions (rigidly following European grammatical categories) from the last two centuries (Magalhães 1876, Rodrigues 1890, Silva 1945, Michaele 1951). The few modern linguistic treatments of Nheengatu include Taylor (1985, 1988), Borges (1991), Grenand and Ferreira (1989), and Rodrigues (1986: ch. 10). The latter work deals explicitly and authoritatively with the diachronic evolution of Nheengatu from Tupinambá; the others are more concerned with phonology than with grammar.

Our own research on modern Nheengatu began in Belém, Brazil, in 1987, initially as a means for teaching field methods. Rather unexpectedly, the research continued sporadically for three years, with a total of ten texts transcribed and analyzed. Emphasis was given to the syntax because of its lack of professional description.

On this basis we present a very brief description of some of the main structural features of the contemporary Nheengatu of the upper Rio Negro, noting obvious resemblances to the structure of its indigenous ancestor or to Portuguese. Unfortunately, no information is yet available on the Nheengatu of other regions and so little can be said about the important question of variation within modern Nheengatu—which may be considerable.

We wish to thank the SSILA for creating an extra session for Brazilian Indian languages at their summer meeting. Travel to the 1993 Linguistic Institute and SSILA summer meetings was made possible by financial support from USAID and CNPq (the Brazilian National Research Council) in the case of Nádia Pires, and by support from the Inter-American Foundation in the case of Sidney Facundes. We thank our Nheengatu informant, Lenir da Silva, for her invaluable assistance with the language and for checking the examples in this paper for accuracy. Spike Gildea removed a number of errors from an earlier version. Support from the Museu Goeldi and from CNPq has been indispensable for our work.
To gain at least a superficial historical perspective on Nheengatu and its evolution, some sources on Tupinambá and on the history of Nheengatu and its relation to socio-political events in Amazonian history were consulted. (Many were not immediately available and could not be consulted.) The linguistically sophisticated work of Freire (1983) was especially useful. On this basis a quick outline of the history of Nheengatu is given, immediately below, focusing on those aspects most relevant for understanding the transmission and the modification of the language during its various phases. After the summary description of the structure of Nheengatu, some final observations are offered about the possible effects of the different types of language contact situations through which Nheengatu passed in different historical periods.

**BRIEF HISTORY OF NHEENGATU**

In 1500, at the time of the first Portuguese contact with what is now Brazil, the eastern coast from São Paulo to the mouth of the Amazon was occupied by native peoples speaking Tupinambá, one of the languages of the Tupi-Guaraní family (of which twenty or so still survive), the most widespread of the ten families of the Tupi linguistic stock. Since there were relatively few European women among the first colonists, many of the Portuguese men married Tupinambá women. Tupinambá was spoken in the household and the mestizo children spoke it natively (Rodrigues 1986:101).

The initial impression of the Europeans was that all the Brazilian Indians spoke the same language, and they thought that knowledge of the language would facilitate the work of conquest and conversion. The Jesuits were active with the indigenous peoples and languages, producing the descriptions by Anchieta (1595) and Figueira (1621). Figueira referred to the language of the coast as the 'Língua Brasileira'. This name was commonly used to refer to it in the Seventeenth Century, though in the second half of that century the name 'Língua Geral' came into use, and in the latter part of the Nineteenth Century the name 'Nheengatu' became common (Rodrigues 1986: 100-103).

The colonization of the Amazon River and its tributaries lagged behind the colonization of the southern regions, where a língua franca with an indigenous base, Língua Geral Paulista, developed and then almost completely disappeared by the 18th century (Rodrigues 1986:102). The Luso Brazilian occupation of the Amazon region began in 1616 with the establishment of Forte do Presépio in the mouth of the Amazon River.

In the Sixteenth Century two expeditions on the Amazon River had been struck by the enormous number of indigenous languages—a very different situation from the coastal uniformity. A Spanish Jesuit who traveled the Amazon River counted more than 150 different languages along the banks of the Amazon and the mouths of its principal tributaries (Acufía 1641:199, cited in Freire 1983: 42).
The European colonists (and the mestizos) depended on Indian labor to extract wealth from Amazônia. A system of slavery and 'aldeias de repartiçao' (resettlement villages) for 'free' Indians was established. Large numbers of indigenous people from many regions, speaking many different languages, were taken from their homes and resettled as laborers for colonists and missionaries. Língua Geral was spoken by the Europeans and mestizos to these Indian laborers.

The use of Língua Geral as a língua franca was favored by the presence of many languages of the Tupi-Guaraní family in the region and by the colonists' desire for a language to communicate with the captured labor force (as well as with their own Tupinambá allies) and by the widespread fluency in Língua Geral that had already been obtained on the coast.

Three years after the Jesuits gained control over the indigenous population through the Regimento das Missões in 1686, Língua Geral was recognized as the official language of Amazônia by the government in Portugal, which endorsed its spread. The Jesuits increased the time that the indigenous inhabitants of the resettlement villages spent in the villages, reducing the time spent in extractive activities. They systematized more the education in Língua Geral. They also increased the expeditions to subjugate and relocate native peoples from more and more remote villages.

Some census figures help understand the sociolinguistic situation at the end of the Seventeenth and the beginning of the Eighteenth Centuries. According to Baena (1831:247, cited in Freire 1983:50), in the four years 1687-1690, just from areas reached by the Tocantins, Amazon, and Negro Rivers, 184,040 Indians were seized and relocated for King and Church. By comparison, the European population was tiny. The 150 Europeans who arrived in 1616 had only grown to 1,000 by 1720, whereas only in Pará, excluding Maranhão, there were 63 resettlement villages with 54,264 Indians, as well as more than 20,000 Indian slaves and a number of mestizos (Raiol 1900:132, cited in Freire 1983:52).

Two facts are noteworthy here. One is that there were massive numbers of new speakers of Nheengatu during the phase of its expansion in the Seventeenth Century and the first half of the Eighteenth Century. Also, these new speakers were from various tribes and spoke various languages; many of which fell into disuse as the speakers' children learned Nheengatu. The existence of a multiplicity of indigenous languages among the captured Indians would favor the spread of Nheengatu, as they turned to it to communicate with each other, just as many Brazilian Indians today speak to other Indians in Portuguese.

The second fact is that there was a large community of native speakers of Nheengatu. Of the classes of people mentioned in the census of 1720, the Whites born in Brazil, the mestizos, the Indian slaves, and the more acculturated Indians in the resettlement villages spoke Nheengatu. While it was the case that the Jesuits used the language as a means of instruction, it would seem, on general grounds, that language learning in the classroom would have been much less significant as a means of transmission than was informal contact with the many native speakers during work, visiting, or religious activities.

By the middle of the Eighteenth Century Nheengatu was nearly universal in colonized Amazônia, even in the capital Belém. This success brought on its decline. Through their
knowledge of the language and control over the Indians, the Jesuits constituted a political force which rivaled that of the State. In the second half of the Eighteenth century the Jesuits were expelled, the State assumed control over Indians and attempted to introduce Portuguese influence into Amazônia. Nheengatu was persecuted and Portuguese was promoted as the language of instruction.

Instruction in Portuguese was ineffective among the Indians. Catastrophic depopulation was already decimating the resettlement villages. Between 1743 and 1750, 40,000 Indians died from diseases in the villages in Pará alone (Freire 1983: 62). In the hands of the State, the Indians continued to fare poorly. Some Portuguese settlers and African slaves were introduced into eastern Amazônia, altering the population balance somewhat in that area.

Brazil became independent in 1822. There had been native insurrections and rebellions previously, all violently surpressed. But the rebellion called the Cabanagem was a large-scale revolt by Indians, caboclos, and negros against the Europeans that lasted ten years, 1837-47, and cost 40,000 lives. The language of the Cabanos was Nheengatu. After the defeat and decimation of the Cabanos, the predominance of Nheengatu was greatly reduced, though it continued in western Amazônia, which still largely depended on Indian labor. The introduction of settlers from the Northeast in the last decades of the Nineteenth Century during the rubber boom reinforced the use of Portuguese.

Freire (1983:73) notes that Correa de Faria, in the mid-Nineteenth Century, compared the Nheengatu he had learned on the Upper Rio Negro with that of the Seventeenth Century, as described by Figueira (1621), and found it to be very different.

In this century, Portuguese has continued to replace Nheengatu, which survives, however, on the Rio Negro, on the Middle Amazon, and probably on the Solimões River.

**PHONOLOGY**

There are some modern treatments of the phonology of Nheengatu, especially the sketches in Taylor 1985 and 1988, and the thesis of Borges (1991). Some observations are offered by Grenand and Ferreira (1989: xiv-xvii). However, many aspects of the phonology are still debatable. We will limit ourselves to a brief, tentative characterization of the phonology of Nheengatu, using the above sources as a point of departure and indicating the details which are unresolved.

One complication is the existence of dialect differences. Another is the problem of separating vocabulary items according to their origins, since there are at least two phonological patterns present: words descended from Tupinambá and words borrowed, more or less recently, from Portuguese. The complexity of the question can be seen from the example of jirimú 'squash', which is listed as part of the vocabulary of Nheengatu by Grenand and Ferreira (62) and which they would consider to be a borrowing from Portuguese since the initial consonant, a
voiced palatal fricative, only occurs, according to them, in such borrowings. However, the word 'jerimum' is itself of Tupi-Guaranian origin, probably borrowed at an early date into Portuguese and then, apparently, borrowed again later into Nheengatu. It is difficult to recognize such examples or even those from other indigenous languages, for example, dákirú 'violin', said by Grenand and Ferreira (xí) to be of Tukanoan origin. The phonological analysis will, of course, change greatly as a function of which vocabulary items it covers.

**BORROWED VOCABULARY**

There are some old borrowings from Portuguese which follow indigenous phonological patterns:

<table>
<thead>
<tr>
<th>Nheengatu</th>
<th>Portuguese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>sorára</td>
<td>'soldado'</td>
<td>'soldier'</td>
</tr>
<tr>
<td>kamíšá</td>
<td>'camisa'</td>
<td>'shirt'</td>
</tr>
</tbody>
</table>

At least for bilingual speakers, recent Portuguese borrowings seem to follow the phonological patterns of Portuguese, with all the consonants and the seven vowels of that language:

<table>
<thead>
<tr>
<th>Nheengatu</th>
<th>Portuguese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>[hó pó]</td>
<td>'roupa'</td>
<td>'clothing'</td>
</tr>
<tr>
<td>[presízu]</td>
<td>'preciso'</td>
<td>'necessary'</td>
</tr>
<tr>
<td>[uí̟stúdál]</td>
<td>'estuda'</td>
<td>'(he) studies'</td>
</tr>
</tbody>
</table>

**NATIVE VOCABULARY**

The surface phonemes of what appear to be non-borrowed words form a more restricted inventory. None of authors cited immediately above agree as to the details of this inventory, though they do agree on its basic components. The analysis adopted here (presented in the table on the left, below) also differs in its details from the others. Marginal or debatable phonemes are enclosed in parentheses. For comparison, the phonemic inventory of Tupinambá, from Lemle 1971 (109), based on Rodrigues 1958, is given below on the right. Some details of the sound system of Nheengatu are discussed and compared with Tupinambá or other languages of the Tupi-Guaranian family.
Modern Nheengatu | Tupinambá
---|---
p | p | t | k | ?
mb | b
nd | s
(g) | š
w | r
m | n
i | i
|m | u| u
e | e
|a | a

' (stress) | ' (stress)
~ (vowel nasализation) | ~ (vowel nasализation)

Most occurrences of the palatal affricate, č, precede i, but a few examples do not:
čá ‘no more’  č-áku- māʔá ‘don’t know…”
not-know-what

While some č before i can optionally be t, indicating a palatalization rule like that of many
dialects of Portuguese (eg. kiči ~ kiti ‘toward’), others cannot (eg. č̪i, *t̪i ‘nose’), and some
examples of t before i cannot be palatalized (eg. ratíwa, *ractíwa ‘uncle’). So, provisionally, č
will be considered a phoneme, with some fluctuation with t before i, at the surface level.

The prenasalized voiced stops, mb, nd, ng, are common and recognized by all authors as
phonemes. They occur initially and intervocally, nasalizing the immediately preceding vowel,
even across morpheme boundaries. They appear to occur only before oral vowels:
mbira ‘offspring’  sê-mbirá ‘my offspring’
ã-amba’ú ‘(I) eat’  u-šendú ‘(he) hears’

Since in many Tupi-Guaranian languages (for example, Urubu-Kaapor, Kakumasu 1986: 401),
the nasal consonant phonemes have prenasalized voiced oral stops as allophones before oral
vowels, one would assume that the nasal series in Tupinambá is the diachronic source of both the
nasal series and the prenasalized series in Nheengatu, though it is not clear what the conditioning
factor for the split was. Interestingly, the principal informant prefers yanê- as the first person
plural prefix of the nominal series and yândé as the free pronoun ‘we’.

Oral voiced stops, b and g, are relatively scarce and are not recognized as phonemes by
Taylor or Borges. However, they do occur in words which are not obvious borrowings, before
oral or nasal vowels:
búya 'snake' se-búya 'my snake' (móya in Tupinambá)
bühbāka 'a palm species' (Grenand and Ferreira: 26)
tibiyara 'a bird species' (Grenand and Ferreira: 166)
garapé 'creek'
apigáwa 'man'

The nasals m and n occur before or after oral or nasal vowels.

mira 'person' se-mira 'my person'
nambah 'ear'
amāna 'rain'

The palatal nasal is analyzed here as (the typically Tupian) ŭ instead of ū because (1) it is usually a glide phonetically and (2) the vowels on either side are obligatorily nasalized, unlike the case of the nasals m and n. It occurs intervocally and (rarely) initially.

ūū 'alone' āūū 'only'
ūāʔā 'that' kūūā 'woman'

Unlike ŭ, the nasal labiovelar glide is rare. Whereas the Tupi-Guaranian languages generally have notable nasalization spreading, this is very marginal in Nheengatu. For example, in yândé 'we', the initial glide is oral, and in aētā 'they', an oral vowel precedes a nasal vowel within the same syllable, at least on the surface.

Two oral glides are generally recognized for Nheengatu, y and w. As analyzed here, these are only slightly reduced high vowels which occur syllable initially and do not carry stress. Examples:

yauči 'turtle' walmī 'old woman'
iwa 'tree' iwā 'fruit'

Unlike Portuguese, Nheengatu, following the indigenous pattern, permits syllables containing two vowels. Note 'turtle' and 'old woman' above and also the following examples:

aētā 'they' pakūa 'banana' múčlu 'bellybutton'
u-likū 'he is' (normal pronunciation, secondary stress on the first vowel)

To avoid sequences of three vowels in one syllable in the example apukwāi 'tie', we tentatively recognize a labiovelar stop, kw, which is probably derived from underlying ku. Some examples of kw (e.g. aikwē 'there is') cannot be ku, though this sequence also exists (e.g. ikwē 'light-colored').
There are only four vowel phonemes in modern Nheengatu, at least in the dialect studied. Nineteenth Century sources often note a fifth vowel, presumably i.

Each Nheengatu morpheme has one primary stress. Within the word, the rightmost stress is maintained and the preceding stresses are successively reduced. Word boundaries can be determined on this basis. Example:

"u-mu-kiri u-mu-kiri 'he causes to sleep' 3-transitivizer-sleep

In our transcription we indicate the stress of each root morpheme with an acute accent mark, though only the rightmost is unreduced. Affixes, except the diminutive, the augmentative and the plural, are stressed on the syllable adjacent to the stem. Affix stress is not marked here.

The status of the glottal stop is not yet clear. Frequently it can occur optionally at morpheme boundaries intervocally, even before an unstressed vowel, e.g., se-?iwá 'my fruit'. It also occurs morpheme internally before stressed vowels, e.g., ka?á 'forest'. It may be fully predictable in this position, but for the time being, it will be transcribed when it is possible morpheme medially.

The syllable pattern of (C)V(V) in Nheengatu differs from that of Tupinambá, which permitted syllable final consonants morpheme finally.

In the transition from Tupinambá to Nheengatu, the principal changes in the inventory of segmental phonemes, pointed out by Rodrigues (1986: 104), were the merger of Tupinambá b (a bilabial fricative) with w, the merger of Tupinambá o with u, and the disappearance of the velar nasal ñ, with accompanying nasalization of the preceding vowel.

**MORPHOLOGY**

**WORD CLASSES**

Nheengatu words fall into eight word classes, approximately equivalent to those of Portuguese: nouns, verbs, adjectives, adverbs, postpositions, pronouns, demonstratives, and particles. Most words in modern Nheengatu texts are of native origin, though there are many borrowings. Almost all borrowings from Portuguese are nouns, verbs, or particles; the other categories seem to be essentially of indigenous origin.

Nouns can be distinguished from adjectives in that the former accept prefixes of the nominal series and the latter do not (though stative verbs homophonous with adjectives do accept these prefixes). Also, adjectives can modify nouns which precede them, but nouns cannot
modify preceding adjectives. Both simple nouns and derived nouns take the same person prefixes, e.g. se-pú 'my hand', se-yasi-wéra 'my cry-baby'.

Verbs follow a typically Tupi-Guaranian pattern, falling into three mutually exclusive subclasses, intransitive, transitive, and stative. All verbs are obligatorily prefixed for subject. Only verbs can constitute complete one word sentences:

<table>
<thead>
<tr>
<th>Type</th>
<th>Prefix</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>intransitive:</td>
<td>a-puraki</td>
<td>'I work.'</td>
</tr>
<tr>
<td>transitive:</td>
<td>a-mūyá (object)</td>
<td>'I make (object).'</td>
</tr>
<tr>
<td>stative:</td>
<td>se-ruri</td>
<td>'I am happy.'</td>
</tr>
</tbody>
</table>

Note that the stative verbs use prefixes of the nominal series, while the intransitive and transitive verbs use prefixes of the verbal series (which occur with no other class). Borrowings from Portuguese seem to enter only the intransitive and transitive subclasses, not the stative subclass. All stative verbs have corresponding adjectives, for example se-ruri 'I am happy' and surí 'happy', but the converse is not true.

Adjectives can be either attributive (maníaka akíra 'green manioc') or predicative (maníaka i-akíra 'the manioc is green). Some predicate adjectives occur with the invariant prefix i-, which is homophonous with the third person of the nominal series. By contrast, stative verbs occur with all the prefixes of the nominal series, showing concordance with the (optional) subject. Adjectives, but not nouns or adverbs, accept the suffix ~'to semi' (puríngá-tó 'almost good', *úká-tó 'almost a house'). Adjectives, unlike transitive and intransitive verbs, cannot accept the prefixes of the verbal series.

Adverbs can be distinguished from nouns and verbs by their lack of person prefixes. They differ from adjectives in that they cannot modify preceding nouns. The free movement of adverbs also distinguishes them from particles and other word classes.

The pronouns are either personal or interrogative. The same set of personal pronouns is used as subject or as object of a verb, as in Tupinambá. Most of the Tupinambá pronouns survived into Nheengatu, but the pronominal system was reanalyzed, converging toward Portuguese. As analyzed by Rodrigues (1990: 420), the Tupinambá system functioned in terms of 'parameters of (a) contrast between speaker and hearer and (b) focality of the 3rd person', rather than the person and number system of today.

Nheengatu Personal Pronouns:

<table>
<thead>
<tr>
<th>Form</th>
<th>Case</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ísé</td>
<td>1 sing</td>
<td>yândé</td>
</tr>
<tr>
<td>ñndé</td>
<td>2 sing</td>
<td>pëyé</td>
</tr>
<tr>
<td>a?é</td>
<td>3 sing</td>
<td>aëtá</td>
</tr>
</tbody>
</table>

Nheengatu also has two interrogative pronouns which are used as question words, both from Tupinambá.
māʔā  ‘what, who, whom’
awá  ‘who, whom’

All the postpositions are of indigenous origin, remaining in Nheengatu even though the basic word order changed from SOV to SVO. Postpositions accept prefixes of the nominal series (e.g. se-irũ ‘with me’), but cannot occur with a free pronoun (*lišé irũ ‘with me’). Some of these postpositions are the following:

<table>
<thead>
<tr>
<th>Postposition</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>rupi</td>
<td>‘through’</td>
</tr>
<tr>
<td>suʔi</td>
<td>‘from’</td>
</tr>
<tr>
<td>irũ</td>
<td>‘with’</td>
</tr>
<tr>
<td>upé</td>
<td>‘in’</td>
</tr>
<tr>
<td>resé</td>
<td>‘in’</td>
</tr>
<tr>
<td>kiti</td>
<td>‘to’</td>
</tr>
</tbody>
</table>

The Portuguese numerals can be used in modern Nheengatu, though at least the lower numerals still exist: yepé ‘one’, mukũʔ ‘two’, and musapiri ‘three’. Unlike in Portuguese, numbers, even borrowed ones, can precede (dózí akayũ ‘twelve years’) or follow (akayũ dózí) a noun.

There are two demonstratives (kwá ‘this’ and ñãʔá ‘that’), which can precede or be the head element in a noun phrase. They cannot occur with pronominal prefixes but can occur with the plural suffix (e.g. kwá-ltá ‘these’). According to Rodrigues (1986: 105) these two elements are the only survivors of the rich Tupinambá system of demonstratives which included forms meaning ‘this (close to the speaker)’, ‘that there (close to the hearer)’, ‘that over there (visible)’, ‘that over there (invisible)’, ‘that physically present’, ‘that we are talking about’, etc.

Particles do not accept inflectional or derivational affixes, though some can form constructions with another free element. Examples:

<table>
<thead>
<tr>
<th>Particle</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ramẽ</td>
<td>‘when’</td>
</tr>
<tr>
<td>či</td>
<td>NEGATIVE</td>
</tr>
<tr>
<td>nẽʔ ~ nẽ</td>
<td>‘nor’</td>
</tr>
<tr>
<td>aráma ~ arã</td>
<td>‘for’</td>
</tr>
<tr>
<td>waʔá</td>
<td>RELATIVIZER</td>
</tr>
<tr>
<td>aikwẽ</td>
<td>‘there is’</td>
</tr>
<tr>
<td>sãyũ</td>
<td>‘just/only’</td>
</tr>
<tr>
<td>sá</td>
<td>‘if’</td>
</tr>
<tr>
<td>ki</td>
<td>‘that (COMPLEMENTIZER)’</td>
</tr>
<tr>
<td>presízo</td>
<td>‘It’s necessary’</td>
</tr>
</tbody>
</table>

Some of the particles are borrowed, such as presízo ‘it’s necessary’ and nẽʔ ~ nẽ ‘nor’, sã ‘if’ and ki ‘that’.
COMPOUNDS

Tupinambá was morphologically complex, with an ample system of incorporation. Examples from Rodrigues (1990: 398-99):¹

ya-y-namí-ʔók-ukár
3-relational-ear-take.off-Caus
'cut the ear off of'

ya-y-pó-pwár-atā
3-relational-hand-tie-hard
'tie up his hands tightly'

Compounding is no longer a very productive process, but a variety of compounds do exist. Examples:

N + N > N  e.g. pi-puópé
foot-nail 'toe nail'

N + N > Adj  e.g. sasi-ára
pain-day 'sad'

N + Adj > N  e.g. maniá-tá-mbéka
manioc-soft 'soft manioc'

V + Adv > V  e.g. kwá-kátú
know-well 'think, believe'

N + N > N  e.g. námbi-púra
ear-part.inside 'earring'

Ptc + Ptc > Ptc  e.g. čí-aráma
not-to 'to not'

AFFIXES

Affixation, as well as compounding, was reduced during the evolution of Nheengatu. Rodrigues (1986: 105) neatly sums up the grammatical changes, observing that (our translation): 'The greatest alterations suffered by Tupinambá in the process of becoming Língua Geral resulted from a progressive simplification of the grammatical forms, accompanied by reorganization of the construction of sentences'. For example, he points out that the Tupinambá verbal morphology, which included a system of five moods (indicative, imperative, gerund, circumstantial and subjunctive) converged to the indicative mood. The noun morphology, which included a system of conjugations in six grammatical cases (nominative, vocative, attributive and three locative cases) was lost in Nheengatu.

¹ Special abbreviations: Caus = causativizer  Cop = copula  V' = complex verb
Neg = negative  Rel = relative  Rlt = relational
Ptc = particle  Foc = focus  Relz = relativizer
Comp = complement  S' = embedded clause  Fut = future
The inflectional and derivational affixes of modern Nheengatu are from Tupinambá; that is, there is no borrowing of any Portuguese affixes. Even recent Portuguese borrowings in Nheengatu can accept person prefixes. Some modern affixes seem to be the result of grammaticization of what were formerly lexical items. The plural suffix, -ítá, a convergence toward Portuguese, was formerly a lexical item, etá, meaning 'many' (Rodrigues, personal communication). Some modern affixes:

<table>
<thead>
<tr>
<th>INFLECTION</th>
<th>DERIVATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Series</td>
<td>TRANSITIVIZER</td>
</tr>
<tr>
<td>a-</td>
<td>mu-</td>
</tr>
<tr>
<td>1 sing</td>
<td>yu-</td>
</tr>
<tr>
<td>re-</td>
<td>'</td>
</tr>
<tr>
<td>2 sing</td>
<td>~to</td>
</tr>
<tr>
<td>u-</td>
<td>mā'ýā</td>
</tr>
<tr>
<td>3</td>
<td>'someone with tendency for...'</td>
</tr>
<tr>
<td>ya-</td>
<td>-sára ~ -góra</td>
</tr>
<tr>
<td>1 pl</td>
<td>'habitual doer of...'</td>
</tr>
<tr>
<td>pe-</td>
<td>-éra ~ -wéra</td>
</tr>
<tr>
<td>2 pl</td>
<td>'characterized by'</td>
</tr>
<tr>
<td>aētā-ú-</td>
<td>-wára</td>
</tr>
<tr>
<td>3 nl.</td>
<td>-íma</td>
</tr>
<tr>
<td></td>
<td>'without'</td>
</tr>
<tr>
<td>Nominal Series</td>
<td>DIMINUTIVE</td>
</tr>
<tr>
<td>se-</td>
<td>-mirí</td>
</tr>
<tr>
<td>1 sing</td>
<td>AUGMENTATIVE</td>
</tr>
<tr>
<td>ne-</td>
<td>-asú</td>
</tr>
<tr>
<td>2 sing</td>
<td></td>
</tr>
<tr>
<td>l-</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>yané-</td>
<td></td>
</tr>
<tr>
<td>1 pl</td>
<td></td>
</tr>
<tr>
<td>pe-</td>
<td></td>
</tr>
<tr>
<td>2 pl</td>
<td></td>
</tr>
<tr>
<td>aētá-</td>
<td></td>
</tr>
<tr>
<td>3 pl</td>
<td></td>
</tr>
<tr>
<td>-ítá</td>
<td>PLURAL</td>
</tr>
<tr>
<td>-āna ~ -wāna</td>
<td>PERFECTIVE</td>
</tr>
<tr>
<td>-rê</td>
<td>IMPERFECTIVE</td>
</tr>
</tbody>
</table>

Reduplication to indicate repetitive action has been retained as a morphological process in Nheengatu, for example, ya-yápi 'throw or shoot repeatedly', pl-píka 'drizzle'. Reduplication was present in Tupinambá, as in most Tupian languages.

**SYNTAX**

**MATRIX CLAUSES**

The matrix clauses show very few borrowings of grammatical morphemes. Their syntax shows convergence toward Portuguese in some aspects and preservation of characteristically indigenous features in other aspects.
Sentence Types

There are three sentence types in Nheengatu, in embedded as well as matrix clauses. Verbal sentences consist of an optional subject followed by one or more VP's containing verbs prefixed for subject. These verbs may be intransitive, transitive, or stative. Transitive verbs are optionally followed by an object, as in Portuguese, in contrast to the Object-Verb order of Tupinambá (and of most Tupian languages). (In the examples below, embedded rather than matrix clauses are given as illustrations if the text examples of the latter are lacking or unclear.) Examples of verbal sentences:

\[
\begin{align*}
[yə-mũũá] & \text{VTrans } cǐmbi?ú & [yə-pinačiška] & \text{VIntr } [yə-mũũá] & \text{VTrans } kaširi \\
1p-make & \text{food} & 1p-fish & 1p-make & \text{chicha} \\
\text{'WE MAKE FOOD, WE FISH, WE MAKE CHICHA.'}
\end{align*}
\]

\[
\begin{align*}
[yə-mbúri] & \text{VTrans } maniáka & \text{paraná } & \text{upé } & [i-mēmbēka] & \text{VStat } aráma \\
1pl-put & \text{manioc} & \text{river in} & \text{3s-be.soft} & \text{to} \\
\text{'WE PUT THE MANIOC IN THE RIVER TO BECOME SOFT.'}
\end{align*}
\]

There are two verbs which might be considered auxiliaries, which occur after the main verb, contrary to the order in Portuguese: putáí 'want' and ikú 'be'. The former can occur without a subject prefix, forming a complex verb. The latter can be preceded by a verb, an adjective, or a postpositional phrase. Examples:

\[
\begin{align*}
[yə-yuwiři] & \text{Vse-retšma kiti} \\
1s-return & \text{want } 1s-land & \text{to} \\
\text{'I WANT TO RETURN TO MY LAND.'}
\end{align*}
\]

\[
\begin{align*}
yàndé & [yə-purĩŋitá] & [yə-ikú] & \text{Aux } yeʃegatú & \text{Vp} \\
\text{we} & 1p-speak & 1p-be & \text{Nheengatu} \\
\text{'WE ARE TALKING NHEENGATU.'}
\end{align*}
\]

\[
\begin{align*}
išé & [se-rúka \text{ upé}] & \text{pp } a-ikú \\
1s-house & \text{in } 1s-be \\
\text{I'M IN MY HOUSE.'}
\end{align*}
\]

The copula sentence type consists of an obligatory subject followed by a predicate noun phrase or adjective phrase. There is no overt copula, unlike in Portuguese. Examples (with inverted order):

\[
\begin{align*}
rē-mbe?ú & \text{ašta-supé } [\text{puráŋga } išé] & \text{S'Cop} \\
2s-tell & 3p-for & \text{good } I \\
\text{'TELL THEM THAT I'M FINE.'}
\end{align*}
\]
The third sentence type consists of a VP with no subject. The VP is composed of a predicative particle followed by a NP or by a clause with an overt subject. These resemble impersonal constructions in Portuguese except that the predicative particle shows no verbal characteristics. At least one of them, presízu (< Port.: É preciso...) 'It is necessary...' is borrowed, and the first syllable of aíkwé 'there is' looks like Portuguese 'aí' 'there'. Example:

[aíkwé]ptc kašuéira
there.be waterfall
'THERE ARE WATERFALLS.'

[presízo]ptc aëtá u-ištudáí pohtugéš upé
need they 3-study Portuguese in
'TIT'S NECESSARY THAT THEY STUDY IN PORTUGUESE'

**Syntactic Processes in Matrix Clauses**

The major syntactic processes affecting matrix clauses look more indigenous than European.

*Negation.* Verb phrases can be individually negated with the particle čí:

čí [a-pítá]yp a-iwiri kwá-kití
not 1s-stay 1s-return this-toward
'I DON'T STAY, I COME BACK TO BELEM.'

Note the structural similarity of this to multiple negation in Tupinambá (Rodrigues 1985: 399):

i-sí' i-memí-r-así-y na s-uwi'-y n i-mara?ár-i...
Rlt-mother not Rlt-son-pain-Neg not Rlt-blood-Neg not Rlt-sick-Neg
'HIS MOTHER DID NOT FEEL ANY CHILDBIRTH PAIN, DID NOT BLEED, WAS NOT SICK...'

The negative particle can occur in the beginning of the clause, negating all of it. It can also form a negative focus construction with a fronted NP:

[čí tapi?íra]ptc apigáwa u-yuká
not tapir man 3-kill
'IT WAS NOT THE TAPIR THAT THE MAN KILLED.' (elicited)

*Topicalization.* Noun phrases can be topicalized, leaving behind third person copies:

[yá?á yawára,]top a?é u-su?ú apigáwa
that dog it 3-bite man
'THAT DOG, IT BIT THE MAN.'
Questions. Polar questions can be formed by intonation.

 índé re-muráí apekátu kwá-su?í tetáma sui?
you 2s-live far this-from city from
'DO YOU LIVE FAR FROM HERE FROM THIS CITY?'

Interrogative word questions are formed using indigenous interrogative words and the particle ta?á.

mā?á ta?á re-wasému pušuéra?
what Q 2s-find ugly
'WHAT DO YOU FIND UGLY?'

As in Portuguese, the interrogative word need not necessarily be fronted.

tána u-mā?á mā?á?
child 3-see what
'THE CHILD SAW WHAT?'

Adverbial movement. Sentence level adverbials can be fronted or placed between phrases.

[kuší?íma]Adv aikwé yepré feiçisíéro a-koñesí wa?á
formerly there.be a shaman 1s-know Relz
'FORMERLY, THERE WAS A SHAMAN WHOM I KNEW.'

Some common syntactic processes in Portuguese, such as passives or clefts, do not occur in Nheengatu.

EMBEDDED CLAUSES

Nheengatu embedded clauses are especially noteworthy in that they show three different patterns:

(1) Subordinate clauses formed on an indigenous pattern
(2) Subordinate clauses formed on a Portuguese pattern, but using indigenous morphemes
(3) Frank borrowings from Portuguese, with accompanying Portuguese grammatical morphemes.

In the first pattern, the clause contains a subordinating particle immediately after the head of the VP, that is, after the main verb, after the predicate nominal or adjectival, or after the predicating particle, according to the type of the VP. These particles include wa?á
RELATIVIZER, ramé TIME, aráma PURPOSE and či-aráma NEGATIVE PURPOSE.
(This last particle occurs clause initially). The relative clauses can have an external head and a corresponding empty internal extraction site:

a-yururé se-máÿá u-pitá aráma iane-réndá upé [se-ratiwa u-šári wa?á yándé ará]S\Rel
1s-ask 1s-mother 3-stay Purpose 1p-farm in 1s-grandpa 3-leave Relz us Purpose
'I ASKED MY MOTHER TO STAY IN OUR FARM THAT MY GRANDPA LEFT FOR US'

(Note that the relative clause modifying 'farm'; has been extrapoased from inside the postpositional phrase to the end of the sentence.)

Or they may be headless, with one missing argument:

aëtá u-kötái [Ø u-akitési wa?á garapé apíra kití]S\Rel
they 3-tell Ø 3-happened Relz stream headwaters toward
'THEY WOULD TELL US WHAT HAPPENED ON THE HEADWATERS OF THE STREAM.'

