





## TUPI STUDIES I

SUMMER INSTITUTE OF LINGUISTICS PUBLICATIONS

IN

LINGUISTICS AND RELATED FIELDS

PUBLICATION NUMBER 29

EDITOR

Irvine Davis

ASSISTANT EDITORS

Alan C. Wares      Iris M. Wares

CONSULTING EDITORS

Doris Bartholomew      Eugene Loos  
Robert E. Longacre      William R. Merrifield  
                                 Kenneth L. Pike

PUBLISHER

Benjamin F. Elson

# TUPI STUDIES I

Edited by:

David Bendor-Samuel

A Publication of the  
SUMMER INSTITUTE OF LINGUISTICS  
of the  
University of Oklahoma  
Norman

1971

Copyright © 1971 by the  
Summer Institute of Linguistics

Printed in U.S.A. by  
The Church Press Inc.  
3915 San Fernando Road  
Glendale, Calif. 91209

## TABLE OF CONTENTS

PREFACE . . . . .	vii
PARINTINTIN PHONOLOGY . . . . . by Helen Pease and LaVera Betts	1
NASALIZATION IN KAIWÁ . . . . . by Carl H. Harrison and John M. Taylor	15
THE MORPHOPHONOLOGY OF ASURINI WORDS . . . . . by Carl H. Harrison	21
COCAMA CLAUSE TYPES . . . . . by Norma Faust	73
INTERNAL CLASSIFICATION OF THE TUPI-GUARANI LINGUISTIC FAMILY . . . . . by Miriam Lemle	107



## PREFACE

The studies which comprise this volume have arisen out of the work of the Summer Institute of Linguistics, and, with one exception, are written by members of that Institute. The majority of the languages dealt with are spoken, and have been studied, within the borders of Brazil, but advantage has also been taken of work done by SIL members in neighboring countries. Within Brazil, SIL members have currently begun work on languages spoken by forty different tribal groups, including representatives of all the major language families. Further published studies resulting from this work will be forthcoming.

The present volume does not attempt a comprehensive description of one language, but presents briefer reports focussing on restricted aspects of a number of languages. The papers were written at different times and places, and reflect differing approaches to linguistic description. Pease and Betts' description of Parintintin phonology was finally polished at the SIL workshop in Belém, February to April, 1968. Taylor and Harrison's paper on Kaiwá is a redraft of one read to the Brazilian Anthropological Association in July of 1963. Harrison's Asurini Morphophonology was submitted in partial fulfilment of the requirements for a master of arts degree at the University of Pennsylvania in 1967.

The content of the papers is varied. 'Parintintin Phonology' provides a brief description of the sound system at all levels, with most detail on nasalization. The treatment is orthodox and handles nasalization from a phonemic point of view. 'Nasalization in Kaiwá', in contrast, deals only with nasalization, and is not closely linked to any one pattern of description either in concept or terminology. Its closest affinity is perhaps with the London school. These two treatments of a feature which is very widespread in Tupi-Guarani languages will be of considerable interest and help to other students of this family. They also provide a contrast in approach which is not without theoretical interest.

'The Morphophonology of Asurini Words' is a generative-transformational description of morphophonology. It is an excellent example of the possibilities, and limitations,

of linguistic description by means of a series of rules. The interest and attention of the reader may well be chiefly in the descriptive technique employed, at least to begin with, and the paper will repay study from this point of view alone. Those already familiar with the technique, or with willingness to master it, will also be grateful for the data provided on another feature common to a number of languages in this family.

'Cocama Clause Types', on the other hand, is a tagmemic study of one aspect of syntax. It deals with surface structure by establishing slots and fillers, and is another good example of the advantages and disadvantages of the descriptive framework chosen. Again, those familiar with the approach will be able to take ready advantage of the data, while those unfamiliar with it must give attention to this first.

It is a pleasure to be able to include in this volume the comparative paper by Miriam Lemle, who is a research student at the Federal University of Rio de Janeiro, and for whom SIL members have a high regard as a colleague. Her work in bringing together and interpreting the data collected by SIL workers and others is a significant contribution to the comparative field that will be of interest to all Tupinologists, and hopefully will provide a stimulus for further study.

David Bendor-Samuel

# PARINTINTIN PHONOLOGY

Helen Pease and LaVera Betts

1. Consonants
2. Vowels
3. Syllables
4. Nasalization
5. Higher Level Features
6. Text

0. The purpose of this paper is to describe the phonology of Parintintin<sup>1</sup> with special focus on the phenomenon of nasalization.

## 1. Consonants.

	Bilabial	Alveolar	Alveo- palatal	Velar	Labio- Velar	Glottal
Stops	p	t	ç	k	k <sup>w</sup>	ʔ
Nasals	m	n	ɲ	ŋ	ŋ <sup>w</sup>	
Fricatives	v					h
Flap		r				

The consonant system consists of fourteen phonemes which contrast at six points of articulation. There are further contrasts with respect to manner of articulation at each of the six points. These are labeled on the chart for convenience as stop, nasal, fricative, and flap articulations. However, these terms do not fully describe the phonetic variants, and further details are given below.

The stops contrast at bilabial, alveolar, alveopalatal, velar, labio-velar, and glottal points of articulation. The nasals contrast at the same points of articulation, except that a glottal member does not occur in the nasal series. The fricatives contrast at bilabial and glottal points of articulation.

The following are examples of the consonants:

Stops

/p/	pira	'fish'	/t/	tata	'fire'
/č/	čãʔă	'okay'	/k/	kaʔa	'leaf'
/kʷ/	kʷara	'sun'	/ʔ/	aʔa	'I fall'

Nasals

/m/	măhă	'where is'	/n/	năhă	'thus'
/ñ/	ñăne	'we inclusive'	/ŋ/	ŋăhă	'they'
/ŋʷ/	ŋʷăʔă	'he mocks'			

Fricatives

/v/	ovava	'it shakes'	/h/	haha	'stalk of bananas'
-----	-------	-------------	-----	------	-----------------------

Flap

/r/	rañiʔi	'long ago'
-----	--------	------------

The stops are voiceless and unaspirated. Except for velar palatalization, which is described later, they have no variant forms. The alveopalatal member of the series, /č/ as in čabeʔe 'old man', is affricated: [tš].

The nasal series has the following variants:

(1) Nasal continuants [m n ñ ŋ ŋʷ] occur utterance initially preceding a nasal vowel and utterance medially between nasal vowels: tupahămă [tupahămă] 'cord', năhă [năhă] 'thus', kũñă [kũñă] 'woman', ɔmɔŋita [ɔmɔŋita] 'he converses', kăŋʷi [kăŋʷi] 'coffee'.

(2) Prenasalized voiced stops [ᵐb, ᵐd, ᵐj, ŋg, ŋgʷ] occur following a nasal vowel and preceding an oral vowel: ɔmoapɪ [ɔᵐboapɪ] 'he cooks', ñănu [ñăᵐdu] 'spider', otĩñipe [otĩᵐjipe] 'he ties right away', ɔŋa [ɔŋga] 'house', ñũŋʷava [ñũᵐgʷabə] 'weeds'. The bilabial and alveolar members of the series also occur in utterance-initial position preceding an oral vowel: marakaʔña [ᵐbarakaʔja] 'ocelot', nehe [ᵐdehe] 'you singular'.

(3) Voiced stops [j g gʷ] occur between oral vowels and in utterance-initial position preceding an oral vowel:

ñihi [ʝihi] 'I', iʔaŋahi [iʔaŋahi] 'it stinks', ŋʷiraʔi [gʷiraʔi] 'bird'.

(4) The phoneme /ñ/ has additional variants. [ʝ] fluctuates freely with [y], and [ñ] with [ɣ̃] utterance medially in unstressed syllables: muñuhu [ᵐbuʝuhu] or [ᵐbuyuhu] 'frog', mőkõñäteʔi [mőkõñäteʔi] or [mőkõɣ̃äteʔi] 'a few'.

Members of the fricative series are respectively a voiced bilabial fricative, /v/, and a voiceless glottal fricative, /h/. Except for palatalization of the glottal fricative, described later, there are no major variants.

The flap is alveolar and has no variant forms.

Palatalization of the consonants /k ŋ h/ may occur following a high front vowel with or without an intervening syllable-final glottal stop. This does not occur, however, preceding another high front vowel. ikatu [ikʷatu] 'it is pretty', akitiʔŋo [akitiʔŋo] 'I soap down', õmoñiheʔa [õᵐboʝihʷeʔa] 'he mixes'.

## 2. Vowels.

	front		central		back	
	oral	nasal	oral	nasal	oral	nasal
high	i	ĩ	ɨ	ɣ̃	u	ũ
low	e	ẽ	a	ã	o	õ

The vowel system consists of twelve phonemes, which contrast with respect to oral and nasal quality; front, central, and back tongue positions; and relatively high and low tongue positions.

The following examples illustrate contrasts within the oral and the nasal vowel series:

/i/	õmõmi	'he stops'	/ĩ/	õmĩ	'he hides'
/e/	õmõme	'he makes cakes'	/ẽ/	ohẽ	'he leaves'
/ɨ/	õmõmĩ	'he fastens'	/ɣ̃/	otɣ̃	'he plants'
/a/	õmõma	'he kills a lot'	/ã/	omã	'he grabs'

/u/	õmõmu	'he pierces'	/ũ/	õmũ	'he spits out'
/o/	õmõmo	'he throws'	/õ/	õmõ	'he sticks'

The following examples show contrasts between oral and nasal vowels:

/i/	opi	'it stops'	/a/	oka	'it breaks'
/ĩ/	opĩ	'he cuts close'	/ã/	ikã	'he is strong'
/e/	ahe	'people'	/u/	akutu	'I puncture'
/ẽ/	ahẽ	'I leave'	/ũ/	ahetũ	'I smell'
/i/	oti	'it catches'	/o/	oñeheʔo	'he cries'
/ĩ/	otĩ	'he plants'	/õ/	opeheʔõ	'he halves'

The 'low front' and 'low back' vowels vary in tongue height from mid open [ɛ ɔ] to mid close [e o]. The front mid close variant is heard only immediately preceding a high front vowel /i/. The back mid open variant tends to occur more frequently between bilabial and alveolar nasals. The 'low central' vowel is normally low open in tongue position, and the 'high' vowels are always high close.

All vowels have voiceless variants which may occur at the end of a phonological sentence (see 5.2). The vowels /a/ and /ã/ have variant forms with generally lax and lenis articulation, in which the tongue is in mid central position. These variants are never stressed.

3. Syllables. Parintintin syllables consist of a nucleus of one or two vowels with an optional initial and/or final consonant. The following patterns have been recorded:

V	ohi	'he falls'	CV	ñihi	'I'
VC	iʔŋʷava	'table'	CVC	kaʔŋʷira	'bush'
VV	aivu	'near'	CVV	aheakãŋã	'person's head'

Syllables of patterns VC and CVC do not occur in utterance final position.

3.1. Complex Segments. A number of sequences of two consonantal sounds are interpreted as single consonants, since nonsuspect syllable patterns allow for no consonant clusters except across syllable boundaries. Complex segments so interpreted include the affricate /ʧ/ (phonetically [tʃ]), which occurs in single C slots as in čaho 'let's go'; the [ʤ] (i.e., [dʒ]) variant of the /ɲ/ phoneme; and the prenasalized stops described in Section 1 as allophones of the nasal series.

Sequences of a consonantal sound plus a nonsyllabic high vocoid are likewise interpreted as single consonants. To interpret the vocoid element as a vowel would result in the recognition of a sequence of three vowels in one syllable (as in [ikʷai] 'it burns'), a pattern that does not otherwise occur. Complex segments of this type include the labio-velar phonemes /k<sup>w</sup> ŋ<sup>w</sup>/ and the palatalized variants of /k ŋ h/ previously described.

Sequences of a consonant plus lenis [ə], and of a consonant plus a voiceless vocalic release are interpreted as CV on the basis of the most economical analysis. Lenis [ə] is recognized as a submember of the phoneme /a/ (see 2), while voicelessness is recognized as a feature of the termination of the phonological sentence (see 5.2). This results in a more economical analysis than the alternative of setting up additional variant forms for all the consonants which may occur with lenis [ə] or vocalic release, as in the examples tupabə [tupabə] 'hammock' and ahereveka [aherebekA] 'belly'.

Sequences of a vowel plus a nonsyllabic high vocoid are interpreted as two vowels because there are examples of other two-vowel sequences occurring in comparable positions in the syllable patterns:

ahēakāŋā 'a person's head' ŋ<sup>w</sup>āivīhī 'old woman'

ŋuitaraga 'frog'

All such sequences are analyzed as comprising a single syllable nucleus because they have only one syllabic pulse, and the first of the two vowels is always the more prominent.

Sequences of a syllabic vowel plus a nonsyllabic high front vocoid plus a syllabic vowel are interpreted as VCV because there are otherwise no examples of a three-vowel sequence: muŋuhu [ᵐbuyuhu] 'frog'.



- V any oral vowel
- $\tilde{V}$  any nasal vowel
- O consonants having only oral forms and variants  
(p, t, č, k, k<sup>w</sup>, ʔ, v, h, r)
- N consonants having only nasal forms and variants  
(m and n)
- G consonants having oral and nasal forms and variants  
(ñ, ŋ, ŋ<sup>w</sup>)
- n and ŋ consonantal variants of N and G which are  
wholly nasal (m, n; ñ, ŋ, ŋ<sup>w</sup>)
- nd and ŋg consonantal variants of N and G which are  
prenasalized (<sup>m</sup>b, <sup>n</sup>d; nɟ, ŋg, ŋg<sup>w</sup>)
- g consonantal variants of G which are wholly oral  
(ɟ, g, g<sup>w</sup>)
- # pause

4.1. Nasalization of Contiguous Elements. Nasal vowels may occur preceding or following any consonant, but oral vowels do not occur preceding N or ŋ. Variant forms of all consonants of the nasal series (N and G) are predictable in terms of contiguous vowels.

The following chart lists all theoretically possible environments and shows patterns and restrictions in co-occurrence of the forms. The chart is to be read as follows: when N is preceded by pause and followed by a nasal vowel, it is articulated as [m] or [n], etc.