The time, purpose and negative purpose clauses formed by ramë', aráma, and či-aráma, respectively, distribute like adverbials or adjectivals:

aëtá u-písika pa?á yándé [ya-ú ramë' čímbiʔú irusánga]S\Adv
they 3p-catch they.say us 1p-eat Time food cold
'THEY WOULD CATCH US WHEN WE ATE COLD FOOD.'

yá-búri maniáká paraná upé [i-mëmbéka aráma]S\Adv
1p-put manioc river in 3-be.soft Purpose
'WE PUT THE MANIOC IN THE RIVER. IN ORDER FOR IT TO BECOME SOFT.'

ya-ú čímbiʔú sakú [či-aráma kurupíra-ítica u-rasú yándé]S\Adv
1p-eat food hot Neg-Purpose kurupíra-Plt 3-take us
'WE WOULD EAT HOT FOOD FOR THE KURUPIRA NOT TO TAKE US AWAY.'

ya-máyá čímbiʔú [apígawa u-ú aráma]S\Adj (elicited)
1p-see food man 3s-eat Purpose
'WE SAW THE FOOD FOR THE MAN TO EAT.'

In the second pattern, a subset of the Nheengatu WH words (MA words) are used in embedded clauses in a manner similar to that of Portuguese. The MA words are awá 'who(m)', mäʔá 'which, that', mairämé 'when', maráma 'because', mamé 'where', and mayé 'as'. The relative clauses with awá and mäʔá cannot have external heads:
[màʔá u-yururé i-tupána u-yúmbuʔé tupána supé...] S'Rel
what 3-asked 3s-god 3-pray god to
'WHAT HE ASKFD (FROM) HIS GOD, PRAYING TO HIS GOD...'

*apigáwa [màʔá u-yururé i-tupána] S'Rel...
man who 3-asked 3-god
('THE MAN WHO/THAT ASKED HIS GOD...')

The clauses formed by the other MA words distribute as adverbials or adjectivals:

aētá u-màʔá úka [mamé a-murái] S'Adj
they 3-see house where 1s-live
'THEY SAW THE HOUSE WHERE I LIVE.'

išé čí a-sasá i-puši [mayé aētá ū-mbeʔú] S'Adv
I not 1s-pass 3-bad how they 3-say
'I'M NOT HAVING A BAD TIME LIKE THEY SAY.'

Embedded questions also follow the Portuguese pattern, but using indigenous MA words:

...čí aētá u-kwá [màʔá kurupíra-itá u-mūyá yane-irú] S'Q
not they 3-know what kurupíra-Pl 3-do 1p-with
'...THEY DIDN'T KNOW WHAT THE KURUPIRA DO TO US.'

Some transitive verbs can take unmarked sentential complements:

...né išé a-mänduʔáí [a-màʔá Índé] S'Comp
nor I 1s-think 1s-see you
'...NOR I THINK OF SEEING YOU.'

In the third pattern, obvious borrowings from Portuguese include:

<table>
<thead>
<tr>
<th>function</th>
<th>Nheengatu</th>
<th>Portuguese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>complementizer</td>
<td>ki</td>
<td>'que'</td>
<td>'that'</td>
</tr>
<tr>
<td>conjunction</td>
<td>i</td>
<td>'e'</td>
<td>'and'</td>
</tr>
<tr>
<td>disjunction</td>
<td>u</td>
<td>'ou'</td>
<td>'or'</td>
</tr>
<tr>
<td>negative disjunction</td>
<td>né</td>
<td>'nem'</td>
<td>'neither/nor'</td>
</tr>
<tr>
<td>conditional</td>
<td>sá</td>
<td>'se'</td>
<td>'if'</td>
</tr>
</tbody>
</table>

...[re-murái iké] vP u [re-murái ñeŋə̱ñi kie̱] vP
2s-live here or 2s-live interior toward
'...YOU LIVE HERE OR IN THE INTERIOR...'
PHRASES

The structure of phrases is rather conservative. Two major changes from the indigenous pattern are the order Verb - Object in the VP and the greater elaboration of adjectival and adverbial phrases as incorporation within the verb declined. Noun phrases retain the order Genitive - Noun:

[yane-ye?énga]NP  'OUR LANGUAGE'
1p-language

[karíwa ye?énga]NP  'WHITE MAN'S LANGUAGE'
White.man language

Also the order Noun - Adjective:

not 1p-can 1p-eat food cold
'WE CANNOT EAT COLD FOOD...'

And Demonstrative - NP:

[[ku?á [se-awá-itá purânga]NP ]NP...
this 1s-hair-Pl pretty
'THIS PRETTY HAIR OF MINE...'

There is a position after the head of the VP which contains aspectual suffixes, subordinating particles, and auxiliaries:

a-mûyâ pâyê mâ?á mamê [a-purakí wa?á a-ikú]VP
1s-do all what where 1s-work Relz 1s-be
'I DO EVERYTHING WHERE I'M WORKING.'

[[a-sika ramë]VP]S' sê-mbíra-itá [u-kiri-âna u-ikú]VP
1s-arrive Time 1s-child-Pl 3-sleep-already 3-be
'WHEN I ARRIVE, MY CHILDREN ARE ALREADY SLEEPING.'

Nheengatu retains postpositions, in contrast to the prepositions of Portuguese, which occur as the head of postpositional phrases which, as is characteristic of Tupian languages, have a strictly adverbial distribution, never modifying nouns.

a-moráí ramë [[se-páya]NP irú]pp...
1s-live Time 1s-father with
'WHEN I LIVED WITH MY FATHER...'
TEXT FRAGMENT OF MODERN NHEENGATU

Conversation in Belém between Two People from the Upper Rio Negro

This is the beginning of a text which was recorded and transcribed in 1988, in the Museu Goeldi in Belém. The two speakers are Lenir da Silva, a young woman in her thirties from the region of the Upper Rio Negro, trilingual in Nheengatu, Portuguese, and Spanish, and Gerson, a somewhat younger man from a Baniwa community who is bilingual in Nheengatu and Portuguese, and who lives in the city of San Gabriel da Cachoeira.

Gerson: __
1. ïndé mu?i akayú ta?á re-moráí iké kwá sidádi upé you how many years Q 2s-live here this city in 'FOR HOW LONG HAVE YOU LIVED HERE IN THIS CITY.?'

Lenir: __
2. išé akayú nóvi akayú-ana a-yuwiri se-retáma su? i I year nine year-already 1s-return 1s-city from 'IT HAS BEEN NINE YEARS THAT I LIVE IN THIS CITY.'

3. išé a-yupukwá iké I 1s-acustom here 'I GOT USED TO THIS PLACE.'

4. išé ci a-mându?áí a-yuwíri se-família-itá rúka kití I not 1s-think 1s-return 1s-family-Pl house to 'I DON'T THINK OF RETURNING TO MY FAMILY'S HOUSE'

5. a-kwakatú išé ci a-yupukwá a-kití 1s-believe I not 1s-acustom there-to 'I THINK I CANNOT ACCustom MYSELF TO THAT PLACE ANYMORE.'

6. a-pitá kuri iké até kumairamé Tupâna-itá kuri u-kwá 1s-stay Fut here until when God-Pl Fut. 3-know 'ONLY GOD KNOWS HOW LONG I'M GOING TO STAY HERE.'

7. mayé ta?á a-sú ñúí a-watá se-retáma kití a-mâ?á ará se-anáma-itá how Q 1s-go only 1s-walk 1s-city to 1s-see Purpose 1-family-Pl 'HOW CAN I GO BACK TO THAT CITY ONLY TO SEE MY FAMILY?'

8. išé cá a-mându?áí a-yuwíri a-kití I not 1s-think 1s-return there-to 'I DON'T THINK OF GOING BACK THERE.'
9. a-yuwi kuri a-ya? a-ma? ará se-anáma-ita
   1s-return Fut only 1s-see for 1s-family-Pl
'I WILL GO BACK THERE JUST TO VISIT MY FAMILY.'

Gerson:

10. kuši?íma re-yuwi ramé kwá kití mayé-ta re-yuwi ará
    formerly 2s-come Time that-to how-Q 2s-come Purpose
    'FORMERLY, WHEN YOU CAME HERE, HOW DID YOU COME?'

11. aikwé áwa u-ruí Índé o re-yuwi putái te ne-rupí ...
    there.be who 3-bring you or 2s-come want even 2s-by
    'WAS THERE ANYBODY TO BRING YOU OR DID YOU YOURSELF WANT TO COME?'

**TEXT FROM THE NINETEENTH CENTURY**

This text is from *Poranduba* (Rodrigues, 1890:87-88). It is reproduced as it was written, in a transcription based on the Portuguese orthography, with no morpheme boundaries indicated and prefixes often written separately. This myth is from the Rio Solimões, about the origin of a bird species, Tinkuan (Cocculus cornutus L.), held to be an omen. The leaves of the carayuru plant produce a red dye. There are few, if any, Portuguese borrowings in the text. The translation is ours.

**UIRA-PAYÉ NHEENGAREÇARA**

*The Spirit Bird Sings*

Uirá payé paá, mocoin tayra tuichaua aitá cuchiyma maarupiara, arecé
bird shaman they.say two sons chiefs they formerly happy for this

cuité aitá tutyra u mutara ima
therefore them uncle 3 hate

'THEY SAY THAT THE SPIRIT BIRDS WERE, FORMERLY, TWO SONS OF A CHIEF, VERY HAPPY, FOR WHICH AN UNCLE HATED THEM.'

U cenõe, paá, aitá, u ayuri u íteca muirá u munhan arama cupichuaa,
3 called they.say them 3 invite 3 cut.down trees 3 make to field

u mucão i cunhambíra etá. Aé uana, paá, u jucá.
3 got.drunk 3 nephew plural Then, they.say 3 killed

'THE CALLED THEM AND INVITED THEM TO CUT TREES, TO MAKE A FIELD AND THEN GOT HIS NEPHEWS DRUNK. THEY SAY THAT THEN HE KILLED THEM.'
Aê uana aitá uiuire i aria pêre, aitá anga iunto ana.
Then they returned 3 grandma with they soul only already.

Aitá u purundú imu çupé:
They 3 asked brother to

'THEN THEY RETURNED TO THEIR GRANDMOTHER WHEN THEY WERE ONLY SOULS. ONE BROTHER ASKED THE OTHER:'

-- Mahy taá ne querpe?
How question 2 dream

'WHAT DID YOU DREAM?'

-- Ce querpe racói, cha yá çuca carayuru irumo.
I dreamed in this way I we washed carayuru with

'I DREAMED THAT WE WASHED WITH CARAYURU.'

-- Yaué tenhen racói iché ce mu.
that manner also that way I my brother

'I DREAMED THE SAME.'

Aintá aria uité u moacó aitá remiú. U neeng cuité aitá:
Their grandma then 3 heated their food 3 speak then they:

'WHEN THEIR GRANDMOTHER HEATED THEIR FOOD THEY SAID:'

-- Ah! ce aria, inti uana yá icó mira arama, yaué anga iunto ana.
Ah! my grandma not already we are people in yes soul only already

'AH! GRANDMOTHER, WE ARE NO LONGER PEOPLE, BUT ONLY SOULS.'

Eré ce aria, cha çu ana ne chii, re cenoe ramé cha neengare,
well my grandma I go already 2 from 2 hear when I sing

cha munhan ramé: "Tincuan! Tincuan!..."
I make when "Tincuan! Tincuan!..."

'SO, GRANDMOTHER, WE WILL LEAVE YOU AND WHEN YOU HEAR ME SING "TINCUART TINCUART! TINCUART!..."'
re iauão oca queté, cha neengare ramé cuité "Titi..ti..ti.." aramé re icuão.
2 flee house for I sing when then "Titi..ti..ti.." then 2 will recognize

'FLEE FOR YOUR HOUSE, AND WHEN I SING "TITI..TI..TI.." THEN YOU WILL RECOGNIZE ME.'

Nhaan piranga uaá ceçá recó çòui cuéra
That red that eyes in blood past.thing

'THE RED IN THEIR EYES WAS BLOOD.'

NHEENGATU AND THE EFFECTS OF LANGUAGE CONTACT

For an adequate account of the modifications in Nheengatu/Língua Geral induced by language contact over the last half millennium, it would be necessary to have a detailed account of its sociopolitical context in each historical period as well as an analysis of the language structure and lexicon as these evolved. The task is not impossible, since relevant documents do exist. Of course, for each of the linguistic descriptions which have been made, it is not immediately obvious what the relation is between that description and the speech of the community of speakers, given the possibility of regional or social dialects, of a prescriptivist attitude on the part of the person making the description, or of common errors and misinterpretations.

One fact is clear: the language called today Nheengatu has changed at a rapid rate: the contemporary form would not be mutually intelligible with its form of 400 years ago. Other Tupi-Guaraní languages have not shown the same changes or the same rate of change. More than natural language change was at work to produce the changes in Nheengatu. At the same time, Nheengatu is far from mutually intelligible with Portuguese, with which it has coexisted for centuries.

A second fact is that there was always a sizable community which spoke Nheengatu or its precursors as a first language; it was never a pidgin. There is a belief among some traditional authors on the subject that Nheengatu was a product of the Jesuits. Rodrigues (1887: x-xi) goes so far as to say that changes occurred in Língua Geral in the Amazon Valley because (our translation), 'There it was great the number of missionaries, all with different accents, who taught the languages to Nheengaiba [non-Tupi-Guaraní speaking] tribes, planting degenerate seeds in terrains of different natures, which resulted in a general corruption, not only in pronunciation, but also in meaning.' No evidence is given that this was the real cause of change, and the patterning of the changes observed points to other processes.

Assuming that a good-sized native speaker population was the main source of transmission of the language, we may look at the historical phases of Nheengatu development and see if the types of sociolinguistic effects one would predict do, in fact, agree with the linguistic record, in so far as it is known to us.
In the first century of Portuguese contact with the Tupinambá on the coast there would have been few Tupinambá who spoke Portuguese, in relation to the large numbers who did not speak it. But intermarriage would increase the proportion of Europeans who spoke the indigenous language, as well as create a group of mestizos who spoke the indigenous language but did not have an indigenous social identity.

Rodrigues (1887: viii) notes differences between the descriptions of Anchieta (1595), who lived in Bahia and Espírito Santo, and that of Figueira (1621) who lived in Maranhão. According to him, 'Anchieta wrote the speech which he learned from the Guayanazes, Tamoyos, and Tupis; Figueira that of the Tabayaras, Potiguaras, and Tupinambás properly speaking; and Montoya that of the Guaranians, Payaguás, Charruas, etc.' (Rodrigues 1887: ix). In this picture it is difficult to separate language change from dialect differences. There were relatively few borrowings from Portuguese in the early period, which is what would be expected if Portuguese was not much used by the indigenous and mestizo populations.

During the period of the expansion of Nheengatu, the Seventeenth Century and the first half of the Eighteenth Century, bilingualism with Portuguese continued at a rather low level. The major factor was, rather, the incorporation of enormous numbers of new speakers into the speech community through slavery and resettlement villages. One would expect extensive substratum effects from speakers of many different indigenous languages undergoing language shift as they are absorbed into the Nheengatu-speaking colonial system.

In fact, in the Eighteenth Century Nheengatu was already recognized as distinct from Tupinambá. It was the language of Amazonian colonial society, not the language of an indigenous tribal group. As would be expected, borrowings from Portuguese were limited, but the grammar was altered by so many new speakers. The simplification of the morphology described above was underway at this time (Aryon Rodrigues, personal communication), though the exact sequence of grammatical and phonological changes during this phase are not yet known to us. It is clear from the Nheengatu documents of the Nineteenth Century (see text above) that the reduction of the morphology had already occurred by then.

After Nheengatu was officially discouraged and many of its speakers killed during the Cabanagem, the proportion of Portuguese speakers in Amazônia increased, as well as bilingualism in Portuguese among those who spoke Nheengatu. Texts and commentaries on Nheengatu from the second half of the Nineteenth Century are readily available. These show increased Portuguese influence, with the speech of Pará, according to Barbosa Rodrigues (xii-xiii) being the most 'corrupt'. He notes the addition of vowels to eliminate closed syllables. As noted above, Correa de Faria was struck by the difference between the Seventeenth and the Nineteenth Century forms of the language.

Still, even in the latter half of the Nineteenth Century, the lexical borrowings one finds (e.g. papêru (<papel) 'paper', murutú (<mulato) 'mulato', kabarû (<cavalo) 'horse') are phonologically marked as older acquisitions. The obvious Portuguese borrowings are lexical items. Only a few grammatical words, such as será 'interrogative', were borrowed. Alongside this very limited lexical diffusion is a far more extensive and more subtle influence from
Portuguese: the many examples in the syntax of what Thomason and Kaufman (1988:351) refer to as structural diffusion without the diffusion of native morphemes. Note in the following example that a native interrogative word, mäʔá (maan in the old transcription, retained here) 'what' occurs as the external head of a relative clause formed by the native relativizing particle waʔá (uaâ) (Rodrigues 1890:37):

Cuere tenhê re u maan [çacu uaâ]S^Rel
now not 2 eat what hot relativizer
'NOW YOU DON'T EAT THAT WHICH IS HOT.'

In modern Nheengatu this relativizer is usually deleted, as in the examples in the syntax section above. This origin explains why such interrogative word relatives in Nheengatu cannot have an external head, which they can have in Portuguese: the 'what' word entered into the relatives as an external head, not as a relative pronoun.

The grammaticization of etá 'many' to become the plural suffix, -ità, was already complete in the Nineteenth Century.

Thomason and Kaufman observe that such cases of structural diffusion are only attested from situations of sustained language contact over centuries. That was the case with Portuguese and Nheengatu. In spite of the limited lexical borrowings, the constant interface with Portuguese produced structural diffusion as shown, for example, in the embedded clauses and also in the reanalysis of the pronominal system.

At the present time most Nheengatu speakers in Brazil also speak Portuguese. There is heavy lexical borrowing from Portuguese, and borrowed words accept native inflectional morphology. As expected, it was only after this extensive bilingualism that syntactic patterns using borrowed morphemes appeared. These are now noticeable in Nheengatu. For example the complementizer li (Port. 'que') now appears, as well as conjunction with l (Port. 'e') and disjunction with u (Port. 'ou'). A number of affixes from the last century listed in Stradelli (1929) are no longer in use.

In the region of the Upper Rio Negro Nheengatu is generally considered by tribal Indians to be a language of the Non-Indians, while among Portuguese speakers Nheengatu is often considered to be an indigenous language. It is certainly a remarkable language, whose further study will enrich our knowledge of language contact processes.
REFERENCES


RECONSTRUCTION OF PROTO-TUPI
CONSONANTS AND VOWELS

Denny Moore and Ana Vilacy Galucio
Museu Paraense Emílio Goeldi, Belém, Brazil

INTRODUCTION

According to the most recent classification by Rodrigues (1984/85), there are ten linguistic families within the large Tupi stock: Tupi-Guarani, Munduruku, Mawé, Juruna, Aweti, Mondé, Ramarama, Arikém, Puruborá, and Tupari. Languages of the last five families are spoken in the Brazilian state of Rondônia (north of Bolivia). Published linguistic descriptions have mainly concerned themselves with the first two, Tupi-Guarani and Munduruku; the other families have come under study more recently. Until now the only proto-language within the Tupi linguistic stock which has been the object of detailed reconstructive work is Proto-Tupi-Guarani (e.g. Lemle 1971, Leite & Facó 1991, Rodrigues 1984/85, Jensen 1984). As part of a long-term comparative Tupi project, linguists linked to the Museu Goeldi in Belém, Brazil, are currently working in all the Tupian families in Rondônia and also in the Juruna family.

This article presents part of the initial results of some of the research which has been initiated on languages of the Tupari family during the last five years. It presents a preliminary reconstruction of the sound segments of Proto-Tupari, the mother language of the four modern languages of the Tupari family: Ayuru (Wayoró, Wayru, Ajuru), Makurap, Mekens (Mequém, Mequens), and Tupari. While the preliminary reconstruction presented here is a modest effort, undertaken mainly to guide further research on the Tupari languages, it is the only study of the four languages or of their prehistory written in English. The reconstruction of Proto-Tupi is important for Tupi comparative studies because Tupari is one of the three Tupian families with enough surviving members (four) sufficiently diverged (the languages are not mutually intelligible) to permit a reliable reconstruction at a considerable time depth, using the comparative method. The other two such Tupian families are Tupi-Guarani and Mondé.

In what follows, information about the speakers of the Tupari languages is given. Data sources and limitations are explained. The sound systems of the four languages are briefly summarized. Then the systematic sound correspondences are presented, along with a tentatively reconstructed proto-segment for each. The reconstructions are justified and the diachronic processes leading to the modern languages are summarized. Lastly, the cognates and their reconstructed forms are presented.

THE AYURU, MAKURAP, MEEKENS AND TUPARI

The peoples speaking the languages of the Tupari family lived traditionally on the headwaters of various rivers, most of which drained south into the Guaporé River, which is the boundary between Brazil and Bolivia. The Makurap (and also the Aruá, of the Mondé family) lived on the headwaters of the Rio Branco, the Ayuru on the Rio Colorado, and the Mekens on the Rio Mequens. The Tupari lived on the headwaters of tributaries of the Rio Machado (Ji-Paraná). Also on this more northerly watershed were the Kepikiriwat, whose language, now apparently extinct, is the fifth language of the Tupari family. From the surviving wordlist it appears to be more remote from the other four languages than they are from each other.
According to Meireles (1989), the existence of the Mekens was reported in the Eighteenth Century, in the region of the Rio Mequens. There were two groups (Amniapá and Guaratágaja), whose speech was very similar.

Sustained contact with national society began in the second quarter of this century for most of these groups as rubber gatherers entered the region. The results of contact were usually economic exploitation and decimation through disease. Descriptions of the indigenous people of southern Rondônia include Caspar 1956, Métraux 1948, Lévi-Strauss 1948, Becker-Donner 1962, and Scolnik 1955.

The survivors of the various tribes were placed on posts of the old Serviço de Proteção do Índio (SPI), precursor of the present Fundação Nacional do Índio (FUNAI). In 1988-90, according to Braga (1992), the Posto Indígena (P.I.) Guaporé was home to approximately seventy-five Makurap, forty-one Ayuru, twenty Tupari, and one Mekens. Accordingly to her, the number of Makurap who actually spoke the language was forty-five, and Ayuru and Tupari only had eight speakers each on this post. The tendency is for young speakers to learn Portuguese as a first language. The largest concentration of Tupari is on the P.I. Rio Branco; there are also a number of Makurap there. The Mekens, with at least two dialect groups, are concentrated on the P.I. Mekens.

On the P.I. Guaporé some members of the older generation still retain the traditional knowledge of their culture. Shamanism is still practiced, sometimes involving hallucinogenic snuff, called 'rapé' in Portuguese, which is consumed in group sessions.

**THE DATA**

As part of an attempt to secure at least some tape documentation of the many languages of southern Rondônia, Moore tape-recorded a standardized list of lexical items in various languages during a field visit to the P.I. Guaporé in 1988. The list, recorded in Dolby stereo using an external microphone, included the Swadesh 200-word list and supplementary lists of animals, plants and material culture items common to the region.

The tapes of the four languages under study here were transcribed by Moore and Galucio independently and then compared. Other sources of data include Moore's field transcriptions of Ayuru and the Master's thesis of Braga (1992) for Makurap and that of Alves (1991) for Tupari (which was based on several hours of tapes recorded by Moore). There is an unpublished description of Tupari by Aryon Rodrigues, as well as an early attempt by Hanke, Swadesh, and Rodrigues (1958) to sketch the phonology of Mekens and relate it to other Tupian languages. These two works were not used as sources of data for the present reconstruction, however.

The transcription and analysis of Makurap by Braga generally agree well with that which is presented here. The analysis of Tupari by Alves differs from ours in several respects, but principally in the labiovelar consonants, which she does not recognize as phonemic.
Because of the very limited nature of the data, tone, length, and other subtle phonetic distinctions cannot be established with certainty. More difficult still is the very preliminary nature of our knowledge of the morphology and morphophonology of the languages.

The informants who furnished the data are the following:

- **Ayuru**: Paulina Macurap, a woman about thirty-two years old in 1988. She was raised by the Ayuru.
- **Makurap**: Sebastião Macurap, a man about twenty-two years old in 1988.
- **Tupari**: Alzira Tupari, a woman about twenty-five years old at the time of recording.
- **Mekens**: Otaviano Mequém, a man about seventy-four years old, perhaps the informant of Hanke.

**SKETCH OF THE SOUND SYSTEMS OF THE TUPARI LANGUAGES**

The approximate segmental phonemic inventories of the four languages are summarized in the table below. Segments whose status is still uncertain are indicated by angled brackets. Significant allophones are indicated in parentheses.

**TABLE 1: PHONEMIC INVENTORY OF CONSONANTS**

<table>
<thead>
<tr>
<th><strong>AYURU</strong> (Ay)</th>
<th><strong>MAKURAP</strong> (Ma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t c k kw</td>
<td>p t c k</td>
</tr>
<tr>
<td>&lt;d&gt; g gw</td>
<td>g</td>
</tr>
<tr>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td>r y (̃ĩ)</td>
<td>r y (̃ĩ)</td>
</tr>
<tr>
<td>m n &lt;nj&gt; (ng)</td>
<td>m n &lt;nj&gt; (ng)</td>
</tr>
<tr>
<td>(mb) (nd) (ng)</td>
<td>(mb) (nd) (ng)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>MEKENS</strong> (Me)</th>
<th><strong>TUPARI</strong> (Tu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t k kw (ps)</td>
<td>p t k kw (ps)</td>
</tr>
<tr>
<td>&lt;d&gt; g &lt;gw&gt; (ts)</td>
<td>&lt;d&gt; g &lt;gw&gt; (ts)</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
</tr>
<tr>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td>r y (̃ĩ)</td>
<td>r y (̃ĩ)</td>
</tr>
<tr>
<td>m n η ηw</td>
<td>m n &lt;nj&gt; η</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 2: PHONEMIC INVENTORY OF VOWELS - in all four languages

<table>
<thead>
<tr>
<th>ORAL:</th>
<th>i</th>
<th>i</th>
<th>u(ə)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASAL:</td>
<td>ũ</td>
<td>ũ</td>
<td>ũ(ə)</td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>ū</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>ā</td>
<td></td>
</tr>
</tbody>
</table>

There are five contrasting vowels in each of the four languages. These show remarkable stability over time. Each of the five vowels can be oral or nasal. The nasality may be autonomous or may be acquired through nasalization spread from some nasal segment. Nasalization spread, e.g. Ay: kwãβa 'partridge', appears to occur in all of the four languages, but the exact conditions for its spread in each of the languages cannot be specified at this time. Nasalization spread is a complex phenomenon in the Tupi-Guarani linguistic family (Harrison & Taylor 1971, Lunt 1973) and in the Mondé family (Moore 1984).

The syllable canon is generally (C)V(V)(C+), where C+ represents a morpheme-final consonant. The exceptions to this are that at least Ayuru and Makurap permit a syllable-final morpheme-medial palatal glide, ţ, and Tupari permits ' (glottal stop) and h in the same position. Syllables with two vowels occur, though many of these seem to span morpheme boundaries or to be the result of diachronic consonant deletion. Braga (1992) reports phonemic vowel length for Makurap.

There is no evidence of contrastive stress in the four languages. The question of tone is unresolved. In Ayuru there are two pitch levels in ascending sequences, but at least three or four levels in descending sequences. Both the Makurap and the Tupari Indians use whistled speech to communicate in the forest. However, it is difficult to find evidence of tone contrasts.

The consonantal inventory is similar in many aspects in the four languages. Each language has a series of voiceless stops and a corresponding series of nasals. In Ayuru and Makurap the nasals have post-oralized allophones before oral vowels (e.g. Ay and Ma: [mo] 'hand') and full nasal ones before nasal vowels (Ay: [o mest] 'my husband', Ma: [nõont] 'other'). We will refer to the post-oralized allophones as prenasalized stops.

The prenasalized voiced palatal stop ŋ is a problem. It cannot be an allophone of the palatal nasal ŋ since this is itself a variant of ţ. Since ŋ is rare and does not occur in the correspondences, its status will be left unsolved for now.

The oral voiced stop series is marginal except for the velars and the labiovelars. The voiced bilabial stop, ð, appears to be always derived from an underlying morpheme-final voiceless bilabial stop, p, before vowels. (See the table of morpheme-final consonant alternations below.) Likewise many examples of g are from underlying k morpheme finally before vowels. There are, however, some examples of g and gw which cannot be explained in this manner, e.g. Ay: o-gotkip 'my neck', gwago 'sweet potato'. The oral voiced dental, ð, is very rare and does not appear in the cognates.
In all the languages the palatal glide, \( y \), can optionally be pronounced as a voiced palatal slit fricative. It acquires nasalization from adjacent nasal vowels, in which case it optionally can be a palatal nasal, \( \tilde{n} \), which is the normal pronunciation morpheme initially. All the languages have the voiced bilabial fricative, \( \beta \), which might also be analyzed as a glide. A flap \( r \) occurs in the four languages. It is optionally \( l \) in Makurap.

The four languages differ in the points of articulation which are distinguished in the voiceless, voiced, and nasal series. Makurap lacks all labiovelar consonants. Labiovelar consonants are recognized in one or more of the series for the other three languages. They are not analyzed as a sequence of stop plus glide because they occur syllable initially, where the syllable canon does not permit consonant clusters. Further, these consonants show very regular correspondences.

Tupari has both the glottal stop, \( ' \), and the glottal fricative, \( h \); Mekens has only the latter. The voiceless palatal stop, \( c \), is phonemic in Ayuru and Makurap. The Tupari voiceless bilabial stop, \( p \), is an affricate, \( ps \), before \( i \), and a bilabial fricative, \( \varnothing \), before \( o \) -- allophone similar to that of the neighboring language, Jeoromitxi. Also the Tupari dental stop, \( t \), is optionally an affricate, \( ts \), before \( i \). In Mekens the dental fricative, \( s \), is optionally an affricate, \( ts \).

In morpheme-final position the contrast between the voiceless, the voiced, and the nasal series is neutralized. After oral vowels only the voiceless stops \( p, t, k \) and the palatal glide, \( y [\gamma^c] \), occur in word-final position. These regularly alternate with their homorganic voiced counterparts (e.g. Ma: \( kip \) 'tree'; \( ki\beta + ot \) 'fruit') when a vowel follows the morpheme boundary:

<table>
<thead>
<tr>
<th>All four languages</th>
<th>Ayuru &amp;</th>
<th>Mekens &amp; Tupari</th>
</tr>
</thead>
<tbody>
<tr>
<td>/( V+V )</td>
<td>/( V+V )</td>
<td>/( V+V )</td>
</tr>
<tr>
<td>( p )</td>
<td>( \beta )</td>
<td>( b )</td>
</tr>
<tr>
<td>( t )</td>
<td>( r )</td>
<td>( r )</td>
</tr>
<tr>
<td>( k )</td>
<td>( g )</td>
<td>( g )</td>
</tr>
<tr>
<td>( y )</td>
<td>( y )</td>
<td>( y )</td>
</tr>
</tbody>
</table>

After nasal vowels only strongly prenasalized oral stops, \([mp, nt, nk]\) and the nasal palatal glide \([\tilde{y}]\) occur word finally. We will analyze these as nasalized allophones of \( p, t, k \) and \( y \), respectively, and transcribe the stop phonemes without the prenasalization. For example, \([\eta\tilde{em}p]\) 'breast' is transcribed as \( \eta\tilde{e}p \). At least in Ayuru and Makurap, word-final \( p, t, k, \) and \( y \) after nasal vowels alternate with \( \beta, n, \eta \) and \( \tilde{y} \), before vowels, for example, Ay: \( m\tilde{eket} 'I vomit', m\tilde{eken}-\tilde{eti} 'I feel like vomiting'.

The data available are insufficient to determine the morpheme-initial morphophonemic alternations, which are more complicated. Some alternations involving dental consonants are worth noting since these help explain one of the sound correspondences, \( nd:ct:th \). This correspondence will be reconstructed as a dental consonant \( *D \) in complementary distribution with \( *r \). At this point we only wish to point out the existence of morphophonemic alternations involving \( r, c, t, \) and \( h \):
Ayuru:
  ek tere  'on top of the house'
ge-a-rere  'up in the sky'

Makurap:
  teret    'name'
o-ceret   'my name'

Tupari:
  het      'name'
e-ret     'your name'

There are also morpheme-initial alternations involving these sounds in cognate words in Tupi-Guarani and in Mawé. Without going into detail, we suggest that the corresponding alternations in the Tupari, Tupi-Guarani and Mawé families will eventually be shown to have a common ancestry. Examples from Gregores and Suáres (1967:223) and Graham, Graham and Harrison (1984:189):

Guarani:  
  tera    'name'       Mawé:
  se-rera 'my name'    -ha     'eye'
  NP rera 'NP's name'  u-heha  'my eye'
  h-era   'his name'   NP eha   'NP's eye'

TRANSCRIPTION

The transcription adopted is basically phonemic, but with certain specified sub-phonemic variation also written. This is the case for the nasal consonants. The prenasalized allophones (mb, nd and ng) of the nasal phonemes m, n, and n are written as such to better illustrate the diachronic process of denasalization. Similarly, the oral and nasal palatal glides, y and ý, and the palatal nasal, h, are distinguished in the transcription.

A few other distinctions which appear to be subphonemic are also written in case they should eventually turn out to be significant: a/i in all the languages, l/r in Makurap, s/t/s in Mekens, and t/s before i in Tupari. Syllable break is indicated by a period (e.g. Ma: ßa.i 'stone'), and vowel length (to the small extent to which it can be determined) is indicated by two identical vowels.

SOUND CORRESPONDENCES:

The systematic sound correspondences among the consonants of the four languages are tabulated below, organized according to the mode of production. Hypothesized reconstructed segments are shown on the left, marked with an asterisk. Conditioning environments hypothesized for the proto-language are listed on the right, when relevant, along with the numbers of the cognate sets in which the correspondence is found. Conditioning environments
for individual languages, when relevant, are given after the sound which occurs in that language, for example in the velar correspondences for cognate set (70), *g gː−gːk (h__).

Consonant clusters spanning morpheme boundaries are often maintained in the daughter languages. When one consonant is lost, as happened in cognate sets (17), (29), (35), (51), (52), (85) and (106), it is always the initial consonant which is lost, except for the Ayuru form for 'knife', ṅnte (52). Metathesis may have occurred in (10), (11) and (80). Rather than list a separate correspondence for each of these deletions, they are simply mentioned now and the cognate set in which each occurs is included as an example of the correspondence which would obtain had not the deletion occurred. For example (29) is included as an example of yːyːyːy although the y has been deleted in Mekens.

**TABLE 3: SYSTEMATIC SOUND CORRESPONDENCES**

<table>
<thead>
<tr>
<th>Proto Tupari</th>
<th>Ay</th>
<th>Ma</th>
<th>Me</th>
<th>Tu</th>
<th>Cognate sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>*p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>7, 14, 15, 18, 19, 20, 27, 30, 31, 32, 33, 38, 40, 41, 46, 48, 49, 52, 55, 57, 58, 62, 67, 69, 70, 73, 76, 79, 90, 93, 94, 103, 112, 114, 120, 122, 123, 34</td>
</tr>
<tr>
<td>*p</td>
<td>β(V+)</td>
<td>p</td>
<td>−</td>
<td>p</td>
<td>(_+V) 31, 122</td>
</tr>
<tr>
<td>*p</td>
<td>Ø</td>
<td>−</td>
<td>b</td>
<td>p</td>
<td>(___+V) 4</td>
</tr>
<tr>
<td>*t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>4, 8, 19, 40, 50, 51, 52, 53, 68, 69, 74, 88, 91, 92, 98, 102, 106, 112, 115, 124</td>
</tr>
<tr>
<td>*t</td>
<td>r +V</td>
<td>l +V</td>
<td>−</td>
<td>t #</td>
<td>73</td>
</tr>
<tr>
<td>*t</td>
<td>t</td>
<td>Ø</td>
<td>−</td>
<td>r +V</td>
<td>118</td>
</tr>
<tr>
<td>*t</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>t</td>
<td>(___+V) 3, 5, 58, 61</td>
</tr>
<tr>
<td>*k</td>
<td>k</td>
<td>k</td>
<td>k</td>
<td>k</td>
<td>1, 8, 15, 22, 25, 26, 27, 29, 32, 33, 35, 37, 44, 48, 49, 56, 60, 62, 69, 71, 72, 74, 83, 89, 92, 96, 99, 102, 103, 111, 114, 115, 118, 119, 123</td>
</tr>
<tr>
<td>*k</td>
<td>k</td>
<td>Ø##</td>
<td>−</td>
<td>k</td>
<td>9</td>
</tr>
<tr>
<td>*k</td>
<td>g</td>
<td>−</td>
<td>g</td>
<td>k</td>
<td>(___+) 116</td>
</tr>
<tr>
<td>*kw</td>
<td>kw</td>
<td>Ø</td>
<td>kw</td>
<td>Ø</td>
<td>78, 87, 107, 109</td>
</tr>
<tr>
<td>*b</td>
<td>Ø</td>
<td>β</td>
<td>b</td>
<td>b</td>
<td>(V−+V) 110</td>
</tr>
<tr>
<td>*g</td>
<td>g</td>
<td>−</td>
<td>k</td>
<td>k</td>
<td>32, 33, 117, 121</td>
</tr>
<tr>
<td>*g</td>
<td>g</td>
<td>−</td>
<td>g</td>
<td>k (h__)</td>
<td>70</td>
</tr>
<tr>
<td>*g</td>
<td>g</td>
<td>−</td>
<td>g</td>
<td>−</td>
<td>38</td>
</tr>
<tr>
<td>*gw</td>
<td>gw</td>
<td>β</td>
<td>kw</td>
<td>β</td>
<td>(___Voral) 2, 5, 6, 24, 80, 105</td>
</tr>
<tr>
<td>*gw</td>
<td>g</td>
<td>β</td>
<td>k</td>
<td>Ø</td>
<td>(___o) 69, 105</td>
</tr>
<tr>
<td>*gw</td>
<td>β</td>
<td>−</td>
<td>kw</td>
<td>−</td>
<td>(#__Vnasal) 77</td>
</tr>
<tr>
<td>*ts</td>
<td>t</td>
<td>t</td>
<td>ts, s</td>
<td>t, s</td>
<td>23, 38, 46, 59, 81, 82, 84, 96, s(__i)</td>
</tr>
<tr>
<td>*nd</td>
<td>nd</td>
<td>s</td>
<td>t,</td>
<td>s(__i)</td>
<td>64, 66</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>---</td>
<td>----</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>*β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>8, 20 76</td>
</tr>
<tr>
<td>*ββ</td>
<td>β</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
<td>(i) 12, 119</td>
</tr>
<tr>
<td>*h</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
<td>h</td>
<td>(V__C) 34, 49, 62, 70, 79, 103, 123</td>
</tr>
<tr>
<td>*'</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
<td>'</td>
<td>3, 5, 10, 19, 28, 30, 31, 44, 58, 61, 91, 101, 116, 122</td>
</tr>
<tr>
<td>*r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>(V__V, __+V) 10, 21, 22, 37, 62, 63, 77, 80, 86, 108</td>
</tr>
<tr>
<td>*r</td>
<td>n</td>
<td>l</td>
<td>-</td>
<td>-</td>
<td>(Vnas__Vnas) 104</td>
</tr>
<tr>
<td>*D</td>
<td>(n)d</td>
<td>c</td>
<td>t</td>
<td>h</td>
<td>(#__Voral) 3, 41, 54, 56, 68, 81, 82, 98</td>
</tr>
<tr>
<td>*D</td>
<td>(n)d</td>
<td>Ø/c</td>
<td>h</td>
<td>h</td>
<td>54</td>
</tr>
<tr>
<td>*D</td>
<td>---</td>
<td>--</td>
<td>s</td>
<td>s/h</td>
<td>(__) 52, 72, 95,</td>
</tr>
<tr>
<td>*y</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>(#) 29, 107, 109</td>
</tr>
<tr>
<td>*y/ñ</td>
<td>-</td>
<td>ñ(_Vnas) -</td>
<td>y(_Voras)</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>*ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>(+__Vnas) 28, 35, 39, 61, 65, 85, 113,</td>
</tr>
<tr>
<td>*ñ́</td>
<td>ṱ́</td>
<td>ṱ́</td>
<td>ṱ́</td>
<td>ṱ́(Vnas__)</td>
<td>1, 44, 83, 97, 99</td>
</tr>
<tr>
<td>*m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>(Vnas) 47, 51, 63, 86, 100</td>
</tr>
<tr>
<td>*m</td>
<td>m</td>
<td>p</td>
<td>-</td>
<td>m</td>
<td>(Vnas) 50</td>
</tr>
<tr>
<td>*mb</td>
<td>mb</td>
<td>mb</td>
<td>p</td>
<td>p</td>
<td>(#__Voral) 36, 43, 67, 116</td>
</tr>
<tr>
<td>*mb</td>
<td>mb</td>
<td>-</td>
<td>mb</td>
<td>-</td>
<td>(Vnas+(?)?) 77</td>
</tr>
<tr>
<td>*n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>(Vnas) 13, 17, 42, 45, 47, 75, 77, 89, 104</td>
</tr>
<tr>
<td>*n</td>
<td>n</td>
<td>Ø</td>
<td>-</td>
<td>Ø</td>
<td>50</td>
</tr>
<tr>
<td>*n</td>
<td>n</td>
<td>t</td>
<td>-</td>
<td>-</td>
<td>111</td>
</tr>
<tr>
<td>*nd</td>
<td>nd</td>
<td>t</td>
<td>t</td>
<td>-</td>
<td>(__Voral) 4, 71</td>
</tr>
<tr>
<td>*nd</td>
<td>nd</td>
<td>-</td>
<td>nd</td>
<td>-</td>
<td>(Vnas+(?)?) 87</td>
</tr>
<tr>
<td>*η</td>
<td>η</td>
<td>η</td>
<td>η</td>
<td>η</td>
<td>(#__Vnas) 14, 89</td>
</tr>
<tr>
<td>*η</td>
<td>η</td>
<td>Vnas</td>
<td>g/k Voral k Vnas</td>
<td>--- 17, 71</td>
<td></td>
</tr>
<tr>
<td>*η</td>
<td>η</td>
<td>ng</td>
<td>ng</td>
<td>k</td>
<td>(Voral) 52, 57, 88, 103, 110, 120, 123</td>
</tr>
<tr>
<td>*η</td>
<td>ng</td>
<td>k</td>
<td>k</td>
<td>-</td>
<td>117</td>
</tr>
<tr>
<td>*η</td>
<td>ngw</td>
<td>β</td>
<td>kw</td>
<td>β</td>
<td>(#__Voral) 10, 101</td>
</tr>
<tr>
<td>*η</td>
<td>ngw</td>
<td>Vor</td>
<td>m</td>
<td>Vnas</td>
<td>m Vnas</td>
</tr>
<tr>
<td>*η</td>
<td>β</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>(Vnas__Vnas) 25, 72, 78</td>
</tr>
</tbody>
</table>

The consonants of the voiceless series, p, t, and k, show near-perfect stability in all positions and are reconstructed as such. There are bilabial correspondences in (31) and (122), dental correspondences in (3), (5), (58), (61), (73), and (118), and a velar correspondence in (116) where exceptionality is due to the morpheme-final consonant alternations before vowels in the four languages. The voiceless labiovelar, kw, is unchanged in two languages, Ayuru and
Mekens, but disappeared in the other two. The modern examples of *kw in Tupari are presumably from some other source.

Of the correspondences reconstructed as voiced oral stops, two, $\emptyset$:b:b and g:--:g:k, are due to morpheme-final consonant alternations before vowels—probably the only source of b. The bilabial is reconstructed as *b instead of $\beta$ since b:$\beta$ is a more natural change than the reverse. There is no clear explanation for the deletion of $\beta$ in Ayuru, though this maybe due to a following glottal stop which was lost.

The velar correspondences indicate the existence of an oral *g and *gw in Proto-Tupari. The correspondence g:--:g:k must be different from k:k:k:k, and is reconstructed as *g. This is in harmony with the very general pattern of devoicing in Mekens and Tupari. There are two exceptions to this. The correspondence g:--:g:hk in cognate set (70) may be from a medial sequence *hg which blocked devoicing of g in Mekens before the h disappeared. The correspondence g:--:g:-- cannot be explained at this time.

The correspondence gw:β:kw:β (which does not occur before o) occurs morpheme initially and medially. It is reconstructed as *gw because (1) labiovelars are more likely to go to bilabial fricatives or glides than the reverse, and (2) *kw was already seen to have different, though parallel reflexes. Before o, *gw seems to have lost its labialization. The correspondence Ay: β:Me: kw is unclear since the forms for the other two languages are missing.

Dental affricates *(n)dz and *ts are reconstructed because no conditioning factor could be found for *(n)d and *t to become affricates. It is not clear from the data whether the voiced affricate is prenasalized or not, which is an important question. The variation /s/ in Tupari is perhaps conditioned by the following vowel.

Of the fricatives, the correspondence β:β:β:β is reconstructed as *β. The correspondence β:O:O:O is reconstructed as *β. It only occurs after i, and the sequence iβ does not appear in any of the Makurap, Mekens, or Tupari forms. Tupari seems to have retained a syllable-final preconsonantal *h, as well as a prevocalic glottal stop, *′.

The correspondences reconstructed as *r and as *D are particularly interesting. The phoneme r occurs only morpheme medially and finally in the middle of words in the four languages, and shows highly regular correspondences. The correspondence n:i:--:-- in (104) is perhaps from a nasalized *r. The correspondence (n)d:c:t:h is reconstructed as a dental segment *D whose exact phonetic shape is unknown and which is in complementary distribution with *r, which never occurs word-initially, whereas *D only occurs in that position. On the basis of this complementarity and also the morpheme-initial morphophonemic alternations given above, we suggest that *D was a desonorantized variant of *r at some point in the past, perhaps in Proto-Tupi, since the characteristic alternations occur in several different Tupian families. The reflex of *D in Surui, a language of the Mondé Family, is i (for example let ‘name’), which also argues for an original liquid source. The correspondence --:--:s:s/h is reconstructed as *D before i. The correspondence nd:O/c:h:h in (54) has no explanation at the present.
The oral palatal glide is stable in final position and reconstructs as *y̞. However it is rare and unstable morpheme-medially. There are three cognate sets, (2), (11), and (100), in which y or y̞ occur in Ayuru or Makurap corresponding to Φ in Mekens and to a glottal stop or syllable break in Tupari. For these correspondences, not listed in Table 3, a cover symbol, *y̞, is given, but no plausible reconstructions can be offered. The palatal nasal is stable, apparently occurring only morpheme initially, and is reconstructed as *n̞. The nasal palatal glide occurs elsewhere and is reconstructed as such.

Looking at the nasalized segments, the simplest and most natural overall explanation for the correspondences observed is that original nasal sonorants were progressively denasalized before oral vowels, and then the denasalized stops were devoiced by the general devoicing change in Mekens and Tupari. This implies a lack of rightward nasalization spread from nasal consonants in Proto-Tupari—otherwise there could not have been oral vowels after nasals.

An alternative which must be rejected is that original voiceless stops before nasal vowels were retained as such in Mekens and Tupari and became nasal sonorants in Ayuru and Makurap. This could not have happened because there are a number of examples of the all voiceless stop correspondences (k:k:k:k, etc.) before nasal vowels (e.g. 13, 44, 78, 83, 99, 106, 115, and 118), and no conditioning factor to explain why these would not also have turned into nasal sonorants in Ayuru and Mekens. The denasalization hypothesis is supported by the existence of similar denasalization in the Gavião language of the Mondé family. Compare, for example, Surui: mēt, Gavião: mēt 'husband' and Surui: mebe, Gavião: bebe 'peccary'.

There are some irregularities in this picture. Makurap sometimes has voiceless stops instead of the expected prenasalized stops. There appears to be fluctuation in the language in this regard, for example, 'wasp' may be either ngap or kap. In Tupari the velar nasal seems to have been eliminated altogether.

Another irregularity is that some nasal vowel correspondences are oddly sporadic. See the discussion of nasal correspondences below.

The correspondence ngw:β:kw:β reconstructs neatly as *ngw, paralleling *gw. The last two correspondences, before nasal vowels, are less clear. One suspects the source to be *n̞w since that would otherwise be missing from the pattern.

**Vowel Correspondences**

Since the vowel correspondences are so regular (a:a:a:a, etc.) we will only list the correspondences which are NOT regular. Two irregular correspondences are not included because they probably are due to transcription errors: in cognate set (26), the Mekens form i-kaar should probably be i-kaa ('water-drink'), and in set (123) the Tupari form should probably be ahkop, as in set (103).
TABLE 4: IRREGULAR VOWEL CORRESPONDENCES

<table>
<thead>
<tr>
<th>Proto-Tupari</th>
<th>Ayuru</th>
<th>Makurap</th>
<th>Mekens</th>
<th>Tupari</th>
<th>Cognate Sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>*i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>(#_pe/βe) 27, 119</td>
</tr>
<tr>
<td>*u</td>
<td>i</td>
<td>o</td>
<td>u</td>
<td>o</td>
<td>(_pi, bi) 30, 110</td>
</tr>
<tr>
<td>*i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>—</td>
<td>metathesis ? 80</td>
</tr>
<tr>
<td>*i</td>
<td>i</td>
<td>Ø</td>
<td>Ø</td>
<td>—</td>
<td>metathesis ? 80</td>
</tr>
<tr>
<td>*a/e</td>
<td>a</td>
<td>e</td>
<td>a</td>
<td>e</td>
<td>11</td>
</tr>
<tr>
<td>*e/a</td>
<td>e</td>
<td>a</td>
<td>a</td>
<td>—</td>
<td>119</td>
</tr>
<tr>
<td>*i/e</td>
<td>i</td>
<td>e</td>
<td>e</td>
<td>—</td>
<td>42</td>
</tr>
</tbody>
</table>

The correspondence *i:i:i is reconstructed as *i on the hypothesis of neutralization of i/i in the specified environment in Mekens. The correspondence *i:u:u is reconstructed as *u on a similar hypothesis of *i:Ay: i in the specified environment. The next two correspondences, *i:i:-- and *i:Ø:Ø:-- are perhaps explicable by postulating the metathesis of the i in *araigwi to after the following consonant in Makurap and its deletion in Mekens. For the last three correspondences, a:e:a:e, e:a:a:--, and i:e:i:--, there is no basis for positing the proto-vowel and these are given as *a/e, *e/a, and *i/e, respectively.

NASAL VOWEL CORRESPONDENCES

Nasal vowels regularly correspond to nasal vowels and are reconstructed as such, for example, 'husband' *mêt, Ay: -mêt, Ma: -mê-picop, Me:-mêt, Tu:mêêt. However, there are some irregularities. At least one of the irregularities is probably due to a transcription error in cognate set (25) 'dog' (cf. 72). A number of irregularities appear to be due to nasal spread after consonant addition or deletion (cognate sets (84) and (85)) or to differing conditions on nasalization spread (cognate set (90) and (97)). In these cases the oral form is regarded as the original form, later affected by nasalization spread.

Some of the other irregularities show a certain degree of systematicity. They are listed below in Table 5.

Table 5: IRREGULARITIES INVOLVING NASAL VOWELS

<table>
<thead>
<tr>
<th>Proto-Tupari</th>
<th>Ayuru</th>
<th>Makurap</th>
<th>Mekens</th>
<th>Tupari</th>
<th>Cognate Sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Vnasal</td>
<td>Vnasal</td>
<td>Voral</td>
<td>Vnasal</td>
<td>—</td>
<td>17, 53, 71</td>
</tr>
<tr>
<td>*Vnasal</td>
<td>Vnasal</td>
<td>Voral</td>
<td>—</td>
<td>Voral</td>
<td>9</td>
</tr>
<tr>
<td>*Vnasal</td>
<td>Voral</td>
<td>Vnasal</td>
<td>Vnasal</td>
<td>Vnasal</td>
<td>1, 74</td>
</tr>
<tr>
<td>*Vnasal/oral</td>
<td>—</td>
<td>Vnasal</td>
<td>—</td>
<td>Voral</td>
<td>115</td>
</tr>
</tbody>
</table>

The first irregular correspondence, with three examples (17, 53, 71), seems to be due to denasalization in Makurap, since the forms in the other two languages are nasalized. In the second irregular correspondence, we will assume the second and third syllables were oral, but there was nasality on the first which shifted in Makurap. In cognate sets (1) and (74) the Ayuru
forms appear to have been denasalized. For the last correspondence, in (115), there is no clear basis for deciding the reconstruction for 'tucan'.

The instability of the irregular nasal correspondences listed above might be explained if nasalization in Proto-Tupari was like that reported for the Tupi-Guaranian language Kaiwá by Harrison and Taylor (1971). In Kaiwá, morphemes are either nasal or oral, but it cannot be predicted which syllable(s) will receive the nasality in the nasal morphemes: both *tupã and *tu̯pa are possible.

**SUMMARY**

The charts of reconstructed consonant and vowel segments is given below in Table 6 and 7. The palatal stops are excluded from the picture because of lack of evidence about their origins, but they should not be forgotten.

<table>
<thead>
<tr>
<th>Table 6: PROTO-TUPARI CONSONANTAL SEGMENTS</th>
<th>Table 7: PROTO-TUPARI VOCALIC SEGMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>p</em></td>
<td><em>i</em></td>
</tr>
<tr>
<td><em>t</em></td>
<td><em>i</em></td>
</tr>
<tr>
<td><em>k</em></td>
<td><em>u(o)</em></td>
</tr>
<tr>
<td><em>kw</em></td>
<td></td>
</tr>
<tr>
<td><em>(b)</em></td>
<td><em>e</em></td>
</tr>
<tr>
<td><em>g</em></td>
<td><em>a</em></td>
</tr>
<tr>
<td><em>gw</em></td>
<td></td>
</tr>
<tr>
<td>*(n)*dz</td>
<td><em>Nasal: ŭ ŭ(ŭ)</em></td>
</tr>
<tr>
<td><em>h</em></td>
<td></td>
</tr>
<tr>
<td><em>(D)</em></td>
<td><em>ē</em></td>
</tr>
<tr>
<td><em>y</em></td>
<td></td>
</tr>
<tr>
<td><em>(mb)</em></td>
<td>* aş*</td>
</tr>
<tr>
<td><em>(nd)</em></td>
<td></td>
</tr>
<tr>
<td><em>(η)</em></td>
<td></td>
</tr>
<tr>
<td><em>(ηw)</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 8: UNDERLYING MORPHOPHONEMES OF PROTO-TUPARI</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>p</em></td>
</tr>
<tr>
<td><em>t</em></td>
</tr>
<tr>
<td><em>k</em></td>
</tr>
<tr>
<td><em>kw</em></td>
</tr>
<tr>
<td>*(n)*dz</td>
</tr>
<tr>
<td><em>h</em></td>
</tr>
<tr>
<td><em>m</em></td>
</tr>
<tr>
<td><em>n</em></td>
</tr>
<tr>
<td><em>(η)</em></td>
</tr>
<tr>
<td><em>(ηw)</em></td>
</tr>
</tbody>
</table>

What would appear to have been the underlying system in Proto-Tupari is presented in Table 8. In this table *D* is considered a variant of *r*, with which it is in complementary distribution. The prenasalized consonants are subsumed under the nasals as allophones. The palatal glide includes its variants. The oral series includes only *g* and *gw*, the bilabial being only derived from *p* morpheme finally before vowels.
There are many details to be verified or altered in this picture. Assuming that this preliminary reconstruction is essentially correct, the major changes operating to produce the daughter languages have been (no ordering implied):

- Denasalization of nasal sonorants before oral vowels, a process perhaps already underway in Proto-Tupari.
- Devoicing of obstruents, mainly in Mekens and Tupari.
- Attrition of the original labiovelars by loss, delabialization, or loss of the velar.
- Loss of preconsonantal *h and of glottal stop, except in Tupari.
- Despirantization of dental affricates in Ayuru and Makurap.
- Desonorantization of *r, a process probably initiated long before Proto-Tupari.

Because of the considerable time depth of the Tupari family of languages, the preliminary reconstruction of Proto-Tupari presented here should eventually help cast some light on Proto-Tupi. More data and more phonological and morphological analysis, as well as data from other Tupian families, are needed to refine and broaden the tentative reconstruction presented here.

It is not possible at this time to do a thorough comparison of Proto-Tupari with the languages (or proto-languages) of the other nine Tupian families. Note however some obvious cognates:

<table>
<thead>
<tr>
<th>Family:</th>
<th>Tupari</th>
<th>Tupi-Guarani</th>
<th>Ramarama</th>
<th>Arikém</th>
<th>Munduruku</th>
<th>Juruna</th>
<th>Mawé</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language:</td>
<td>Proto-Tupari</td>
<td>Proto-T-Guarani</td>
<td>Karo</td>
<td>Karitana</td>
<td>Munduruku</td>
<td>Xipaya</td>
<td>Mawé</td>
</tr>
<tr>
<td>Amadillo</td>
<td>*ndayto</td>
<td>*tatu</td>
<td>yayo</td>
<td>sosi</td>
<td>day³do²</td>
<td>dusa</td>
<td>sa'ho</td>
</tr>
<tr>
<td>Peccary</td>
<td>*Daotse</td>
<td>-----</td>
<td>yate</td>
<td>sotysa</td>
<td>da³je²</td>
<td>uza</td>
<td>—</td>
</tr>
</tbody>
</table>

Some of the reconstructed items give a small sample of Proto-Tupari material culture: 'ax', 'basin', 'basket', 'canoe', 'hammock', 'knife', 'salt', and 'seat'. ('Clothing' is an extension of 'skin'.) Domesticated plants include 'cotton', 'maize', 'pepper', 'sweet potato' and perhaps 'tobacco'. ('Banana' probably refers to a wild species which is similar in appearance.)

**LIST OF COGNATES AND RECONSTRUCTED FORMS**

In the following list, some forms are included, in parentheses, even though they are doubtful as cognates. They are included since some part of them may eventually prove to be cognate or at least be useful for clarifying the segmentation of the cognates. Note for example, that in (35) 'flea', the Mekens form, *iq-tsap, supports the segmentation of the Tupari form, *ŋe-tap.

Extraordinary segments may be included without being separated by hyphens if the segmentation is obvious, as in, for example, (29) 'earth'. Where it is useful to indicate
segmentation, as in (34) 'fish', this is done by hyphens, which do not necessarily indicate morpheme boundaries. Morpheme boundaries are indicated (by a +) only when they are relevant to reconstruction and there are strong reasons to posit them, especially, (i) when a known morpheme (such as the prefix *ki- 'first person plural' in Mekens) is involved, (ii) when considerations of canonical form indicate a morpheme boundary (such as between most consonant clusters), or (iii) when the sound correspondence is what would be predicted by well-attested morpheme-final or initial morphophonemic alternations (such as in (58) 'macaw'). Note in (58) that morpheme boundaries are indicated in the Tupari form, *pet+'a, and in the reconstructed form, *pet+'a, but not in the forms for the other three languages, *pera, since fusion may have rendered the morpheme boundary undetectable in these three languages. There are several morphemes of the form */' word-finally in Tupari which may be classifiers, e.g. pep+'o 'wing, feather'.

The reconstructions provided aim at accounting for the forms in the daughter languages as much as possible. Inevitably, there are cases such as (24) 'distant' in which there is some irregularity which cannot be reliably distinguished from transcription errors at this point. In these cases a degree of arbitrariness in the reconstruction is unavoidable.