Vocalic Environment	Consonantal Variant		
	N	G	O
# - $\tilde{V}$	n	ŋ	O
$\tilde{V}$ - $\tilde{V}$	n	ŋ	O
$\tilde{V}$ - V	nd	ŋg	O
# - V	nd	g	O

Vocalic Environment	Consonantal Variant		
	N	G	O
V - V	-	g	O
V - $\tilde{V}$	-	-	O

In two-vowel sequences only  $VV$  and  $\tilde{V}\tilde{V}$  occur, except in the case of reduplicated morphemes (see 3.2).

The patterns and restrictions of co-occurrence shown above are herewith summarized. Nasal vowels precede both nasal and prenasalized, but not oral, variants of N and G; and they follow only their nasal variants. Oral vowels follow both oral and prenasalized, but not nasal, variants of N and G; and they precede only the oral variants. Prenasalized variants of N occur only before oral vowels, and nasal variants occur elsewhere. Oral variants of G occur when there is no contiguous nasal vowel; nasal variants occur when there is no contiguous oral vowel; and prenasalized variants occur following a nasal and before an oral vowel.

Categories and rules may therefore be set up to account for patterns of co-occurrence between contiguous elements. N and  $\tilde{V}$  are inherently nasal, and never have oral variants. G and V are potentially nasal but occur as oral variants except when influenced by a contiguous nasal form. O is inherently oral, and never represents nasal variants.  $\tilde{V}$  exerts a nasalizing influence, progressively and regressively, upon contiguous N and G, resulting in the selection of nasal or prenasalized variants (the latter when an oral vowel follows). It also regressively nasalizes a contiguous vowel. V exerts an oralizing influence, regressively, upon contiguous N and G, resulting in the selection of prenasalized variants when they follow  $\tilde{V}$ . N and  $\eta$  exert a nasalizing influence, regressively, upon contiguous vowels, resulting in the selection of nasal forms.

#### 4.2. Nasalization of Noncontiguous Elements.

(1) Nasalization across glottal stop. Sequences of two vowels separated by a glottal stop are always either both oral or both nasal, e.g.,  $k^w\tilde{i}^?i$  'porcupine'.

When a nasal vowel is followed by a sequence of  $?N$  or  $?G$  the nasal or prenasalized variants of N or G always occur:

n and ŋ before nasal vowels and nd and ŋg before oral ones, e.g., mārāʔŋu [mārāʔŋu] 'who knows if'. Sequences of ʔN and ʔŋ are always preceded by nasal vowels, e.g., okāʔñĩ [okāʔñĩ] 'he forgets'.

Ũ, N, and ŋ may therefore be said to exert a nasalizing influence on phonemes within the same morpheme which are separated from them by only a glottal stop, just as if they were contiguous.

(2) Regressive nasalization within the word. Any syllable which begins with N or G is fully nasal when it precedes an inherently nasal element in the same grammatical word, i.e., the vowel in that syllable is nasal and the consonant occurs in its nasal variant n or ŋ.<sup>2</sup> Examples are: ñaʔŋwara [jaʔgʷarə] 'jaguar' plus -pīnīmā [-pīnīmā] 'spotted' becomes ñāŋwāpīnīmā [ñāŋwāpīnīmā] 'spotted jaguar'; ñi [ji] 'I' plus -pirivutū [-pirivutū] 'brown skinned' becomes ñīpirivutū [ñīpirivutū] 'I am brown skinned'.

This phenomenon may be accounted for by the following additional rule: an inherently nasal element exerts a regressive nasalizing influence upon any syllable preceding it within the same grammatical word which contains N or G.

(3) Progressive nasalization within the word. Certain suffixes may be regarded as potentially nasalized and others as inherently oral, because the former occur with nasal vowels following syllables with nasal vowels, while the latter always occur with oral vowels. This may be stated as a type of progressive nasalization, the operation of which is restricted to certain suffixes, e.g., nā/nān- [nā/nān-] 'like this' plus -uhu [-uhu] 'augmentative' becomes nānūhū [nānūhū] 'really like this', while nā/nān- plus -veʔe [-veʔe] 'thing' becomes nāveʔe [nāveʔe] 'a thing like this'.

Some suffixes are frequently, but not invariably, nasalized or may be nasalized only on the first syllable, e.g., ōñĩʔĩŋ- [ōñĩʔĩŋ-] 'he talks' plus -ahĩ [-ahĩ] 'forcefully' becomes ōñĩʔĩŋāhĩ [ōñĩʔĩŋāhĩ] 'he scolds'.

4.3. An Alternative Analysis to that presented here is also possible; nasalization may be treated on a level higher than that of the syllable. Although the phenomena are not entirely parallel, a detailed discussion of this problem may be seen in the paper by Harrison and Taylor which also appears in this volume.

It should be noted that there is some free variation in the degree of nasalization which occurs, differing according to both occasion and speaker. Inherently nasal vowels are generally more nasalized than potentially nasal ones.

5. Higher Level Features. The higher level features of Parintintin have not yet been fully analyzed; the following, however, is presented as a tentative analysis.

5.1. Phonological Phrase. A phonological phrase is postulated to account for certain features of stress, pitch, and length which are commonly associated with groups of syllables. The phonological phrase is composed of one or more syllables, generally more than one. In a series of syllables comprising a phonological phrase one syllable occurs with higher pitch and with more stress than the others. This is termed 'phrase stress'.

Stress. Phrase stress generally occurs on the first syllable of the phrase except in initial phrases in a sentence, in which case one or two unstressed syllables may precede the stressed syllable (see Sentence 5 of the text in Section 6).

Pitch. Although the pitch normally drops following the phrase stress, it is sometimes not noticeably higher on the phrase stressed syllable than on the other syllables of the phrase. The last phrase in a phonological sentence ends on a low pitch unless it carries sentence stress (see below). A sentence-medial phrase may also end on a low pitch or on a pitch which is slightly rising, leading up to the phrase stress of the following phrase. In a sentence-initial phrase the syllables preceding the stressed syllable normally occur with a pitch lower than that of the stressed syllable, but some have been found with a higher pitch.

Length. Stressed syllables tend to be slightly lengthened. When a CV or V syllable is followed by a syllable beginning with the phoneme h, the h may drop and the surrounding vowels fuse into one slightly lengthened syllable, e.g., ěmuhu may be pronounced [ẽ<sup>m</sup>bu·]. This reduction may even occur across phrase boundaries.

5.2. Phonological Sentence. The phonological sentence is set up to account for certain features of stress, intonation, voicelessness, and pause commonly associated with groups of phrases. The phonological sentence consists

of one or more phonological phrases, generally more than one.

Stress. The heaviest stress of the phonological sentence is termed 'sentence stress', and the syllable on which it occurs has the highest pitch. Sentence stress falls on a phrase stressed syllable, and may occur initially, medially, or finally in the sentence, depending upon what part of the grammatical construction seems to be in focus. In an interrogative sentence containing a question word or phrase, for example, the sentence stress normally occurs on that word or phrase (see Sentence 11 in the text, Section 6). The sentence stress most often occurs on one of the first three phrases of the sentence.

In some sentences heavy stress occurs at two points. The second of these is always on the last or next to last syllable of the sentence and does not necessarily coincide with phrase stress. In such cases, the second stress appears to mark the addition of a construction to a sentence which would be grammatically complete without it. This added element frequently serves to clarify or add to information already given in the preceding part of the sentence. This is illustrated in Sentences 1 and 3 of the text.

Intonation. The most common intonation pattern for the sentence is a gradual rise in pitch to the sentence stress followed by a gradual drop in pitch. Thus, each phrase preceding the one containing the sentence stress is successively higher in pitch than the one before, while each phrase following it is successively lower.

Voicelessness. Vowels may be voiceless at the end of phonological sentences, and this voicelessness includes at times as many as the last three syllables. When this is the case, words with the consonants v, r, or the [g] allophone of  $\eta$  in an unstressed final syllable generally occur in variant forms in which these consonants are replaced by p, t, and k, respectively.

Pause. Phonological sentences are normally followed by pause. Occasionally, two phonological sentences may be juxtaposed with normal intonation on each and no pause separating them. This is considered a complex sentence, of which Sentence 5 of the text is an example.

Emphasis. Emphasis may also occur in a sentence and is

marked by extra stress and higher pitch on one syllable, plus one or more of the features described below. An emphasized syllable always has the sentence stress, and the phrase containing it will be referred to as the stressed phrase.

(1) A sharp drop in pitch may occur over one or two of the first syllables of the stressed phrase and/or over the entire phrase immediately preceding it. This pitch drop has been found to occur on each of two phrases immediately preceding a stressed phrase with cases of morpheme reduplication.

(2) A glottal stop may be introduced at the end of the phrase preceding the stressed phrase, e.g., Sentence 6 in the text. The glottal stop may separate the syllables of a morpheme.

(3) Lengthening may occur on the final vowel or glottal stop of phrases with lowered pitch preceding the stressed phrase, or on the first consonant of the stressed phrase, e.g., Sentence 6. Some cases of lengthening of the final vowel of the stressed phrase have also been observed.

6. Text. The following transcription consists of two excerpts from a tape recording made by Catarina, the chief's wife at Canavial. It gives a brief account of the pacification that took place years ago of the Jupá Indians, a group related to the Parintintin who were living on the Machado River at that time.

The text is written phonemically, except for the indication of voicelessness and length which are significant to the sentence. The following symbols are used: / indicates phrase division; // pause; ' phrase stress; " sentence stress; V progressive devoicing of the preceding vowel; A, E, etc. voiceless vowels; · lengthened vowels or consonants. The phonological sentences are numbered consecutively. The numbering within each sentence of the text correlates with the numbers of the free translation. Words in parentheses in the free translation are added to give sense to the English.

1. ǎẽ"rẽ<sub>1</sub> ɲã<sub>2</sub> hoi<sub>3</sub>/ 'ñupa<sub>4</sub> ɲã<sub>4</sub>/ 'pĩri<sub>5</sub> hako<sub>6</sub> /  
 'ñire<sub>7</sub> "ɲãV<sub>8</sub> // 2. 'ñire<sub>1</sub> / 'ɲã<sub>2</sub> nã/"ẽ<sub>3</sub> ko<sub>1</sub> o/'ho<sub>4</sub> ɲã<sub>5</sub> /

'piri<sub>6</sub> ha/'koV<sub>7</sub> // 3. he''?ii<sub>1</sub> na<sub>2</sub> o/'ŋovo<sub>3</sub> nã<sub>4</sub> / 'piri<sub>5</sub>  
 ei<sub>6</sub> nã<sub>7</sub> / 'nã<sub>8</sub> ne/'piakA<sub>9</sub> // 4. 'kiro<sub>1</sub> ore<sub>2</sub> / ''ruri<sub>3</sub> pe<sub>4</sub>  
 mōñi/'rōmō<sub>5</sub> ei<sub>6</sub> / 'nãV<sub>7</sub> // 5. oro<sub>1</sub> 've<sub>2</sub> na<sub>3</sub> / 'imō/'mori<sub>4</sub>  
 ŋwe/'revi<sub>5</sub> ei<sub>6</sub> / 'nã<sub>7</sub> / 'čave<sub>8</sub> / ''na<sub>8</sub> po ōmō/'mo<sub>9</sub> ŋwerevi<sub>10</sub> /  
 nã<sub>11</sub> nehe<sub>12</sub> raka/'eV<sub>13</sub> // 6. 'teremō?/'mori<sub>1</sub> ei<sub>2</sub> ore<sub>3</sub> nã<sub>4</sub>  
 / 'pe<sub>5</sub> ei<sub>6</sub> // 7. 'po<sub>1</sub> ahe<sub>2</sub> ore<sub>3</sub> ñu/'ka<sub>4</sub> te<sub>5</sub> ei<sub>6</sub> nã<sub>7</sub> // 8.  
 'noro/'koi<sub>1</sub> ko<sub>1</sub> / 'ore<sub>2</sub> / 'pe<sub>3</sub> ne/'he<sub>4</sub> e<sub>5</sub> po / 'ñire<sub>6</sub> / 'kA<sub>6</sub>  
 // 9. ei<sub>1</sub> 'ñi<sub>2</sub> ei<sub>3</sub> / 'ñire<sub>4</sub> / 'naV<sub>4</sub> // 10. 'noro/'koi<sub>1</sub>  
 ore<sub>2</sub> / 'pe<sub>3</sub>ne/'he<sub>4</sub> // 11. mō'mě<sub>1</sub> pe<sub>2</sub> ne/'koi<sub>3</sub> ei<sub>4</sub> // 12.  
 pai''vo<sub>1</sub> o/'re<sub>2</sub> re/'koiV<sub>3</sub> // 13. 'ore<sub>1</sub> rero/'ho<sub>2</sub> ō/'na<sub>3</sub> pe<sub>4</sub>  
 hE/'tE<sub>5</sub> // 14. 'torohe/'pia<sub>1</sub> peña/'piña<sub>2</sub> ei<sub>3</sub> o're<sub>4</sub> //

Translation of the Text. (The word po has only been translated in Sentence 7. It indicates a certain degree of indefiniteness or uncertainty with regard to the particular thing in question.)