<table>
<thead>
<tr>
<th>English</th>
<th>Proto-Tupari</th>
<th>Ayuru</th>
<th>Makurap</th>
<th>Mekens</th>
<th>Tupari</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Agouti</td>
<td>*ῃwak'ηyā</td>
<td>ƞgwak'ηyā</td>
<td>māk'ηyā</td>
<td>māk'ηyā</td>
<td>——</td>
</tr>
<tr>
<td>2. Alligator</td>
<td>*gwaYto</td>
<td>gwayco</td>
<td>βato</td>
<td>kwato</td>
<td>βa.o</td>
</tr>
<tr>
<td>3. Ant, big</td>
<td>*Dat+'a</td>
<td>ndara</td>
<td>——</td>
<td>——</td>
<td>hat+'a</td>
</tr>
<tr>
<td>4. Armadillo</td>
<td>*ndayto</td>
<td>ndato</td>
<td>tayto</td>
<td>tato</td>
<td>——</td>
</tr>
<tr>
<td>5. Assai (palm)</td>
<td>*gwit+'i</td>
<td>gwiri</td>
<td>βirica</td>
<td>kwiri</td>
<td>βi+t+i</td>
</tr>
<tr>
<td>6. Ax</td>
<td>*gni</td>
<td>——</td>
<td>βi</td>
<td>kwi</td>
<td>βii</td>
</tr>
<tr>
<td>7. Banana</td>
<td>*ehpilip</td>
<td>epiip</td>
<td>——</td>
<td>——</td>
<td>——</td>
</tr>
<tr>
<td>8. Basin</td>
<td>*βāk'tit</td>
<td>βāk'tit</td>
<td>——</td>
<td>——</td>
<td>βāk'tit</td>
</tr>
<tr>
<td>9. Basket, big</td>
<td>*ängerek</td>
<td>ängerek</td>
<td>akēnē</td>
<td>——</td>
<td>ip-akerek</td>
</tr>
<tr>
<td>10. Bat</td>
<td>*nwar+i+a</td>
<td>ƞgwaria</td>
<td>βa-ca-ria-y</td>
<td>kwari-sa</td>
<td>βa+i+a</td>
</tr>
<tr>
<td>11. Blood</td>
<td>*a/eYi</td>
<td>o+ya</td>
<td>c+eyi</td>
<td>ki+ai</td>
<td>e.i</td>
</tr>
<tr>
<td>12. Blow</td>
<td>*βa</td>
<td>y+βa</td>
<td>β-ii-ka</td>
<td>s-eb-ii.a</td>
<td>ia</td>
</tr>
<tr>
<td>13. Brazil nut tree</td>
<td>*kānā, *arao</td>
<td>kānā</td>
<td>araokiee</td>
<td>kānā</td>
<td>ara.o.a</td>
</tr>
<tr>
<td>14. Breast</td>
<td>*nēp</td>
<td>nēp</td>
<td>nēp</td>
<td>——</td>
<td>kēp</td>
</tr>
<tr>
<td>15. Canoe</td>
<td>*kip-pe</td>
<td>kipe</td>
<td>kipe</td>
<td>——</td>
<td>kipe</td>
</tr>
<tr>
<td>16. Capibara</td>
<td>(loan)</td>
<td>caβi</td>
<td>caβi</td>
<td>——</td>
<td>——</td>
</tr>
<tr>
<td>17. Cicada</td>
<td>*nōtēnā</td>
<td>nōtēnā</td>
<td>koko.ī</td>
<td>kōtkōnā</td>
<td>——</td>
</tr>
<tr>
<td>18. Clothing</td>
<td>*pe</td>
<td>pe</td>
<td>——</td>
<td>ki+pe</td>
<td>pēe</td>
</tr>
<tr>
<td>19. Coat</td>
<td>*pīt</td>
<td>pīt</td>
<td>——</td>
<td>pīt</td>
<td>pīt</td>
</tr>
<tr>
<td>20. Cockroach</td>
<td>*a/βape</td>
<td>aβape</td>
<td>——</td>
<td>eβape</td>
<td>(paba'pairu)</td>
</tr>
<tr>
<td>21. Cotton</td>
<td>*ororo</td>
<td>ororo</td>
<td>ororo</td>
<td>ororo</td>
<td>ororo</td>
</tr>
<tr>
<td>22. Crab</td>
<td>*kera</td>
<td>(koro)</td>
<td>——</td>
<td>kera</td>
<td>kera.a</td>
</tr>
<tr>
<td>23. Deer</td>
<td>*iṣiṣi</td>
<td>iṣiṣi</td>
<td>iṣiṣi</td>
<td>iṣiṣi</td>
<td>——</td>
</tr>
<tr>
<td>24. Distant</td>
<td>*gwetsok</td>
<td>gweeto</td>
<td>βetok</td>
<td>kwesop</td>
<td>(tōg-o)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>25.</td>
<td>Dog</td>
<td>*šjwēko</td>
<td>ābēko</td>
<td>amēko</td>
<td>amēko</td>
</tr>
<tr>
<td>26.</td>
<td>Drink</td>
<td>*ka</td>
<td>kap</td>
<td>ka</td>
<td>i-kaa</td>
</tr>
<tr>
<td>27.</td>
<td>Duck</td>
<td>*ipek</td>
<td>ipek</td>
<td>-----</td>
<td>ipek</td>
</tr>
<tr>
<td>28.</td>
<td>Dust</td>
<td>*kūdči</td>
<td>kiyt-ūdči</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>29.</td>
<td>Earth</td>
<td>*kiy</td>
<td>kiy</td>
<td>kiy</td>
<td>kimākāy</td>
</tr>
<tr>
<td>30.</td>
<td>Egg</td>
<td>*upi+a</td>
<td>iapia</td>
<td>c+upi</td>
<td>s+upi</td>
</tr>
<tr>
<td>31.</td>
<td>Feather</td>
<td>*pep+o</td>
<td>peo</td>
<td>-----</td>
<td>pebo</td>
</tr>
<tr>
<td>32.</td>
<td>Fire</td>
<td>*agopkap</td>
<td>agopkap (?)</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>33.</td>
<td>Firewood</td>
<td>*agopkap</td>
<td>agopkap</td>
<td>(ocatpot)</td>
<td>(otat)</td>
</tr>
<tr>
<td>34.</td>
<td>Fish</td>
<td>*pot</td>
<td>ī-βoy</td>
<td>pot-kap</td>
<td>-----</td>
</tr>
<tr>
<td>35.</td>
<td>Flea</td>
<td>*nōk</td>
<td>nōk</td>
<td>nōk</td>
<td>-----</td>
</tr>
<tr>
<td>36.</td>
<td>Foot</td>
<td>*mbi</td>
<td>mbi</td>
<td>mbi</td>
<td>pi</td>
</tr>
<tr>
<td>37.</td>
<td>Fowl</td>
<td>*ōkira</td>
<td>ōkira</td>
<td>ōkira</td>
<td>-----</td>
</tr>
<tr>
<td>38.</td>
<td>Genipap</td>
<td>*tsigaap</td>
<td>tigaap</td>
<td>(mēncaap)</td>
<td>-----</td>
</tr>
<tr>
<td>39.</td>
<td>Give</td>
<td>*nūa</td>
<td>nūa</td>
<td>nūa</td>
<td>nūa</td>
</tr>
<tr>
<td>40.</td>
<td>Good</td>
<td>*poat</td>
<td>poatēp</td>
<td>-----</td>
<td>(isāmēp)</td>
</tr>
<tr>
<td>41.</td>
<td>Hair</td>
<td>*Dap</td>
<td>ndap</td>
<td>-----</td>
<td>onē-tap</td>
</tr>
<tr>
<td>42.</td>
<td>Hammock</td>
<td>*či/ni</td>
<td>ŋin</td>
<td>ŋin</td>
<td>ŋin</td>
</tr>
<tr>
<td>43.</td>
<td>Hand</td>
<td>*mbo</td>
<td>mbo</td>
<td>mbo</td>
<td>ki+po-pi</td>
</tr>
<tr>
<td>44.</td>
<td>Hawk</td>
<td>*keβ+y+ā</td>
<td>-----</td>
<td>-----</td>
<td>keβa</td>
</tr>
<tr>
<td>45.</td>
<td>Heart</td>
<td>*ānōa</td>
<td>m+ānōa</td>
<td>-----</td>
<td>ki+ānōa</td>
</tr>
<tr>
<td>46.</td>
<td>Heavy</td>
<td>*potsi</td>
<td>poti</td>
<td>poti</td>
<td>i-potsi</td>
</tr>
<tr>
<td>47.</td>
<td>Honey Marten</td>
<td>*āmānā</td>
<td>āmānā</td>
<td>āmānā</td>
<td>-----</td>
</tr>
<tr>
<td>48.</td>
<td>Horn</td>
<td>*apikip</td>
<td>-----</td>
<td>apikip</td>
<td>-----</td>
</tr>
<tr>
<td>49.</td>
<td>Hot</td>
<td>*ahkop</td>
<td>y+akop</td>
<td>-----</td>
<td>s+akop</td>
</tr>
<tr>
<td>50.</td>
<td>Humming bird</td>
<td>*mēnē</td>
<td>mēnē</td>
<td>mēnē</td>
<td>-----</td>
</tr>
<tr>
<td>51.</td>
<td>Husband</td>
<td>*mēt</td>
<td>ō+mēt</td>
<td>mē-picop</td>
<td>ō+mēt</td>
</tr>
<tr>
<td>52.</td>
<td>Knife</td>
<td>*ngōte</td>
<td>ngōte</td>
<td>-----</td>
<td>kipet sēt</td>
</tr>
<tr>
<td>53.</td>
<td>Know</td>
<td>*toa</td>
<td>kiua toa</td>
<td>opoe toa</td>
<td>-----</td>
</tr>
<tr>
<td>54.</td>
<td>Leaf</td>
<td>*Dep</td>
<td>kānā-nde</td>
<td>ep/cep</td>
<td>hep</td>
</tr>
<tr>
<td>55.</td>
<td>Liver</td>
<td>*pia</td>
<td>pia</td>
<td>piat</td>
<td>o+pia</td>
</tr>
<tr>
<td>56.</td>
<td>Lizard</td>
<td>*Dako</td>
<td>-----</td>
<td>cako</td>
<td>tako</td>
</tr>
<tr>
<td>57.</td>
<td>House</td>
<td>*aŋgip</td>
<td>aŋgip</td>
<td>aŋgip</td>
<td>kip</td>
</tr>
<tr>
<td>58.</td>
<td>Macaw</td>
<td>*pet+a</td>
<td>pera</td>
<td>pera</td>
<td>pera</td>
</tr>
<tr>
<td>59.</td>
<td>Maize</td>
<td>*atsiis</td>
<td>atiti</td>
<td>atiti</td>
<td>asisi</td>
</tr>
<tr>
<td>60.</td>
<td>Mandi (fish)</td>
<td>*mōko</td>
<td>-----</td>
<td>mōko</td>
<td>ōko</td>
</tr>
<tr>
<td>61.</td>
<td>Meat</td>
<td>*nēt+a</td>
<td>nērā</td>
<td>nērā</td>
<td>ō-nērā</td>
</tr>
<tr>
<td>62.</td>
<td>Monkey, Capuchin</td>
<td>*sahkirap</td>
<td>-----</td>
<td>-----</td>
<td>sakirap</td>
</tr>
<tr>
<td>63.</td>
<td>Monkey, Spider</td>
<td>*ārīmē</td>
<td>-----</td>
<td>(alēbo)</td>
<td>ārīmē</td>
</tr>
<tr>
<td>64. Mortar</td>
<td>*ëndzi</td>
<td>ëndë</td>
<td>-------</td>
<td>ësi</td>
<td>si-ka'</td>
</tr>
<tr>
<td>65. Mother</td>
<td>*ñä</td>
<td>ñä</td>
<td>ñä</td>
<td>ñä</td>
<td>ñä</td>
</tr>
<tr>
<td>66. Mountain</td>
<td>*(n)dzo</td>
<td>ndoo</td>
<td>ndoa</td>
<td>soo</td>
<td>tuh-tet</td>
</tr>
<tr>
<td>67. Nail</td>
<td>*mbo-ape</td>
<td>mbo-ape</td>
<td>o+po-ape</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>68. Name</td>
<td>*Det</td>
<td>ndet</td>
<td>o+cer-et</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>69. Neck</td>
<td>*gworthkip</td>
<td>o+gotkip</td>
<td>ßotkip</td>
<td>o+ko-tkip</td>
<td>õtkip</td>
</tr>
<tr>
<td>70. New</td>
<td>*pahgop</td>
<td>pagop</td>
<td>-------</td>
<td>i-pagop</td>
<td>pahkop</td>
</tr>
<tr>
<td>71. Night</td>
<td>*ñindaK</td>
<td>ñindaK</td>
<td>gitak</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>72. Ocelot</td>
<td>*ñweKo Dïñt</td>
<td>-------</td>
<td>-------</td>
<td>ñmëko siñt</td>
<td>ñmëko hiñt</td>
</tr>
<tr>
<td>73. Old</td>
<td>*poot</td>
<td>poor+ia</td>
<td>puul+e</td>
<td>-------</td>
<td>poot</td>
</tr>
<tr>
<td>74. One</td>
<td>*kët</td>
<td>kiet</td>
<td>-------</td>
<td>-------</td>
<td>kët</td>
</tr>
<tr>
<td>75. Other</td>
<td>*nëë</td>
<td>ño</td>
<td>nöët</td>
<td>-------</td>
<td>oasë-nööt</td>
</tr>
<tr>
<td>76. Owl</td>
<td>*popoëba</td>
<td>(ibao)</td>
<td>popoëba</td>
<td>popoëba</td>
<td>-------</td>
</tr>
<tr>
<td>77. Paca</td>
<td>*gwânëmbiro</td>
<td>ãbnëmbiro</td>
<td>-------</td>
<td>kwânëmbiro</td>
<td>-------</td>
</tr>
<tr>
<td>78. Partridge</td>
<td>*kwânëwa</td>
<td>kwâba</td>
<td>-------</td>
<td>kwâmëa</td>
<td>-------</td>
</tr>
<tr>
<td>79. Path</td>
<td>*pee</td>
<td>pee</td>
<td>pee</td>
<td>pe</td>
<td>ahpe</td>
</tr>
<tr>
<td>80. Peanut</td>
<td>*araigwi</td>
<td>araigwi</td>
<td>araïëik</td>
<td>arakwi</td>
<td>-------</td>
</tr>
<tr>
<td>81. Peccary</td>
<td>*Daotse</td>
<td>caote</td>
<td>caote</td>
<td>tuase</td>
<td>-------</td>
</tr>
<tr>
<td>82. Peccary, collart</td>
<td>*Daotsey</td>
<td>caote</td>
<td>caote</td>
<td>tuase</td>
<td>aote'iri</td>
</tr>
<tr>
<td>83. Pepper</td>
<td>*kôëy</td>
<td>---------</td>
<td>kôöy</td>
<td>pe-kôöy</td>
<td>-------</td>
</tr>
<tr>
<td>84. Person</td>
<td>*antse</td>
<td>aote-nëp</td>
<td>-------</td>
<td>aote</td>
<td>-------</td>
</tr>
<tr>
<td>85. Piranha</td>
<td>*ipëny</td>
<td>ipëny</td>
<td>ipëny</td>
<td>-------</td>
<td>ÿëny</td>
</tr>
<tr>
<td>86. Push</td>
<td>*móra</td>
<td>ô-móra</td>
<td>-------</td>
<td>-------</td>
<td>i+móra</td>
</tr>
<tr>
<td>87. Rotten</td>
<td>*ënde, *ëkwë</td>
<td>( explanations)</td>
<td>ët</td>
<td>(s+ënde)</td>
<td>ë.ë</td>
</tr>
<tr>
<td>88. Salt</td>
<td>*ëgiët</td>
<td>ëgiët</td>
<td>-------</td>
<td>kiët</td>
<td>kiët</td>
</tr>
<tr>
<td>89. Scorpion</td>
<td>*kitëñëã</td>
<td>kitëñëã</td>
<td>-------</td>
<td>kitëñëã</td>
<td>-------</td>
</tr>
<tr>
<td>90. Seat</td>
<td>*âbë-pe</td>
<td>âbë-pe</td>
<td>ëbëpe</td>
<td>ëbëo</td>
<td>ëëap</td>
</tr>
<tr>
<td>91. See</td>
<td>*to'a</td>
<td>toa</td>
<td>toa</td>
<td>-------</td>
<td>to'a</td>
</tr>
<tr>
<td>92. Seed</td>
<td>*kit</td>
<td>(asti)</td>
<td>(tambët)</td>
<td>ikit</td>
<td>kit</td>
</tr>
<tr>
<td>93. Shell</td>
<td>*ape'</td>
<td>y+ape</td>
<td>ape</td>
<td>-------</td>
<td>s+ape'</td>
</tr>
<tr>
<td>94. Skin</td>
<td>*pe'</td>
<td>pe</td>
<td>pe-cët</td>
<td>ki+pe</td>
<td>pe</td>
</tr>
<tr>
<td>95. Small</td>
<td>*Dïñt</td>
<td>-------</td>
<td>-------</td>
<td>sët</td>
<td>sët</td>
</tr>
<tr>
<td>96. Smooth</td>
<td>*atsik</td>
<td>y-atik</td>
<td>-------</td>
<td>s-asik</td>
<td>asik</td>
</tr>
<tr>
<td>97. Snail</td>
<td>*ëyë</td>
<td>ëyë</td>
<td>ëmëyë</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>98. Snake</td>
<td>*Dat</td>
<td>ndat</td>
<td>cat</td>
<td>-------</td>
<td>hat</td>
</tr>
<tr>
<td>99. Sour</td>
<td>*këy</td>
<td>këy</td>
<td>-------</td>
<td>pe-këy</td>
<td>-------</td>
</tr>
<tr>
<td>100. Speak</td>
<td>*māYā</td>
<td>māYā</td>
<td>————</td>
<td>————</td>
<td>mā'ā</td>
</tr>
<tr>
<td>101. Stone</td>
<td>*ŋwa+i</td>
<td>ŋngwai</td>
<td>ba,i</td>
<td>kwai</td>
<td>β+’i</td>
</tr>
<tr>
<td>102. Straight</td>
<td>*ki:t</td>
<td>kīt</td>
<td>kīt</td>
<td>————</td>
<td>————</td>
</tr>
<tr>
<td>103. Sun</td>
<td>*ŋgiakhop</td>
<td>ŋgiakop</td>
<td>————</td>
<td>————</td>
<td>kiahkop</td>
</tr>
<tr>
<td>104. Surubim (fish)</td>
<td>*ānōrē</td>
<td>ānōnē</td>
<td>ānōlē</td>
<td>————</td>
<td>————</td>
</tr>
<tr>
<td>105. Sweet Potato</td>
<td>*gwagwo</td>
<td>gwago</td>
<td>baibo</td>
<td>kwako</td>
<td>βa.o'</td>
</tr>
<tr>
<td>106. Swim</td>
<td>*tīpťpānāa</td>
<td>tīpťpānāa</td>
<td>tītīnāa</td>
<td>————</td>
<td>————</td>
</tr>
<tr>
<td>107. Tail</td>
<td>*okway</td>
<td>okway</td>
<td>c+oay</td>
<td>s+okway</td>
<td>oay</td>
</tr>
<tr>
<td>108. Take</td>
<td>*ara</td>
<td>————</td>
<td>————</td>
<td>y+ara</td>
<td>ts+ara</td>
</tr>
<tr>
<td>109. Tapir</td>
<td>*ikwāy</td>
<td>ikwāy</td>
<td>iay</td>
<td>ikwāy</td>
<td>————</td>
</tr>
<tr>
<td>110. Termite</td>
<td>*ŋgub+i</td>
<td>ŋgii</td>
<td>ŋgub +i</td>
<td>kubi</td>
<td>kubi</td>
</tr>
<tr>
<td>111. Timbo</td>
<td>*nīk</td>
<td>nīk</td>
<td>tīk</td>
<td>(kitkit)</td>
<td>————</td>
</tr>
<tr>
<td>112. Tobacco</td>
<td>*pitoa</td>
<td>pitoa</td>
<td>(bita)</td>
<td>pitoa</td>
<td>(kipa)</td>
</tr>
<tr>
<td>113. Tooth</td>
<td>*nāāy</td>
<td>nāāy</td>
<td>nāāy</td>
<td>ki+nāāy</td>
<td>i+nāāy</td>
</tr>
<tr>
<td>114. Tree</td>
<td>*kip</td>
<td>kip</td>
<td>kip</td>
<td>kip</td>
<td>kip</td>
</tr>
<tr>
<td>115. Tucan</td>
<td>*yo/ŋōkāt</td>
<td>————</td>
<td>ŋōkāt</td>
<td>————</td>
<td>yokāt</td>
</tr>
<tr>
<td>116. Turtle</td>
<td>*mbok+a</td>
<td>mboga</td>
<td>————</td>
<td>poga</td>
<td>pok+a</td>
</tr>
<tr>
<td>117. Urucum</td>
<td>*ŋgop</td>
<td>ŋgop-gaap</td>
<td>iko</td>
<td>kob+a kaap</td>
<td>————</td>
</tr>
<tr>
<td>118. Vomit</td>
<td>*ēkēt</td>
<td>m+ēkēt</td>
<td>n+ēkē</td>
<td>————</td>
<td>ēkē+ā</td>
</tr>
<tr>
<td>119. Vulture</td>
<td>*iβe/ako</td>
<td>ibeko</td>
<td>iako</td>
<td>iako</td>
<td>————</td>
</tr>
<tr>
<td>120. Wasp</td>
<td>*ŋagap</td>
<td>ŋagap</td>
<td>ŋagap</td>
<td>kap</td>
<td>kap</td>
</tr>
<tr>
<td>121. Water</td>
<td>*i+i</td>
<td>i+i</td>
<td>i+i</td>
<td>i+i-k’a (?)</td>
<td>i+i+k’a (?)</td>
</tr>
<tr>
<td>122. Wing</td>
<td>*pep+’o</td>
<td>peo</td>
<td>————</td>
<td>ipebo</td>
<td>pep+’o</td>
</tr>
<tr>
<td>123. Year</td>
<td>*ŋgiakhop</td>
<td>ŋgiakop</td>
<td>————</td>
<td>————</td>
<td>ikhop</td>
</tr>
<tr>
<td>124. You</td>
<td>*ēt</td>
<td>ēt</td>
<td>ēt</td>
<td>ēt</td>
<td>ēt</td>
</tr>
</tbody>
</table>

1. We wish to thank Paulina Macurap, Sebastião Macurap, Otaviano Mequém, Alzira Tupari, and other helpful people of the P.I. Guaporé for the information they provided about their languages. The SSILA kindly added an extra session for papers on Brazilian Indian languages to their summer meetings. Financial support for the participation of Ana Vilacy Galucio in the SSILA meetings and in the 1993 LSA Linguistic Institute was provided by a tuition grant from the Linguistic Society of America, by a grant for short-term study abroad from USAID, and by travel support from the organization People-to-People. The 1988 field visit was supported by National Science Foundation grant BNS-8619117. The comparative Tupi project is currently supported by Wenner-Gren research grant no. 5578 and by the Brazilian Financiadora de Estudos e Projetos (FINEP) grant no. 63.92.0468.00. The field research was conducted with authorization from the Brazilian National Council for the Development of Science and Technology (CNPq) and from the Função Nacional do Índio (FUNAI). Both the regional FUNAI staff in Guajará-Mirim and the personnel of the P.I. Guaporé were very helpful. Our work has only been possible through the excellent cooperation offered by CNPq and the Museu Paraense Emílio Goeldi. Spike Gildea and Francisco Queixalós offered useful comments on an early draft of this article; remaining errors are our own.
REFERENCES


LOUKOTKA, C. 1950. La parenté des langues du bassin de la Madeira. *Lingua Posnaniensis* (Poznan) 2:123-144


Basic word order in Karitiana (Arikem family, Tupi Stock)

Luciana R. Storto
Pennsylvania State University

Introduction
The aim of this paper is to propose a preliminary analysis of the phenomenon of constituent order variation in Karitiana, a language spoken today by 154 people who live in "Area Indigena Karitiana", in the state of Rondônia, Brazil. The language has previously been studied by David and Rachel Landin, missionaries of the Summer Institute of Linguistics who lived among the Karitiana for five years. In an article based on his masters thesis D. Landin poses SVO as the basic word order of Karitiana (1984:221). However, no structural argument was ever given to support this hypothesis. D. Landin bases his choice on Greenberg's suggestion that the basic word order in a language should be elicited from the most frequent order used in declarative transitive sentences with overt nonpronominal arguments. Considering that SVO is only one of the 6 word orders found in Karitiana and since it is not the most common one, we have no reason to accept D. Landin's description. Also, SVO is the basic word order in Portuguese, the Brazilian national language, spoken fluently by most of the Karitiana for at least half a century. It is not clear that the use of this word order in translations of transitive sentences out of context is free from influences from Portuguese.

Another problem with posing SVO as the basic constituent order in Karitiana has to do with the parametric characteristics of the language, which are consistent with OV and not VO order: noun-postposition, genitive-noun, noun-adjective. Even though the latter is slightly inconsistent with OV order, it follows a pattern found in genetically related Tupi-Guaraní languages, which are OV (Moore 1991:1).

Furthermore, most other Tupian languages of which reliable studies are available are consistently OV. The word order in some of the best studied languages of the Tupi-Guaraní family (by far the larger of the Tupian families) are: Kaapor: SOV; Kamayura: SOV; Tupinamba: SOV; Asurini of the Trocara: OVS; Guajajara: SOV in embedded clauses and VSO in matrix clauses. Other Tupian families show the same pattern: Munduruku (Munduruku family): SOV; Gaviao (Monde family): OV; Xipaya (Juruna family): possibly nonconfigural, but presents OV constructions; Karo (Ramarama family): SOV; Ayuru (Tupari family): SOV. According to Moore, these facts "tend to confirm some earlier speculation that in their past stages the Tupian languages had the basic word order

---

1It is worth mentioning that my definition of word order is the underlying order of constituents in a language. This idea is certainly closer to Chomsky's notion of D-Structure than to Greenberg's criteria of frequency of occurrence.
Object-Verb (OV) rather than Verb-Object (VO)". He points out that Landin's analysis of Karitiana as an SVO language poses the most serious claim to VO order for a Tupian language (Moore 1991:2-3).

Methodology
In an attempt to solve the problem, I elicited and analyzed from a corpus of texts all sentences which presented at least a two-place predicate with overt nonnominal arguments. Pronominal arguments were not considered because I am not completely certain about the rules that regulate their distribution. However, whenever needed in the explanation, I used data which exemplifies the occurrence of free pronouns and personal verb prefixes. The texts utilized consist of mythological tales, historical narratives and dialogues.

Results
From a total of 62 sentences analyzed, the proportion of word order variation found was:

VOS: 27
OVS: 13
SVO: 9
VSO: 9
SOV: 2
OSV: 2

VOS and OVS: The majority of the texts analyzed are mythological tales or reported stories. This fact seems to have influenced the results above, since transitive subject final clauses are much more frequent in narratives than in conversations. A more careful analysis of the data might prove that the sentence final position has a semantic function of agenticity in story telling. R. Landin (1982:3-8) suggested that this position is reserved for the discourse theme. Her evidence, however, is not conclusive since what she describes as the discourse theme almost always coincides with the subject of the sentence.²

1. Na-pisorok-ŋ’ mijo Botyj
   erg-gather-nf nut Botyj
   'Botyj gathered the nuts'

2. Ga Y-ti-m-‘a-t
   field 1p-top-caus-make-nf
   'I made (prepared) the field'

3. Sal na-pitan-ta‘at opok
   salt erg-share-evid white man
   'The white man shared the salt'

²The abbreviations used on this paper are: nf: non-future tense; f: future tense; top: topicalizer; erg: ergative case; abs: absolutive case; 1p: 1st person singular prefix; 1p.pron: 1st person singular free pronoun; caus: causative; evid: evidential marker; co.3p.poss.: 3rd person singular anaphoric possessive pronoun; asp: aspect; asp.sup: aspect (supine); neg: negation; progr: progressive.
VOS and VSO are very common word orders in texts, in contexts where the action is being emphasized. Sentences 4-7 are extracted from a myth where the activities performed by the main character are being described:

4. Na-petet-Ø Botýj mijo ket
   erg-cook-nf Botýj nut unripe
   'Botýj cooked the unripe nuts'
5. Dok Byjyty
   seat (ideophone) Byjyty
   'Byjyty sat'
6. Na-ambo-t kendo ohyn Byjyty
   erg-seat-nf coconut above Byjyty
   'Byjyty sat on the coconut'
7. Na-mynira-t ta-iso Byjyty
   erg-lit-nf co.3p.poss.-fire Byjyty
   'Byjyty lit his fire'

OVS and OSV word orders sometimes present the verb prefix ti- which Rachel Landin describes (1982:15) as a marker of topicalization of the object. Whenever this prefix occurs on a verb, the object is the first constituent in the sentence. However, not every instance of topicalization of the object is marked by this morpheme, as first noted by Rachel and David Landin. R. Landin correctly noted that this topicalizer occupies the same position in the verb that the ergative/absolutive markers do. She also noted that ti- never occurs in the narrative portion of a text, being restricted to "monologues, conversation, and speech quotes in narratives" (1982:11). D. Landin did not make use of this information in his thesis, posing an optional rule (1984:233) that deletes the case prefixes and inserts the topicalizer ti- when there is uncertainty as to the syntactic functions of the arguments in a sentence. However, from the examples below it is clear that ti- is not a disambiguator of syntactic function, since it is present even when the subject is dropped:

8. Moramon a-ti-m-'a-tykat, y-ta'it
   what 2p-top-caus-do-asp 1p-uncle
   'What are you doing, uncle?'
9. Tyký ti-m-'a-tykat, y-saka'et
   palm heart top-caus-do-asp 1p-nephew
   'I am taking (gathering) palm hearts, my nephew'
10. Pom ememo ti-m-'a-t
    nambu black top-caus-make-nf
    '(He=pro) created the black nambu'
11. Ese i-ti-m'-a-t Ora
    water ?-top-caus-make-nf Ora
    'Ora created the water'
12. Y'it kyry y-tyi'-y-tysypak?
    1p-son liver 1p-top-eat-asp.sup.
    'Am I eating my son's liver?'  
13. Atykiri naka-sot-Ø [esety Ora ti-m'-a]
    then erg-exist-nf [water big Ora top-caus-do
    'Then there was the river which Ora created'
Also, the characterization of this morpheme as a topicalizer seems to be inadequate because it marks the fronting of elements whose referent is "unknown" such as WH words generated in object position. It is clear that in WH sentences movement is not motivated by pragramatic factors similar to the ones that characterize topicalizations.

Furthermore, ti- is restricted to a certain sentential type used to express direct speech or embedded sentences and it indeed occupies the same structural position filled by the morphemes which indicate ergative/absolutive case in the narrative sentential type. These morphemes (na(ka)/ta(ka)) were shown by D. Landin (1984:223-227) to be in ergative/absolutive distribution, where ta(ka) marks transitive verbs preceded by objects or intransitive verbs preceded by subjects while na(ka) marks everything else. In a parallel fashion ti- marks movement of accusative, as opposed to nominative arguments (which are unmarked) to sentence initial position in direct speech and embedded sentences. That is, the presence of the morpheme ti- reflects a nominative/accusative system in the direct speech sentential type while the morphemes na(ka)/ta(ka) reflect an ergative/absolutive system in narrative sentences. In face of this evidence I suggest that Karitiana has a split ergative system.

14. Sosy i-ti-oky-t pōrāsi
   armadillo 3p-top-kill-nf trap
   'The trap killed the armadillo'

15. Pōrāsi i-oky-t sosy
    trap 3p-kill-nf armadillo
    'The trap killed the armadillo'

Data recently elicited in the field shows that the ergative/absolutive markers also occur in contrast in identical environments:

16. Taso na-oky-t ombaky
    man erg-kill-nf jaguar
    'The man killed the jaguar'

17. Taso ta-oky-t ombaky
    man abs-kill-nf jaguar
    'The man killed the jaguar'

According to two fairly sophisticated informants tested independently, sentences 16 and 17 convey the same meaning, but the latter is used as "a warning, when you know something will happen as a consequence of the action". In my opinion the absolutive marking is demoting the subject in its characteristics of agentivity (control and intentionality) in order to emphasize the action. The process above could perhaps be described as semantic ergativity.

The obvious conclusion concerning the consituent orders OSV and OVS is that since they get marked for object movement, they must be deviations from the basic word order.

---

3 The allomorphs maka and taka prefix stress-initial verbs while na and ta prefix verbs with all other stress patterns.
SVO: This is the typical word order used in translations of transitive declarative sentences from Portuguese:

18. Taso na-oky-j ombaky
man erg-kill-f jaguar
'The man will kill the jaguar'

19. Y-hay naka-kip-∅ sosy
1p-old brother erg-open-nf armadillo
'Vemy older brother opened the armadillo'

This order also occurs in both the narrative and the direct speech sentential types. Semantically, it appears to be unmarked.

The most striking fact we observed about the word order variation, is that the SOV and OSV word orders were found to occur only in subordinate clauses. It is a widely accepted fact that the word order in embedded clauses has a tendency to be more conservative than the one in main clauses, since the former are less subject to influences of pragmatic nature than the latter. This generalization could be used as an argument for the hypothesis that a sequence with the form NP NP V is the basic word order in Karitiana. I am inclined to think that SOV is the basic word order in the language due to the fact that it is morphologically less marked or simpler than OSV, where ti-sometimes indicates movement of the object (as in example 13). However, the latter seems to be by far the most used word order in subordinate clauses, SOV being more common in mythological tales. Other examples of embedded clauses are given below:

20. [ombaky tas oky tykiri] y-taka-hyryp-∅ yn
[jaguar man kill when] 1p-abs-cry-nf 1p.pron.
'When the man killed the jaguar I cried'

21. [Ahoy byhip tyki-oott] na-pa'ira-t jonso
[arroz cook when-progr] erg-angry-nf woman
'The woman was angry when she was cooking rice'

22. Atykiri Botyj naka-m-'a'-ot hyryp [Ora ta-'it then Botyj case-caus-do-first cry [Ora co.3.poss.-son byhot tykiri]
transform when]
'Then Botyj cried first when Ora transformed his son'

23. [Ambi Joana ama tykiri] naka-tat-∅ Maria
[house Joana buy when] case-go-nf Maria
'When Joana bought the house, Maria left'

24. [Dinheiro y-ahit-iki tykiri] y-taka-tat-∅ yn
[money 1p-get-neg when] 1ps-case-go-nf 1ps
'[When I did not receive the money] I left'

In an analogy with the analysis proposed for German and all other Germanic languages other than English, I suggest that Karitiana might be a verb second (V2) language. It is a widely accepted fact that the word order in German is that of the embedded sentences (SOV) and that the tensed verb (verb/Infl) has to raise to the COMP position in main clauses (the second structural position in the sentence) in case there is no lexical item occupying that position for the sentence to be well-formed. The COMP node in V2 languages such as German is interpreted as the head of the sentence - an inherently tensed position which needs to be lexically realized in order to assign nominative case to the subject (Platzack 1986).
A consequence of the analysis proposed for constituent order in Karitiana indicates that verbs will always be in final position in embedded sentences and never so in main clauses. This is indeed the distribution I have found in my preliminary analysis of the data.

The V2 hypothesis for Karitiana will not be fully developed in this paper because I do not completely understand the distribution of certain crucial morphemes such as pronominals and case, topicalization and tense markers. However, an interesting fact which seems to support this hypothesis is the absence of several inflectional morphemes in subordinate sentences. These clauses present either no tense marker at all or the unmarked verb suffix which indicates present or past tense. The presence of a future tense suffix in subordinate clauses is considered ungrammatical, which could indicate that COMP is the structural position where future tense is assigned to the verb since narrative main clauses always require a tense marker. Unlike main clauses, embedded sentences do not present aspect markers suffixed to the verb. Subordinate sentences which semantically bear aspect have it expressed in VP internal adverbials as in example 21.

If the V2 phenomenon is a reality in Karitiana, then the structural description of sentences will be such that the SPEC of CP position can be occupied by WH words or any argument noun phrase, while the COMP position is always occupied by either the verb or the tensed auxiliary in main clauses. The morpheme ti- seems to mark exactly this fronting of arguments generated in object position to SPEC of CP. A pre-sentential position has to be posed, where conjunctions which function at a level above the clause and ideophone (onomatopoetic) phrases occur. Also, there has to be a clause-final position where the subject can move in certain discourse environments.

Although I am not in a position to give compelling evidence for the V2 phenomenon in Karitiana, the data seems to point to SOV as the basic constituent order in the language. Finally, I hope to have persuaded the reader that at least further analysis is needed before we accept David Landin's claim that SVO is the basic constituent order in Karitiana.
AKNOWLEDGMENTS

I thank Wenner-Gren Foundation for Anthropological Research and The Pennsylvania State University for their financial support of my graduate studies and the Museu Emilio Goeldi/FINEP (Brazil) for travel grants which enabled my field research among the Karitiana. I also thank Samuel Cruz from FUNAI (The Brazilian Indian Affairs Bureau) for logistic support. I am grateful to Daniel Everett, Denny Moore, Jeffrey Pullum, David Wible, and my fellow junior researchers at the Goeldi Museum for helpful commentaries and discussion of earlier versions of this paper. Finally, I thank the Karitiana people for their cooperation with this research, specially Nelson and Meireles Karitiana who have been particularly interested and dedicated.

REFERENCES


'How', and 'Thus' in UA Cupan and Yuman: A Case of Areal Influence

Eric Elliott
UCSD

Prefixation onto verbal roots is unusual in Cupan Luiseño². Excepting prefixes resulting from reduplication, Luiseño normally only allows pronominal prefixation onto verbs. Although the other Cupan languages, notably Cahuilla, allow some prefixation on verbs, there is within Cupan, and most notably in Luiseño, a general predilection for suffixation onto verbal roots. The Cupan verb yâx-, 'be, do, say' is unusual in that it allows the affixation of a number of prefixes. These prefixes include a proximal definite 'i-', a distal definite 'á-', and two indefinite prefixes m- and h-. In this paper we will see that both Yuman and Cupan show evidence of multiple prefixation onto a verb meaning 'be, do' in Yuman, and 'be, do, say' in Cupan. In particular, it will be demonstrated that the morphological analyses of the Cupan words for 'how' and 'thus' closely parallel those for Yuman Diegueño 'how' and 'thus'. In both Cupan and Yuman Diegueño, 'how' and 'thus' are derived from a verb meaning 'be, do' (or 'say' in the case of Cupan). It will then be shown that while prefixes 'á-', 'i-' and h- are of family-internal UA origin, m- is an intrafamiliar areal phenomenon of Southern California, attested most strongly in Yuman Diegueño and UA Cupan.

In Section I of this paper, we will examine the prefixation processes possible in conjunction with the Cupan verb yâx-. We will then explain how the Cupan languages have developed causative verbs from yâx-. These derived causative verbs bear the meaning of 'do somehow', or 'do thus', and are used in forming the Cupan words for 'how' and 'thus' respectively.

In Section II we will discuss the interrogatives of a non-Cupan UA language, Serrano. It will here be shown that the common element in all Serrano interrogatives is the indefinite prefix h-.

In Section III we will discuss the interrogatives of the Cupan languages. Here we will see that both m- and h- are present.

In Section IV we will concentrate on Diegueño, Cupan's Yuman neighbor. It will be seen that Diegueño possesses an indefinite (interrogative) morpheme whose onset is /m/. Diegueño further possesses a verb yuu, 'be'. It will be shown that yuu is used in forming the Diegueño words for 'how' and 'thus'. A parallel development for 'how' and 'thus' may be observed between Yuman Diegueño and UA Cupan. Evidence from UA Serrano will be provided to demonstrate that this parallel development for 'how' and 'thus' may not be mirrored outside UA Cupan.
In Section V we will provide evidence from other Yuman languages to show that the pervasiveness of \textit{m-} as an interrogative marker is unique to Diegueño within Yuman.

In Section VI we will draw our conclusions based on the evidence provided in Sections I through V.

\textbf{Section I: Prefixation Onto \textit{yáx-}}

All three Cupan languages possess a verb \textit{yáx-} with the meaning 'be, say, do'.

When unprefixed (except with pronominal prefixes), the verb \textit{yáx-} generally means 'say':

(1)
\begin{align*}
\text{Luiseño:} & \quad \text{yáx-wun-pum} \\
& \quad \text{be/say-PR:PL-3pl} \\
& \quad \text{They say}
\end{align*}
\begin{align*}
\text{Cahuilla:} & \quad \text{hém-yax-we(n)} \\
& \quad \text{3pl-be/say-PR:PL} \\
& \quad \text{They say}
\end{align*}

Common to both surviving Cupan languages is a series of four verbs derived from \textit{yáx-}. Note that Non-Cupan UA Serrano\textsuperscript{4} diverges from its Cupan relatives:

(2)
\begin{tabular}{llll}
 & \textbf{CUPAN} & \textbf{Luiseño} & \textbf{NON-CUPAN} \\
\text{Cahuilla} & \text{yáx-} & \text{yáx-} & \text{řah, kíy} \\
\text{Luiseno} & \text{؟iyax-} & \text{؟iyx*-} & \text{pana’ řah} \\
\text{Cahuilla} & \text{؟áax*-} & \text{؟áax*-} & \text{pana’ řah} \\
\text{Cahuilla} & \text{miyax-} & \text{miyix*-} & \text{hamin řah} \\
\text{Luiseno} & \text{hiyax-} & \text{hi(y)x*-} & \text{kíy hiiti’}
\end{tabular}

The Cupan verbs (B) through (E) are all derived from verb (A). We will now examine more closely the morphology of the Cupan verbs (B) through (E), using Cahuilla examples as representative for Cupan in general.

Verb (B) is analyzable as:

(3) 'í-yax-
PDEF-be
be this way

The prefix 'í-' is nearly identical with one form of the proximal demonstrative in Cahuilla:

(4) 'í-
this
Verb (C) is analyzable as:

(5) 'á-yax-
    DDEF1-be
    be that way, seem, resemble

An independent distal demonstrative *'á* is not extant in Cupan. Serrano, however, has:

(6) 'i-p               'a-p
    PDEF-LOC          DDEF1-LOC
    here              there

'i-m               'aa-m
PDEF-PL            DDEF1-PL
these              those

Verb (D) is analyzable as:

(7) mi-yax-
    IND1-be
    to be somehow

Prefix m(i)- is abundantly attested elsewhere in Cupan (see Section III below).

Verb (E) is analyzable as:

(8) hi-yax-
    IND2-be
    say what

The indefinite prefix h(i)- is also well attested in other UA languages, including Cahuilla, Luiseno and Serrano (see Section II below).

**Causatives of 'i-yax- and mi-yax-**

Cupan Luiseno and Cahuilla have developed causative verbs 'cause to be' (= 'do') by affixing a causative suffix to a derivative of yáx-:

(9)

<table>
<thead>
<tr>
<th>Cahuilla</th>
<th>Luiseno</th>
</tr>
</thead>
<tbody>
<tr>
<td>'éxan-</td>
<td>'axán'i-</td>
</tr>
<tr>
<td>do thus</td>
<td>do thus, cause to resemble</td>
</tr>
</tbody>
</table>

'axán'a-
be done thus, be done likewise
These causative verbs are analyzable as:

(10)  
\('e-x-an-\)  
DDEF1-be-CAUS  
do thus  
\('a-x-\ddot{a}n-'i-\)  
DDEF1-be-CAUS-TRANS  
do thus, cause to resemble  
\('a-x-\ddot{a}n-'a-\)  
DDEF1-be-CAUS-INTR  
be done thus, be done likewise

Cahuilla 'éxan- is presumably from:

(11)  
\(*'a-yax-an-\)  
\(*'a-yx-an-*\)  
\(*'e-yx-an-*\)  
\('e-x-an-\)

Cahuilla further developed the following from yáx-:

(12)  
mé-x-an-  
IND1-be-CAUS  
do somehow, do something ('cause to be like something').

Similarly, Luiseño developed from yáx-:

(13)  
ma-x-án-'a-  
IND1-be-CAUS-INTR  
become, happen (i.e. 'be caused to be somehow')

\(ma-x-\ddot{a}n-'i-\)  
IND1-be-CAUS-TRANS  
bring, gather, get. (i.e. 'cause to be somehow')

Luiseño maxán-'a- and maxán-'i- presumably underwent historical changes something similar to:

(14)  
\(*m-yax-\ddot{a}n-\)  
\(*m-a-x-\ddot{a}n-\)  
\(*m-a-x-\ddot{a}n-\)  
ma-x-án-

The /a/ of ma- is epenthetic and may be harmonic with the /a/ of yáx-. Initial consonant clusters are almost non-existent in Luiseño.
Note that Luiseño maxán'a/i- have undergone considerable semantic shifts. Whereas a native speaker of Cahuilla would be prone to translate méxan- as 'do something' or 'do somehow', the native speaker of Luiseño would not at all be likely to translate maxán'a- as anything close to 'be somehow'. For the Luiseño speaker, maxán'a- simply means 'become' or 'happen', and maxán'i- 'bring', 'gather' or 'get'.

Although no longer common in everyday speech, Luiseño retains in archaic texts two words which are much closer in meaning and form to Cahuilla méxan- and 'éxan-:

(15) mí-yx-an-
    IND1-be-CAUS
    do somehow, do something

    'i-x-án-'i-
    PDEF-be-CAUS-TRANS
    say, do again

Returning to Cahuilla as the (in this instance) semantically more conservative Cupan language, we see that méxan- is analyzable as:

(16) mé-x-an-
    IND1-be-CAUS
    do somehow, do something (cause to be like something),

Cahuilla méxan- may have undergone historical changes something like:

(17) *m-yax-an-
    *mí-yax-an-
    *mi-ax-an-
    mé-x-an-

We see that from a base verb yáx-, Cahuilla has developed the following derived forms:

(18) BASE BE THUS BE SOMEHOW DO SOMEHOW DO THUS
    yáx-  'i-yax-  'é-x-an-
    mi-yax  mé-x-an-
Section II: Serrano Interrogatives

Representative of the majority of UA languages*, the interrogatives of Serrano all begin with /h/:

(19)
who   hamī’
what  hiit
where haayp
when haaypa’n
which hye’chat
why  hamin Raawnk
how  hamin
how much hami’qat

Serrano hamī’ and hamin may share a common base, *ami(‘):

(20) h-ami’  h-ami-n
IND2-?  IND2-?-INST
who       how

This base *ami(‘) may be related to the third person singular pronoun ‘ama’.

Serrano haayp probably contains the locative suffix -p:

(21) h-aay-p
IND2-?-LOC
where

Serrano hamin Raawnk, ‘why’ translates literally as ‘by means of doing what’:

(22) h-ami-n  Maa-wnk
IND2-?-INST  do-SSP
why (‘by means of doing what’)

Serrano hamya’qaT ‘how’ contains both the prefix h- and probably the verb ‘be’ qaT:

(23) h-amya’-qaT
IND2-?-be
how
Section III: Cupan Interrogatives

Like Serrano, the Cupan languages have interrogatives beginning with /h/:

(24) Luiseño Cahuilla Cupéñó
who háx háxí’ háx
what hiýcha hícha’ hish

Other Cupan interrogatives begin with m-:

(25) Luiseño Cahuilla Cupéñó
where michá’ míva’ míví’aw
when miikínga mípa’ mípa
why (hiyngayÁ) miyaxwe mivingáxÁ
how michá’ ‘axáninik méxannukÁ méxanukÁ

Thus far, we have seen in Cupan examples of both the indefinite prefixes, h- and m-. In each instance, Luiseño and Cahuilla have concurred in their choice of either h- or m- for a given interrogative. For ‘how much’, however, we find in Luiseño and Cahuilla:

(26) Luiseño Cahuilla
how much hík mik

The existence of both h-ík and m-ík suggests that one of the two indefinite prefixes, h- or m-, is an innovation which partially supplanted the other, original indefinite prefix. As h- is almost universal elsewhere in UA, the innovation in Cupan is more likely to be m-.

Section IV: Mesa Grande Diegueño Interrogatives

Mesa Grande Diegueño possesses an indefinite (interrogative) morpheme whose onset is /m/.

Mesa Grande Diegueño (Couro: 1973) reveals the following set of interrogatives:

(27) who maap
what ‘uuch
where maayÁ
when ma’yum kumyum
why mu(u)yuu/mu’yuuÁ
how mu(u)yuu/mu’yuuÁ
how much muuyum/mu’yum
We see that Mesa Grande Diegueño interrogatives largely begin with /m/, the only exceptions being 'uuch and kumyum.

Mesa Grande Diegueño kumyum may be composed of the common nominalizing prefix, ku-, added to a form beginning in m-. Langdon (PC) suggests for kumyum the following possible analysis:

(28) ku-m-yu-m
   NOM-INDF2-be-DS?
   when

Mesa Grande Diegueño 'uuch is closely related to the word for '(some)thing':

(29) 'uuchuch
   (some)thing

Langdon (PC) points out that most Mesa Grande Diegueño interrogatives may also be used as indefinites".

Based on the evidence in (27), we can posit an indefinite prefix m- for Diegueño.

We see that Diegueño 'how' is mu(u)yuu. Compare:

(30) mu(u)yuu                  puyuu
    (be, do) how               (be, do) thus

Note that mu(u)yuu means both 'how' and 'be/do (some)how', and puyuu means both 'thus' and 'be/do thus'.

Langdon (1978:104) confirms that #yuu, 'be', may also be reconstructed for Proto-Yuman.

The Diegueño verb puyuu, '(be, do) thus' is also prefixed with a reduced form of the independent demonstrative:

(31) p-uu
    DEM-3sg
    that (he, she, it)

According to Langdon (PC), mu(u)yuu and puyuu may be analyzed as follows:

(32) m-u-yuu"                  p-u-yuu
    IND2-3sg-be/do            DEM-3sg-be/do
    (be/do) somehow           (be, do) thus
Besides both showing evidence of the prefix m-, Yuman Mesa Grande Diegueño and UA Cahuilla, representative of all Cupan, also show some morphological similarities in rendering 'how' ('what way') and 'thus' ('that way'):

(33)  
<table>
<thead>
<tr>
<th>How:</th>
<th>Thus:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diegueño:</td>
<td></td>
</tr>
<tr>
<td>mu(yu)u</td>
<td>p(yu)u</td>
</tr>
<tr>
<td>Cahuilla:</td>
<td></td>
</tr>
<tr>
<td>méxannuk</td>
<td>'éxannuk</td>
</tr>
</tbody>
</table>

The initial morphemes of Cahuilla méxannuk and 'éxannuk correspond nicely to the Diegueño mu(u)yuu and p(yu)u (with the exception of a third person marker, absent in Cahuilla, and present in Diegueño -u-):

(34)  
<table>
<thead>
<tr>
<th>Cahuilla:</th>
<th>Diegueño:</th>
</tr>
</thead>
<tbody>
<tr>
<td>mé-x-an-nuk</td>
<td>m-u-yuu</td>
</tr>
<tr>
<td>IND1-be-CAUS-SSP</td>
<td>IND1-3sg-be</td>
</tr>
<tr>
<td>how?</td>
<td>how?</td>
</tr>
<tr>
<td>'é-x-an-nuk</td>
<td>DDEF1-be-CAUS-SSP</td>
</tr>
<tr>
<td>DDEF1-3sg-be</td>
<td>thus</td>
</tr>
</tbody>
</table>

UA Serrano also shows some evidence of alternating a demonstrative prefix (or root) with an indefinite prefix (or root):

(35)  
| h-aay-p IND2-?-LOC | 'i-p PDEF-LOC | 'a-p DDEF-LOC |

In set expressions, Serrano reveals a series of two-word constructions, the first indefinite, signalled by h-, the second definite, signalled by 'i-'. These set expressions all translate as 'some X or another':

(36)  
| h-amya'-qaT IND2-?-be PDEF-?-be |
| somehow or another          |
| h-ami  IND2-? PDEF=this     |
| someone or another          |
| h-aay-p  IND2-?-LOC PDEF-LOC=here |
| somewhere or another        |
The Serrano definite forms 'iv'i and 'ipe may be used independently to translate 'this' and 'here' respectively. The native Serrano speaker is, however, adamant about 'ivya'qat occuring only in this set expression following hamya'qat. The form 'ivya'qat cannot be used independently to express 'thus' or 'this way'.

Although we have some evidence of Serrano alternately prefixing an indefinite h- and a definite 'i- to the same base qat ('be'), in the set expression 'somehow or another', there is no evidence in Serrano of prefixing a definite affix to the existential verb qat to form 'thus'. Recall that Cahuilla (34) prefixes a definite affix and suffixes a causative affix to the existential verb yâx to form 'thus':

(37) 'é-x-an-nuk
DDEF1-be-CAUS-SSP
thus
'Thus' in Serrano is either an independent word, or a compound:

(38) tingîk pana' Maa-wnk
thus be/do-SSP
thus ('by being/doing thus')

There are further examples of indefinite Cupan m- alternating with a definite prefix 'i-'. Note the Cahuilla:

(39) mi-ngki-sh 'i-ngki-sh
IND1-kind-ABS PDEF-kind-ABS
some kind this kind

In a few set expressions, Cahuilla even has forms which alternate between m- indefinite, and p- definite, reminiscent of the Diegueño mu(u)yuu, puyu in form, and of the Serrano hamya'qat 'ivya'qat, 'somehow or another', in meaning:

(40) me-m-ik me-p-ék
RDUP-IND1-amount IND1-DEM-amount
a few (i.e. 'some amount or another')

mi-Niki' pè-Niki'
IND1-kind DEM-kind
some relative or another

Santa Ysabel Diegueño (Ponchettí; PC) also has a set expression reminiscent of Serrano hamya'qat 'ivya'qat, 'somehow or another':

(41) m-u-yuu p-u-yuu kinemi'
IND1-3sg-be DEM-3sg-be perhaps
somehow or another
The \( p(i) \)- of these Cahuilla set expressions (40) is probably related to the Cahuilla third person singular pronoun \( p\acute{e}^{'}, \) corresponding to Luiseño \( p\acute{e}'. \)

In contrast to Cupan and Mesa Grande Diegueño, recall that Serrano reveals for 'now' and 'thus':

\[
\begin{array}{ll}
\text{(42) Serrano} & \text{how:} & \text{thus:} \\
\text{h-ami-n} & \text{pana'} & \text{(maa-wnk)} \\
\text{IND2-?-INST} & \text{thus} & \text{be/do-SSP} \\
\text{how} & \text{thus} & \\
\text{h-amy\texttext{á}'-qaT} & \text{t\texttext{lin}ik} & \\
\text{IND2-?-be} & \text{thus} & \\
\text{how} & \text{thus} & \\
\end{array}
\]

Section V: Yuman Interrogatives

We have seen that \( m- \) is absent (or at least marginal) in Serrano interrogatives, as well as being almost universally unattested in all other UA languages.

It is common knowledge that the indefinite prefix \( h- \) of Cupan and Serrano is of Proto-Uto-Aztecan origin. If the indefinite prefix \( m- \), pervasive only in Cupan within UA, were universal in Yuman, the existence of \( m- \) in UA Cupan could be explained away as a borrowing from Yuman which would have entered Cupan through the Yuman language representative spoken directly adjacent to the Cupan-speaking territory. This Yuman language is Diegueño.

We find, however, that \( m- \) as an indefinite marker is pervasive only within one sub-branch of Yuman\textsuperscript{29}, namely among the California (Diegueño) languages of Delta-California:
(43) Delta-California:

<table>
<thead>
<tr>
<th></th>
<th>M Grande</th>
<th>Barona</th>
<th>Jamul</th>
<th>Campo</th>
<th>Cocopa</th>
</tr>
</thead>
<tbody>
<tr>
<td>who</td>
<td>maap</td>
<td>maap</td>
<td>ma'ap</td>
<td>ma'ap</td>
<td>lu:ny</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>lu:i:p</td>
</tr>
<tr>
<td>what</td>
<td>'uuch</td>
<td>'uuch</td>
<td>maayit</td>
<td>maalich</td>
<td>lu:ny</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>lu:i:p</td>
</tr>
<tr>
<td>when</td>
<td>kumyum</td>
<td>nyamu'yumchu</td>
<td>maayum</td>
<td>maayuum</td>
<td>nyimuulum</td>
</tr>
<tr>
<td></td>
<td>nyakumyum</td>
<td>ma'yuum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>where</td>
<td>maay</td>
<td>maay</td>
<td>ma'ay</td>
<td>ma'ay</td>
<td>0</td>
</tr>
<tr>
<td>why</td>
<td>muuyuu</td>
<td>muuyuu</td>
<td>0</td>
<td>muluum</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>mu'yuu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>how</td>
<td>muuyuu</td>
<td>0</td>
<td>0</td>
<td>me-m-luu te-m-paa (how are you?)</td>
<td>km'ap (how can it be?)</td>
</tr>
<tr>
<td></td>
<td>mu'yuu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>which</td>
<td>maayvech</td>
<td>maay</td>
<td>0</td>
<td>ma'ay</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>maayvu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h.much</td>
<td>muuyum</td>
<td>muuyum</td>
<td>chu'i</td>
<td>mu' lum</td>
<td>kla'âm</td>
</tr>
<tr>
<td></td>
<td>mu'yum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The River branch of Yuman languages shows evidence of both \( m^- \) and \( k^- \) as an indefinite prefix:

(44) River:

<table>
<thead>
<tr>
<th></th>
<th>Mojave</th>
<th>Yuma</th>
<th>Maricopa</th>
</tr>
</thead>
<tbody>
<tr>
<td>who</td>
<td>maka-ch (SJ)</td>
<td>maka-ny (OJ)</td>
<td>mki-sh (SJ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ki-sh (SJ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mki-ny (OJ)</td>
</tr>
<tr>
<td>what</td>
<td>kuch</td>
<td>ka-</td>
<td>kawit-sh (SJ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>kawish (OJ)</td>
</tr>
<tr>
<td>when</td>
<td>kanya:du:m</td>
<td>kanya:mathú:-</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>cem (When-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-does-he)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>where</td>
<td>maki</td>
<td>0</td>
<td>mki-ly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mki-ii</td>
</tr>
<tr>
<td>why</td>
<td>ka:du:k i-m</td>
<td>ka:thún-tek</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(he-does-why)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>how</td>
<td>0</td>
<td>ka:'étám</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(What-does-it say?)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ka'athóm3k</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>'athú:wú:m (How</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>shall I do it?)</td>
<td></td>
</tr>
<tr>
<td>which</td>
<td>makapch</td>
<td>makyip</td>
<td>kip-sh (SJ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mkip-sh (SJ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mkip (SJ)</td>
</tr>
<tr>
<td>h.much</td>
<td>k+ly3vi</td>
<td>ka:lyaví:m</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(What-does-it-resemble)</td>
<td></td>
</tr>
</tbody>
</table>

Further to the north, we have the Pai branch of Yuman languages, in no known direct historic contact with UA Cupan. Here we find that Walapai, Yavapai, and Havasupai exhibit only \( k^- \) as an interrogative prefix. Paipai, although a member of the Pai branch of Yuman, is now located directly south of the California-Delta branch of Yuman in Baja California and is known (Langdon: PC) to have borrowed a number of Diegueño traits. Paipai also exhibits both \( k^- \) and \( m^- \). Note that Hualapai and Havasupai /g/ is an orthographic convention for an unaspirated \( [k] \):
(45) Pai Branch:

<table>
<thead>
<tr>
<th></th>
<th>Hualapai</th>
<th>Yavapai</th>
<th>Havasupai</th>
<th>Paipai</th>
</tr>
</thead>
<tbody>
<tr>
<td>who</td>
<td>ga vga</td>
<td>vka</td>
<td>gaj ga’a</td>
<td>mka</td>
</tr>
<tr>
<td>what</td>
<td>gwəgəyuː</td>
<td>kweθ</td>
<td>gwee’e</td>
<td>0</td>
</tr>
<tr>
<td>when</td>
<td>ganyům</td>
<td>0</td>
<td>gaŋyum</td>
<td>qwas=ka-v1+wi-m-e</td>
</tr>
<tr>
<td>where</td>
<td>ge vge</td>
<td>vke</td>
<td>ge’e vge’e</td>
<td>mke</td>
</tr>
<tr>
<td>why</td>
<td>gavyuːjim</td>
<td>kavyum</td>
<td>gavyuim</td>
<td>kavyuyum</td>
</tr>
<tr>
<td></td>
<td>gavyuːjim</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gavyuːjim</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>how</td>
<td>ga ge gav</td>
<td>kavwik kavwim</td>
<td>ga’a</td>
<td>0</td>
</tr>
<tr>
<td>which</td>
<td>gayuːj (2 or more)</td>
<td>kavkyuchiva</td>
<td>gavyuj</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>gàv(gi)yuːj</td>
<td></td>
<td>gayuj</td>
<td></td>
</tr>
<tr>
<td>h.much</td>
<td>0</td>
<td>kavolik kavolim</td>
<td>gavlwig</td>
<td>kablui</td>
</tr>
</tbody>
</table>
Finally, the Kiliwa branch, represented historically only by Kiliwa in Baja California, shows evidence of $p-$ and $?-$ as interrogative prefixes:

<table>
<thead>
<tr>
<th>(46)</th>
<th>Kiliwa</th>
</tr>
</thead>
<tbody>
<tr>
<td>who</td>
<td>0</td>
</tr>
<tr>
<td>what</td>
<td>$?kwit</td>
</tr>
<tr>
<td>when</td>
<td>$?mat $p+$?im</td>
</tr>
<tr>
<td>where</td>
<td>$?ap-</td>
</tr>
<tr>
<td>why</td>
<td>$p+yu-m$</td>
</tr>
<tr>
<td>how</td>
<td>$p+yu-m$</td>
</tr>
</tbody>
</table>
| which    | $p+$?kwiiit  
|          | $?apu-win |
| how much | $p-+$yuu-(t) |

**Section VI: Conclusion**

Note the following map charting the distribution of $k-$, $m-$ and $h-$ among Yuman and Takic (UA) in the Southwest:
The map reveals that the California sub-branch of Yuman is probably the epicenter of the \textit{m}- indefinite morpheme. The California sub-branch of Yuman California-Delta is the only language group in either UA Cupan or Yuman where \textit{m}- is the interrogative marker par excellence. From this epicenter, \textit{m}- appears to have migrated out to both UA Cupan (Luiseño, Cupeño and Cahuilla), where it partially supplanted the UA indefinite marker \textit{h}-, to fellow Yuman Mojave, Yuma, Cocopa and Paipai in adjacent areas, and perhaps into non-Cupan UA Serrano where we have some traces of an indefinite \textit{m}-. The interrogative \textit{m}- therefore most likely originated in the territory occupied by Diegueño speakers. The ultimate source of \textit{m}- is still a mystery.

Further evidence of intra-familiar influence between Yuman Diegueño and UA Cupan is found in the parallel morphological analyses of Cupan and Diegueño ‘how’ and thus:

(47)

\textbf{UA Cupan (Cahuilla):} \hspace{1cm} \textbf{Yuman Diegueño:}

\begin{tabular}{ll}
\textit{mé-x-an-nuk} & \textit{m-u-yuu} \\
\textit{IND1-be-CAUS-SSP} & \textit{IND1-3sg-be} \\
how? & how? \\
\textit{é-x-an-nuk} & \textit{p-u-yuu} \\
\textit{DDEF1-be-CAUS-SSP} & \textit{DEM-3sg-be} \\
thus & thus
\end{tabular}

Intra-familiar influence between UA Cupan and Yuman Diegueño is thus two-fold:

1) In the sharing of the indefinite morpheme \textit{m}-

2) In the sharing of the concept of affixation of definite and indefinite prefixes onto a verb meaning ‘be, do’ (or in the case of Cupan, also ‘say’) to form words for ‘thus’ and ‘how’ respectively.
Notes:

1. The abbreviations used are the following:

CAUS  Causative
DDEF  Distal Definite
DEM  Demonstrative
DS  Different Subject
HAB  Habitual
IND1  Indefinite Type 1
IND2  Indefinite Type 2
INST  Instrumental
INTR  Intransitive
MC:PR  Modifying Clause Present
NOM  Nominalizer
OJ  Object
PC  Personal Communication
PDEF  Proximal Definite
PL  Plural
PR:PL  Present Plural
PR:SG  Present Singular
Q  Question Marker
RDUP  Reduplicated
SJ  Subject
SSP  Same Subject Participle
TRANS  Transitive
UA  Uto-Aztecan

The orthographic conventions for the Takic languages discussed in this paper are the following:

'ch' for /c/  'qw' for /q /
'kw' for /k/  'sh' for /s/
'l' for /l/  '$' for /s/
'ng' for /n/  'T' for /c/

'th' for / /

Long vowels are represented by two adjacent identical vowels. The acute accents signals primary stress.

2. Luiseño is a member of the Cupan branch of Uto-Aztecan. The Cupan languages are all indigenous to Southern California. There were originally three Cupan languages: Cupeño, Luiseño, and Cahuilla. Cupeño has recently become extinct. The Cupan languages are not the only Uto-Aztecan languages of Southern California; a higher sub-classification of Uto-Aztecan, Takic, includes other Uto-Aztecan languages of Southern California:
Uto-Aztecan
(So. California Branch)

Takic

Kitanemuk  Serrano

Cupan

Gabrielino-Fernandeño

Luiseño

Cupéno  Cahuilla

The above tree has been adapted from Jacobs (1975).

3. Cupéno has:

mixe-  happen
ixe-  to do a certain way, to say
hixe-  say (Hill 1973)

4. Langacker (PC) notes that the Numic branch of UA shows evidence of /y/ becoming nasalized to /ŋ/ through the influence of certain prefixes. It is possible that the onset /M/ of Serrano Mah is also the result of the nasalization of /y/. The meaning of this possible nasal prefix cannot be ascertained at this point. If Serrano Mah was once historically /yah/, then the divergences between Serrano and Cupan would be greatly reduced. If Mah = N-yah, then Serrano yah would be a cognate of Cupan yax. Serrano pana' Mah is already analyzable as:

pana' Mah
thus  be/do
be (do) thus (i.e. 'this way' or 'that way')

Serrano pana' Mah would therefore differ from Cupan (Cahuilla) 'i-yax-' and 'á-yax-' only by not directly prefixing a demonstrative, and by not distinguishing between the proximal and distal demonstrative.

Serrano would furthermore differ from Cupan in not having for Mah the meaning 'say', and by not having indefinite forms for Mah meaning 'be somehow' (Cahuilla mi-yax-) and 'say what' (Cahuilla hi-yax-).
5. Whereas Cahuilla has maintained the more original meanings given for (A) through (D) above, the Luiseño counterparts have undergone semantic shifts.

    In Luiseño, the form \(\text{iyx-}\) is now most commonly used to translate '(be) also', as in:

\[
\begin{align*}
\text{Nōon} & \quad \text{kūp-lowut} & \quad \text{iyx-lowut.} \\
1sg & \quad \text{sleep-gonna} & \quad \text{also-gonna} \\
\text{I'm going to sleep too.}
\end{align*}
\]

    The form \(\text{ááx-}\), now normally means 'resemble', or 'seem that way', as in (note the deletion of /x/ before /q/, as described in Elliott (1993)):

\[
\begin{align*}
\text{Né-yk} & \quad \text{áá-qat} & \quad \text{áá-q.} \\
1sg-to & \quad \text{seem-MC:PR} & \quad \text{seem-PR:SG} \\
\text{It seems that way to me.}
\end{align*}
\]

    The form \(\text{míyx-}\) is now used most often in the copulative sense (note once again deletion of /x/ before /q/):

\[
\begin{align*}
\text{Nóónil} & \quad \text{soldado} & \quad \text{míy-qu$.} \\
1sg & \quad \text{soldier} & \quad \text{be-PAST} \\
\text{I used to be a soldier.}
\end{align*}
\]

6. We saw above (2) that the verb \(\text{á-yax-}\) is still attested in Cahuilla. In some instances the second /a/ of \(\text{á-yax-}\) also deletes synchronically:

\[
\begin{align*}
\text{á-yx-anuk} \\
\text{DDEF-be-SSP} \\
\text{being that way, resembling}
\end{align*}
\]

    The second syllable of \(\#\text{á-yax-an-}\) was apparently also eliminated historically through the deletion of its nucleus /a/).

7. We posit an intermediate \(\#\text{á-yx-an-}\). There is at least one synchronic example where the Desert Cahuilla dialect has /ay/, and the Mountain Cahuilla dialect /ey/. Note:

Seiler (1979:250) gives kút yáy'al, 'whirlwind' for Desert Cahuilla. The Mountain Cahuilla equivalent is kút yéy'al.
8. Listed below are the UA interrogatives containing a bilabial (Langacker: PC):

<table>
<thead>
<tr>
<th>WHAT</th>
<th>WHERE</th>
<th>WHICH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono</td>
<td>pl’a’i</td>
<td>-(mi)y’i’i</td>
</tr>
<tr>
<td>Hopi</td>
<td>pl’</td>
<td></td>
</tr>
<tr>
<td>Hopi</td>
<td>him’i</td>
<td></td>
</tr>
<tr>
<td>Tarahumara</td>
<td>piri</td>
<td></td>
</tr>
<tr>
<td>Kawaiisu</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Papago     | hibai       |             |
| Tubatulabal| maa         |             |
| Hopi       | haqam       |             |
| Tarahumara | kami, kabu, kumi, kamu, kabe |     |

| Northern Tepehuan | ma-k’idl | ke-mi-’aane |
| Tubatulabal       | maa-l     | maa-t-wa-n |

| HOW MUCH          |          | HOW        |
| Tarahumara        | kipu, chu kipu, mapu kipu, kipu chi | kis kom |
| Huichol           | keepalme |            |
| Tubatulabal       | manhigish|            |
| Pochula           | kiskom   |            |

| Pochula           | ke(m)     | WHY        |
| Aztec             | keenin    |            |
| Tubatulabal       | mash      |            |

Note the high frequency of a possible prefix ma- in Tubatulabal.

9. Some Serrano words expressing uncertainty do however have /m/ initially: mit, ‘perhaps’, mermerher’, ‘any (old) X’, ‘any X at all’. Serrano mit is probably analyzable as:

mi-t
IND1-Q
perhaps

Whether or not the /m/ of such words is related to the indefinite m- of Cupan cannot be determined yet.
Further note that the /m/ of hamì’, hamìn, and hamya’qaT is also reminiscent of the /m/ of the indefinite Cupan m-. Note that Cahuilla has mi’, ‘which?’. If we accept the /m/ of the Serrano forms as related to the indefinite Cupan m-, we could analyze the three /m/-bearing Serrano interrogatives as:

| ha-mi’    | ha-mi-n       | ha-m-ya’-qaT       |
| IND2-IND1 | IND2-IND1-INST | IND2-IND1-be-be    |
| who       | how           | how                |

10. Luiseño hiyngay is analyzable as:

hiy-ngay
what-ABL
why (i.e. ‘from what’)

11. /nn/ of méxannuk reduces to /n/ in actual speech. We write /nn/ to stress the morphological make-up of méxannuk, for which see (34).

12. Cupeño has for both mivingax and mixanuk the alternate hingax (Hill: 1973).

13. Compare with Serrano h-aay-p:

h-aay-p
IND2-?-LOC
where

14. Langdon (PC) notes that mu(u)yuu means ‘how’ when used with the same subject suffix, ‘why’ when used with the different subject suffix.

15. Luiseño hiycha also does double duty as ‘what’ and ‘(some)thing’: in fact, like Mesa Grande Diegueño, all Luiseño interrogatives double as indefinites: michá’, ‘where?’, ‘somewhere’ etc.

16. Langdon (PC) notes that mu(u)yuu is a full-fledged verb which can be inflected for person:

me-me-yuu te-me-waa?
IND2-2-be AUX-2-sit
How are you?

17. Such morphologically related pairs for ‘how’ and ‘thus’ are by no means rare in other languages, for example Russian kak and tak.

18. Compare this construction to Cahuilla mì’miki’ péhiki’ below in (40).
19. Recall from Note 9 the alternative analysis of hamya'qaT:

ha-m-ya'-qaT
IND2-IND1-be-be
somehow

In this analysis, we would accept the sequence /ya'/ as a vestige of the verb *yah hypothesized above in Note 4, to which both the indefinites m- and h- would have been prefixed.

20. For all Yuman data I have used the source orthography.

21. Cahuilla also has an indefinite prefix qa- used in forming indirect questions, such as:

Kill pen'é'nanqa qa-méxannuk pish pe'külve'.
not I:know indirect-how that you:did:it
I don't know how you did it.
References:


Hyde, Villiana. *Personal Communication*.


Langdon, Margaret. *Personal Communication*.

Miller, Amy. ms. nd. *Dictionary of Jamul Diegueño*.


----. Personal Communication.


Ramon, Dorothy. Personal Communication.

Sauvel, Katherine. Personal Communication.


Kroeber and Harrington on Mesa Grande Diegueño (Iipay)

Margaret Langdon, UCSD

Introduction

At the First J.P. Harrington Conference held in Santa Barbara in 1992, I reported that I had been unable to locate in the available guides to the field notes of J.P. Harrington (Mills 1981, Walsh 1976) any reference to material on Diegueño prior to about 1925, even though Kroeber and Harrington (1914)—the first description of the phonetics of Diegueño—states that Harrington "... had a brief opportunity to hear Diegueño..." I was then informed by Mr. John Johnson of the Museum of Natural History in Santa Barbara that the Museum's archives contain the manuscript of Kroeber and Harrington (1914) and other documentation relevant to that article. I am grateful to Mr. Johnson for allowing me access to this archive and giving me the opportunity to make xerox copies of the relevant sections. A major portion of this paper relates my observations on this material.

Background

In the spring of 1963, just over thirty years ago, I started working on Diegueño in San Diego County. At the time, with the exception of a few unpublished wordlists collected by non-linguists, there was only one published source on any aspect of the structure of the language, Kroeber and Harrington's (1914) 'The Phonetic Elements of the Diegueño Language' (hereafter KH). The paper was essentially written by Kroeber and its purpose was to compare some Diegueño data collected by him with the equivalent forms in Mojave,¹ a language Kroeber had had considerable exposure to and had described phonetically in Kroeber (1911). Notes on the comparative phonetics of these two related languages² and a comparative word list of 75 words

¹ I follow the most common current practice of spelling the name of this language as Mojave, except when quoting sources where the spelling Mohave is used.

² The place of Mojave and the various Diegueño languages in the Yuman family is given in the following subgrouping. Note I now consider there to be at least three Diegueño languages (Langdon 1990). Dialects relevant to this paper are listed in parentheses.

California-Delta:

Diegueño:
   Iipay (Mesa Grande, San Pasqual)
   Kumeyaay (Campo)
   Tiipay (Jamul)

Cocopah

River:

Yuma
Maricopa
Mojave

170
are found in KH.

The unexpected discovery of unpublished Harrington notes supplementing KH, including Harrington's re-eliciting of the same 75 words provided by Kroeber gives me a unique opportunity to compare their transcriptions to each other and to my own transcriptions begun some fifty years later. Particularly fortunate is the fact that Kroeber's consultant, but not Harrington's, was Mr. Rosendo Curo, the father of my own main consultant. There can therefore be no doubt that we are dealing with the same subdialect of fipay Diegueño.

The history of KH

The contribution of Harrington to KH is described by Kroeber as follows:

Independent observations on the phonetics of the language courteously furnished by Mr. J.P. Harrington, who has had a brief opportunity to hear Diegueño, have been added as notes initialled by him. (KH:177)

The manuscript of the paper in the archives at Santa Barbara consists of a carbon copy of a typewritten version obviously prepared by Kroeber. The author is listed as Kroeber only and no reference to Harrington is made. The text is much like the published article, but the footnotes by Harrington are not present. There are a few typographical errors, which were mostly corrected in the published version. The implication is that this represents the copy of the manuscript sent by Kroeber to Harrington, presumably to prompt his comments. It is not clear in view of the lack of co-author on this version of the paper that there was at the time any plan to write the paper jointly, although something of the sort must have been agreed to (maybe later) in view of a covering note by Harrington transmitting his comments to Kroeber, the relevant part of which is quoted below.

Here is the paper. On reviewing my notes and trying to incorporate my cold material with your cold material, I became so dissatisfied [the word is crossed out in the original ML] discouraged that I decided to adopt the present form [i.e. footnotes ML], which I hope will prove satisfactory. [I have tried my best and hope that you will appreciate my endeavor.]

I am so rushed to death with a lot of things that simply must be done that I hardly know where I am at...

I presume that by "here is the paper" Harrington meant that he had annotated a copy of the paper to be returned to Kroeber, although this is not sure, since there are handwritten notes by Harrington of more or less the material that appeared as footnotes in the article. There seems to

Pai:
1. Upland:
   Havasupai
   Hualapai
   Yavapai (Tolkepaya)
2. Paipai
Kiliwa (KI)

Sources of data used in this paper other than KH are as follows. Mesa Grande: Couro and Hutcherson (1973); Cocopa: Crawford (1989); Maricopa: Gordon (ms); Yuma: Halpern (1946-47); Mojave: Munro, Brown, and Crawford (1992); Tolkepaya: Munro and Fasthorse (1993ms); Jamul: Walker (ms).
have been the intention of a closer collaboration on the writing of the paper at one time which was subsequently abandoned by Harrington. I rather suspect that Harrington would not have been an ideal co-author.

Most of the footnotes actually appeared in print. The major difference is in the wording of the paragraph describing Harrington's contribution, the printed version of which has already been quoted above. The Harrington text is as follows:

Certain observations on the phonetics of the language by Mr. J.P. Harrington, have been added. Mr. Harrington had a brief opportunity to hear Diegueño in July, 1908 and [crossed out in the original ML] February and again in September, 1913, while making collections for the Panama-California Exposition.

The crossed out reference to 1908 refers to the collection of material, parts of which were published in a footnote to Harrington (1908), which cites numerals in several Yuman languages, including Diegueño. The Diegueño numerals are interesting in their own right but I suspect they are not from the same Diegueño language as the 1913 notes, as they seem to represent a dialect with some Yuma influence. The collection of the 1913 material was obviously commissioned by the organizers of the Panama-California Exposition of 1915, which took place in San Diego. Why mention of this was omitted from the published version of the paper is not clear, but it is possible that it might not have been politically wise to mention that Harrington was collecting information under the sponsorship of the Exposition for other than his commissioned task.