1. After this<sub>1</sub> they<sub>2</sub> went<sub>3</sub> to be with<sub>5</sub> the Jupás<sub>4</sub>, a long time ago<sub>6</sub>--Jiré<sub>7</sub> and the others<sub>8</sub> (went). 2. Jiré<sub>1</sub> and the others<sub>2</sub> went<sub>4</sub> first<sub>3</sub> to be with<sub>6</sub> them<sub>5</sub> long ago<sub>7</sub>. 3. "A lot of<sub>1</sub> them<sub>2</sub> went<sub>3</sub> to be with<sub>5</sub> them<sub>4</sub>," they<sub>7</sub> said<sub>6</sub>, "to see<sub>9</sub> them<sub>8</sub>." 4. "Now<sub>1</sub> we<sub>2</sub> have come<sub>3</sub> to tame<sub>5</sub> you<sub>4</sub>," they<sub>7</sub> said<sub>6</sub>. 5. "And<sub>2</sub> then<sub>1</sub> he<sub>3</sub> almost<sub>5</sub> threw<sub>4</sub> (an arrow)," they<sub>7</sub> said<sub>6</sub>. "The old man<sub>8</sub> almost<sub>10</sub> threw<sub>9</sub> (an arrow) on<sub>12</sub> them<sub>11</sub> long ago<sub>13</sub>." 6. "'Don't throw<sub>1</sub> (it)!' we<sub>3</sub> said<sub>2</sub> to<sub>5</sub> them<sub>4</sub>," (they) said<sub>6</sub>. 7. "A person<sub>2</sub> could<sub>1</sub> really<sub>5</sub> kill<sub>4</sub> us<sub>3</sub>," they<sub>7</sub> said<sub>6</sub>. 8. "We<sub>2</sub> aren't fighting<sub>1</sub> with<sub>4</sub> you<sub>3</sub>," Jiré<sub>6</sub> said<sub>5</sub>. 9. "I<sub>2</sub> said<sub>1</sub> (it)," Jiré<sub>4</sub> said<sub>3</sub>. 10. "We<sub>2</sub> aren't fighting<sub>1</sub> with<sub>4</sub> you<sub>3</sub>." 11. "Where<sub>1</sub> do<sub>3</sub> you<sub>2</sub> live<sub>3</sub>?"

(he') said<sub>4</sub>. 12. "We<sub>2</sub> live<sub>3</sub> a long ways away<sub>1</sub>." 13. "Take<sub>2</sub> us<sub>1</sub> right<sub>5</sub> to<sub>4</sub> the houses<sub>3</sub>. 14. We want to see<sub>1</sub> your houses<sub>2</sub>," we said<sub>3</sub>.

## FOOTNOTES

<sup>1</sup>The Parintintin language is spoken by approximately one hundred and fifty Indians living principally in the State of Amazonas, Brazil. There are groups living at three places on the Ipixuna River, a tributary of the Madeira; others are living in the area of Três Casas; and a group of some seventy Indians are near the headwaters of the Marmelos River. The family group living at Canavial, one of the above-mentioned locations on the Ipixuna River, consists of the chief, Paulino Neves; his immediate family; and his sons-in-law.

Most of the data for this paper were collected during two field trips to Canavial. The first visit extended from February through April, 1961, and the second from July, 1961 through January, 1962. Assisting steadily as informants were Catarina, the chief's wife, and Ida, his married daughter. Their approximate ages were forty-five and twenty-seven respectively.

We are deeply indebted to Dr. David Bendor-Samuel and to Eunice Burgess for their invaluable help in the preparation of this paper and to others of our colleagues in the Summer Institute of Linguistics, Inc. who have also assisted in various ways. Thanks are also due to the Museu Nacional of Rio de Janeiro under whose auspices the field work has been done, and to the Serviço de Proteção aos Índios for permission granted in making the field trips.

<sup>2</sup>This regressive nasalization does not occur on certain common morphemes which would be rendered homophonous with other morphemes if nasalized, e.g., *ŋa* [ga] 'he' contrasts with *ŋã* [ŋã] 'they'. Both of these may be prefixed with similar grammatical functions.

# NASALIZATION IN KAIWÁ

Carl H. Harrison and John M. Taylor

The purpose of this paper is to describe the feature of nasalization as it occurs in Kaiwá, a Tupi-Guarani language of Brazil.

An early tentative analysis of the language which posited an oral and nasal series of vowels with resulting oral and nasal syllables did not account for all of the complexities perceived. It was also noted that the question of the phonemic status of pure versus postoccluded nasals was related to the question of nasalization. Although at first it appeared that one could simply consider the pairs  $m-m^b$ ,  $n-n^d$ ,  $\tilde{n}-d^v$ ,  $\eta-\eta^g$ , and  $\eta^w-g^w$  as allophones (the first member of each pair occurring in nasal syllables, the second in oral syllables), subsequent research revealed further interesting relationships among the nasalized parts of words, and consequently the picture became less clear.

It was soon discovered that in many cases the nasal allophones of such consonants occurred occasionally with oral vowels, and vice-versa. In the words  $t\ddot{u}p\ddot{a}\eta^w\text{wasu}$  'important gods' and  $k\ddot{x}^?x\eta^w\text{wasu}$  'large peppers', the vowel  $a$  which follows  $\eta^w$  is unambiguously oral. This is in spite of the fact that, according to the analysis referred to above, this vowel should have been nasal, and so account for the selection of the allophone  $\eta^w$  (since  $\eta^w$  was to occur with nasal vowels and  $g^w$  with oral). In the example  $\ddot{o}m\ddot{a}n\ddot{o}m^b\text{a}-m\ddot{a}$  'they all died', it is impossible to decide which of the last two vowels, in actual fact, is the more nasalized, though the aforementioned analysis would call for the prejunctional  $a$  to be clearly oral. Sometimes, in a given instance, it is quite difficult to label a particular syllable of a form as definitely either nasal or oral, as for example in:

$t\ddot{u}p\ddot{a} \sim t\ddot{u}p\ddot{a} \sim t\ddot{u}p\ddot{a}$  'gods, spirits'

$o\ddot{k}\ddot{e} \sim \ddot{o}k\ddot{e} \sim \ddot{o}k\ddot{e}$  'opening'

$\ddot{o}r\ddot{o}i\ddot{k}\ddot{o}t\ddot{e}v\ddot{e} \sim \ddot{o}r\ddot{o}i\ddot{k}\ddot{o}t\ddot{e}v\ddot{e} \sim \text{oroik}\ddot{o}t\ddot{e}v\ddot{e}$  'we (exclusive)  
are in need'

Yet we can make the generalization that at least one of the

syllables of (certain parts of) the words will be nasalized in a given instance. This leads us to the conclusion that nasalization needs to be described in terms of something other than the syllable. A hypothesis which attributes intrinsic nasality to certain whole morphemes in the language turns out to be the most fruitful, although certain other qualifications must be added in order to develop a truly comprehensive descriptive system for nasalization in Kaiwá. Throughout the rest of this paper, underlining will indicate intrinsically nasal morphemes. A til over a vowel is used to mark nasalization which results from the application of certain rules about combinations of morphemes at least one of which is nasal, as well as to mark nasalization caused by other factors. The examples cited above could then be written tupa, oke, and õrõ ikoteve (with space marking the morpheme boundary). In the last case õrõ ikoteve would indicate that oro has become nasalized by association with ikoteve.

Certain other definitions are necessary before the rules can be given. A STRESS GROUP is a group of syllables which can occur immediately before and after a syllable which has primary stress. PRIMARY STRESS is primary intensity with concomitant features of length of the whole syllable and rising pitch. A stress group has a nucleus and a margin (separated by a hyphen in this paper). We define NUCLEUS to be the syllable with primary stress and all syllables that precede it in a stress group. The MARGIN consists of any syllables that may follow the syllable which has primary stress. The division between stress groups, that is the point at which the margin leaves off and another nucleus begins, is determined by a combination of the following: potential pause; an audible rhythm change in which there is a relative speeding up of the first part of the nucleus in relation to the last part of the preceding margin; and syntactical divisions (the description of which is beyond the scope of this paper).

Positing the two units of nucleus and margin greatly facilitates the description of certain differences of effect of the nasality of one morpheme on that of another. It also helps in describing the limits to the operations of some but not all of the rules, whereas stress-group boundaries form the limit of all rules described here. Nasality, that is intrinsic nasalization, may occur on the nucleus, or the margin, or both, or neither. A margin is either completely nasal or completely oral. By contrast, a nucleus may be completely nasal, completely oral, or partly oral

and nasal.

Rule 1

Within the nucleus, if an intrinsically nasal morpheme occurs, it will potentially nasalize morphemes that precede it within the limits of the stress group boundary.

šě apij g<sup>w</sup>a-rupi

my nose cavity in 'in my nostril'

Note that e of še becomes nasalized.

ǒ ñě no-ta

he reflexive lay desiderative 'he intends to lie  
down'

Note that if the reflexive were not followed by a nasal morpheme, it would have the form d<sup>w</sup>e.

ǒ mano 'he died'

Qualification 1a

The occurrence of a glottal stop will tend to inhibit the nasalization effect of intrinsically nasal morphemes on successive syllables preceding it in the nucleus. Thus, although all the morphemes preceding an intrinsically nasal morpheme may become nasalized in normal speech, in careful (slower) speech syllables preceding a glottal stop will be oral (subject to Qualification 2a below).

base form: d<sup>w</sup>u<sup>?</sup>i ŋ<sup>w</sup>e 'frog (classifier)'

normal speech: ñũ<sup>?</sup>ĩ ŋ<sup>w</sup>e

slow speech: d<sup>w</sup>u<sup>?</sup>ĩ ŋ<sup>w</sup>e

Qualification 1b

If the last morpheme of a nucleus is intrinsically nasal, and if it is not followed by a margin, the last syllable(s) of that morpheme will generally be oral. If fol-

lowed by a margin, it (they) will be nasal.

base form: oro ikoteve

with margin: őrő ikōtēvē-vaʔe 'we (exclusive) are  
the ones who are in need'

without margin: őrő ikōtēve 'we (exclusive) are in  
need'

### Rule 2

The postoccluded nasals m<sup>b</sup>, n<sup>d</sup>, ŋ<sup>g</sup> (which occur only in nuclei) will nasalize all syllables which precede them in a nucleus in normal speech and often in slow speech as well.

tēm<sup>b</sup>iʔu 'food'

ő mō ñē m<sup>b</sup>o asi 'he makes him sorry for himself'

### Qualification 2a

In careful speech, nasalization will be inhibited before a glottal stop as in Qualification 1a. However, any nasal morpheme or postoccluded nasal which precedes the glottal stop will reinitiate the nasalization of preceding syllables according to Rules 1 and 2.

base form: še rem<sup>b</sup>iʔu ra 'my food in the making'

slow speech: šē rēm<sup>b</sup>iʔū ra (Rules 1 and 2; Quali-  
fication 2a)

normal speech: šē rēm<sup>l</sup>iʔū ra (Rule 1)

### Rule 3

A nasal margin does not affect an oral nucleus apart from a possible slight nasalization of the final vowel of the nucleus.

ő mano m<sup>b</sup>ā-ma 'they all died' (Rules 1 and 3)

Rule 4

Certain margins manifest nasal allomorphs if the last morpheme of the nucleus is nasal.

o ke-pi

his sleep in 'in his sleep'

oke-mĩ 'in the doorway' (Rule 4)

Examples are here given to illustrate various combinations of nasal and oral morphemes in nuclei and margins. Nasalization (written as before with a til) will be referred to one of the rules in each case. Some additional comments are made where relevant.

nãnde 'we (inclusive)' (Rule 2)

nãñẽ apeku 'our (inclusive) tongue' (Rule 1)

Note that full nasalization of nãnde forces the choice of the pure nasal as over against the postoccluded variant.

o ñẽ mō n<sup>d</sup>ij-ma 'he got a scare' (Rule 2)

hasi-maramo 'the moment they got sore'

There is no effect of margin nasalization on the nucleus.

õ mano-tama 'he is at the point of dying' (Rule 1)

o ke-ta 'he intends to sleep'

o me?e 'he gave' (base form)

one variant: o mẽ?ẽ is less common.

another: õ mẽ?e is more common (Qualification 1b).

another extreme variant: õ me?e occurs, and in such cases a slight postocclusion is sometimes heard on the nasal consonant: õ m<sup>(b)</sup>e?e.

ama 'rain'

<u>tupa</u>	'gods, spirits'
tupa	'lying place'
<u>oke</u>	'opening'
o ke	'he slept'
õ <u>ha'aro</u>	'he waits' (Rule 1)
<u>taj</u>	'teeth'
õm <sup>b</sup> o poši	'he makes angry' (Rule 2)
õrõ <u>ŋ<sup>w</sup>ahe-vi</u>	'as we (exclusive) arrive' (Rule 1)

In the example above, the syllable he is nasalized (Qualification 1b).

õrõ ŋ<sup>w</sup>ahe 'we (exclusive) arrive'

In this case the syllable he is often oral (Qualification 1b).

õrẽ rẽm<sup>b</sup>i'ʉu 'our (exclusive) food' (Rule 2)  
 tẽm<sup>b</sup>i'ʉu 'food' (Rule 2)

### Notes

<sup>1</sup>The Kaiwá analyzed here is that spoken at the Francisco Horta Post, Dourados, Mato Grosso, Brazil. Field work for this paper was done during 1961 and 1962 under the auspices of the Summer Institute of Linguistics in cooperation with the Museu Nacional do Rio de Janeiro.

The segmental types of Kaiwá are: p, t, k, k<sup>w</sup>, ʔ, m, m<sup>b</sup>, n, n<sup>d</sup>, ñ, d<sup>v</sup>, j, ŋ, ŋ<sup>g</sup>, ŋ<sup>w</sup>, g<sup>w</sup>, r, v, s, š, h, i, e, a, o, u, i. The segments k<sup>w</sup>, m<sup>b</sup>, n<sup>d</sup>, d<sup>v</sup>, ŋ<sup>g</sup>, ŋ<sup>w</sup>, and g<sup>w</sup> are interpreted to be simple unit phones on the basis of non-suspect syllable patterns. The question of the taxonomic phonemic status of such pairs as m-m<sup>b</sup> and n-n<sup>d</sup> is not considered in this paper.

# THE MORPHOPHONOLOGY OF ASURINI WORDS

Carl H. Harrison

1. Introduction
2. Symbols Used and the Form of Rules
3. Morphophonological Rules for Asurini Words
4. The Rules Applied to Selected Text

1. Introduction. The purpose of this paper is to present a number of important rules for representing the contrastive pronunciation of slow-speech Asurini words.<sup>1</sup> The form for presentation of the rules is related to that of synchronic item-and-process descriptions of this century in American descriptive linguistics. Certain notational conventions and the overall conceptual framework of transformational-generative linguistics are used.<sup>2</sup>

The form for individual rules is  $L \Rightarrow R$ , where L ranges over any substring of a string of a derivation under consideration.  $L \Rightarrow R$  is to be read "replace L by R", and R is a string of symbols.<sup>3</sup>

A rule of this form needs an input. In this paper, input will be given as text written morphophonemically. This text will be divided into sentences, bounded by space; and words, bounded by #. The words will be referred to in this paper as input strings.