The fieldnotes of Harrington's trip in 1913 are also available from the set of materials at the Santa Barbara Museum of Natural History and consist of 34 pages of mostly ethnographic notes (with appropriate lexical forms for the items discussed). It is likely that the contribution for the KH paper was based on the notes of September 1913 (although no date except 1913 is given in the notes), to which is appended, almost as an afterthought, a wordlist which clearly was elicited from Kroeber's Diegueño wordlist and thus allows direct comparison of Kroeber's and Harrington's transcriptions. Since this list was not published in KH, I reproduce it here as Table 2. This list was apparently obtained from Mr. Isidro Nejo at Mesa Grande, who is identified by name in Harrington's footnote 9 (KH:179). According to my consultants, Mr. Nejo was originally from San Pasqual where an Iipay dialect is also spoken. With only minor differences, his list matches that of Mr. Rosendo Curo, Kroeber's Mesa Grande consultant.

Kroeber's transcription

Kroeber's transcription practices need comment. I have reproduced the comparative wordlist of Mojave and Diegueño from KH as Table 1, which should be consulted while reading this section.

A rather puzzling practice is the inclusion of dashes in certain forms. In the most obvious cases, this is done when a segment in one of the languages is not present in the other. So 6, the word for 'five', has a final -k in Mojave not present in Diegueño. 7 'salt' is the other way round, Diegueño has final -ly, which is absent in Mojave, just as in 52 'leg'. This final -ly is a peculiarity of some words in a few varieties of Diegueño and its origin is obscure. kwe- is missing in Mojave 15 'white'. 31 'hot' has -k in Mojave and nothing in Diegueño. Of a different order are the dashes in 9 'ocean' which is a compound of the words for 'water' and 'salt' in both languages.
What to make of dashes in 13 'star', 14 'mockingbird' and 41 'mortar' is not so clear, but probably implies that only one part of the words is being compared; in 32 'skunk' the Diegueño form contains a typographical error already present in the manuscript, the word is *kallytzwiiw* and does not contain an *r*, as is duly pointed out by Harrington in footnote 31.

The interpretation of vowel symbols also calls for comment. A good deal of discussion is devoted to the "slurring of unaccented vowels", the famous schwa problem of Yuman languages, but no mention is made of vowel length, even though "lengthened consonants" are discussed to some extent. Note on Table 1 that the only vowel diacritic used in Diegueño is the macron, whereas Mojave also has some grave accents which are unexplained in the paper. For their interpretation it is necessary to refer to Kroeber (1911). It turns out that the macron for Kroeber means not only length, but simultaneous "close" pronunciation, while the grave accent means the vowel is long and "open". Kroeber indicates stress in Mojave by the acute accent following the stressed vowel. No indication of stress in found in KH either in Mojave or Diegueño, and the only discussion of "accent" states "the stress and pitch accents of Diegueño seem to be identical with those of Mohave" (KH185). This is probably not correct since Mojave has (as discussed in Kroeber 1911:63-64) a number of exceptions to the general Yuman final stem syllable stress rule. An example of contrasting forms given by Kroeber (KH64) is *a'ha* 'water' vs *ah'ata* 'cotton-wood'. This is still the case in present-day Mojave, as described in Langdon (1977), where it is demonstrated that the perception of stress in non-stem-final syllables is a manifestation of underlying final-syllable stress conditioned by the distribution of long and short vowels in the word. Because this is essentially predictable, modern recordings of Mojave (e.g. Munro, Brown, and Crawford 1992) do not indicate this. In my exposure to various dialects of Diegueño, I have observed no such alternations in the place of stress, as all words have stress on the stem-final syllable.

The only indication of length in the Diegueño words is the macron and length does not necessarily match between Mojave and Diegueño.

In the case of unstressed vowels, i.e. any vowel not in stem-final syllable, Kroeber uses mostly the same vowel symbols as for stressed vowels, except for the additional symbol *E* apparently representing a central schwa-like unrounded vowel. It also appears once in stressed position in 28 'tongue', a word in which I heard the same vowel as in 29 'ear', i.e. *a/. In stressed position, Kroeber's practice implicitly recognizes the basic 3-vowel system of Diegueño (as opposed to the 5-vowel system of Mojave), with the exception of 28 discussed above, and also 58 'two', the only instance of *o* in the entire Diegueño list; this captures quite nicely the backed and rounded allophone of */a/* in that word.

Palatal consonants are written as clusters, *tc* [ʦ], *ly* [ʃ], etc; voiceless laterals are written *L* and *Ly* or *Li*. Initial glottal stops are not marked, but true vowel-initial words begin with the symbol denoting the slightly aspirated onset of these vowels, so the contrast is captured adequately. Long vowels are discussed in more detail below.

There is one serious typographical error in the Table of Consonants (KH:183) where the sound described as "affricate, half sonant or aspirated surd" is listed as *t*, but should have been *tc*.

---

3 Throughout this paper, I cite forms in the orthography of the sources; some are standard phonemic transcriptions, some are practical orthographies also based on phonemic principles, the conventions of which should not cause any serious confusion.
Harrington's transcription

Table 2 should be consulted while reading this section. Harrington's consonant notation is pretty straightforward, but a few comments are needed in special cases. Voiceless laterals use the notation \( l \); palatal consonants are as follows: \( j[y], \breve{l}, n_y, \dot{t}_\varepsilon, l_\varepsilon \). Words written with initial vowels should be interpreted as beginning in glottal stop, though in a few cases (24, 26, 66, 67, 68, 75), a glottal stop initial is marked as \( \breve{a} \). The aspirated onset of true vowel-initial words is indicated as \( h \), an unambiguous notation, since the language does not have an /h/ phoneme.

Stress is marked on most forms; Harrington uses the acute accent over the vowel. It always appears in the stem-final syllable, as is still the case in the language today. In stressed position Harrington (like Kroeber) implicitly recognizes the basic three-vowel system. In addition to \( i, u \), \( a \), in stressed position, we also find \( \Lambda \) in 49, 58, 62, 68, which correctly captures the allophone of /a/ in these environments. The \( \varepsilon \) in 11 correctly represents the phonetic quality of the stressed vowel in this word. Its quality is actually predictable if it is recognized that the final consonant should have been \( \breve{a}_\varepsilon \) and not \( \dot{a}_\varepsilon \).

Harrington was acutely aware of the difficulty in transcribing unstressed vowels. His statement is particularly apt as it describes my own predicament when I started fieldwork on the language. In fact, it has sustained me through the nightmare of trying to make sense of them. It is reproduced below.

The determination of the quality of the vowel in these unaccented syllables proved so baffling that I determined to operate with a large number of characters. I found myself using nearly all the symbols for mixed vowels provided by the alphabet of the International Phonetic Association. (KH:184, fn 20)

Most of the vowels in non-final syllable position in the Diegueño material are of this type. These are the vowels that I eventually phonemized as schwa. Attested in Harrington's hand are: \( i, e, a, y, \breve{y}, \breve{\breve{y}}, \Lambda, \varepsilon, \breve{\varepsilon}, a, u \). \( y \) seems to stand for \( [\tilde{I}] \). What he meant by \( \breve{y} \) in 19 as distinct from \( y \) elsewhere, I don't know. Note that some words, i.e. 35 'roadrunner' and 38 'moon', have what appears to be a marker of secondary stress; one, 13 'star', even seems to have primary stress on the first syllable, something I have never heard.

The s/s problem

On the whole, Harrington and Kroeber's transcriptions, allowing for differences in notation, are remarkably similar, suggesting that Kroeber's was a pretty accurate transcription, capturing the state of the language on the Mesa Grande reservation where both interviews were conducted from two different speakers. In view of this, it is all the more surprising that some aspects of both Kroeber's and Harrington's transcription do not discriminate contrasts attested in my own recordings. One is the omission of the sound \( s \), a post-alveolar fricative often informally called California \( s \) (discussed in detail in Bright 1978). California \( s \) or its reflex clearly contrasts with dental \( s \) in all varieties of Diegueño. Harrington heard it sporadically in some of his later Diegueño transcriptions around 1925, but not in 1913. There are, however, very few instances of words that should contain it in the KH wordlist. They are 17 'buzzard', 14 'mockingbird', and 70 'bird', all containing the word for 'bird', in two cases as first member of a compound. It is nevertheless surprising that Kroeber did not distinguish a second s-sound, since in discussing fricatives, he notes:

Mohave surd interdental \( \theta \) is always \( s \) in Diegueño (4-10, 51). Diegueño \( s \), however, corresponds also to Mohave \( s \) (11-16).
<table>
<thead>
<tr>
<th>English</th>
<th>Mohave</th>
<th>Diegueno</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>alte</td>
<td>zx-i'in²⁴</td>
</tr>
<tr>
<td>good</td>
<td>nēhāt</td>
<td>zxann₂⁵</td>
</tr>
<tr>
<td>sex</td>
<td>marhó</td>
<td>parzau,</td>
</tr>
<tr>
<td>fly</td>
<td>čılı:nəmo</td>
<td>mesa-pu:ly</td>
</tr>
<tr>
<td>woman</td>
<td>čuna:čika</td>
<td>sīnyəx</td>
</tr>
<tr>
<td>sève</td>
<td>čaráp-k</td>
<td>saray²⁶</td>
</tr>
<tr>
<td>salt</td>
<td>sə'1</td>
<td>zə:ly</td>
</tr>
<tr>
<td>drink</td>
<td>səi</td>
<td>səi</td>
</tr>
<tr>
<td>ocean</td>
<td>(a)ha-čo'ilya</td>
<td>xə:nə:l</td>
</tr>
<tr>
<td>raccoon</td>
<td>nəmaθa</td>
<td>nəməs</td>
</tr>
<tr>
<td>hand</td>
<td>lεsλya</td>
<td>ə:ku</td>
</tr>
<tr>
<td>liver</td>
<td>tipasa</td>
<td>tə:ipəzəl</td>
</tr>
<tr>
<td>star</td>
<td>ha-muse</td>
<td>kwili-mesap²⁷</td>
</tr>
<tr>
<td>mockingbird</td>
<td>sakwa:da'əlyə</td>
<td>sakwi-laui</td>
</tr>
<tr>
<td>white</td>
<td>nyaməsəm</td>
<td>kwə:nimsap</td>
</tr>
<tr>
<td>buzzard</td>
<td>əcəi</td>
<td>ə'li</td>
</tr>
<tr>
<td>eagle</td>
<td>nεpə</td>
<td>kə:n-paşi</td>
</tr>
<tr>
<td>sleep</td>
<td>hum-kə</td>
<td>xə:mat</td>
</tr>
<tr>
<td>tooth</td>
<td>tə:nyu</td>
<td>xya:n</td>
</tr>
<tr>
<td>eye</td>
<td>tə:lyu</td>
<td>lyu</td>
</tr>
<tr>
<td>sweet</td>
<td>maruły-k</td>
<td>miyuły-k</td>
</tr>
<tr>
<td>where</td>
<td>makı</td>
<td>małyə²⁸</td>
</tr>
<tr>
<td>stone</td>
<td>avl</td>
<td>kwi²⁹</td>
</tr>
<tr>
<td>house</td>
<td>avə</td>
<td>awa</td>
</tr>
<tr>
<td>south</td>
<td>kəvək</td>
<td>kawak</td>
</tr>
<tr>
<td>rattlesnake</td>
<td>kəv</td>
<td>kwi</td>
</tr>
<tr>
<td>you (pl.)</td>
<td>máteva</td>
<td>miwayapte³⁰</td>
</tr>
<tr>
<td>tongue</td>
<td>tɨpιya</td>
<td>tɨnəple</td>
</tr>
<tr>
<td>ear</td>
<td>ɬaməya</td>
<td>ɬamət</td>
</tr>
<tr>
<td>asb</td>
<td>həmmyəle</td>
<td>empil</td>
</tr>
<tr>
<td>hot, day</td>
<td>tɨpιly-k</td>
<td>tɨpιl</td>
</tr>
<tr>
<td>skunk</td>
<td>ɬiyənəυ</td>
<td>kwi:lyəxwiru³¹</td>
</tr>
<tr>
<td>rabbit</td>
<td>həλya:avə</td>
<td>xə:laui</td>
</tr>
<tr>
<td>spider</td>
<td>həλyətə</td>
<td>xə:tət</td>
</tr>
<tr>
<td>roadrunner</td>
<td>təlypo</td>
<td>təlpu</td>
</tr>
<tr>
<td>gourd</td>
<td>ɬəbəyalə</td>
<td>axma:l</td>
</tr>
<tr>
<td>rat</td>
<td>amalyka</td>
<td>malk</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>38 moon</td>
</tr>
<tr>
<td></td>
<td></td>
<td>39 arrow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40 bead</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41 mortar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>42 so</td>
</tr>
<tr>
<td></td>
<td></td>
<td>43 sun</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44 Pleiades</td>
</tr>
<tr>
<td></td>
<td></td>
<td>45 metate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>46 mountain-sheep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>47 dog</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48 night</td>
</tr>
<tr>
<td></td>
<td></td>
<td>49 earth</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50 sky</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51 medicine man</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52 leg, foot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>53 cloud, rain</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54 belly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55 nose</td>
</tr>
<tr>
<td></td>
<td></td>
<td>56 nail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>57 white man</td>
</tr>
<tr>
<td></td>
<td></td>
<td>58 two</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59 mouth</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60 knee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>61 old man</td>
</tr>
<tr>
<td></td>
<td></td>
<td>62 blood</td>
</tr>
<tr>
<td></td>
<td></td>
<td>63 snow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>64: fire</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65 dance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>66 bow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>67 tobacco</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68 deer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>69 Badger</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70 bird</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71 raven</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72 no</td>
</tr>
<tr>
<td></td>
<td></td>
<td>73 this</td>
</tr>
<tr>
<td></td>
<td></td>
<td>74 bad</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75 canoe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>76 tæila</td>
</tr>
</tbody>
</table>

Table 1. Kroebber's comparative wordlist

²² With the alternance Mohave -o - Diegueno -au; cf. Mohave -o - Diegueno -o in no. 51. — J. P. H.
²³ Also given as kə:n-ik, it is one. — J. P. H.
²⁴ It is stated that the proper Diegueno term for sève is ɬə:nəkäl. — J. P. H.
²⁵ The informant gave kwa:nmesap, evidently a variant form. — J. P. H.
²⁶ ɬə:pa! — J. P. H.

²⁷ Given as maipə. — J. P. H.
²⁸ Given also as kwili. — J. P. H.
²⁹ Given as minyawap. — J. P. H.
³⁰ Given as kə:nkə:nwiru. — J. P. H.
³¹ Recorded as kwaziyalan. — J. P. H.
³² Mohave bə< kwel! — J. P. H.
³³ The pronunciation is practically identical with the Cocopa. — J. P. H.
| 1  | one       | exyn'k        |
| 2  | good      | axán         |
| 3  | fox       | parxáu       |
| 4  | fly       | mýsxapúk     |
| 5  | woman     | šín'g        |
| 6  | five      | xásál xáká' |
| 7  | salt      | šínš         |
| 8  | drink     | èst'         |
| 9  | ocean     | axa(h) asílį́k|
| 10 | raccoon   | namás        |
| 11 | hand      | hasát        |
| 12 | liver     | týypšį́        |
| 13 | star      | kwán'msą́p    |
| 14 | mockingbird | sakwyláu   |
| 15 | white     | kwynam'șą́p   |
| 16 | buzzard   | sa'į́        |
| 17 | eagle     | axpá         |
| 18 | sleep     | kχýmá'(kxma')|
| 19 | tooth     | hýjá'        |
| 20 | eye       | hós'ú        |
| 21 | sweet     | myjukk       |
| 22 | where     | máipę́če     |
| 23 | stone     | èwilę̊́      |
| 24 | house     | λwá        |
| 25 | south     | kwék̕        |
| 26 | rattlesnake | 'lwi̊́  |
| 27 | you (pl.) | minawáp      |
| 28 | tongue    | hanápá̊́      |
| 29 | ear       | xamał        |
| 30 | ash       | źąmpılı̊́     |
| 31 | hot, day  | hupį́        |
| 32 | skunk     | kałyxwį́́̊́     |
| 33 | rabbit    | xalašą́        |
| 34 | spider    | xalštút      |
| 35 | roadrunner | týlį́pů́    |
| 36 | gourd, turtle | axnáš̕     |
| 37 | rat       | amaš̕į́        |
| 38 | moon      | xýty'á̊́      |
| 39 | arrow     | apál (not k) |
| 40 | beard     | halamį̊́      |
| 41 | mortar    | kalxwá'       |
| 42 | see       | ayů̊̀       |
| 43 | sun       | yńą̊́       |
| 44 | Peiades   | xỳxą̀        |
| 45 | metate    | axpi̊́        |
| 46 | mountain sheep | šíntą́  |
| 47 | dog       | axát         |
| 48 | night     | tìŋą̊́m      |
| 49 | earth     | amat         |
| 50 | sky       | amáj̕        |
| 51 | medicine man | kwasijáí  |
| 52 | leg, foot | hemił̕         |
| 53 | cloud, rain | akwį́  |
| 54 | belly     | hétú̊́       |
| 55 | nose      | haxú̊́       |
| 56 | nail      | sišą́xwáą̊́     |
| 57 | whiteman  | xaló̊́ (Mexican) |
| 58 | two       | xawák        |
| 59 | mouth     | ha           |
| 60 | knee      | hamąPATunǻn |
| 61 | old man   | kwyą̊́į́ (not trilled) |
| 62 | blood     | haxwát       |
| 63 | snow      | halá́p       |
| 64 | fire      | áú̊́         |
| 65 | dance     | hýmá́       |
| 66 | bow       | 'atýį́       |
| 67 | tobacco   | 'úp           |
| 68 | deer      | 'akwák       |
| 69 | badger    | maxwá́       |
| 70 | bird      | asá         |
| 71 | raven     | axák (crow)  |
| 72 | no        | humáą̊́       |
| 73 | this      | piją́       |
| 74 | bad       | wyńti̊́       |
| 75 | cane      | 'ąxtá́       |
In fact, the latter part of that statement is correct only in a couple of exceptional cases (involving sound symbolism), whereas the regular correspondence is Mohave ʂ: Diegueño ṣ. I don’t know how to account for this deficiency. It is truly doubtful that both Kroeber’s and Harrington’s consultants had no contrast between these two segments, especially since Kroeber’s Diegueño consultant was Rosendo Cuero, of Mesa Grande, the father of Ted Cuero.\(^4\) In the speech of Ted Cuero, as well as in that of all other speakers of Diegueño I have worked with, there is a contrast between two kinds of s’s, and the Yuman comparative evidence in general makes clear that the contrast is archaic. Perhaps this was both Kroeber’s and Harrington’s first exposure to California s contrasting with another s-sound and they consequently interpreted ṣ as a variant of s as in some American dialects of English.

The long vowel problem.

The most startling feature of both Kroeber’s and Harrington’s notation is the paucity of long vowels, which are clearly distinctive in the language in both stressed and unstressed position. They are harder to hear in unstressed position and only after repeated exposure did I consistently record them there, so it is understandable that upon a first short exposure to the language this contrast might have been missed. In modern Mesa Grande speech, unstressed long vowels are found in 4 ‘fly’ mes-haapwuly where Harrington, but not Kroeber, correctly heard the length of the stressed final syllable, but neither heard the unstressed aa as long. Other instances of long unstressed vowels occur in 14 ‘mockingbird’ aasakaawilsaaw, 27 ‘you (pl)’ memyaawap, 48 ‘night’ tiinyaam, 63 ‘snow’ aalap, 65 ‘dance’ limaa, 66 ‘bow’ aatim, 70 ‘bird’ aashaa. These are quite distinct from the a, i, u colored short vowels I analyze as schwa, and I am at a loss to explain the lack of distinctions, especially in Harrington’s transcription, where he is at such pains to use many different symbols for unstressed vowels.

Length is also contrastive for stressed vowels in closed syllables; Harrington’s transcription of these is better than Kroeber’s, as in 4 ‘fly’ mes-haapwuly, 25 ‘south’ kewak; he does not hear vowel length consistently in this environment, even in a near minimal pair like 36 ‘gourd’ ahnall, vs 37 ‘woodrat’ emallk.

Diegueño also has contrastive short and long diphthongs. In the KH wordlist, the only diphthongs are what I transcribe as long ones, there being no instances of short diphthongs in the list. Note, however, that 42 ‘see’ Ewu for Kroeber, has a long diphthong in modern Mesa Grande where the word is ewuuw.

The problem of vowel length is most acute in the case of final open stressed syllables, where in modern Diegueño dialects, at least in the Lipay area where Mesa Grande is located, but also in Kumeyaay varieties, final stressed vowels are all long. In Kroeber’s list, only four are long: 20 ‘eye’, which in fact is a (long) diphthong and is accurately so transcribed by Harrington (see Table 2); 26 ‘rattlesnake’, also long for Harrington; 59 ‘mouth’, short for Harrington but long for Kroeber; and 75 ‘cane’, long for Kroeber but short for Harrington. So it’s not that Kroeber and Harrington could not hear final long vowels at all, and something more must be going on.

Before engaging in speculation about the meaning of these discrepancies, it is necessary to discuss what is known about both Kroeber and Harrington’s practice and reliability with respect

\(^4\) The difference in spelling of the last name is not significant.
to vowel length. It is often asserted that Kroeber was not a good phonetician so that his transcriptions may not be overly reliable. Some examples of this have been noted above with respect to the two varieties of s in Diegueño. However, the majority of forms relevant here are fairly accurately transcribed for a first exposure to the language. Taking first his Mojave transcriptions, it should be noted that vowel length is not altogether easy to hear in that language. Thus, Pamela Munro\(^5\) states:

There is no question that Mojave has a clearcut contrast between short and long vowels, in both stressed and unstressed positions, as attested by the existence of numerous minimal pairs for vowel length. However, researchers from Kroeber to Crawford to myself have often found it difficult to determine the length of particular stressed vowels. (Munro, Brown, and Crawford 1992:4)

When comparing Kroeber’s Mojave wordlist with modern recordings of the language (Munro, Brown, and Crawford 1992) there are a total of 27 mismatches of vowel length in 75 words, each of which contains at least two vowels (many have three). There are thus many more agreements than disagreements. 10 of the mismatches are in unstressed position, the most difficult position in which to hear vowel length distinctions in Yuman languages, leaving only 17 mismatches in stressed vowels. Given the fact that the data were collected from different speakers at an interval of more than 50 years, and that vowel length is notoriously hard to hear in Mojave, these discrepancies are not altogether surprising.

Admittedly, Kroeber had much less exposure to Diegueño than to Mojave, but he certainly had non-trivial phonetic expectations of Yuman languages, which would in general be an asset rather than a liability. Kroeber’s transcriptions of Mesa Grande forms should therefore not be disregarded, and the fact that he heard both short and long vowels in final open stressed syllables must be taken into account.

As for Harrington’s accuracy in transcribing vowel contrasts, the following observations are relevant. His transcription of vowel length in Mojave in early fieldnotes suffers from the same problems as Kroeber’s, but is not particularly more inaccurate (Pamela Munro, p.c.) In the case of Diegueño, in the 1913 wordlist, he heard some long vowels (as discussed above, and see Table 2). In later notes on the language, he heard long final stressed vowels quite regularly in 1925, when he writes them as sequences of two vowels, e.g. xəp'əa ‘moon’ a word in which he heard a short vowel in 1913. It is my conclusion that, as with Kroeber, Harrington’s perception of vowel length cannot be discounted out of hand.

I conclude that both Kroeber and Harrington’s transcriptions of long vowels, while imperfect, reflect something significant about the data. The conclusion seems warranted that the Mesa Grande dialect in 1913 showed both short and long final stressed vowels, while in Mesa Grande in the early 1960’s, these were all long.

The Yuman comparative evidence

Iipay and Kumeyaay are the only languages of the whole Yuman family to consistently have long vowels in open stressed final syllables at the present time; all others have both short and long vowels in that position. Vowel length distinctions are obviously archaic in Yuman,

\(^5\) I am grateful to Pamela Munro for discussing the issues raised in this paper, particularly with respect to the vowel length problem, and for providing recently rechecked data from Tolkapaya and Maricopa.
playing an important role in the morphology, as length alternations pervade the derivational morphology and also distinguish unrelated lexical items. It is nevertheless the case that not all cognate sets show full agreement for vowel length and determining whether a reconstructed Proto-Yuman form should have a long or a short vowel is often quite difficult.

As an illustration of this state of affairs, I have gathered in Table 3 data from all Yuman subgroups (except Kiliwa which is considerably divergent) for the words in KH ending in stressed vowels. I have included data from Jamul Diegueño, a dialect which does not regularly have long final stressed vowels. Also included are cognates from Cocopa, in the same subgroup as Diegueño, the three River languages Yuma, Mojave, and Maricopa, and one Pai dialect of Yavapai, Tolkapaya. These were chosen because the data available for them in the database for the Comparative Dictionary of the Yuman Languages\(^6\) are most complete and have been checked for accuracy.

These forms are compared to Harrington's 1913 Mesa Grande Diegueño list.

The results are, if not totally conclusive, at least of considerable interest. What is abundantly clear is that all languages listed, including Lipay in 1913, have vowel length distinctions in final stressed vowels, although the vowel length does not necessarily match across languages. While it is certain that vowel length was distinctive in Proto-Yuman, Table 3 shows that there are some problems deciding which forms should be reconstructed with final long vowels in Proto-Yuman, since the distribution of long vowels acrosss the languages is erratic.

Of the 24 sets in Table 3, only 7 agree in the length of the final vowel, and all of them are consistently short. They are 24 'house', 35 'roadrunner', 41 'mortar', 45 'metate', 46 'sheep', 56 'white man', and 65 'dance'. They can presumably be reconstructed securely as short in Proto-Yuman. Not only do all the others show a mixture of long and short vowels across the languages, but no distinct pattern of length systematically clustering in the same languages emerges.

Obviously, the reconstruction of these forms in Proto-Yuman raises a number of problems, and this topic needs further study.

**Chronology of changes within Diegueño**

I have presented evidence above that the rule of lengthening of stressed final vowels in Lipay and Kumeyaay is a recent development having taken place sometime between 1913 and as early as 1925 in some areas (as attested in fieldnotes of Harrington at that time) and most certainly was complete before 1960 when Bright conducted his dialect survey of Diegueño and found essentially the same situation.

Since there are other traits that distinguish the various varieties of Diegueño, it is tempting to attempt a chronology of these distinctions.

First, there is the odd fact that a few words which in some varieties of Diegueño and all other Yuman languages end in a short high front vowel, end in others in -ily. The source of this distinction is far from clear. Its distribution is limited to Lipay, Sycuan, and San Miguel, the latter two being dialects intermediate between Lipay and Tiipay. San Miguel is attested only in a wordlist collected in 1856 (Langdon 1992). Whatever its origin, this trait obviously was well-

---

\(^6\) Available on computer at UCSD and supported by NSF Grant No. BNS 8317837.
<table>
<thead>
<tr>
<th></th>
<th>JPH</th>
<th>Jamul</th>
<th>Cocopa</th>
<th>Yuma</th>
<th>Mojave</th>
<th>Maricopa</th>
<th>Tolkapaya</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 drink</td>
<td>ṭsí</td>
<td>si</td>
<td>ṣi</td>
<td>asĩ(-)</td>
<td>ithii</td>
<td>sii</td>
<td>thii</td>
</tr>
<tr>
<td>12 liver</td>
<td>ṭjy̲pyś́</td>
<td>ṭjpsí</td>
<td>ṭpsuš</td>
<td>—</td>
<td>chavusii</td>
<td>xhlyuusii</td>
<td>chvsii</td>
</tr>
<tr>
<td>17 eagle</td>
<td>ʔaxpá</td>
<td>shpa</td>
<td>—</td>
<td>ʔaspa</td>
<td>ṭaspa</td>
<td>hshpaa</td>
<td>'sa</td>
</tr>
<tr>
<td>18 sleep</td>
<td>ḫxa'má</td>
<td>shma</td>
<td>ṭma</td>
<td>asmá</td>
<td>isma</td>
<td>shmaa</td>
<td>smaa</td>
</tr>
<tr>
<td>24 house</td>
<td>ʔxwá́</td>
<td>wa</td>
<td>wa</td>
<td>ʔavá</td>
<td>'ava</td>
<td>va</td>
<td>'wa</td>
</tr>
<tr>
<td>26 rattlesnake</td>
<td>ʔwí</td>
<td>a'wi</td>
<td>awi</td>
<td>ʔavé</td>
<td>'ave</td>
<td>'ave</td>
<td>'lwi</td>
</tr>
<tr>
<td>35 roadrunner</td>
<td>ʔtípý</td>
<td>tìlypu</td>
<td>clpu</td>
<td>talypo</td>
<td>talypo</td>
<td>talypo</td>
<td>tłu</td>
</tr>
<tr>
<td>36 moon</td>
<td>ʔx̂ý̖ 'a</td>
<td>x̂ly̖'aa</td>
<td>x̂p̂a</td>
<td>x̂l̂p̂a'</td>
<td>haly'aa</td>
<td>hly'aa</td>
<td>hala</td>
</tr>
<tr>
<td>40 beard</td>
<td>halemí</td>
<td>alemi</td>
<td>liyalmiś</td>
<td>yavumé</td>
<td>yavumone</td>
<td>yav'umoone</td>
<td>yav*nymi</td>
</tr>
<tr>
<td>41 mortar</td>
<td>kalg mú́</td>
<td>xmu</td>
<td>xmu</td>
<td>ʔaxmó</td>
<td>'ahmo</td>
<td>hmučhe</td>
<td>—</td>
</tr>
<tr>
<td>43 sun</td>
<td>ʔn̪já</td>
<td>nyaa</td>
<td>n̪'a</td>
<td>ʔan̪'á</td>
<td>'anya</td>
<td>'nyaa</td>
<td>'nyaa</td>
</tr>
<tr>
<td>44 Pleiades</td>
<td>xat̪̞a</td>
<td>xsa</td>
<td>xacá</td>
<td>hachaa</td>
<td>hachaa</td>
<td>hachaa</td>
<td></td>
</tr>
<tr>
<td>45 metate</td>
<td>ʔaxpí</td>
<td>xpi</td>
<td>xpi</td>
<td>ʔaxpé</td>
<td>'ahpe</td>
<td>'hpe</td>
<td>hapi</td>
</tr>
<tr>
<td>46 sheep</td>
<td>ʔčímu</td>
<td>—</td>
<td>mu</td>
<td>ʔamó</td>
<td>'amo</td>
<td>mo</td>
<td>'mu</td>
</tr>
<tr>
<td>53 cloud</td>
<td>akwí</td>
<td>kwii</td>
<td>kwi</td>
<td>ʔak̚'e</td>
<td>'iikwe</td>
<td>kwe</td>
<td>'kwi</td>
</tr>
<tr>
<td>54 stomach</td>
<td>h̕en̪</td>
<td>tu</td>
<td>icu</td>
<td>i-tó</td>
<td>iito</td>
<td>iito</td>
<td>tuu-</td>
</tr>
<tr>
<td>55 nose</td>
<td>haxú</td>
<td>xu</td>
<td>ʔixu</td>
<td>i-xtó</td>
<td>iihu</td>
<td>iihu'u</td>
<td>huu</td>
</tr>
<tr>
<td>57 white man</td>
<td>xaikú</td>
<td>xayku</td>
<td>xayku</td>
<td>—</td>
<td>hayiko</td>
<td>hiko</td>
<td>hayko</td>
</tr>
<tr>
<td>59 mouth</td>
<td>ha'aa</td>
<td>ʔiya</td>
<td>i-yā</td>
<td>iiyā</td>
<td>iiyaa</td>
<td>ya</td>
<td></td>
</tr>
<tr>
<td>65 dance</td>
<td>hý̑má</td>
<td>ʔiima</td>
<td>ʔi-má</td>
<td>iima</td>
<td>iima</td>
<td>iima</td>
<td></td>
</tr>
<tr>
<td>69 badger</td>
<td>maćwa</td>
<td>—</td>
<td>mx̂a</td>
<td>max̂'á</td>
<td>mahwa</td>
<td>mhwaaw</td>
<td>mhwaaw</td>
</tr>
<tr>
<td>70 bird</td>
<td>ʔasá</td>
<td>aasha</td>
<td>ʔa</td>
<td>ʔasé</td>
<td>'ase</td>
<td>shee</td>
<td>'ichs</td>
</tr>
<tr>
<td>73 this</td>
<td>ʔpi̞já</td>
<td>piya</td>
<td>ʔpí̞</td>
<td>védá̞</td>
<td>vida̞-</td>
<td>vda̞-</td>
<td>—</td>
</tr>
<tr>
<td>75 cane,reed</td>
<td>'axta</td>
<td>—</td>
<td>xca</td>
<td>—</td>
<td>'ahta</td>
<td>hta</td>
<td>'hta</td>
</tr>
</tbody>
</table>

NOTE: In a few cases where the length of the stressed vowel is not the same for Kroeber as for Harrington, the Kroeber form is given just below the Harrington form, preceded by the notation K. The Maricopa list has been checked against recent reelicitations by Pamela Munro and corrected accordingly.
installed by the mid 1800's. Examples of final -ily in KH are 7 'salt' and 52 'leg, foot' (see Table 1).

Second, there is an odd correspondence between instances of ingleton in all Yuman languages including all Diegueño varieties except Lipay, and x in Lipay, only in certain words and in odd environments. This was clearly well installed in Mesa Grande Lipay in 1913 as shown in both Kroeber and Harrington's data, and in fact one example of it dates from the 1908 data collected by Harrington. Examples from Table 1 are 1 'one', 17 'eagle', 18 'sleep' (Mojave s is the regular reflex of Proto-Yuman *ingleton).

Third, there is the final stressed vowel lengthening rule of Lipay and Kumeyaay which was introduced after 1913 and had spread considerably by 1925 (Harrington's notes). It was fully installed by 1960.

I propose that the relative chronology of these changes is in the order listed above, even though the distribution of the more recent trait is wider than the other two. Without the evidence of Kroeber and Harrington such a relative chronology could not have been proposed.

References


Harrington, John P. 1913. Unpublished notes at the Museum of Natural History, Santa Barbara, CA.


Munro, Pamela and Molly Star Fasthorse. 1993ms. *Tolkapa-yaa-English dictionary*.


Conjunctions and Reference Tracking in Yuma

Amy Miller

1. Conjunctions

Langdon (1985:491-492) observes that many Yuman languages have function words that appear between clauses and are translated 'and', 'so', or 'then' and thus can be considered "some type of 'conjunction'". She finds "a diversity of sources" for these conjunctions, noting that most of them come from either (i) a nominal source consisting of a sequence of demonstratives followed by a case marker, or (ii) a verbal source, consisting of a verb, often an auxiliary verb, sometimes marked with a temporal prefix, and typically followed by a switch reference marker (1985:493).