The input strings are made up of lexical forms, or morphemes, written in a symbol system which carries both lexical and phonological information in most cases. In a few cases only lexical information is carried. Lexical information is conveyed by writing each dictionary entry with the same sequence of symbols every time it occurs in an input string. Thus, all phonological variations of a form are suppressed at this stage so that what is felt to be a morpheme or formative<sup>4</sup> is stored in the dictionary as one form. Phonological information is present in the following manner. The alphabet used to symbolize formatives in input strings conveys just that phonological structure which is common to all of the (phonological) variants of a given form.

A full set of redundancy rules as required in a generative phonology is not attempted here, although they could be formulated if they were felt to be within the scope of

this paper. Here we are basically concerned with both syntactically and phonologically conditioned alternation between, in the main, contrastive segment types corresponding in most cases to those types which would result from traditional segmental phonemic analysis.

A writing system, then, which includes lexical and phonological generalizations as outlined above satisfies our intuition that a form to which we assign the same (context determined) set of dictionary glosses in its various occurrences quite often has a substantial substratum of recurring phonological properties in all or most of these occurrences. Thus, the graphic representation of a lexical form in the input string symbolizes what is phonologically constant in that form, and the morphophonological rules of this paper account for a large part of the further regularities of pronunciation that are introduced by the context in which the form appears. As it turns out, the total number of segmental symbols needed to represent all lexical forms in the input strings is relatively small, supporting the generalization that linguistic communication depends on a small set of features and segment types for transmission of messages.

A number of intuitive constraints are put on the input alphabet. We have come to expect certain properties in phonological systems; among them, exploitation of a small number of articulatory and acoustic features, and preservation of certain kinds of syllable structure. An attempt has been made here to pay respect to such properties.

The author arrived at the input representations for the forms by what might be called synchronic reconstruction, a process which tries alternate writing systems until one appears that is useful for showing all of the generalizations that are felt to be possible for the system. The description then takes the form of a historical grammar where items from one stage (output) replace items for another (input).

Although the reconstructed morphophonological alphabet may have certain relationships to the phonological structure of earlier stages of the language, such problems are not considered in this work, though it is hoped that the results will prove useful to investigators working in related languages.

The actual phonological variants (allomorphs) of

lexical forms are written as strings of symbols from an output alphabet and the set of rules can be thought of as a theory about the relations between the two alphabets which may be modeled in Asurini slow speech.

The symbols of the output alphabet are taken from articulatory phonetics, one symbol per contrastive segment type, each representing a bundle of contrasting articulatory features. The input alphabet is somewhat larger, containing some ad hoc cover symbols for smaller sets of contrastive features as well as for possible null realization ( $\emptyset$ ).

The texts used as illustrations were elicited from speakers of Asurini. Neither noncontrastive phonetic detail of segments nor information about phonological units larger than the stress group or word (where such may contain more than one primary stress) is given.

2. Symbols Used and the Form of Rules. The symbols used in the rules are of the following types: segmental alphabet symbols, juncture marks, symbols for relations among segments, accent marks,  $\emptyset$ ,  $\infty$ , symbols for syntactic form classes, and one lexical symbol D which, as such, has no phonological content. The rules are ordered in such a fashion that the whole set is run through once for each word (#...#), with the added convention that upon arriving at a rule, that rule is to be applied as often as necessary for the word in question before moving on to the next rule. Ordering will be discussed in the presentation of the rules.

2.1. The following output alphabet reflects the general phonetic framework under which the field work was done. The symbol q is used for [k<sup>w</sup>]. Certain problems of non-uniqueness occur: c could have y as an allophone. Yet there is another rule in which initial i is slightly shortened before vowels, making it a candidate for representation as y and bringing it into quasi-contrast with c. Such assignments involve the type of arbitrary judgments which are not necessary to the understanding of contrastive word pronunciation. Because of the continuous nature of speech, assignments to discrete segment types are indeed necessary at least for purposes of analysis. Most token assignments (once a phonological system is known) are not arbitrary. Some, however, may be. We will assume that such choices as are necessary for the purposes of description have been

made, and that certain segment types which are similar (w to o; y to i) may both be instantiated in any single phonetic segment token, even though we have chosen to consistently symbolize them as one or the other for descriptive purposes. Thus the rule for I says that I is more i-like in some environments, y-like in others, though evidence could be marshalled to defend assigning certain instances to either.

An articulatory phonetic apparatus is used because, though not so economical or so universal in its contrastive segment specification as the acoustic distinctive feature description, it has the advantage of being established in linguistics as a tool for field work and description, and of not requiring sophisticated equipment to practice.<sup>5</sup>

The characteristics of the output alphabet as over against the input alphabet are that the elements of the output alphabet are uniquely specified by their features. The input alphabet has symbols for which some of these features are not specified. Such feature values are filled in by the rules. The purpose of the following matrix is simply to show the general universal phonetics under which the field work was accomplished, not to see to what minimum the actual features can be reduced, nor to show the complete structure of Asurini phonology. A numeral 1 in the matrix indicates that the feature at the head of the column is present in the segment which labels the row. 0 indicates that it is absent.

#### Consonants<sup>6</sup>

	stop	sib	nas	res	lab	alv	alp	vel	glo
p	1	0	0	0	1	0	0	0	0
t	1	0	0	0	0	1	0	0	0
k	1	0	0	0	0	0	0	1	0
q	1	0	0	0	1	0	0	1	0
ʔ	1	0	0	0	0	0	0	0	1
m	0	0	1	0	1	0	0	0	0
n	0	0	1	0	0	1	0	0	0

	stop	sib	nas	res	lab	alv	alp	vel	glo
ŋ	0	0	1	0	0	0	0	1	0
c	0	1	0	0	0	0	1	0	0
w	0	0	0	1	1	0	0	0	0
r	0	0	0	1	0	1	0	0	0
y	0	0	0	1	0	0	1	0	0
h	0	0	0	1	0	0	0	0	1

## Vowels

	front	central	back	high	low
i	1	0	0	1	0
e	1	0	0	0	1
a	0	1	0	0	1
o	0	0	1	1	0
ɨ	0	1	0	1	0

2.2. Included in the input alphabet is the output alphabet. When symbols from the input alphabet appear in the input string, unless they are enclosed in parentheses (and hence disappear in some environments) they are carried through the rules without being replaced. Along with those symbols there are certain upper-case letters. Such symbols have at least one articulatory feature that is unspecified. In the following matrix those features that are unspecified for any symbol are marked with an empty pair of parentheses.

Notice that there are at least two unspecified features in each case. The purpose of the rules is to give the kinds of environments that will evaluate the features one way or the other. Any feature columns from the previous matrix that do not appear in this one are all 0 for the elements in this matrix. I have chosen to use cover symbols instead of feature matrices in the rules. Thus we retain readability at the cost of increasing our alphabet slightly. In the following matrix, 1 and 0 have the same

values as in the preceding matrix; V means 'vowel'; and C means 'consonant'. The rules disambiguate pairs or triples of mutually incompatible parentheses in any row.

Note that for Ń there is no corresponding resonant. This represents a lacuna in the phonological pattern. An ŋ fills in for the missing resonant in the pattern of alternation.

An upper-case symbol, then, can be thought of as a variable to which the rules assign an appropriate constant or lower case symbol for any given environment.

	C	V	lab	alv	vel	stop	nas	res
P	1	0	1	0	0	( )	( )	0
T	1	0	0	1	0	( )	( )	0
K	1	0	0	0	1	( )	( )	0
M	1	0	1	0	0	( )	( )	( )
N	1	0	0	1	0	( )	( )	( )
Ń	1	0	0	0	1	( )	( )	( )

	C	V	high	front	central	low	back
I	( )	( )	1	1	0	0	0
O	( )	( )	0	0	0	1	1
A	0	1	( )	0	1	( )	0

	alp	resonant	sibilant	C
Y	1	( )	( )	1

2.3. A part of a form which is enclosed in parentheses in the input strings stands for a phonological segment (or segments) which is lost (or zeroed, or has null realization) in certain environments, retained in others. The parentheses assert that, for whatever historical reasons, the enclosed segment is different from its corresponding segment without parentheses. Such a convention was chosen to

contrast with upper-case symbols since in the case of parenthesized elements a different type of phonological relationship is involved, namely, alternation with  $\emptyset$ .

Along with the alphabets, certain other upper case symbols are used.<sup>7</sup>

C is a variable for any consonant. It is used to mark the place in a substring where any one of the following occurs: p,t,k,q,ʔ,m,n,ŋ,c,w,r,y,h,P,T,K,M,N,Ń,Y plus any of these enclosed in ( ). This symbol is then an abbreviation for a number of rules, one for each consonant.

V is a variable for any vowel and functions in the same manner as C in the rules. The vowels are i,e,ɪ,a,o,I,O,A.

S stands for a syllable in the string. It is used in the rules that indicate accent placement. In order to determine accent, one needs to know how many syllables there are, which ones have inherent accents, and the kinds and locations of junctures in the combinations. We will introduce the following convention for determining the number of syllables in a string. For any string, every vowel is a syllable center and can be assigned an S; consonants may be disregarded. Accent placement is then determined according to the inherent accent marks, syllable symbols, and juncture marks in the string by applying Accent Rules 24-29. Syllables in Asurini have the structure V, CV, VC<sub>1</sub>, and CVC<sub>1</sub>, where C<sub>1</sub> is a variable for m,n,ŋ, and y. There are no other restrictions or combinations, except that y never occurs word medially before V. Syllable boundary shifts occur to bring the structure of a string as closely as possible into a series of CV syllables. If the syntax gives a CVC<sub>1</sub> followed by a V, the structure imposed is CV.C<sub>1</sub>V. Some of the rules for null realization seem to reflect this pressure.

The following junctures are used. # marks the beginning and end of a word. // marks the coming together of two syntactical elements of the following types: (1) Two nouns joined to form a possessor-possessed relationship; (2) a noun stem followed by another noun stem in a (r)amo or (r)ire phrase. = indicates a complex stem juncture and - indicates all other boundaries between lexical forms. Junctures are important to the system, since a juncture will often form part or all of the relevant environment in which a certain replacement takes place. Although they

have no phonological feature content of their own, to formulate the rules without them would be a difficult job. Juncture symbols stand for grammatical boundaries which often bring about certain processes in segments which border on them. In fact, there is no single rule in the set which does not deal with the changes involving boundaries of forms. With the exception of Rule 13 which deals with an (apparently archaic) vowel harmony, and some accent rules, those changes take place right at the boundaries. It seems that syntactic information is important to understanding Asurini word pronunciation.<sup>8</sup>

The justification for the different kinds of juncture symbols is twofold. On the one hand, they represent different grammatical processes and could each be given a formal characterization in the syntax. On the other hand, each represents a distinction from the others in the phonological activity around it in at least one of the rules. No doubt there are more juncture types that would need to be recognized for higher-level phonology.

There are two inherent kinds of accent. ` is weaker, disappearing in a word when preceded by a strong accent. ' is strong inherent accent which is only zeroed in a word where it is followed by a strong accent, unless // juncture should come between them. Accent is phonetically an intensification of the signal in relation to surrounding syllables, plus a higher pitch, plus slight lengthening of the vowel and surrounding consonants.

A problem of interest in morphophonology is the determination of the minimal recurring part of all the possible environments in which a certain alternation occurs. Since the possible set of environments to be accounted for is potentially infinite by the very nature of the recursive processes within a language, a trial and error method based on ad hoc hypotheses has been used. These hypotheses come from a set of typological possibilities learned elsewhere. Two possible related goals concerning the choice of environments for rules giving slightly different results are worthy of note. The first is to state all and only the invariantly recurring elements of the set of possible environments for a process. The second is to state a minimum number of recurring elements in the set of all possible environments that it is necessary to state in order to distinguish that set from all the other environments where the process would not take place. In general, the second of these goals has been adopted. One case will illustrate.

-(t) is used as an environment symbol. However, -(t) only occurs in one form and hence is always followed by a. If the first goal were chosen, -(t)a would be used. Such a choice would involve a favoring of lexical conditioning for the process. Yet the choices are phonologically related, even if the environment itself is not phonological.<sup>9</sup> Hence, except in the case of Aff 1, Aff 2, and Aff 3 (explained below) where there is no doubt about the nature of the conditioning, it is hoped that something of the (possibly lost) phonological conditioning may be reflected. This is a notational extension of the implicit assumption under which linguists often operate, that if a process is describable as both grammatically and phonologically conditioned, the latter type is used in the description.

Invariant environments for the Asurini rules are of the following types:

- (1) Single parenthesized letters of the phonological alphabets are used. These constitute phonological conditioning for the most part.
- (2) Two or more concatenated symbols of the alphabets are used. Such a substring is often part of one and only one form and hence is not a strictly phonological environment, though historically it may have been.
- (3) Syllable structure dominates a number of the processes. It must be remembered that such symbols as C and V in a rule represent an abbreviation for a number of rules.
- (4) Junctures appear in all of the rules and can be thought of as orienting the rule so that it applies in the proper place in the string. In the sense that junctures mark syntactical boundaries, none of the rules represents a purely phonological process. A possible qualification would be to say that once a grammatical boundary is established, such and such a process is phonological.
- (5) The symbol  $\infty$  is used to indicate that other environments have been mentioned in the same rule, and that if the string being scanned contains the symbol to be replaced as specified in the rule but does not match any of the other environment specifications in the rule, then the part of the rule containing  $\infty$  is to be applied.
- (6) When the name for a class of lexical forms appears, the conditioning is clearly syntactic. The

classes are as follows. (Note that the following numerical notation is economical for specifying the dictionary meanings of these forms. It will be used throughout the remainder of the paper.)

- 1 first person singular
- 2 second person singular
- 12 first person plural, inclusive of addressee
- 13 first person plural, exclusive of addressee
- 22 second person plural
- 3(3) third person, singular or plural

Aff 1 (affix class 1)

ce- '1', ne- '2', cene- '12', ore- '13',  
pe- '22', I- '3(3)'.

Aff 2

we- '1', e- '2', cere- '12', ore- '13',  
pece- '22', O- '3(3)'.