This paper examines a set of conjunctions found in Yuma, a Yuman language spoken along the Colorado River. Yuma has a variety of conjunctions which come from verbal sources; most remain analyzable as verbs synchronically. Of particular interest are a set of three conjunctions which Halpern (in prep.) translates 'so': adúm, awím, and aʔím. These conjunctions are composed of the verb stems adú 'be', awí 'do', and aʔí 'say' respectively, plus a suffix =m, and they come from a set of auxiliaries known in the Yuman literature as "behavioral" or "existential". The relation of the 'be', 'do', and 'say' conjunctions to the corresponding auxiliaries is discussed in section 2. In section 3 I demonstrate that the choice among the 'be', 'do', and 'say' conjunction forms is useful in the tracking of action and subject reference. In the present section I describe some of the ways in which the three conjunctions as a group are used in discourse.

First, 'be', 'do', and 'say' can indicate a return to the main thread of the narrative after a digression of some sort, which might include descriptive or background material. Consider (1.a): in the first line, we are told that a messenger comes to a particular place to collect clothes. The second and third lines constitute a digression describing the scene which the messenger encounters at his destination. At the end of the third line we find the conjunction adúm, and in the fourth line the speaker returns to the main thread of the story.

1.a ná:-s-i-k adí:-k
that-far-loc-from come.after-K
he comes from the distance to get [the clothes],

paʔi:pá: ?ac-k”a-madá:v-n”c
people things-rel-grieve-def-sj
and the people who are sorrowing,

u:vá-k u:vá-k vi: dáw-k adú=m
be.locd-SS be.locd-SS be.locd-SS be.here.col-SS be=m
they are here and here and here [in little groups], and so,

paʔi:pá ?aʔent-k av’á:-k vi:df:-k n’a:-vá:-k cakak’á-k
person be.one-SS walk-SS come-SS when.arrive-SS ask.question-K
[this] one person comes walking up, and when he arrives he asks ...' (TK 1)
In (1.b), the first line refers to the gathering of certain feathers needed for ceremonial purposes. Lines 2 through 6 constitute a digression which provides details about the types of feathers needed. The conjunction aʔfː=m in the seventh line marks the end of the digression, after which the speaker returns to the main thread of the story.

\[ n'a:=-stůː=m \quad ꞯa:va:va-km \]
\[ \text{when-gather=m \ when.be.here-KM} \]
\[ \text{when he has collected [the feathers that he needs], then —} \]

?aspá: aʔé-t-m-a
\[ \text{eagle say-assrt-M-end} \]
\[ \text{it's eagle, they say.} \]

?aspá: aʔfː=m
\[ \text{eagle say=m} \]
\[ \text{They call it eagle,} \]

tal pó aʔfː=m
\[ \text{roadrunner say=m} \]
\[ \text{they call it roadrunner,} \]

k'ak xoː: aʔé-t-nti-m-a
\[ \text{woodpecker.sp say-assrt-again-M-end} \]
\[ \text{and they call it woodpecker, too.} \]

aʔfː=m
\[ \text{say=m} \]
\[ \text{So,} \]

n'a:va
\[ \text{this} \]
\[ \text{as for these [feathers],} \]

awfː=m
\[ \text{do=m} \]
\[ \text{he does it,} \]

astůː=m
\[ \text{gather=m} \]
\[ \text{he gathers them,} \]

awfː vu:nóː-k
\[ n'a:vi:r-k \]
\[ \text{do be.around.here.pl-SS when-finish-K} \]
\[ \text{he goes on doing [this], and when he finishes, ...’} \quad \text{(HC 2)} \]

A second function of the 'be', 'do', and 'say' conjunctions is to mark the point at
which the speaker begins to summarize, rephrase, or elaborate on what he has been saying. In (2.a), for instance, the speaker is discussing a cycle of songs and explaining how they fit into the mourning ceremony. The first six lines of (2.a) identify and describe the contents of two songs in the cycle. The conjunction a?Ím appears in line 7, and the material which follows this conjunction is a summary of the material which precedes it.

2.a   vadá-n’
     this.nr-def
     ‘as for these [next two songs],

     amák
     behind
     after [the two i:maf.y songs],

     u:maf    kaná:v-k
     nom.weep describe-SS
     they describe weeping,

     ”amf:=m” a?tf:=m
     weep=m say=m
     ”He cries,” [the songs] say,

     donasá    kaná:v-k
     tears describe-K
     [the songs] describe tears,

     a?tê-t
     say-assrt
     they say,

     a?tf=m
     say=m
     and so,

     ʃa:vár xavîk  n’l:dâw-mti-xa
     song be-two be.there.col-again-irr
     there will be two more songs.’     (IC 20)

In (2.b), the conjunction awf:m occurs in line 2. The material which follows this conjunction (lines 2-4) elaborates on the material which precedes it (line 1).

2.b   i:?ê   n’â:-dâw-k
      scalp when-take-SS
      ‘When he takes the scalp,
**awf:=m**
do:=m
so,

\[
\begin{align*}
\text{a:k'q-t-k} & \quad \text{vu:nó:-k} & \quad \text{vu:nó:-k} & \quad \text{vu:nó:-k} \\
\text{cut-SS} & \quad \text{be.around.here-SS} & \quad \text{be.around.here-SS} & \quad \text{be.around.here-SS} \\
\text{he goes on and on and on cutting it}
\end{align*}
\]

\[
\begin{align*}
n'a:-dáw-k & \\
\text{when-take-SS} & \\
\text{and when he takes it,}
\end{align*}
\]

\[
\begin{align*}
tasót-k & \quad n'a:-dáw-k & \quad \text{awf:=m} \\
\text{pull.off-SS} & \quad \text{when-take-SS} & \quad \text{do:=m} \\
\text{when he pulls it off and takes it, ...} & \quad (HC \ 21)
\end{align*}
\]

The 'be', 'do', and 'say' conjunctions also appear at a change in subject, as exemplified in (3). In (3.a), the subject of the first line is understood to refer to the people who make ceremonial shields. A conjunction appears in the second line. The third line has a new subject: the singer.

3.a \[
\begin{align*}
\text{?ak*él-p-va} & \quad n'i:ca:mán-k & \quad \text{acéw-k} \\
\text{shield-this} & \quad \text{start.there-SS} & \quad \text{make-SS} \\
\text{'as for these shields, they start there and make them,}
\end{align*}
\]

\[
\begin{align*}
\text{awf:=m} & \\
\text{do:=m} & \\
\text{and so,}
\end{align*}
\]

\[
\begin{align*}
k*-a:svá:r-n²-c & \quad a:svá:r-nti-m \\
\text{rel-sing-def-sj} & \quad \text{sing-again-DS} \\
\text{the singer sings again, ...} & \quad (IC \ 12)
\end{align*}
\]

While in (3.a) the new subject is lexically specified, in other cases, such as (3.b), the new subject is understood. The first two lines of (3.b) describe the actions of a group of people known as image handlers. The conjunction awf:m appears in the third line. The subject of the fourth line is a different group of people, the shield handlers, who are present in the house at the same time as the image handlers; this line tells us that they too gather their up gear and take it out of the house. The conjunction awf:m in the third line marks a shift in focus from one set of participants to another within the same scene.

3.b \[
\begin{align*}
\text{aštú:} & \quad \text{vu:nó:-k} & \quad \text{vu:nó:-km} \\
\text{gather} & \quad \text{be.around.here.pl-SS} & \quad \text{be.here.pl-KM} \\
\text{'They go on and on gathering [the images],}
\end{align*}
\]
cACPá:cm-k
take.them.out-SS
and they take them out,

awi=m
do=m
and so,

?qak*éél?  avá-n'  awi=m  awi=m
shield  this.md-def  do=m  do=m
they [the shield handlers] do it [the same thing] with these shields, and so,

awi=m
do=m
and so,

n'í:ca:mán-k
start.there-SS
they begin there,

cACPácm-k  vi:wá:-k
take.out-SS  go.towards.here-K
they take them out and go along,'  (IC 28)

Notice that more than one conjunction can be found in this example. A second instance of awim occurs at the end of the fourth line and a third in the fifth line (these two conjunctions are italicized). Immediately following them, in lines 6 and 7, the image handlers and the shield handlers are referred to jointly. The italicized conjunctions mark the shift from a narrow focus on one or the other sets of participants to a wider perspective in which the two sets are described as a single group.

Finally, the 'be', 'do', and 'say' conjunctions appear at changes in temporal setting, as exemplified in (4). As background to (4.a), people have been making a ceremonial procession around the house. The second and third lines of (4.a) tell us that they stop during this procession. A conjunction awim appears at the end of line 3, and in line 4 we are told that the procession resumes again. In fact, some time passes between the time the procession stops and the time it resumes; several songs are sung during the interval. The conjunction at the end of line 3 is a signal of the temporal shift that takes place between the events described in lines 1-3 and those in lines 4-5. Likewise, in (4.b), an unspecified amount of time passes between the event of the first line and that of the third line, and the conjunction in line 2 marks the point at which the temporal setting changes.

4.a  ?avá  Pu:vé:v-k
house  be.halfway-SS
'They are halfway to the house [and halfway to the pyre],
n'u:v?ó:-km
stand.there.col-KM
and they [stop and] stand there,

n'i:namák  awí=m
quit.there  do=m
they quit there, and so,

n'd:n² naman-t-k
that  start-assrt-SS
they begin [again],

n'a:-nakʷf:n-k
when-go.around.col-K
and when they go around [the house], ...' (IC 22)

4.b  n'a:n¹-a cacpac-k
that-A  bring.out-SS
'they bring out those [shields],

awí=m
do=m
and so,

cakxáv-nti-k
take.in-again-K
they take them in again,' (TK.SHI 4)

Thus far I have shown that the 'be', 'do', and 'say' conjunctions are used (i) at the end of a digression, (ii) at the point at which the speaker begins to summarize or rephrase or elaborate on what he has been saying, (iii) at a change in subject, and (iv) at a change in temporal setting. While these are not the only uses of the conjunctions, they are the most prevalent. They have in common the fact that each marks a type of discontinuity: for instance, the conjunctions exemplified in (3) mark discontinuity of subject, while those in (4) indicate temporal discontinuity in the flow of events. Others mark discontinuity in the flow of the narrative itself, as in (1.a), where the speaker shifts from background to foreground, or in (1.b), where he ends a digression, or in the examples in (2), where the speaker summarizes or elaborates on what he has been saying.

Strikingly similar functions have been found for morphemes in certain other languages, including the Voltaic language Suyiré (Carlson 1987) and the Uto-Aztecan languages Pima and Papago (Scancarelli 1988). In these languages, the morphemes in question sometimes are or have been analyzed as different-subject markers; however, they mark not just change in subject but a more general discontinuity which might also include a change in place or a change in time or a shift from description or elaboration back to action (Carlson 1987:16, Scancarelli 1988:137-142). Curiously, the morphemes which mark general discontinuity in Yuma are not putative different-subject markers. Yuma does have
the remains of a switch reference system (see section 4), but what marks general discontinuity is not this system but the 'be', 'do', and 'say' conjunctions.

2. Auxiliaries and conjunctions

While these three conjunctions as a group signal various types of discontinuity, the choice among the 'be', 'do', and 'say' forms can be manipulated to track action and subject reference. In order to describe this special function, I must first discuss the 'be', 'do', and 'say' auxiliaries that are the source of the conjunctions.

Like most auxiliaries in Yuma, the 'be', 'do', and 'say' auxiliaries occur in a grammaticalized multicausal construction. This construction consists of a clause containing a main verb followed by a clause containing an auxiliary. Both main verb and auxiliary inflect for person of subject (with third person zero-marked). They always have the same subject. Typically the main verb takes the suffix -k (often analyzed as a same-subject switch reference marker). Main verb and auxiliary occur together under a single intonation contour.

Exactly what information the 'be', 'do', and 'say' auxiliaries contribute to the construction is not entirely clear. What is known is this: First, 'say' has several uses not shared by 'be' and 'do'; for instance, as quotative and as a modal. Second, in a comparative study of Yuman auxiliaries, Norwood (1981:140-141) found that in Yuma the auxiliary adúm 'be' systematically followed stative verbs, verbs of motion, and active verbs affecting only the subject. She found that awfim 'do' followed active transitive verbs, and that a?fim 'say' followed verbs of communication and verbs of internal state. Her analysis accurately reflects the use of auxiliaries in texts collected in the 1930's. Numerous examples supporting Norwood's analysis can be found in a text collected in 1935 and published as Halpern (1976); a sample is given in (5).  

5.a asáy-t-k adú-t'a
be.fat-assrt-SS be-truly
'indeed he was fat.' (Halpern 1976:19)

grown.ones-def be.like.that-again-SS be-truly when-be.big.col-SS
'They resembled the grown-up ones when they got big.' (ibid p.13)

5.c ná:i:pák-t-k adú-t'a
arrive.there-assrt-SS be-truly
'indeed he arrived there.' (ibid p.5)

5.d ?amáy-k maç-ca?ór-t-ntí-k adú-t'a
top-on refl-go.around.on-assrt-again-SS be-truly
'indeed they squirmed on top of each other again.' (ibid p.13)
5.e  ?a?f:  adáw-t-k  awf-t’a
     wood  pick.up-assrt-SS  do-truly
     'Indeed they picked up wood.' (ibid p.8)

5.f  n’á:vi  cásmá:-t-k  a?l-t’a
     here  insult-assrt-SS  say-truly
     'Here indeed he insulted him.' (ibid p.18)

Texts which Halpern collected in 1978, however, reveal changes in the way auxiliaries are used. First, the choice between ‘be’, ‘do’, and ‘say’ is based on different criteria in the 1970’s than it was in the 1930’s. The criteria relevant in the 1970’s are not yet fully understood, and the matter merits careful investigation; however, one tendency stands out: active verbs, whether transitive or intransitive, generally take the auxiliary ‘do’. In particular, both verbs of motion and reflexive verbs (which in the 1930’s took ‘be’) now take ‘do’. This is exemplified in (6.a,b); compare (5.c,d). ‘Be’ still tends to be used with stative verbs (transitive as well as intransitive), as it was in the 1930’s, and ‘say’ to be used with verbs of vocal activity and internal state. However, in the 1970’s these patterns are just tendencies, whereas in the 1930’s they were the rule.

6.a  vi:yá:-n’pat-k  awf-t’a
     go-in.turn-SS  do-truly
     'they go too.' (HC 22)

6.b  mat-tavér-k  awf-t’a
     refl-chase-SS  do-truly
     'they chase one another.' (HC 21)

An even more striking change in the use of these auxiliaries concerns frequency of use. A text collected in 1935 (Halpern 1976) averages four ‘be’, ‘do’, or ‘say’ auxiliaries per page, while a text collected in the 1970’s averages one such auxiliary per 2-1/4 pages; thus auxiliary use appears to have decreased by about 90% in forty years. Coinciding with this decline is the innovative use of ‘be’, ‘do’, and ‘say’ as discourse conjunctions.

The ‘be’, ‘do’, and ‘say’ conjunctions are similar in many respects to the corresponding auxiliaries. First, both are based on the same three stems: adú ‘be’, awf ‘do’, and a?l ‘say’. Second, the conjunctions, like the auxiliaries, inflect for person of subject. First and second person subject is overtly marked, while third person subject is zero-marked. The texts in my database are for the most part third person narratives, so most conjunctions in them occur in third person contexts and take the zero third person subject prefix. (All the examples in 1-4 were of this type.) First- or second-person passages are sufficiently rare that I had to venture outside my database to find an example containing a conjunction without unwanted distractions. That in (7) comes from a text collected in 1976.

7.  n’á:va  ?-a:svá:r-k  ?-uxág-m-t-k  ?-a?f=m  ?-a:svá:r  ap?a?ém-xa
     this  1-sing-SS  1-know.how.neg-assrt-SS  1-say=m  1-sing  1.not.say-irr
     ‘I don’t know how to sing that [song], so I won’t sing it.’ (Emerson and Halpern 1978 ¶ 26)
The 'be', 'do', and 'say' conjunctions also resemble the corresponding auxiliaries in their syntactic behavior. The conjunction has the same subject as the lexical verb which precedes it (a fact that is most clearly demonstrated in examples where person of subject is overtly marked, such as 7), and the lexical verb is marked with the suffix -k, as may be seen in many of the examples above, for instance, in (3.a).

Especially important to the remainder of this paper is the fact that the conjunctions, like the corresponding auxiliaries, typically reflect the active, stative, or vocal (or internal-state) nature of the lexical verb with which they are associated. This may be seen in (1.a), where the 'be' conjunction follows a chain of stative verbs; in (3.b), where the 'do' conjunction follows an active verb; and in (2.a), where the 'say' conjunction follows a verb of vocal activity.

These morphological, syntactic, and semantic facts argue that the 'be', 'do', and 'say' conjunctions have their source in the corresponding auxiliaries. It should be pointed out that conjunctions are nonetheless easily distinguished from auxiliaries on the basis of function and in that they exhibit subtle syntactic, morphosyntactic, and intonational differences. Because conjunctions link one bit of discourse with another, they occur either in medial position (cf. 1.a), or at the beginning of a new syntactic "sentence", where they form a link with the preceding one (e.g. 1.b). Auxiliaries, on the other hand, never begin syntactic "sentences". They are sometimes found in medial position, as in the fifth line of (2.b), but more often they appear at the end of a "sentence", as in the examples in (5). Consequently, an auxiliary can take either medial suffixes (usually =m) or suffixes which identify the ends "sentences" (such as -ta), while conjunctions take only those suffixes that are appropriate in medial position. A final difference between auxiliaries and conjunctions is that auxiliaries are always intonationally bound to their main verb, while conjunctions can — and often do — occupy an intonation unit distinct from that of the associated lexical verb; cf. (2, 3, 4.b).

For the sake of completeness it must be mentioned that 'be', 'do', and 'say' are used not only as auxiliaries and conjunctions but also as lexical verbs. An example of 'say' used as a lexical verb may be seen in the second line of (1.a), and examples of lexical verbs 'be' and 'do' are found in the first line of (1.b). Lexical uses of 'be', 'do', or 'say' are easily identified on the basis of their meaning.

3. Conjunctions and reference tracking

While the choice between the 'be', 'do', and 'say' conjunction forms generally reflects the active, stative, or vocal nature of the preceding verb, it is nonetheless quite common to find instances of 'be', 'do', and 'say' conjunctions which deviate from this norm. Examples are given in (8). In (8.a), the conjunction 'do' appears in the third line. The material in the third line, however, is descriptive in nature, leading one to expect 'be' rather than 'do'.

```
8.a  cá:m n'am-caPví:-k acéw-k
     all thereby-make.even-SS make-SS
     'they make it all even [in length],

     arúv-m
     be.dry-DS
     and it's dry,
```
awf=\text{m}
do=\text{m}
and so,

r:\text{a}:n^r
that
that [stuff],

a:menamén-k awf=\text{m}
make.a.roll-SS do=\text{m}
they make it into a roll,' (TK.SH 9)

Let's consider (8.a) more closely. As background, one should know that certain people are in the process of making ceremonial shields. They have acquired a special kind of grass and are going to wrap bundles of it around the perimeter of each shield frame. These people are the subject of line 1, which tells us that they make the bundles of grass all even in length. Line 2 gives a brief description of the grass. Line 3 contains a conjunction which marks the end of the descriptive passage, and the action resumes in lines 4 and 5. When the meaning of each line is considered, it becomes clear that the conjunction is semantically associated not with the descriptive verb in line 2 but rather with a verb further back in the discourse — specifically, with the active verb acéw in the first line. By virtue of this semantic association, the ‘do’ conjunction creates a link between the active material in line 1 and that in lines 4-5, thus bridging the descriptive digression in line 2. In other words, while the mere presence of a conjunction marks the end of the descriptive digression, the use of the ‘do’ form (rather than ‘be’ or ‘say’) specifies the active nature of the material interrupted by the digression and explicitly connects the pieces.

Another example can be seen in (8.b). In the first line we are told that people put down poles. The second line provides descriptive information about the poles, followed by the conjunction awfim ‘do’, and the thread of the story resumes in the third line. The verb with which the conjunction is semantically associated is not one of the descriptive verbs in line 2; rather, it is the active verb camf:im in the first line. The choice of the ‘do’ form of the conjunction explicitly links the action in line 1 with the action in line 3.

8.b \text{ aqf: camf:=m}  
wood put.down.long.obj=\text{m}  
‘and they put down poles,

\text{ aqf: lóq vi:dfs-k-awf=m}  
be.long be.real.long lie.here-DS do=\text{m}  
there are long ones, real long ones, lying here, and so,

\text{ aqf: tata:t-k va:wf:=m}  
wood set.up.many-SS do.thus=m  
they set up poles like this,' (IC 6)

‘Say’ is used in a similar capacity to indicate continuity of vocal activity across a
digression. In (8.c), vocal activity is described in lines 1-2. Lines 3 and 4 provide background information about other people present on the scene. The conjunction ‘say’ appears in line 5, where it marks the end of this digression, and vocal activity resumes in line 6. Notice the semantic mismatch between the ‘say’ conjunction and the stative verbs in lines 3 and 4. The verbs with which the ‘say’ conjunction are semantically associated are ‘sing’ in line 1 and ‘go on [singing] from here’ in line 2. The choice of the ‘say’ form of the conjunction specifies the vocal nature of the events interrupted by the digression, thus explicitly connecting lines 1-2 with line 6.

8.c  a:svá:r-k
     sing-SS
     he sings,

     ná:vi:yém-k
     go.from.here-SS
     he goes on [with his song] from here,

     ?u:tf:s  k*:awíc-n*:c  u:v?ó:-k
     bow  rel-nom.use-def-sj  stand.col-SS
     and the bow handlers stand,

     k*:i:má:č  u:v?ó:-k
     rel-nom.dance  stand.col-SS
     and the dancers stand,

     a?lé:=-m
     say=m
     and so,

     a:stu:vár-k  va:?é  van*a:wa:-k  van*a:wa:-k
     sing.dp-SS  do.thus  when.go.towards.here-SS  when.go.towards.here-K
     they sing like this, and as they go along and go along, ... ’ (TK.CIR 7)

Marking continuity of physical or vocal action appears to be the primary function of conjunction choice in (8.a-c). Notice, however, that in each case continuity of action implies continuity of subject: in (8.a), the understood subject of the verbs in the first line is the same as that of the active verbs in line 5; in (8.b) the understood subject of the active verb in the first line resumes as that of the active verbs in line 3; and in (8.c), the subject of the verbs in lines 1-2 is included among the referents of the plural subject of the verbs in line 6. This is not surprising, given that subject continuity is normal when action is continuous. However, other data in which continuity of action is crucially lacking show that subject continuity alone is sufficient to motivate conjunction choice. Examples are given in (8.d,e).

The first line of (8.d) and the first clause in the second line describe the actions of a group of people: they bundle clothes into a shawl and put them down. The second clause in line 2 provides information not about these people’s activities but about the bundles: here they are, as a result of having been put down. The third clause in line 2 is the ‘do’
conjunction. ‘Do’ semantically links the active verbs in line 1 and in the first clause in line 2 with the material in line 3. Notice that the use of ‘do’ here does not indicate a return to action; what follows ‘do’ is not physical action at all but a passage in which the people who were the subject of the action in the first line and in the first clause of the second line talk. Instead, the choice of ‘do’ serves to indicate that the talking in line 3 is somehow connected with the actions in line 1: specifically, that the people who do the talking are the same as those who performed the actions in line 1 and in the first clause of line 2. Thus it indicates subject continuity across a descriptive digression.

8.d maːʃkɑː-r əɡɔː-ːk vuːnɔː-ːk
    shawl-in wrap-SS be.around.here.pl-SS
    ‘they go about wrapping [the clothes] in a shawl,

    acɛː viːdaw-m    awɛː=m
    put.them.away be.here.col-DS do=m
    they put them aside and here they [the clothes] are, and so,

    vanˈaːːdaw-m
    when.be.here.col-M
    then,

    "vadɛː-k ːPɛvː-ːk-ːdɔː-ːa"
    this.nr-sj be.ready-SS-be-evid
    "This is ready for them," [they say];’ 
    (IC 5)

In (8.e), the conjunction used is aduːm ‘be’. In the first two lines of (8.e), stative verbs are used to describe people who carry images in the mourning ceremony. In line 3 there begins a long quotation addressed to these people, followed in line 8 by the lexical verb ‘say’. The quotation and the lexical verb ‘say’ together constitute a digression which I will refer to as the speech report. The lexical verb ‘say’ is in turn followed by the conjunction ‘be’. The presence of a conjunction here signals the end of the speech report, and the choice of the ‘be’ conjunction form recalls the stative clauses in line 2, creating a link between these clauses and with the material which follows the conjunction. Notice that ‘be’ does not indicate a return to stative description: the material which follows it in line 9 is active. Instead, ‘be’ tells us that line 9 picks up where line 2 left off — and has the same subject as line 2. In this example, then, conjunction choice is used to indicate continuity of subject across a speech report.

8.e paʔiːpːaː ːuːcɛw kʰ-ːaːvɛkːɛw vadɛː-c
    person nom.make rel-carry this.nr-sj
    ‘these who carry the "constructed people" [i.e. the image carriers]

    nuːnɔː-ːk    vuːnɔː-ːntiː-m
    be.around.there.pl-SS be.around.here.pl-also-DS
    they are here and there too;
"pa?i:pá: ma:-k*-ašf:nt-c
  person  2-rel-be.one.pl-sj
"You people,

n'a:yú:
thing
well,

mask*e ma:-k*-a:vkéw
image  2-rel-carry
you who carry the images,

k-antamák  ak'ka?ém-k  vi:kadáw-k
imp-leave.col imp.not.do-K  imp.be.here.col-K
don't leave, stay here,

k-a:cvf:r-k*
imp-finish.dp-K
and finish,"

a?ti: =m  adúi: =m
say =m  be =m
he says, and so,

katán-k  vi:dáw-k  vi:dáw-k
arrive.col-SS  be.here.col-SS  be.here.col-K
they [the image carriers] come and stay here and stay here,'  (IC 13)

In summary, the examples in (8) demonstrate that the choice among the 'be', 'do', and 'say' forms of the conjunction may be manipulated to indicate continuity of action and subject reference. While reference tracking may have originated as a derivative of the action-tracking function of conjunction choice, it has become an important function in its own right.

For the sake of completeness it should be pointed out that in some cases the choice among 'be', 'do', or 'say' appears to be motivated by neither the immediately preceding clause nor material further back in the discourse; in fact, I have yet to find any motivation for it at all. Some such data tempt me to conclude that 'be' is being grammaticalized as the default conjunction form, while other such data convince me that 'do' or 'say' is being grammaticalized as the default conjunction form. I hope that further work will lead to an explanation of these apparent exceptions.

4. A note about switch reference marking

The Yuma switch reference system is often unreliable in reference tracking, and Miller (1992) argues that the suffixes -k and -m, often analyzed as 'same-subject' and 'different-subject' respectively, are in the process of being reanalyzed as verb class markers. This reanalysis is ongoing, however, and some occurrences of -k and -m still appear to be motivated by reference tracking considerations. For this reason it is not clear to what extent
the switch reference system plays a part in reference tracking in the examples in (8). Let us for the sake of argument follow the conservative course of analyzing -k and -m as switch reference markers whenever possible. (Towards this end I have glossed -k and -m as 'SS [potentially analyzable as same-subject marker]' and 'DS [potentially analyzable as different-subject marker]' whenever consistent with the interpretation of the examples.) Under such an analysis, switch reference marking alone proves sufficient to track reference in (8a) and perhaps in (8c). In (8b) and (8d), however, the lexical verb with which the conjunction is semantically associated either obligatorily takes the verb class marking suffix =m (cf. 8b) or is unsuffixed (cf. 8d), so that switch reference marking fails to track reference across the descriptive digressions found in these examples. In (8e), furthermore, switch reference marking provides no help at all in tracking the reference of the image carriers across the speech-reporting digression.14

Thus, even when -k and -m are analyzed as switch reference markers wherever possible, they fail to track subject reference in a majority of the examples in (8). It is perhaps because of their unreliability that conjunction choice has emerged as an important device for tracking subject reference.

5. Conclusion

In this paper I have attempted to make two points. First, the presence of a 'be', 'do', or 'say' conjunction signals a discontinuity in the focus of attention, in the flow of events, or in the flow of the narrative itself. Second, the choice among the 'be', 'do', and 'say' forms may be manipulated to track action and subject reference -- and in some cases subject reference alone -- across digressions and other discontinuities. There is irony in this situation: Since the presence of a conjunction marks discontinuity, conjunctions are functionally parallel to different-subject marking in certain other languages. However, since the choice of conjunction is useful in tracking subject reference across intervening material, conjunction choice is functionally similar to same-subject marking in certain other languages.

It is also ironic that in a language which is often considered to have a switch reference system, functions such as reference tracking and general marking of discontinuity should be accomplished not by the switch reference markers but by conjunctions. This, I believe, has to do with the fact that the switch reference system is breaking down and being reanalyzed as a verb class marking system. It is not surprising that, coincident with this breakdown, the 'be', 'do', and 'say' conjunctions should take over functions typically associated with switch reference systems.

Notes

1. Unless otherwise indicated, data come from texts collected by Abraham M. Halpern in 1978 and presently being prepared for publication as Halpern (in prep). The texts were transcribed and originally translated by Halpern. I am responsible for the analysis; I have also divided the text into lines which coincide with intonation units and I have slightly revised the translation.
2. Several conjunctions meaning ‘then’ are found, including \textit{vān'a:vāk}, composed of the auxiliary verb stem \textit{vī:vā} ‘be here’ plus the temporal prefix \textit{nā}- and a suffix -\textit{k}, and its plural counterpart \textit{vān'a:dāw}/\textit{k}, composed of auxiliary verb stem \textit{vī:dāw} ‘be here (col)’ with the same affixes. There is also a conjunction \textit{a?ː}s (or \textit{a?ːs:s}) ‘but’ composed of the auxiliary verb ‘say’ plus dubitative suffix -\textit{s} (and, sometimes, the emphatic suffix -\textit{ā}.

3. ‘Be’, ‘do’, and ‘say’ belong to a small class of verbs which always take the suffix =\textit{m}, unless an intervening suffix is present; for discussion see Halpern (1947:157-158) and Miller (1992:71-73). When suffixed with =\textit{m}, members of this class are subject to alternations in the length of the stressed vowel (these alternations are not yet understood); see Halpern (1947b:157 note 9). ‘Be’, ‘do’, and ‘say’ are also also subject to alternations in the quality of the stressed vowel before certain suffixes; see Halpern (1947a:21, 1947b:157-159).

4. This is similar to a use of English ‘so’ described by Schiffirin (1987:191-201).

5. A code following each example identifies the speaker and section of text from which it is drawn. The following abbreviations are used in examples: \textit{assert} ‘assertive’; \textit{col} ‘collective plural’; \textit{dp} ‘distributive plural’; \textit{def} ‘definite’; \textit{DS} ‘potentially analyzable as different-subject marker’; \textit{evid} ‘evidential’; \textit{imp} ‘imperative’; \textit{irr} ‘irrealis’; \textit{md} ‘middle distance’; \textit{nom} ‘nominalized form’; \textit{nr} ‘near’; \textit{pl} ‘plural’; \textit{rel} ‘subject relative clause’; \textit{sj} ‘subject’; \textit{sp} ‘species’; \textit{SS} ‘potentially analyzable as same-subject marker’. =\textit{m} glosses the verb class marker =\textit{m}, and -\textit{k} and -\textit{m} gloss suffixes which cannot be analyzed as switch reference markers and may be verb class markers. The functions of -\textit{Δ} and -\textit{KM} are not yet clear.

The character \textit{d} represents the phoneme \textit{ð}, \textit{\v{t}} represents \textit{t}, and \textit{s} represents the post-alveolar fricative \textit{s}, which Halpern transcribed as \textit{\v{s}}.

6. See also Dahlstrom (1982), who argues that in Lakhota a set of conjunctions previously analyzed as switch reference markers does not mark switch reference at all but rather continuity vs. discontinuity of action.

7. Auxiliary constructions of this type are widespread in the Yuman family were first described by Langdon (1978). The Yuma auxiliary construction is actually somewhat more complicated than I have described it here. In fact it may consist of a chain of clauses all having the same subject, and it is possible for one or more clauses in this chain to separate the clause containing the main verb from that containing the ‘be’, ‘do’, or ‘say’ auxiliary.

8. In the examples in (6), I have replaced Halpern’s (1976) literal glosses with morpheme-by-morpheme glosses. The translations are Halpern’s.

9. This generalization applies more to the speech of two of Halpern’s three consultants (IC and HC) than to the speech of the third (TK).

10. In one narrative randomly chosen from Halpern (in prep.), I found an average of one auxiliary per 5-1/2 pages. However, pages formatted for Halpern (in prep.) hold about half as much text as the pages of Halpern (1976). In an attempt to achieve comparability, I counted each page of Halpern (in prep.) as half a page.
11. It should be noted that the text collected in 1978 was tape recorded, while the 1935 text was probably dictated. It is likely that dictation required the text to be broken up into more and smaller "sentences" than are typical in ordinary narration. The larger number of "sentences" would result in an increase in the number of unit-final auxiliaries. I thank Geoff Kimball for pointing this out.

12. It is not clear that the traditional notion of sentence corresponds to a genuine structural category in Yuma discourse. I use the term in scare quotes to refer to an often long syntactic unit that is perhaps more like a paragraph than a traditional sentence.

13. See note 3. In addition to =m, conjunctions sometimes take the assertive suffix -t followed by -k or -m.

14. If treated as a switch reference marker, the -m in line 2 of (8.10) must be analyzed as marking different subject with respect to the lexical verb ‘say’ in line 8; thus it tracks reference into, rather than across, the speech report.

References


The Shifting Status of Initial Glottal Stop in Barbareño Chumash

Marianne Mithun, Tsuyoshi Ono, and Suzanne Wash

University of California, Santa Barbara

The phonetic difference between words beginning with a vowel (#V-) and those beginning with a glottal stop plus vowel (#ʔV-) is a fine one, one that we might not expect to be exploited in many languages. Indeed, many languages simply contain no vowel-initial words at all. Initial glottal stops, like other consonants, are an integral part of the morphemes in which they occur, and remain in all contexts. Such glottals can be seen in the Central Pomo verb root ṭa- ‘gather’.