To the person and number indication for this last set should be added 'reflexive orientation', roughly translated 'my own', 'your own', etc.

2.4. The Domain and Form of the Rules. Single rules may have the form

$$XAY \Rightarrow XBY^{10}$$

where X and Y stand for any substrings which remain invariant for the process and A stands for the symbols to be replaced by B. If nothing is supplied for X and Y in the rule, it is understood that it is indifferent for the rule what occurs in those positions. A is null in one rule, B is null in a number of rules.  $\emptyset$  is used for null B. Null realization means that no segments occur in the position indicated in the sequence.

If two rules have one or more structurally analagous parts in common, they may be put together into one rule by using the symbol(s) in common only once and using brackets in such a fashion that the other relationships are not lost. Given the rules

$$AX \Rightarrow BX$$

$$AY \Rightarrow CY$$

we can conflate them in the following fashion

$$A \begin{bmatrix} X \\ Y \end{bmatrix} \Rightarrow \begin{bmatrix} BX \\ CY \end{bmatrix}$$

The brackets indicate that if the particular substring of the string of the derivation under consideration is of the form AX, one replaces it by BX and if the substring has the form AY it is to be replaced by CY.

Given the rules

$$AX \Rightarrow BX$$

$$AY \Rightarrow EY$$

$$CX \Rightarrow DX$$

$$CY \Rightarrow FY$$

we can use subscripted brackets to reduce them to

$$\begin{bmatrix} A \\ C \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \Rightarrow \begin{bmatrix} B \\ D \\ E \\ F \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

$\begin{matrix} 1 & 1 & 2 & 2 \\ & 1 & 1 & \\ & & 1 & 1 \\ & & & 2 \end{matrix}$

The motivation for the use of such brackets is not necessarily economy since the number of letters lost is often offset by the number of brackets and numbers gained. The important gain is in being able to put what are intuitively the same kinds of generalizations into one rule.

The symbol  $\emptyset$  occurring in a rule is an instruction to delete its corresponding element on the left side.

An alternative form for the presentation of context-dependent replacement rules has been used in other phonologies. Thus a rule such as

$$A \begin{bmatrix} X \\ Y \end{bmatrix} \Rightarrow \begin{bmatrix} BX \\ CY \end{bmatrix}$$

has often been given as

$$A \rightarrow \left\{ \begin{array}{l} B/_Y \\ C/_Z \end{array} \right\}$$

In the case of  $C/_Z$ , if a convention is allowed that the parts are ordered,  $_Z$  can be omitted since the last replacing symbol is understood to mean in the environment 'elsewhere', and need not be specified. The choice of rule form is a matter of taste. Braces are often kept for abbreviating context free rules such as

$$A \rightarrow \left\{ \begin{array}{l} B \\ C \end{array} \right\}$$

where either B or C may be chosen without regard to context. Brackets, on the other hand, are usually reserved for context-sensitive rules or transformations. Such is the motivation for the choice of form in this paper.

3. The Morphophonological Rules. In this section, the rules are presented with a discussion and examples for each. Words used for examples under any rule will be given to the left of the arrow with all of the changes they have undergone as a result of the application of preceding rules. To the right of the arrow, they will show the result of the application of the rule under discussion. Remaining changes for any example can be discovered by applying subsequent rules in order.

The segmental rules (1-23), in all but two cases (#1 for epenthesis and #13 for vowel harmony), contain either an upper case letter or a parenthesized segment or both. They can thus be grouped into six distinctive sets:

- (1) epenthesis: 1

(2) vowel harmony: 13

(3) possible null realization: 2,3,6,7,9,11, (part of 12),15,17,18,19,21.

(4) possible realization by more than one lower-case element: 4,5,8,10,12,14,20,22,23.

(5) accent rules: 24-29

(6) a rule for a lexical element with two unrelated forms: 16.

A full range of applications of the rules is also presented in Section 4 by giving texts in their input and output form. In some cases full or partial derivations are given to illustrate rule application procedures.

### Rule 1

$$I - \begin{bmatrix} (o) \\ i \end{bmatrix} \Rightarrow I - c - \begin{bmatrix} o \\ i \end{bmatrix}$$

This is a rule for epenthesis.

Examples:

$$\#I-iwirápaN-A\# \Rightarrow \#I-c-iwirápaN-A\#$$

3(3)-bow-noun marker

'his/her/their bow(s)'

$$\#I-(\delta)pe\# \Rightarrow \#I-c-\delta pe\#$$

3(3)-to/at

### Rule 2

$$\begin{bmatrix} \text{Aff 1} \\ \infty \end{bmatrix} - (o) \Rightarrow \begin{bmatrix} \text{Aff 1} - o \\ \infty - \emptyset \end{bmatrix}$$

This describes a certain lexically conditioned alternation.

Examples:

#ce-(ð)pe# ⇒ #ce-ðpe#

1-to/for

#ʔi-A-(ð)pe# ⇒ #ʔi-A<sup>h</sup>pe#

water-n-to/for

Rule 3

$$\begin{bmatrix} V \\ C \end{bmatrix} - A - pe \Rightarrow \begin{bmatrix} V - \emptyset - pe \\ C - A - pe \end{bmatrix}$$

Words are characteristically marked with -A when functioning syntactically as nouns. They lose this when joining with other stems to form complex noun or verb stems. However, there is also one surface instance where A is lost. That is before the post-positional -(o)pe, when the last segment of the preceding stem is a vowel.

Example:

#ʔi-A<sup>h</sup>pe# ⇒ #ʔi<sup>h</sup>pe#

water-n-to

Note the ordering restriction with Rule 2. Were this rule to precede 2 the environment would have to be -(o)pe.

Rule 4

$$\begin{bmatrix} e \\ \infty \end{bmatrix} - A \Rightarrow \begin{bmatrix} e - e \\ \infty - a \end{bmatrix}$$

The only occurrences of the morphophoneme A are in occurrences of the noun marker mentioned previously. This form does not carry its own innate accent.

Examples:

#máʔe-A# ⇒ #máʔe-e#

something-n

#coówi-A# ⇒ #coówi-a#

grass-n

Rule 5

I - (e) ⇒ h - (e)

Rule 5 involves the commitative -(e)ro which has an idiosyncratic effect on I.

Example:

#I-(e)ro'ʔaN-D# ⇒ #h-(e)ro'ʔaN-D#

3(3)-com-fall-dependent

Rule 6

$$(V) - \begin{bmatrix} V \\ C \end{bmatrix} \Rightarrow \begin{bmatrix} \emptyset - V \\ V - C \end{bmatrix}$$

Parenthesized elements in general seem to come from vestiges of pressure to preserve acceptable syllable structure, although it is often the case that non-application of such a rule as this would give a syllable structure acceptable to Asurini as spoken now. A sequence such as na-V would not violate the norms.

Examples:

#n(a)-O'KeN-ih# ⇒ #n-O'KeN-ih#

negative-3(3)-sleep-negative

#t(e)-pe'ʔo# ⇒ #te-pe'ʔo#

in order for-22/3(3)-eat

'in order for you all to eat it'

Rule 7

e - (e) ⇒ e - ∅

This rule is an exception to the general Rule 9. Therefore the ordering of this rule with respect to 9 is important.

Example:

#cene-(e)ro'(to)N-D#  $\Rightarrow$  #cene-ro'(to)N-D#  
 x/12-committative-come-dependent verb marker

### Rule 8

$$C - \begin{bmatrix} I \\ 0 \end{bmatrix} \Rightarrow C - \begin{bmatrix} i \\ o \end{bmatrix}$$

This rule is partially repeated in Rule 23 for convenience since we are concerned here with activity near the beginning of words and there with the end. This repetition saves us a few ordering problems. The C referred to in the rule are the ones given to us by Rule 6.

Example:

n(a)-I  $\Rightarrow$  n-I

This is clearly a syllable structure rule. It is importantly ordered with respect to 10 since 10 under certain conditions gives y and w for I and O. This ordering reflects the fact that the previous segment dominates.

Example:

#n-o'ʔaN-ihī#  $\Rightarrow$  #n-o'ʔaN-ihī#  
 neg-3(3)-fall-neg

### Rule 9

$$\begin{bmatrix} \begin{bmatrix} - \\ \# \\ 1 \end{bmatrix} & CV \\ \infty & \\ 2 & 2 \end{bmatrix} - (e) \Rightarrow \begin{bmatrix} \begin{bmatrix} - \\ \# \\ 1 \end{bmatrix} & CV - \emptyset \\ \infty & - e \\ 2 & 2 \end{bmatrix}$$

Examples:

#pe-(e)ro-áta#  $\Rightarrow$  #pe-ro-áta#

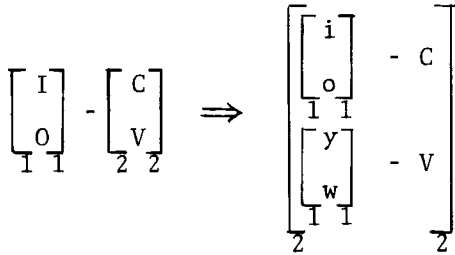
22-commitative-walk

#oro-(e)ro-áta#  $\Rightarrow$  #oro-ero-áta#

13-com-walk

Notice the absence of pre-CV juncture in the last case.

Rule 10



This accounts for regular shortening and heightening of third person prefixes before vowels.

Examples:

#O-áPaM#  $\Rightarrow$  #o-áPaM#

3(3)-finished

#O-áta#  $\Rightarrow$  #w-áta#

3(3)-walked

#I-áPaM-D#  $\Rightarrow$  #i-áPaM-D#

3(3)-finish-dep

#I-áta-I#  $\Rightarrow$  #y-áta-I#

3(3)-walk-post relational marker

Rule 11

$$\left[ \begin{array}{c} \text{Aff 3} \\ \infty \end{array} \right] - (\text{to}) \Rightarrow \left[ \begin{array}{c} \text{Aff 3} - \text{to} \\ \infty - \emptyset \end{array} \right] \quad (\text{Aff 2} - \text{Aff 3} \\ = \text{o-})$$

The verb root  $-(\text{to})\text{N}^{11}$  'to come' is subject to an irregularity not found elsewhere: syntactically conditioned loss of part of the form (after certain affixes). The parenthetic statement is to be read: Aff 3 is a subset of Aff 2 which contains all of the elements of Aff 2 except o-.

Examples:

#we'-(to)N-D#  $\Rightarrow$  #we'toN-D# (eventually wétota)

1-come-dependent marker

#o'-(to)N#  $\Rightarrow$  #o'N# (eventually ón)

3-come

Rule 12

$$\left[ \begin{array}{c} \# \\ \text{Aff 2-} \\ \text{i-} \\ \text{pe-} \\ \text{o-} \\ \text{mo-} \\ \infty \\ // \end{array} \right] \quad \text{N} \Rightarrow \quad \left[ \begin{array}{c} \#t \\ \text{Aff 2-t} \\ \emptyset\text{-h} \\ \text{pe-n} \\ \text{o-}\emptyset \\ \text{w-}\emptyset \\ \infty\text{-r} \\ //r \end{array} \right]$$

Aff 2 is discussed in Section 2. The collapsing of i-N to h may be a slightly artificial solution since it is not clear where the h comes from. It does seem that it is related to the h in Rule 5. We could formulate this rule so that h would replace i and N would go to  $\emptyset$ . The choice is arbitrary.<sup>12</sup>

## Examples:

#konómi-a//Néha-a#  $\Rightarrow$  #konómi-a//réha-a#

boy-n//eye-n

#we-Néha-a#  $\Rightarrow$  #we-téha-a# (Aff 2)

1 (own)-eye-n

#Néha-a#  $\Rightarrow$  #téha-a#

eye-n

#i-Néha-a#  $\Rightarrow$  #héha-a#

3(3)-eye-n

#pe-Néha-a#  $\Rightarrow$  #pe-néha-a#

22-eye-n

#ce-Néha-a#  $\Rightarrow$  #ce-réha-a#

1-eye-n

#o-NópiN#  $\Rightarrow$  #O-ópiN#

3(3)-lift

#o-mo-Náhi#  $\Rightarrow$  #o-w-áhi#

3(3)-causative-hurt

Rule 13

$$\left[ \begin{array}{c} \text{ero} \\ \infty \\ \text{I} \quad \text{I} \end{array} \right] - \left[ \begin{array}{c} \text{ha} \\ \text{ka} \\ \text{2} \quad \text{2} \end{array} \right] \Rightarrow \left[ \begin{array}{c} \left[ \begin{array}{c} \text{era} \\ \infty \\ \text{I} \quad \text{I} \end{array} \right] - \text{ha} \\ \left[ \begin{array}{c} \text{ere} \\ \infty \\ \text{I} \quad \text{I} \end{array} \right] - \text{ka} \\ \text{2} \quad \text{2} \end{array} \right]$$

This is a rule for two verbs whose paradigms are irregular,

possibly because they are among the most commonly used forms and hence more resistant to analogic levelling: -ha 'go', -ka 'to be present, now'. Several syntactic processes place o before the forms mentioned.

Examples:

#o'ha#  $\Rightarrow$  #a'ha#

3(3)-go

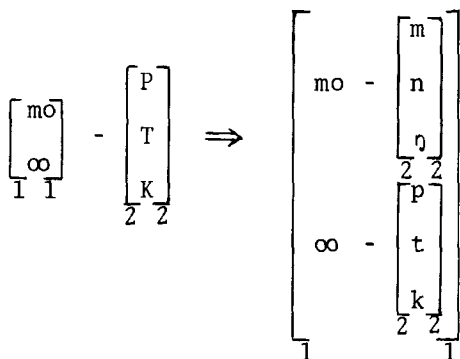
#w-ero'ha#  $\Rightarrow$  #w-era'ha#

3(3)-committative-go

#h-ero'ka#  $\Rightarrow$  #h-ere'ka#

3(3)-com-be present

#### Rule 14



Certain verb stems which under most conditions begin with stops, begin with corresponding nasals when preceded by mo-'causative'.