(1) Central Pomo (Frances Jack p.c.)

\[
\begin{align*}
\text{ʔaw} & \quad [\text{ʔáw}] \quad \text{‘gather (something)’} \\
\text{ša-ʔdw} & \quad [\text{šaʔáw}] \quad \text{‘gather up (in sweeping motion)’}
\end{align*}
\]

In many other languages, vowel-initial words are automatically pronounced with a laryngeal onset, like the glottal stop in the Mohawk below. The fact that the initial glottal is not a basic part of the word is shown by the fact that it fails to appear utterance-internally.

(2) Mohawk (Josie Horne p.c.)

\[
\begin{align*}
\text{ʔhúʔí-neweʔ} & \quad [\text{ʔhúʔí-neweʔ}] \quad \text{‘they will arrive’} \\
y-ʔhúʔí-neweʔ & \quad [\text{yʔhúʔí-neweʔ}] \quad \text{‘they will arrive there’}
\end{align*}
\]

In the Chumash languages, glottalization plays a number of roles. One intriguing question is its status in word-initial contexts. In his recording of Barbareño Chumash over the first half of this century, John Peabody Harrington systematically distinguished vowel-initial words from those with a glottal onset.

(3) Barbareño Chumash: John Peabody Harrington from Mary Yee

\[
\begin{align*}
\text{awis} & \quad \text{‘to fix’} & \text{ʔawax} & \quad \text{‘jug basket’} \\
\text{aciš} & \quad \text{‘beard’} & \text{ʔap’an} & \quad \text{‘to build’} \\
\text{eqwel} & \quad \text{‘to make’} & \text{ʔel} & \quad \text{‘necklace’} \\
\text{isawus} & \quad \text{‘sweat’} & \text{ʔip} & \quad \text{‘to say’} \\
\text{uʔo} & \quad \text{‘to leave’} & \text{ʔuwas} & \quad \text{‘grapes’}
\end{align*}
\]
Since Harrington was a careful phonetician, and extremely knowledgable about Chumash, we cannot attribute the variation to inaccurate transcription nor to inconsistencies in level of representation. In fact, he often draws attention in his notes to the specific presence or absence of glottal stop. Individual roots are transcribed consistently. (Glottal stop does not appear word-initially before consonants in the language.)

It is interesting that a phonetic distinction as subtle as initial glottalization before vowels could carry such a functional load. In what follows, we will show that its distribution is not random. In fact, the pattern suggests possible explanations of how it came to be established in the language. All material cited here is drawn from Harrington's fieldnotes taken during work with Mary Yee.

1. The Verb-Noun Asymmetry

   The majority of vowel-initial words in the Chumash languages, including Barbareño, are verbs, while the majority of glottal-initial words are nouns.

(4) Verb roots versus noun roots

<table>
<thead>
<tr>
<th>verb</th>
<th>meaning</th>
<th>noun</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>awiš</td>
<td>'to fix'</td>
<td>?awax</td>
<td>'jug basket'</td>
</tr>
<tr>
<td>eqwel</td>
<td>'to make'</td>
<td>?eneq</td>
<td>'woman'</td>
</tr>
<tr>
<td>ixip</td>
<td>'to finish'</td>
<td>?ixpaniš</td>
<td>'acorn'</td>
</tr>
<tr>
<td>oqmol</td>
<td>'to spit'</td>
<td>?oʔ</td>
<td>'water'</td>
</tr>
<tr>
<td>ušpak</td>
<td>'to pick up'</td>
<td>?uwaš</td>
<td>'pipe'</td>
</tr>
</tbody>
</table>

Reasonable explanations for this modern state of affairs can be inferred from a reconstruction of the development of the morphology.

2. Verbs

   In the modern Chumash languages, verb roots rarely occur in word-initial position. They are normally preceded by a pronominal prefix specifying their subject, or by a prefix that nominalizes or subordinates the verb.

(5) Subject pronominal prefixes: root eqwel 'make'

<table>
<thead>
<tr>
<th>prefix</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-eqwel</td>
<td>'I made'</td>
</tr>
<tr>
<td>p-eqwel</td>
<td>'you made'</td>
</tr>
<tr>
<td>s-eqwel</td>
<td>'he or she made'</td>
</tr>
<tr>
<td>?al-eqwel</td>
<td>'(that) made'</td>
</tr>
</tbody>
</table>
Verb stems thus always appear with a prefix, except in imperatives. Most noun roots, by contrast, usually appear unprefixed. The correspondence between lack of initial glottal stops in verb roots and the usual word-internal position of the roots suggests two possible scenarios.

One possibility is that vowel-initial words were automatically given a default glottal onset. Since verb roots seldom appear word-initially, due to the usual presence of a pronominal prefix or nominalizer, they would rarely occur with the default glottal. Since noun roots usually do appear word-initially, they would generally be pronounced with the default glottal. At a certain point, the default glottal stops could have been reanalyzed as an integral part of the basic form of the noun roots with which they usually occurred.

A second possibility is more complex but also more compelling. At an earlier time, before the development of pronominal prefixes in Chumash, all words may have been characterized by a consonantal onset, as in Central Pomo. Among the consonants was glottal stop. The glottal may itself have originated as a default onset, but acquired the status of a basic consonantal element of the lexical items by this point.

(6) Hypothesized point of departure

*ʔawiš ‘to fix’  
*ʔeqwel ‘to make’  
*ʔixip ‘to finish’  
*ʔogmol ‘to spit’  
*ʔušpak ‘to pick up’  
ʔawax ‘jug basket’  
ʔeneq ‘woman’  
ʔixpaniš ‘acorn’  
ʔoʔ ‘water’  
ʔuwaš ‘pipe’

The grammaticization of the pronominal prefixes k-, p-, and s- would have created clusters kʔ-, pʔ-, and sʔ- at the beginning of verbs like those in (6). We know that frequently occurring forms are typically subject to phonological attrition. The grammaticization of the obligatory pronominal prefixes with verbs would have set the stage for the simplification of the marked obstruent-glottal clusters to plain obstruents. Once the simplified forms had become well established, the verb roots were reanalyzed as vowel-initial, yielding the modern forms.

(7) Cluster simplification and reanalysis

*kʔeqwel > *kʔeqwel > keqwel > k-eqwel  ‘I made’  
*pʔeqwel > *pʔeqwel > peqwel > p-eqwel  ‘you made’  
*sʔeqwel > *sʔeqwel > seqwel > s-eqwel  ‘s/he made’

Reanalysis need not have occurred simultaneously in all forms. Sally Thomason (p.c. 1993) points out that the initial sʔ cluster might have been less stable than the initial kʔ and pʔ, which could have been pronounced as ejectives k̂ and p̂. The cluster simplification and subsequent reanalysis could thus have originated in the third person singular forms, then
been generalized to other persons and numbers. In any case, a parallel process would not have occurred with nouns like those in (6), because they usually appeared unprefixed.

There is, however, a small set of verb roots that do contain initial glottal stops. The glottal stops appear not only in the word-initial imperative forms, but word-externally, following the pronominal prefixes.

(8) Glottal-initial verb roots I: phonology

\[

g\text{ip} \quad \text{‘say, call, think, guess’} \\
k\text{-gip} \quad \text{‘I said’} \\
p\text{-gip} \quad \text{‘you said’} \\
s\text{-gip} \quad \text{‘s/he said’} \\

\text{g\text{es}} \quad \text{‘to weave’} \\
\text{g\text{uq\text{š}}} \quad \text{‘to stink, be rotten’} \\
\text{g\text{u\text{w}}} \quad \text{‘to eat, suck, bite’}
\]

These exceptions share a notable phonological characteristic: they all consist of a single syllable. This exception is difficult to explain under the first scenario described above, whereby glottal onsets were assigned to full words complete with pronominal prefixes. It is more easily understood under the second scenario, involving cluster simplification and subsequent reanalysis; the initial erosion would have occurred only in words with a substantial phonetic substance: two or more syllables.

A second set of glottal-initial verbs share a different characteristic.

(9) Glottal-initial verb roots II: morphology

\[

g\text{ap\text{‘an}} \quad \text{‘to build’} \\
g\text{a\text{iš\text{š}}} \quad \text{‘to sit on’} \\
g\text{aw\text{ini}} \quad \text{‘to be half’} \\
g\text{a\text{ič}} \quad \text{‘to wage war’} \\
g\text{ele\text{č}} \quad \text{‘to wear a necklace’} \\
g\text{i\text{w\text{in}}} \quad \text{‘to cut’} \\
g\text{o\text{č\text{č}}} \quad \text{‘to be wet’} \\
g\text{unu\text{‘w\text{ine\text{č}}} \quad \text{‘to get married’}
\]

These verbs share a morphological property; they are derived from nouns.
Nominal bases of derivation

\( ?\text{ap} \)  ‘house’
\( ?\text{as} \)  ‘mat, seat, chair’
\( ?\text{awi}n\text{ı} \)  ‘side’
\( ?\text{axi}ç \)  ‘war’
\( ?\text{el} \)  ‘necklace’
\( ?\text{iwi}ç \)  ‘knife’
\( ?\text{o}ç \)  ‘water’
\( ?\text{uni}wi \)  ‘spouse’

The existence of the glottal-initial derived verbs indicates that the erosion of root-initial glottals from verbs was a historical process, one no longer active in the language.

Finally, a third set of glottal-initial verb roots shares a quite different characteristic.

Glottal-initial verb roots III: origin

\( ?\text{akawayu} \)  ‘to be on horseback’
\( ?\text{alasâ}l \)  ‘to pray’
\( ?\text{alè}l \)  ‘to read’
\( ?\text{untal} \)  ‘to grease’

These verbs are Spanish loans, which would have been first heard without pronominal prefixes. Both verbs (\text{untar} > ?\text{untal} ‘to grease’) and nouns (\text{aguja} > ?\text{awuxa} ‘needle’) were systematically borrowed from Spanish with initial glottals, sometime after the founding of the Santa Barbara mission in 1786.

Spanish sources of glottal-initial verb roots

\begin{itemize}
  \item \textit{a caballo}
  \item \textit{a rezar}
  \item \textit{a leer}
  \item \textit{untar}
\end{itemize}

3. Nouns

As noted earlier, most glottal-initial words in Barbareño are nouns. A small set of nouns appear with initial vowels, however.
(13) Vowel-initial noun roots: semantics

\[
\begin{align*}
\text{acīs} & \quad \text{‘beard’} \\
\text{ahaš} & \quad \text{‘soul’} \\
\text{antik} & \quad \text{‘spirit’} \\
\text{ašhunač} & \quad \text{‘ruler, boss’} \\
\text{ičč} & \quad \text{‘younger sibling’} \\
\text{isawus} & \quad \text{‘sweat’}
\end{align*}
\]

As can be seen, these nouns share a semantic characteristic: they refer to entities that would normally be possessed, or inalienable. They take on full meaning only in relation to another entity. Such nouns are normally preceded by a pronominal prefix (of the same form as the subjective prefixes on verbs) indicating the possessor. Such inflected nouns would have been subject to the same process of cluster simplification as the verb roots with subject prefixes, resulting in the loss of initial glottal stop: *k-\text{acīs} > kacīs ‘my beard’.

Interestingly, Harrington provides glottal-initial counterparts to some of these vowel-initial inalienable nouns. These are just the kind of objects that occur as either alienable or inalienable possessions.

(14) Doublets

\[
\begin{align*}
\text{acīspa čis} & \quad \text{‘beard, silk (of corn plant)’} \\
\text{ahaš\textipa{a}haš} & \quad \text{‘spirit, soul (of a person)’} \\
\text{ašhunač\textipa{a}šhunač} & \quad \text{‘rule, ruler, boss’}
\end{align*}
\]

Harrington does not always specify differences in meaning between the forms, but their use indicates that they differ in alienability. The meaning of a derived form of ‘spirit’ was described as in (15).

(15) \text{ka\textipa{a}š\textipa{a}š}\textipa{a}š

\[
\begin{align*}
\text{k-\textipa{a}haš\textipa{a}š}\textipa{a}š & \\
1\text{-spirit}
\end{align*}
\]

‘my ghost that I see out of the window walking around but not my ghost, just as one says our moon, though we do not really own it’

Like monosyllabic verbs, monosyllabic noun roots retain their original glottals.
(16) Monosyllabic inalienable nouns

\(\text{nık} \quad \text{‘mouth’} \\
\text{nīl} \quad \text{‘leg, foot, paw’}

Just as noun roots retain their basic glottal-initial forms under verbalization, we might expect vowel-initial verb roots to retain their basic forms under nominalization. Interestingly, this does not appear to be the case.

(17) Nominalization

\begin{align*}
\text{apīt} & \quad \text{‘to climb’} & \text{apīt}^{i?} & \quad \text{‘ladder’} \\
\text{astipil} & \quad \text{‘to be thick’} & \text{astīpilaš} & \quad \text{‘thickness’} \\
\text{axīye} & \quad \text{‘to cure’} & \text{axīye} & \quad \text{‘medicine, a cure’} \\
\text{axtiawayan} & \quad \text{‘to feel refreshed’} & \text{axtiawayanpi} & \quad \text{‘shade’} \\
\text{axtkay} & \quad \text{‘to contain’} & \text{axtkay}^{i?} & \quad \text{‘a container’} \\
\text{eqwel} & \quad \text{‘to make’} & \text{eqwèleš} & \quad \text{‘shape, appearance’} \\
\text{išō} & \quad \text{‘to roast’} & \text{išoš} & \quad \text{‘roast, barbecued meat’} \\
\text{iqip} & \quad \text{‘to close, lock’} & \text{iqip} & \quad \text{‘cover, lid’} \\
\text{ixīl} & \quad \text{‘to cover’} & \text{ixiš} & \quad \text{‘roof, awning’} \\
\text{iwon} & \quad \text{‘(animal) to sound’} & \text{iwonuš} & \quad \text{‘a sound’} \\
\text{oxoxon} & \quad \text{‘to cough, catch cold’} & \text{oxoxonuš}^{i?} & \quad \text{‘a cold’} \\
\text{uškal} & \quad \text{‘to be strong’} & \text{uškališ} & \quad \text{‘strength’}
\end{align*}

This pattern could be the result of two possible sequences of events.

i. The nominalization may have occurred before the fusion of pronominal prefixes, and thus before the erosion of verb-internal glottal stop.

\begin{align*}
\text{a.} & \quad \ast\text{išō-} & = & \text{išoš} \\
& \text{to.roast-nom} & \quad \text{‘a roast, barbecued meat’}
\end{align*}

\begin{align*}
\text{b.} & \quad \ast\text{-išō} & \quad \text{išō} \\
& 3.\text{subject-to.roast} & \quad \text{‘s/he is roasting’}
\end{align*}

ii. Alternatively, speakers may have begun to notice the general pattern of vowel-initial verbs versus glottal-initial nouns at some point, perhaps even reanalyzing the initial glottals of nouns as nominalizers.
Sequence ii in fact seems likely to us, given the apparent productivity of nominalization following this pattern in Harrington's notes. Not only do derived nouns appear with the glottalization, but also nominalized clauses.

(18) Nominalized clause

... isį̄expenuša?š
   hi   s-iy-ʔ-expen-us-sa?š
   DET  3-PL-NOM-sing-3.BEN-NOM

'[and they responded] to their being sung to.'

Of course it is quite possible that both sequences of events took place, the first creating earlier nominalizations, the second the more modern nominalizing process.

4. Conclusion

The Chumash distinction between vowel-initial and glottal-initial roots can thus be understood as the consequence of a development in the morphology: the grammatization of pronominal prefixes. The fact that verbs and inalienable nouns normally appear with prefixes allowed the distinction to remain robust in the language, since those roots normally occur word-externally and automatic onsets would not appear. Other nouns appear most often utterance-externally as well, since basic word order is verb-initial, and nouns are often preceded by an article. The bond between the article and the noun is not as strong as that between pronominal prefixes and roots, however, so erosion of initial glottals is less likely to occur.

There was probably little phonetic distinction between basic vowel-initial words and basic glottal-initial words. As in many languages, there may have been an automatic utterance-initial glottalic onset even in modern Barareño. In his earlier transcriptions with Mary Yee's grandmother, Luisa Ygnacio, Harrington included initial glottals before all vowels. In the later work, the only context in which vowel-initial verb roots occur phrase-initially, in imperatives, they were recorded with a glottal: ʔeqwélus 'make it for him!'. Finally, a note from Harrington is indicative:

(19) The glottal stop occurs:

Before the initial vowel of many words, but not all words. A convenient test to prove its occurrence is to call for a possessive form, (1st, 2nd or 3rd person sing.) of the word. Thus:

ʔap⁸ 'house';  kap⁸ 'my house'
ekwel 'to make';  kekwel 'I make'

The majority of the vowel-initial roots they provide are verbs, while the majority of glottal-initial roots are nouns: Yuma a̱ḏ ‘to gather greens’, ʔa̱vé ‘snake’ (Halpern 1946: 275, 252). Among nouns, the glottal-initial roots tend to be alienable, while vowel-initial roots tend to inalienable (ʔi̱·ʔó̱· ‘willow’, i̱·ʔó̱· ‘tooth’ (1946: 264, 263). The mystery here, seems deeper. Although Yuman languages contain pronominal prefixes referring to their subjects, third person subjects of verbs and third person possessors of inalienable possessions are unmarked, or represented by zero.

Halpern remarks, however, that ‘initial vowels are pronounced with an aspirated attack’ (1946: 252), rather than glottalization. Munro (1974: 2) notes a similar automatic aspirated onset in Mojave. The aspirated onset would have served to distinguish basic glottal-initial words from vowel-initial forms. Gordon and Miller note that in Maricopa and Jamul respectively, however, initial vowels are automatically preceded by initial glottal onsets. Miller reports that ‘When words are spoken in isolation, it is difficult to distinguish an initial sequence of glottal stop followed by vowel from an initial vowel not preceded by glottal stop. (This is because initial vowels lack the aspirated onset found in some other Yuman languages.) In connected speech, however, initial glottal stops are usually heard. Stem and initial glottal stops are always recoverable when a stem is prefixed (1990: 10).’ In the absence of a third person pronominal prefix, and an aspirated onset, it could be difficult to maintain the glottal distinction. Miller provides an interesting note: ‘In Mrs. Dumas’ speech, word-initial glottal stops appear to have been lost’ (Miller 1990: 10).

Note
*Our work on Barbareño Chumash has been made possible by grant BNS90-11018 from the National Science Foundation. Chumash material cited here comes from microfilms of the fieldnotes of John Peabody Harrington, kindly made available to us by the Santa Barbara Museum of Natural History.

References
Final Glottalization in Barbareño Chumash and Its Neighbors*

Tsuyoshi Ono, Suzanne Wash, and Marianne Mithun

University of California, Santa Barbara

0. Introduction

Final glottalization in Barbareño Chumash appears in a variety of environments which include reduplication, imperative, and emphasis.1 Interestingly, we have found that final glottalization occurs in similar environments in neighboring languages which include Uto-Aztecan, Yokuts, and Yuman languages. Below, we will present final glottalization data from Barbareño and report the results of a small survey of the functions of final glottalization in neighboring languages. We hope our paper will stimulate discussion and further investigation by interested individuals.

1. Data

Barbareño data come from microfilms of John Peabody Harrington’s manuscripts.2 Barbareño transcription has been regularized because Harrington used a variety of symbols for individual segments over the course of his work. The data for other languages are taken from published sources such as grammars and dictionaries. We have generally retained the transcriptions of the sources.

2. Barbareño Chumash Final Glottalization

In this section, we will present final glottalization data from Barbareño Chumash. We will first discuss its phonetic

*We would like to thank Dale Kinkade, Margaret Langdon, Herb Luthin, Amy Miller and Mauricio Mixco for sharing with us their expertise on the languages they have worked on. Our work on Barbareño Chumash has been made possible by grant BNS90-11018 from the National Science Foundation.

1Final glottalization probably appears in similar environments in other Chumash languages. For instance, all the environments described in this paper are reported in Ineseño (Applegate 1972). Final glottalization is also found with transitivization and nominalization in Barbareño. Interestingly, these other uses also seem to be shared by neighboring languages. This awaits for future investigation.

2The microfilms were kindly made available to us by the Santa Barbara Museum of Natural History.

208
characteristics then present the three different, but related, functions associated with it.

2.1. Form

Several instances of the reduplicated form of *ku* 'person', recorded by Harrington, can be seen in (1):

(1)

kukū?   kuhkū?  'people'
kuhkū?  kuhkū?
kuhku?  kuhku?
kuhku?  kuhku?
kuhku?  kuhku?
kuhku?  kuhku?  kuhku?  "harku?"

Harrington transcribed final glottalization either as a glottal feature of a particular consonant or as an independent glottal stop. There is also some variation in Harrington’s transcription. However, since this glottalization always occurs on the last syllable, we would like to take it essentially as a feature associated with the entire last syllable. As the transcriptions may indicate, final glottalization is usually accompanied by some kind of stress and lengthening.

2.2. Functions

There are three different uses associated with final glottalization in Barbareño. First, it is used for emphatic purposes:

(2) Emphasis

sumọ’wōn  'sweeten'  sumowōn  'sweeten a lot'
ʔanaqipnās 'be fine'  ʔanaqiphāŋ  'be of such fineness'
kẹ’pú  'now'  kẹ’pú  'right now'

naʔalsaʔéʔ nhuq histáʔnlw
When it is a female baby ...

ʔikhù naʔalsaʔih électrique histaʔtáʔnlw
But if it’s going to be a male baby ...

The glottalized member of each pair has a more emphatic meaning.

The examples show that this use of final glottalization is
observable across different parts of speech. The most notable is the last example where the male baby is contrasted with the female baby and is marked by final glottalization. The second use of final glottalization is in imperatives:

(3) Imperatives

\[ ꠟ/ifugu ꠟ/ifugu \ 'tell (someone)' ꠟ/ifugu ꠟ/ifugu \ 'tell him!' \]

\[ ꠟ/ifugu ꠟ/ifugu \ 'grasp' ꠟ/ifugu ꠟ/ifugu \ 'grasp it!' \]

\[ ꠟ/ifugu ꠟ/ifugu \ 'put in' ꠟ/ifugu ꠟ/ifugu \ 'put (it) in!' \]

\[ ꠟ/ifugu ꠟ/ifugu \ 'sing' ꠟ/ifugu ꠟ/ifugu \ 'sing!' \]

The final glottalization forms imperatives from stems. Finally, final glottalization accompanies noun reduplication:

(4) Accompaniment to reduplication

\[ ꠟ/ifugu ꠟ/ifugu \ 'trail' ꠟ/ifugu ꠟ/ifugu \ 'trails' \]

\[ ꠟ/ifugu ꠟ/ifugu \ 'horse' ꠟ/ifugu ꠟ/ifugu \ 'horses' \]

\[ ꠟ/ifugu ꠟ/ifugu \ 'skin' ꠟ/ifugu ꠟ/ifugu \ 'skins' \]

\[ ꠟ/ifugu ꠟ/ifugu \ 'dish' ꠟ/ifugu ꠟ/ifugu \ 'dishes' \]

In the above pairs, reduplication indicates plurality, and final glottalization appears with the reduplicated forms. In all of these uses, final glottalization can be considered to function as an intensifier. Thus, there are two characteristics which unite these different uses of glottalization: position and semantics. For these reasons, we will treat this final glottalization as a single morpheme with one larger function 'intensification'. This is the analysis adopted by Harrington (n.d., Reel 33 Frame 273R and elsewhere). All three of its uses, emphasis, imperative, and accompaniment to reduplication, seem to be highly productive.3

---

3It is of interest that the three different uses of final glottalization can be aligned along a kind of continuum of degree of grammaticization:

Emphasis < Imperative < Accompaniment to reduplication

less grammaticized <----------> more grammaticized

The emphatic use occurs with a variety of parts of speech and seems to be the most discourse-related, two facts that indicate that grammaticization is still under way. Imperatives are (continued...)
3. Neighbors

Languages from three distinct language families, which are neighbors of Barbareño, exhibit similar uses of final glottalization. In this section, we will present data from these languages.

3.1. Uto-Aztecan

We first examine data from Uto-Aztecan languages from several different branches of the family.

Kawaiisu

In Kawaiisu, a Numic language, final glottalization is used to mark imperatives: 3

(5) Imperatives (Zigmond et al. 1990:35)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Stem</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>naake-</td>
<td>naake?e</td>
<td>'Listen!'</td>
</tr>
<tr>
<td>'to hear'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ka?a-tii-</td>
<td>ka?atii?i=ni</td>
<td>'Feed me!'</td>
</tr>
<tr>
<td>eat -CAUS-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'to feed'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-gwee-</td>
<td>magatii-gwe?e</td>
<td>'Go feed it!'</td>
</tr>
<tr>
<td>'go in order to'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>feed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-go:to:MOM</td>
<td></td>
</tr>
</tbody>
</table>

3 (...continued)

restricted to verbs, and final glottalization is the sole segmental indicator of this function, facts that suggest that final glottalization has become part of the verbal morphology. Finally, in the third use, final glottalization accompanies reduplication: it occurs with nouns and does not have independent status as a meaningful marker. In other words, its occurrence is an automatic consequence of noun reduplication. This may indicate that final glottalization is more fully integrated into the morphology than in the case of imperative.

4 Zigmond et al. (1990:96) also give the following set of data where the addition of glottal can be said to indicate some type of intensification. Zigmond et al. treat it as the momentaneous aspect and the position of the glottal is not on the final syllable:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>yozo-</td>
<td>'to fly'</td>
</tr>
<tr>
<td>?aga-zi-ki-</td>
<td>'to be a light'</td>
</tr>
<tr>
<td>?aaka-ci?i-ki-</td>
<td>'to be a flashing light'</td>
</tr>
<tr>
<td>yo?o?i-</td>
<td>'to jump'</td>
</tr>
</tbody>
</table>
Mono

Mono, another Numic language, seems to use final glottalization to indicate emphasis: 5

(6) Emphasis (Lamb 1958:220-221)

cawu 'good'  cawu' 'very good'
pyty 'after a while'  pyty' 'pretty soon'

Tübatulabal

Final glottalization seems to have some intensification function in Tübatulabal, which by itself makes up a separate subgroup of the Uto-Aztecan family:

(7) Iterative (Distributive) (Voegelin 1935:110-111)

atatdaha 'it bursts open'
a'tatda- 'they (pl.subj.) burst open'

ītībīha 'it breaks'
ī'tība- 'it broke in many places'

ītītdīha 'it is sawed up'
ī'tīdī- 'the sticks (pl.subj.) got sawed up'

atsabaha 'it is torn'
a'tsaba- 'it got torn in many places'

a'amaha 'tree is felled (through an impersonal agent)'
a'ama- 'the trees (pl.subj.) got felled (in the wind)'

Voegelin uses the terms 'iterative', but the term 'distributive' seems to capture the function better. 6 Notice in these pairs

5 The following data are also given by Lamb (1958:220-221), which suggests that glottalization may not necessarily be on the final syllable:

kywapaah 'beside'  kywa'paah 'beside (and very close to)'

ma -ni -hi  mani'hi
that-like-DEM  'just like that'
'like that'

6 Kroeber and Grace (1960:138) present the following set of data from Luiseno, a Takic language (another branch of Uto-Aztecan). Though they describe these data in terms of pluralization, it looks very much like the distributive use of final glottalization in Tübatulabal:

(continued...
that final glottalization is accompanied by lengthening and stress as in Barbareño.

**Kitanemuk**

Similar examples are found for Kitanemuk, a Takic language:

(8) Accompaniment to reduplication (Anderton 1988:59, 61)

<table>
<thead>
<tr>
<th>?a-mukpi 'his nose'</th>
<th>?a-mu-mukpiʔ 'points of land'</th>
</tr>
</thead>
<tbody>
<tr>
<td>hayhaʔy 'bird sp.'</td>
<td>huyhuʔy 'bird sp.'</td>
</tr>
<tr>
<td>caycaʔy 'bird sp.'</td>
<td></td>
</tr>
</tbody>
</table>

Though there are no simplex forms for the last three examples, the pattern is still very suggestive. In fact, these forms may never have had simplex counterparts; the final glottalization could have been due to an analogy to regular reduplicated forms which may also have had automatic final glottalization.

Kitanemuk also shows the imperative use of final glottalization:

(9) Imperative (Anderton 1988:698)

wiloy 'play instr.'  wiroʔyi imp.

This may not be a very productive process because it is the only example we found in the source.

**Cahuilla**

Cahuilla, another Takic language, shows the imperative use:

6(...continued)

<table>
<thead>
<tr>
<th>koti 'cover it!'</th>
<th>kotiʔi 'cover them!'</th>
</tr>
</thead>
<tbody>
<tr>
<td>hol 'spread it'</td>
<td>holîʔi 'spread them'</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>huʔyaki 'pull out two'</th>
<th>húʔiʔi/huʔyíʔiʔ 'pull them out, up'</th>
</tr>
</thead>
<tbody>
<tr>
<td>huʔyaq 'is left over, exceeds'</td>
<td></td>
</tr>
</tbody>
</table>

7A closely related language, Serrano, exhibits an intriguing pattern where the glottal seems to be moved to the end to form imperatives (Hill 1967:217):

<table>
<thead>
<tr>
<th>čiʔa</th>
<th>čiáʔ 'pick it up!'</th>
</tr>
</thead>
<tbody>
<tr>
<td>piʔa</td>
<td>piáʔ 'throw it!'</td>
</tr>
</tbody>
</table>
(10) Imperatives (Langacker 1977:54-55)

ne-' -tee-
me-you-see-IMP
'Look at me'

paxa -ni -' e -'aś
enter-CAUS-IMP your-pet
'Stable your horse'

Cupeño

Imperative glottalization also appears in Cupeño, another Takic
language:

(11) Imperatives (Hill 1966:164)

yéləcə 'clean' yéləcɨʔ 'clean (it)!'
cēlə 'snip' cēlɨʔ 'Snip (it)!'
kʷəwə 'be hollering' kʷəwəʔ 'holler!'
qāʔayə 'speak Luiseño' qāʔayəʔ 'Speak Luiseño!'

3.2. Yokuts

Wikchamni

One Yokuts language, Wikchamni, uses final glottalization to mark
imperatives:

(12) Imperatives (Gamble 1978:19, 64)

tiʔi: 'to sink' tiʔiʔ 'sink!'
ṭaka: 'to stay' ṭakaʔ 'stay overnight!'
čutu: 'to urinate' čutuʔ 'urinate!'

3.3. Yuman

Diegueño

Finally, one Yuman language, Diegueño, uses final
glottalization to mark imperatives:

Hill (1966:164) also gives the following data, which show
that glottalization is not always on the final syllable.
According to her, the position depends on the verb class:

kúpe 'sleep' kúʔpe 'sleep!'
cáṣpele 'mend' cáṣpeʔe 'mend!'
(13) Imperatives (Langdon 1970:74)

- a. 'to go'  kaʔ  imp.
- mi. 'to cry'  kəmiʔ  imp.

4. Possible Explanations

Thus we have seen that the intensification function of final glottalization in Barbareño is shared by neighboring languages from distinct families. There could be four possible explanations for this phenomenon. A first possibility is common genetic inheritance. However, Chumash, Yokuts, Uto-Aztecan, and Yuman are all separate families. A further grouping has not been successful.

A second possibility is a universal human propensity. It seems reasonable to entertain the idea that human vocalization naturally involves glottalization in its expressive modes. That is, glottalization may be produced naturally when humans are vocalizing in certain expressive (or excited) modes. It is not impossible that this could lead to the grammaticization of glottalization to encode intensification.

This suggestion would predict that the type of morphemes we described above should then be common in languages of the world. However, our informal survey of other Native American languages does not seem to support this prediction. That is, though there are some languages which have the same kind of morphemes, they are still not very common and there is certainly nothing like the kind of concentration of languages which we discussed above.

A third possibility is contact. There are several points which should be made here. First of all, these languages are all spoken in a relatively small area and, in fact, contact among the speakers of these languages has been amply demonstrated (Heizer 1978). It is especially noteworthy that, compared with these groups, Barbareño had a relatively small amount of contact with its northern neighbor, Salinan and that we did not find any relevant patterns in Salinan. An incompleteness of the Salinan database is probably not the reason for this because the main portion of the Salinan data was collected by Harrington in 1920s and 1930s. By that time, he was quite familiar with the grammars of Chumash languages. If he had found anything similar in Salinan, he would not have missed it.

Another pertinent fact comes from the Yuman family. Diegueño, which uses final glottalization to mark imperatives, is spoken in the most northern part of the Yuman speaking area, the closest among many Yuman languages to the languages in this paper. These facts regarding Salinan and Diegueño may point to contact as the ultimate factor. However, Mauricio Mixco (p.c.) has pointed out that Kiliwa, another Yuman language spoken in Mexico, also uses glottalization in similar environments. Intriguingly, Mauricio Mixco also has pointed out that Diegueño and Kiliwa are the only Yuman languages which have this feature and that the migration history of the speakers of these languages is not very clear.
A forth possibility is chance. First of all, Langacker (1977:54) states that the imperative suffix -'u is attested in both Northern and Southern branches of the Uto-Aztecan family: the two divisions of the family. It seems possible to think this morpheme was in the proto-language and it has resulted in a simple final glottalization through the reduction of the vowel in the Uto-Aztecan languages we discussed above.

Similarly, some Yokuts languages including Wikchamni seem to use a glottal suffix to mark the future tense (Newman 1944). A semantic change from future to imperative seems to be common cross-linguistically. These facts can suggest that Uto-Aztecan and Yokuts may have independently developed the intensifying final glottalization, and thus the similarity among the languages of these two families may be due to chance.

However, even if the scenarios given above turn out to be correct, we still have to account for the situations in Barbareño and Diegueño, especially the former, which uses the final glottalization extensively. Here we should remind ourselves that the segment in question, glottal stop, is rather small, and cross-linguistically such segments as /k/, /q/, and /h/ are often reduced to it over time. If this type of sound change accounts for the Barbareño and Diegueño cases, we should see similar morphemes in many languages. However, as we noted above regarding the universal human propensity possibility, this does not seem to be the case. This fact thus may also suggest that the phenomenon we found may not be simply due to chance. Thus we are again left with the fascinating concentration of languages in Santa Barbara and its surrounding area all of which seem to use final glottalization for the intensification function. A much more in depth investigation will be necessary to evaluate all these facts and possibilities fully.

Finally, Margaret Langdon (p.c.) reminds us of the possibility of more than one factor responsible for the sharing of features among languages. Accordingly, the glottalization could have originated in one family as the result of natural expressive tendencies or phonological reduction, then spread through contact. We hope what we have presented in this paper will stimulate discussions and further investigations by interested scholars.

References