Examples:

#o-mo'KeN#      #o-mo'ŋeN#

3(3)/3(3)-causative-sleep

'she put him to sleep (usually said of a child)'

#o'keN#  $\Rightarrow$  #o'keN#

3(3)-sleep

### Rule 15

$C_3 = \Rightarrow \emptyset = (C - C_3 = y)$

This rule is concerned with compound or complex stem formation and is one of the justifications for establishing = juncture. (The other is related to accent.) It states that when, and only when, two stems are joined, if the first ends in a consonant other than y, and the second begins in a consonant, the first consonant will drop. The parenthetic notation means that  $C_3$  contains the same elements as C except for y. In a complete grammar this juncture mark would be inserted by the syntax.

Examples:

#o-manáhaN≠paM#  $\Rightarrow$  #o-manáha≠paM#

3(3)/3(3)-cut-finish

#o'kay≠paM#  $\Rightarrow$  #o'kay≠paM#

3(3)-burn-finish

Reduplication of stems brings this rule into effect if the stem ends in a consonant other than y and the stem has at least two consonants.

Example:

#i'qaN≠qaN#  $\Rightarrow$  #I'qa≠qaN#

3(3)-has several holes

### Rule 16

$$\begin{bmatrix} C \\ V \end{bmatrix} - \underline{D} \Rightarrow \begin{bmatrix} C - (t)a \\ V - w \end{bmatrix}$$

This rule contains the lexemic symbol discussed in Section 2. The symbol (D) refers to no phonological content as

such but stands for two phonologically unrelated forms which are clearly assignable to the same dictionary entry. It is a grammatical marker of dependency, a verb suffix whose occurrence relates the embedded clause in which the verb occurs to the sentence or paragraph in which the clause occurs. Although the variants are phonologically unrelated, their occurrences are phonologically conditioned.<sup>13</sup>

Examples:

#a'ha-D# ⇒ #a'ha-w#

3(3)-go-dependent

#o'keN-D# ⇒ #o'keN-(t)a#

3(3)-sleep-dep

Rule 17

$$\begin{array}{c} \left[ \begin{array}{c} (N) \\ (N) \end{array} \right] - \left[ \begin{array}{c} \# \\ \infty \end{array} \right] \Rightarrow \left[ \begin{array}{c} \left[ \begin{array}{c} \emptyset \\ \emptyset \end{array} \right] - \# \\ \left[ \begin{array}{c} N \\ N \end{array} \right] - \infty \end{array} \right]$$

$\begin{array}{cc} \text{I} & \text{I} \\ \text{2} & \text{2} \end{array}$

This rule is concerned with a small number of stems that drop final consonants in word final positions.

Examples:

#o-píhi(N)# ⇒ #o-píhi#

3(3)/3(3)-catch

#o-píhi(N)-(r)áme# ⇒ #o-píhiN-(r)áme#

-continuative

The reader is referred to Rule 20 for a discussion of the verb -pótaN 'like, desire' and the suffix -(pó)ta(N) 'future, desiderative'.

Rule 18

$$\begin{bmatrix} C \\ V \end{bmatrix} - (r) \Rightarrow \begin{bmatrix} C - \emptyset \\ V - r \end{bmatrix}$$

Certain suffixes begin with the segment -(r). This is one of the eleven segmental rules where syllable structure plays a part.

Examples:

$$\#a'ʔaN-(r)áme\# \Rightarrow \#a'ʔaN-áme\#$$

1-fall-now (eventually aʔaráme)

$$\#o'ʔo-(r)áme\# \Rightarrow \#o'ʔo-ráme\#$$

3(3)/3(3)-eat-now

Rule 19

$$\begin{bmatrix} C \\ V \end{bmatrix}_{\substack{1 \\ 1}} - \begin{bmatrix} (po) \\ (c) \end{bmatrix}_{\substack{2 \\ 2}} \Rightarrow \begin{bmatrix} C - \begin{bmatrix} \emptyset \\ \emptyset \end{bmatrix}_{\substack{2 \\ 2}} \\ V - \begin{bmatrix} po \\ c \end{bmatrix}_{\substack{2 \\ 2}} \end{bmatrix}_{\substack{1 \\ 1}}$$

There is only one form in the language which contains (po). It is the suffix -(pó)ta(N) 'future, desiderative', appearing often in text. As one can see by the base form, it has the following variants:

potan potat potar pota

tan tat tar ta

(c) occurs in a small number of suffixes.

Examples:

$$\#o'keN-(pó)ta\# \Rightarrow \#o'keN'ta\#$$

3(3)-sleep-future

#a'ha-(pó)ta#  $\Rightarrow$  #a'ha-póta#

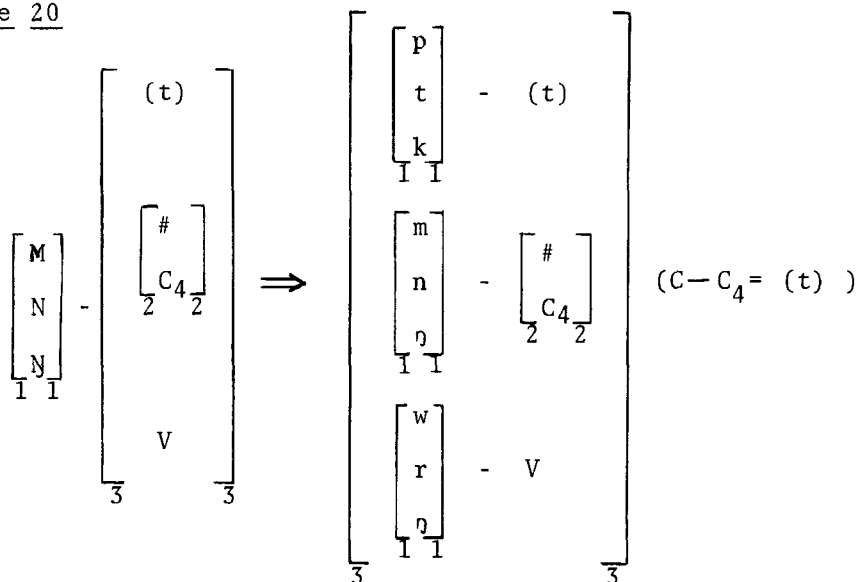
1-go-future

#o'keN-(c)awa#  $\Rightarrow$  #o'keN-awa#

3-sleep-perhaps

#a'ha-(c)awa#  $\Rightarrow$  #a'ha-cawa#

Rule 20



The parenthetic notation indicates that  $C_4$  has the same elements as C except for (t) which belongs only to C. A large number of stems end in one of the three upper-case nasals, and hence the rule will appear often in actual derivations. A real problem in ordering occurs here. If we allow the rule for deletion of (t) to precede this one, we are left with the vowel that follows it as an environment which would then invalidate the generalization we wish to make about the nature of change before -V. If we put the (t) rule after this one we must indicate that it is not to be considered one of the consonants of the language for this rule. Clearly the conditioning is grammatical for p, t, k and phonological for the rest (with the exception of the environment #). Hence ordering problems are a result

of the attempt to present fundamentally different kinds of conditioning in the same set of rules. Such a practice is justified, however, by the striking systematic phonological relations that obtain among the variants. The environment \_\_\_-(t) is quasi-phonological but basically lexico-syntactic.

Examples:

#o'paM-(t)a# ⇒ #o'pap-(t)a#

3(3)-finish-dep

#o'paM# ⇒ #o'pam#

#o'paM'ta# ⇒ #o'pam'ta#

-future, desiderative

#o'paM-áme# ⇒ #o'paw-áme#

-now

#o'keN-(t)a# ⇒ #o'ket-(t)a#

-sleep-

#o'keN# ⇒ #o'ken#

#o'keN'ta# ⇒ #o'ken'ta#

#o'keN-áme# ⇒ #o'ker-áme#

#i'ʔaŋ-(t)a# ⇒ #i'ʔak-(t)a

x/3(3)-dig up-dep 'someone then dug it (them) up'

#o'ʔaŋ# ⇒ #o'ʔaŋ#

3(3)/3(3)-dig up

#o'ʔaŋ-áme# ⇒ #o'ʔaŋ-áme#

The verb -potaN 'desire, like', related to the suffix -(pó)ta(N), retains the final consonant in all cases, hence is not subject to Rule 17 as is the suffix, but follows the alternations described in Rule 20.

#o-pótaN# ⇒ #o-pótan#

#i-pótaN-(t)a#  $\Rightarrow$  #i-pótat-(t)a#

#o-pótaN-áme#  $\Rightarrow$  #o-pótar-áme#

Rule 21

$$\begin{bmatrix} Y \\ C_3 \end{bmatrix} - (t) \Rightarrow \begin{bmatrix} Y - \bar{t} \\ C_3 - \emptyset \end{bmatrix}$$

Note that  $C_3$  was used earlier in 15. In at least two types of instances  $y$  does not behave like a  $C$ . One possible solution is to create another classification such as semivowel. The alternative is to do as we have done and recognize that under certain conditions its vocality asserts itself. One should not fail to notice that the symbol  $(t)$  was created for the case  $y$ -. The only other justification is the fact that when  $\bar{(t)}$  is used as an environment, it may be thought of as imparting its stop quality to the preceding  $M$ ,  $N$  and  $\bar{N}$  (see Rule 20).

Examples:

#o'kaY-(t)a#  $\Rightarrow$  #o'kay-ta#

3(3)-burn-dep

#i'ʔak-(t)a#  $\Rightarrow$  #i-ʔák-a#

x/3(3)-dig up-dep

Rule 22

$$Y - \begin{bmatrix} \begin{bmatrix} C \\ \# \end{bmatrix} \\ 2 \quad 2 \\ V \\ 1 \quad 1 \end{bmatrix} \Rightarrow \begin{bmatrix} y - \begin{bmatrix} C \\ \# \end{bmatrix} \\ 2 \quad 2 \\ c - V \\ 1 \quad 1 \end{bmatrix}$$

This describes an interesting alternation which is not characteristic of any other pair in the language. Were it not for the  $y$ -like quality of prevocalic  $I$ , one would be tempted to call  $c$  and  $y$  allophones in phonemic terminology.

Examples:

#o'kaY# ⇒ #o'kay#

3(3)-burn

#o'kaY'ta# ⇒ #o'kay'ta#

-future

Note that (po) disappears after y, which means that in this case y is behaving like a consonant environment (see Rule 20).

#o'kaY-óho# ⇒ #o'kac-óho#

-fiercely

Rule 23

$$\begin{bmatrix} C \\ V \end{bmatrix} - I \Rightarrow \begin{bmatrix} C - i \\ V - y \end{bmatrix}$$

See Rule 8 for this. It is repeated here to solve ordering problems. The purpose is to take care of two instances of the shortening of i in suffixes. One is -Ihi 'negative' and the other is -I, a grammatical agreement form indicating that either a relator word (adverb) or a postpositional phrase or both precede the main verb in the sentence.

Examples:

#i'ha-I# ⇒ #i'ha-y#

3(3)-go-agreement

#i'ker-I# ⇒ #i'ker-i#

3-sleep-agreement

The following rules are concerned with the placement of accent in words. Such rules must account for the fact that a word will only have one accent mark in the terminal string (unless there is no inherent accent in any of the constituent forms). The placement of that accent is a function of the inherent accents and their relationship to

each other and junctures.

The following conventions will be observed. In ...ś...ś#, three dots indicate that there may be from 0 to n unaccented syllables between the S's, and 0 to n syllables, accented or unaccented, before the first S (that is, to the left of it). In #...ś...S...# the three dots will indicate from 0 to n unaccented syllables.

Rule 24

$$\begin{bmatrix} \#S... \# \\ //S... \# \end{bmatrix} \Rightarrow \begin{bmatrix} \#\acute{S}... \# \\ //\acute{S}... \# \end{bmatrix}$$

Occasional single syllable forms may appear. Innate accent for all forms is on the penult. Hence a form with only one syllable is given by the dictionary in the form 'S. If such a form is given by the syntax directly following a # or // juncture, accent is progressive to the only syllable of the form.

Examples:

$$\#\acute{?}a\eta-a\# \Rightarrow \#\acute{?}á\eta-a\#$$

house-n

$$\#\acute{c}áwar-a//\acute{c}i-a\# \Rightarrow \#\acute{c}áwar-a//\acute{c}í-a\#$$

dog-n//nose-n

//juncture has been set up for the express purpose of accounting for the fact that in constructions such as noun possessing noun, accent remains on both nouns. Although a certain morphophonological change (Rule 12) does take place at this boundary which makes it distinct from # juncture, it does not allow the zeroing of accents which precede it in the construction as would = and - junctures (Rule 29).

Rule 25

$$S \neq S \Rightarrow \acute{S} = S$$

= juncture, though essentially affecting the bordering

segments differently than does - juncture, allows accent to regress to the final syllable of the preceding form.

Example:

#o-ápo≠pam# ⇒ #o-ápó=pam#

3(3)/3(3)-make/do-finish

### Rule 26

s ' s ⇒ ś - s

Accent behaves the same way here as in Rule 25 when the following morph consists of a single syllable.

Example:

#o'ken# ⇒ #ó-ken#

3(3)-sleep

### Rule 27

...ś...śś# ⇒ #...ś...śś#

Post-positionals which occur with weak innate accent lose their accent if preceded in the word by another accented syllable.

Example:

#ʔáŋ-a-pipe# ⇒ #ʔáŋ-a-pipe#

house-n-in

### Rule 28

#...śś# ⇒ #...śś#

The innate accent of a postpositional becomes the word accent if not preceded by an accented syllable elsewhere in the word.

Example:

#i-pipe# ⇒ #i-pípe#

3(3)-in

Rule 29

...ś...ś...  $\left[ \begin{array}{c} // \\ \# \end{array} \right] \Rightarrow \dots S \dots \acute{S} \dots \left[ \begin{array}{c} // \\ \# \end{array} \right]$

This is the rule for zeroing all remaining innate accents within a word but the last one. If there are more than two accents this rule is mechanically applied the necessary number of times, taking pairs of accents two by two, until only one accent is left between the pairs of juncture marks #...//, //...# and #...#, with the understanding that either // or # defines the limit of the operation. That is, it is understood that neither of these marks appears elsewhere in the string indicated to the left of the rule, except at the very left of the substring under consideration.

Examples:

#o-ápo-póta# ⇒ #o-apo-póta#

3(3)/3(3)-do-will

#o-ápo-pám=tar-áme#

Rule 25 #o-ápo-pám=tar-áme#

Rule 26 #o-ápó-pám=tar-áme#

Rule 29 #o-ápó-pam=tar-áme#

Rule 29 #o-ápo-pam=tar-áme#

Rule 29 #o-apo-pam=tar-áme#

3(3)/3(3)-do-finish-fut-contin

#ce-rá?ir-a//raη-a#

1-son's-n//house-n

Rule 24 #ce-rá?ir-a//ráη-a# (no other change)

There is a further phonetic rule which is concerned with secondary accent. Once the primary accent has been established one can place secondary accent regressively on every other syllable. This involves phonetic detail which is automatic once primary accent is established, and for this reason it is not formalized here. Two interesting notes concerning secondary accent deserve mention. For finding the appropriate syllable structure of a combination, V?V is to be counted as one syllable. Also, secondary accent may accidentally fall on the same syllable as the original inherent accent.

Example:

#iwo?oetimakiŋŋawa# (secondary accent underlined)

soap-n

#### Rule 30<sup>14</sup>

Simply for purposes of easier reading we may apply this orthography rule for the deletion of lower level junctures.

$$\begin{bmatrix} - \\ = \end{bmatrix} \Rightarrow \begin{bmatrix} \emptyset \\ \emptyset \end{bmatrix}$$

#### 4. The Rules Applied to Selected Texts.

##### 4.1. Introduction to the Presentation of Selected Texts.

The purpose of this section is to give a broad sample of possible rule application combinations, as well as to give the reader some idea of other features of the language for their own sake. The texts presented were obtained from Asurini informants, and transcribed and analyzed with their help.

In some cases, complete or partial derivations are given. In other cases, the input form is given and the output form is given. The reader is free to compare the two in order to determine which rules have been applied.

With each sentence in the text the following information is presented:

- (1) Given with each form is a gloss directly above it

so as not to interfere with the continuity of the derivation below it. Numerals are used for persons as explained before. A slant line with numerals on both sides indicates that the numeral to the left of the line is the person of actor and the numeral to the right is the person of goal of the verb. Such a combination also indicates that the verb is transitive.

Examples:

1/3(3) 'I acted upon him, her, it, them'

x/3 'someone (unspecified) acted upon me'

22/13 'you (plural) acted upon us (myself and another)'

(2) Above the line containing dictionary glosses there will be a line with labels for word or phrase types. Such a labelling represents a surface analysis of the sentences.

Symbols:

I: independent	part: particle
D: dependent	(r)ire phrase: phrase ending in
i: intransitive	(r)ire
t: transitive	(r)amo phrase: phrase ending in
N: noun phrase	(r)amo
V: verb phrase	Dsc: descriptor
Aux: auxiliary verb	dep: dependence marker
com: commitative	PP: post positional
n: noun marker	Aff: affix

There is one further procedure that must be followed. Occasionally one may reach a point where it is possible to apply a single rule twice within one word. The rules are designed so that all possible applications of a rule within a word (# #) should be made before moving on to the next rule. Failure to observe this may result in an incorrect derivation. An example is given here to illustrate the convention.

```

#3-sleep-future-continuative#
input #O'KeN-(pó)ta(N)-(r)áme#
10 #o'KeN-(pó)ta(N)-(r)áme#
14 #o'keN-(pó)ta(N)-(r)áme#
17 #o'keN-(pó)taN-(r)áme#
18 #o'keN-(pó)taN-áme#
19 #o'keN'taN-áme#
20 #o'ken'taN-áme#
20 #o'ken'tar-áme#
26 #ó-ken'tar-áme#
26 #ó-kén-tar-áme#
29 #ó-ken-tar-áme#
29 #o-ken-tar-áme#
30
—
—
—
30 #okentaráme#

```

A full derivation for a sentence in this part of the grammar would begin with the terminal string of the syntax as the first line, then apply all of the rules in order, one at a time, to a word. It would then move on to the next word, and so forth for the whole sentence. In such a derivation, if a symbol is not being replaced by another through application of a rule, it is carried on to the next line of the derivation. Only one rule is applied on each line. The purpose of such a derivation-oriented rule system is to provide an explicit mechanical method for checking to see if the rules are correct.

To give full derivations for all examples would take too much space. Therefore, one full derivation is given for purposes of illustration and thereafter they will be progressively shortened. Some of the final examples will simply give the input and output for whole sentences, leaving the reader to observe the rules that have been applied.

In the following complete derivation, the number of the rule being applied to a line  $n$  which gives its results on line  $n + 1$  will be put to the left of line  $n + 1$ .

```

#IiV#Aux iV#
#3-walk-future#3-go#

#0-áta-(pó)ta(N)#O'ha#
10 #w-áta-(pó)ta(N)#O'ha#
17 #w-áta-(pó)ta#O'ha#
19 #w-áta-póta#O'ha#
29 #w-ata-póta#O'ha#
30 #wata-póta#O'ha#
30 #watapóta#O'ha#
10 #watapóta#o'ha#
13 #watapóta#a'ha
26 #watapóta#á-ha#
30 #watapóta#áha#

```

4.2. The Texts. In the following derivations, only the word in question will be carried to the following line. Some lines will show more than one rule application.

Text 1

1.1. #NP# ItV# AiV# DtV# DtV#

#grass-n#1/3-cut#i/now#X/3-cut-dep#X/3-rake-dep#

#coówi-A#a-manáhaŋ#we´ka# I-manáhaN-D# I-péhiŋ-D#

4,30 #coówia#

20,30 #amanáhaŋ#

26,30 #wéka#

10 #i-manáhaŋ-D#

16 #i-manáhaŋ-(t)a#

20 #i-manáhak-(t)a#

21 #i-manáhak-a#

30,30 #imanáhaka#

10 #i-péhiN-D#

16 #i-péhiN-(t)a#

20 #i-péhit-(t)a#

21,30,30 #ipéhita#

#coówia#amanáhaŋ#wéka#imanáhaka#ipéhita#

1.2. #ItV# ItV# Aux iV# DtV#

#1-rake#1-scrape#1-here#X/3-cut-dep#

#a-moíriŋ#a-péhiN#we´ka# I-manáhaŋ-D#

20 #A-moíriŋ#

20 #a-péhin#

26 #we´ka#

- 10 #i-manáhaŋ-D#
- 16 #i-manáhaŋ-(t)a#
- 20 #i-manáhak-(t)a#
- 21 #i-manáhak-a#
- 30,...,30. #amoíriŋ#apéhin#wéka#imanáhaka#
- 1.3. #ItV#Part#N#
- #1/3-cut-future#again#grass-n#
- #a-manahaŋ-(po)ta(N)#no#coowi-A#
- 17 #a-manáhaŋ-(pó)ta#
- 19 #a-manáhaŋ' ta#
- 20 #a-manáhaŋ' ta#
- 26 #a-manáháŋ-ta#
- 29 #a-manaháŋ-ta#
- 4 #no#coówi-a#
- 30... #amanaháŋta#no#coówia#
- 1.4. #Adv#(r)amo-phrase#IiV#Aff-PP#
- #there#sun-n-go down-when#I-leave-future#3-from#
- #pe#qaN-A-ha-(r)amo#a-pohin-(po)ta(N)#I-(o)hi#
- 24 #pé#
- 4 #qaN-a'ha-(r)amo#
- 18 #qaN-a'ha-ramo#
- 20 #qar-a'ha-ramo#
- 24,etc. #qár-a'ha-ramo#
- 29 #qar-á-ha-ramo#
- 30,- #qaráharamo#

17	#a-póhin-(pó)ta#
19	#a-póhin-´ta#
26	#a-póhín-ta#
29	#a-pohín-ta#
30,-	#apohínta#
1	#I-c-(ð)hi#
2	#I-c-ðhi#
10	#i-c-ðhi#
28,30,-	#icóhi#
	#DiV#DtV#N#
	#1-eat-dep#X/3-eat-dep#game meat#
	#we-káro-D#I´?o-D#ma?eá?a-A#
16	#we-káro-w#
30,-	#wekárow#
10	#i´?o-D#
16	#i´?o-w#
26	#í-?o-w#
30	#í?ow#
4	#ma?eá?e-a#
30	#ma?eá?aa#
	#wekárow#í?ow#ma?eá?aa#

In this section the sentences will be given in the following form:

Line 1: word types

Line 2: dictionary glosses

Line 3: input form

Line 4: output form (terminal string)

Line 5: numbers of rules applied for each word under that word.

1.5. #DiV#Part#

#I-bathe-fut-dep#also#

#we-cáhoŋ-(pó)ta(N)-D#we#

#wecahóŋtata#we#

#16,17,19,20,21,28,29,30...# — #

1.6 #DSC#Part#Part#

#I-hot-fut#emphatic#again#

#ce-pirifʔay-(pó)ta(ŋ)#ca#no#

#cepiriʔáyta#ca#no#

#17,19,28,29,30# — # — #

1.7. #(r)ire-Phrase#N-PP#IiV#

#X/3-cut-finish-when#soap-n-by means of#I-bathe-fut.#

#I-manáhaŋ<sup>4</sup>paM-(r)íre#iwoʔoetimakiŋáŋaM-A-po#

#imanahapawíre#iwoʔoetimakiŋáŋawa-po#

#10,15,18,20,28,29,30...# 4,20,30#

#IiV#

#I-bathe-future#

#a-cáhoŋ-(pó)ta(N)#

#acahóŋta#

#17,19,20,28,29,30#

Free translation of Text 1.

1. I was cutting grass over there. 2. I rake it and cut it. 3. I'm going to cut some more. 4. When the sun goes down there I'm going to quit and dine; I'll eat some meat. 5. I'll take a bath too. 6. I'll be very hot. 7. When I finish cutting (the grass), I'll take a bath with soap.

The following text will be given in this form:

Line 1: word/phrase names

Line 2: dictionary glosses

Line 3: input form

Line 4: output form

2.1. #Iiv#Aff=PP#DiV#N#

#I=talk#2-to#I-sit-dep#I-n#

#a-poroŋéta#ne-(ð)pe#we'toyn-D#íce-A#.

#aporoŋéta#neópe#wétoyna#ícee#.

2.2. #ItV#N#DiV#Aff=PP#

#orienter-2/3-write on#paper-n#2-sit-dep#1-facing#

#t(e)-ere-mo-cíwaŋ#pápe-A#e'ʔiyn-D#ce-(e)ro-wáke#

#teremocíwaŋ#pápee#é'iyina#cerowáke#

pápee < papel (Ptg.)

2.3. #IiV#Aff=PP#DiV#N-PP#

#orienter-1-look-orienter#3-at#I-sit-dep#2-write-  
-instr.-n-at#

#t(e)-a-máʔe-ne#I-Nèhe#we'toyn-D#ne-mo-cíwaŋ'(t)aM-A-  
-Nèhe#

#tamáʔene#héhe#wétoyna#nemociwákawarehe#

2.4. #ItV#DiV#N#DtV#

#2-write-write-fut#2-sit-dep#paper-n#X/3-write-dep#  
 #ere-kánaM=kánaM-(pó)ta(N)#e'ˈtoyn-D#pápe-A#I-mo-cíwaŋ-D#  
 #erekanakanámta#étoyna#pápee#imocíwaka#

note: toyn may go to ?iyn following e-.

2.5. #IiV#DiV#DiV#Aff=PP#  
 #I-sit-fut#I-sit-dep#I-look-dep#2-at#  
 #a-ápiŋ-(pó)ta(N)#we'ˈtoyn-D#wemá?e-D#ne-Nèhe#  
 #aapíŋta#wétoyna#wemá?ew#neréhe#

2.6. #IiV#Aff=PP#DiV#  
 #I-see-fut#2-at#I-sit-dep#  
 #a-má?e-(pó)ta(N)#ne-Nèhe#wé'ˈtoyn-D#  
 #ama?epóta#neréhe#wétoyna#

#### Text 2 Free Translation

1. I'm sitting talking to you. 2. You're writing on paper seated there across from me. 3. I'm sitting looking at it, at your pen. 4. You're writing and writing on the paper sitting there. 5. I'm going to sit down and look at you. 6. I'll sit here and look at you.

#### Text 3

1. I'm going to cut grass. 2. I'm going to cut just a little bit. 3. When the sun has only a little way to go down I'm going to leave it (quit) and take a bath. 4. That's all.

#### Text 4

1. I'm going hunting. 2. I'm going to kill a deer and bring it. 3. There is no meat around to eat. 4. Because of this (3) I'm going hunting to kill deer and wild pig and bring them back. 5. That's what I'm going to kill and bring.

- 3.1. #N#ITV#Aux iV#  
 #grass-n#i-cut-fut#I-go#  
 #coówi-A#a-manáhaŋ-(pó)ta(N)#we'ha#  
 #coówia#amanaháŋta#wéha#
- 3.2. #Adv#ItV#Aux iV#  
 #a little#I-cut-fut#I-go#  
 #pípi#a-manáhaŋ-(pó)ta(N)#we'ha#  
 #pípi#amanaháŋta#wéha#
- 3.3. #Dsc#IiV#Aff=PP#DiV#  
 #sun-small#I-leave-fut#3-from#I-bathe-dep#  
 #qaráhi-pípi#a-póhin-(pó)ta(N)#I-(ò)hi#we-cáhoŋ-D#  
 #qarahipípi#apohínta#icóhi#wecáhoka#
- 3.4. #IiV#  
 #3-finish#  
 #O'paM#  
 #ópam#
- 4.1. #IiV#Aux iV#DtV#  
 #I-walk-fut#I-go#X/3-kill-dep#  
 #a-áta-(pó)ta(N)#we'ha#I-cóka-D#  
 #aatapóta#wéha#icókaw#
- 4.2. #N#part#DtV#ItV#  
 #deer-n#indef.#X/3-com.-come-dep.#12/3-eat#  
 #mícaN-A#mo#I-(e)ro'(to)N-D#ca'ʔo#  
 #mícara#mo#hérota#cáʔo# ( (e)ro is not Aff 3)

## 4.3. #anohi#N#N#

#there is no#food#game meat#

#anóhi#kóme-A#ma?eá?a-A#

#anóhi#kómee#ma?eá?aa# kómee < Ptg. comer 'to eat'

## 4.4. #(r)amo Phrase#IiV#Aux iV#DtV#DtV#N#N#

#this-because of#I-go-fut#I-walk#X/3-com-come-dep#X/3-

kill-dep#deer-n#wild pig-n#

#á?e-(r)amo#a'ha-(pó)ta(N)#we-áta-D#I-(e)ro'(to)N-D#

#á?eramo#ahapóta#weátaw#hérota#

#I-cóka-D#mícaN-A#ciwá?a-A#

#icókaw#mícara#ciwá?aa#

## 4.5. #Adv#ItV#DtV#

#these 1/3-kill-fut#X/3-com-come-dep#

#eómi#a-cóka-(pó)ta(N)#I-(e)ro'(to)N-D#

#eómi#acókapóta#hérota#

Text 5

## 5.1. #Adv#N=PP#particle#IiV#particle#

#there#Trocara-n-to#rec-part#l2-go#again#

#qey#tokára-A-Nòpi#raka#ca'ha#no#

#qéy#tokáraaropi#raka#cáha#no#

## 5.2. #N=PP#part#IiV#DtV#N#part#DtV#

#it=to#hortative#l2-go#X/3-kill-dep#thing-n#indef#

#á?e-A-Nòpi#ke#ca'ha#I-cóka-D#má?e-A#mo#I-(e)ro'(to)N-D#

#á?eeropi#ke#cáha#icókaw#má?ee#mo#hérota#

- 5.3. #Adv#part#N#IiV#  
 #there#rec-past#motor-n#12-com-go#  
 ʔpe#raka#móto-A#ca-(e)roʔha#  
 #pé#raka#mótoa#caráha# mótoa < Ptg. motor
- 5.4. #N=PP#part#ItV#DtV#N#  
 #that-n-at#hortative#12-com-arrive#X/3-com-go#motor-n#  
 #áʔe-A-(ð)pe#ke#ce-(e)roʔhem#I-(e)roʔha-D#móto-A#  
 #áʔepe#ke#ceróhem#heráhaw#mótoa#
- 5.5. #DtV#DtV#N#part#DtV#  
 #X/3-kill-dep#X/3-com-come-dep#game meat#indef#X/3-eat-  
 dep#  
 #I-cóka-D#I-(e)roʔ(to)N-D#maʔeáʔa-A#mo#Iʔo-D#  
 #icókaw#hérota#maʔeáʔaa#mo#íʔow#
- 5.6. #N#part#part#DtV#  
 #fish-n#indef#also#X/3-catch-D#  
 #ipíra-A#mo#we#I-píhi(N)-D#  
 #ipíraa#mo#we#ipíhika#

## Text 5. Translation

1. We went off to Trocará. 2. We waited there to kill something and bring it back. 3. We took the motor boat there. 4. We arrived there taking the motor. 5. We killed game and brought it to eat. 6. We also caught fish.

Text 6

- 6.1. #N#part#ItV#Aux tV#N#

#Garotão-n-unattested#3/3-make#X/3-com-here/now#  
 manioc-finisher-n#

#karótaM-A-(r)imo#O-ápo#I-(e)ro-ka#maní?aηόpaM-A#  
 #karótawarimo#oápo#heréka#maní?aηόpawa#

(karótawa <Ptg. Garotão)

6.2. #N#ItV#Aux tV#(r)amo phrase#

#stream-n#3/3-divide#3/3-committative-here/now#  
 manioc-n//holder-to become#

#iharápe-A#O-mánaŋ#I-ero'ka#maní?aη-A//Níro-(r)amo#  
 #iharápee#ománaη#heréka#maní?aηa//ríroramo#

6.3. #N#It#Aux tV#N#(r)amo phrase#

#stream-n#3/3-divide-indef#X/3-committative-here/now#  
 Teteηoa#manioc-n//holder-become#

#iiqápi-A#O-mánaŋ-(r)imo#I-(e)ro'ka#tetéηo-A#maní?aη-  
 A//Níro-(r)amo#

#iiqápia#ománaηimo#heréka#tetéηoa#maní?aηa//ríroramo#

6.4. #DtV#Aff-PP#N#

#X/3-throw-throw-fut-dep#X/3-in#manioc#

#I-mána=mána-(pó)ta(N)-D#I-pípe#maní?aη-A#

#imanamanapótata#ipípe#maní?aηa#

6.5. #Adv#DtV#Adv#(r)amo Phrase#

#far off#X/3-roast-fut-dep#tomorrow#3-soft-when#

#qeY#I-píkoY-(pó)ta(N)-D#océ?iwe#I-táwa-(r)amo#

#qéy#ipikóytata#océ?iwe#itáwaramo#

6.6. #IiV#Aff-PP#(r)ire Phrase#

#3-come-fut-indef#3-from#X/3-do-finish-when#

#O'(to)N-(pó)ta(N)-(r)imo#I-(ð)hi#I-ápo'paM-(r)ire#

#óntarimo#icóhi#iapopawíre#

6.7. #Itv#N#IiV#N#

#3/3-do-finish-fut-indef#manioc-n//holder-n#3-come-fut-  
indef#Garatao-n#

#O-ápo'paM-(pó)ta(N)-(r)imo#maní'ãaη-A//Níro-A#O'(to)N-  
(pó)ta(N)-(r)imo#karótaM-A#

#oapopámtarimo#maní'ãaηa//ríroa#óntarimo#karótawa#

#### Text 6. Translation

1. Garotão is making a manioc processor. 2. He's dividing the stream to make a manioc-holder. 3. Tetenoa is dividing the stream to make a manioc holder. 4. They'll throw manioc into it. 5. Then (way off) they'll roast it tomorrow when it's soft. 6. They'll come back when they're finished. 7. When Garotão finishes the manioc holder, he'll come back.

The following are some words taken from text which contain interesting rule application combinations. The form of presentation will be apparent in each case.

#DtV#

#X/3-pierce(sew)-fut-dep#

#I-kótoN-(pó)ta(N)-D#

#ikotótata#

#N-PP-part.#

#what-for-indefinite#

#máʔe-A-Nèhe-(r)imo#

#máʔeereherimo#

#N-PP#

#3-hair-with#

#Iʔam-A-Nèhe# 4,10,12,20

#íʔawarehe#

(Note that íʔawa 'his hair'; íʔam 'he has much hair' from #Iʔam-A# and #Iʔam# respectively.)

#night-during#

#ipf̄ton-(r)imo#

#ipf̄tonimo#

#13-dance=finish-when#

#oro-poráhay=paM-(r)íre#

#oroporahaypawíre# (See Rule 22)

#far-off#he-went-function marker#

#qeY#I'-ha-I#

#qéy#íhay# (See Rule 23)

#I-come-dep#

#we-(to)N-D#

#wétota# 11,16,20

#12-sit=sit=dep#

#cere-ápiN=ápiN-D# (reduplication)

#cereapiápika# (See Rule 15)

#X/3-lift-X/3-lift-dep#

#I-NópiN=I-NópiN-D# (reduplication)

#hopihópita# (See Rule 15)

#### FOOTNOTES

<sup>1</sup>Asurini is the name given by Brazilians to an indigenous group which made its first peaceful contact with Brazilian society some fifteen years ago. Before the contact theirs was chiefly a hunting and gathering society. They wandered in the area between the Tocantins and Pacajá rivers at the latitude of the town of Tucuruí (formerly Alcobaça). Since their pacification, they have lived on the Indian Protection Service post established eighteen kilometers downriver from Tucuruí. Some have learned Portuguese, and loanwords from that language are beginning to appear in Asurini conversation. The Asurini call themselves áqawa (also used to refer to other indigenous groups) and their language áqawa cé?eña. The language belongs to the Tupi-Guarani stock, which has affinities over a great part of the South American continent.

Field work for this thesis was done in 1962 and 1963 under the auspices of the Summer Institute of Linguistics in cooperation with the National Museum of Rio de Janeiro, Brazil.

<sup>2</sup>Some especially useful devices developed in transformational-generative grammars are the use of brackets and braces to collapse a number of related rules into one, as well as the idea that the rules of the phonological component apply to the output of the syntactic component. See the following bibliography items: Bach (1964), Chomsky (1957, 1964, 1965), Chomsky and Halle (1965), Fodor and Katz (1964), Halle (1962, 1964), Hockett (1942, 1954), Koutsoudas (1966), Lounsbury (1953).

<sup>3</sup>In a more complete grammar for Asurini, these would be the terminal strings of the syntactic component of a generative

grammar. The advantages of a generative grammar are clear. The rules seem reasonably easy to write and manipulate, and the concept of the derivation gives a mechanical procedure for checking the correctness of a set of interrelated rules. That is, by taking the input as given (#S# if it is a sentence grammar; strings of lexical forms if it is a morphophonology), and if conventions are specified concerning the order and manner of application of rules, one can apply a set of rules mechanically to see if they give the desired output and adjust accordingly. Where possible, of course, one hopes to mirror some of the intuitions about the structure by choosing the best set of rules according to a mixture of declared as well as innate criteria.

<sup>4</sup>A formative, taken as one of the primitives in this paper, corresponds to a lexeme in the terminology of Sydney Lamb (1964). It might be thought of as a dictionary entry. Although it is not semantically invariant, it is hoped that the number of semantic readings of a lexeme will be limited and that the variants can be predicted in most cases by context.

<sup>5</sup>The practice of specifying phonologies in terms of acoustic distinctive features without using a spectrograph is somewhat questionable, and must eventually fall prey to any weaknesses of articulatory phonetics. Acoustic specification of Asurini phonology would have involved certain specific kinds of recordings which I did not make at the time of the field work for lack of a precise knowledge of the nature and demands of such a machine.

<sup>6</sup>In the text material the wedge is omitted from ř and č for typographical ease. Abbreviations, respectively, stand for sibilant, nasal, resonant, labial, alveolar, alveopalatal, velar, and glottal.

<sup>7</sup>The form -D 'dependent verb marker' is a lexeme or formative which is clearly a single dictionary entry yet has phonologically unrelated realizations in the text. This is something like the case of good-bet-bes in English.

<sup>8</sup>This is an old chestnut. See Pike (1947a). Much of recent work in generative phonology emphasizes the relationships between syntax and phonology.

<sup>9</sup>The case of the auxiliary have in some English dialects has the alternation /hæf/ before infinitives (have to go) and /həv/ elsewhere (have taught). In the event that the

alternation conditioning cannot be attributed to accent, this is a candidate for consideration as syntactically conditioned phonological alternation.

<sup>10</sup>Such a rule may also be written

$$A \rightarrow B / \underline{X} \underline{Y}$$

We will adopt the other form. We will also adopt the convention that when we find that environments for one side of the symbol to be rewritten do not need to be specified, no symbol need appear on that side and 'all possible environments' is understood for the blank (lack of symbol).

<sup>11</sup>Subsequent analysis has led this investigator to recognize the mutually exclusive distribution of the roots -(to)N and -caN, both meaning 'to come'. -caN occurs when the verb is intransitive and the person is non-third, -(to)N elsewhere. As in the case of -D (see Rule 16 and accompanying footnote), it is likely to be a case of person and number based suppletion as a result of syncretism.

<sup>12</sup>The fact that a choice is necessary is a consequence of the form of the descriptive apparatus. See Lamb (1966).

<sup>13</sup>Historical linguistics generally recognizes this as suppletion or replacement by a different morph in certain environments in which "...the phoneme discrepancies between suppletive co-allomorphs (e.g., E go;wend) are not significantly repeated in allomorphs making up the other morphemes of the language." (Hoenigswald 1960.48). This comes about as a result of syncretism. "...morphs which originally contrast recur, at a later stage, in phonemically corresponding shape but in complementary distribution and therefore as co-allomorphs..." (Hoenigswald 1965.36).

<sup>14</sup>Rule ordering has been discussed in a number of cases. Some pairs of rules are not ordered with respect to each other since they do not concern the same subject matter. Some pairs of rules have an order imposed on them for the purpose of getting correct results in the terminal string. This is done because the members of a given pair may involve the same subject matter and the output of one rule affects the input of the next. The type of grammar chosen here entails ordering in such cases. Often this order is arbitrary. In general, an order is chosen which will rewrite the symbols from left to right in a word, and if this is not at issue then that ordering which gives the most general rule last is chosen because it often saves writing a good bit of detail in the rules.

Some pairs of rules seem to be intrinsically ordered, that is, to change the obvious order would be to vastly complicate the description.

In the following scheme we will use 'NOT' for pairs that are not ordered, IMP for pairs with imposed order and INT for pairs with intrinsic ordering. Since the number of pairs to consider among thirty rules is high, we will take the rules by sets and in some cases talk about the sets as wholes. Upper case letters represent the sets (A,B,C...).

A 1,2,3,4 ordered within the group  
1,2 IMP; 2,3 IMP; 3,4 INT

B 5,6,7,8,9,10 ordered within the group  
5,6 IMP; 6,7 IMP; 7,8 INT;  
8,9 NOT; 9,10 INT

C 11 highly idiosyncratic

D 12

E 13

F 14

G 15

H 16,17,18,19,20,21 (ordering important for the whole set)  
16,17 NOT; 17,18 INT; 18,19 NOT; 16,17,18,  
19,20,21 INT

I 22

J 23

K 24-29 24,25 NOT; 25,26 NOT; 26,27 IMP; 27,28 IMP;  
28,29 INT; 24-27,29 INT

A,B IMP; B,C IMP; A,C IMP; C,D NOT; D,E NOT; E,F NOT;  
C,E NOT; C,F NOT; F,G NOT; G,H INT; H,I IMP;  
I,J INT; J,K NOT

## Bibliography

Bach, Emmon

1964 An Introduction to Transformational Grammars.  
Holt, Rinehart and Winston, Inc.

Bloch, Bernard

1941 Phonemic Overlapping. American Speech 16.278-84.

Chomsky, Noam

1957 Syntactic Structures. Mouton & Co.

1964 Current Issues in Linguistic Theory. Mouton & Co.

1965 Aspects of the Theory of Syntax. The MIT Press.

Chomsky, Noam and M. Halle

1965 Some Controversial Questions in Phonological  
Theory. Journal of Linguistics, Vol. 1, No. 2.

- Fodor, Jerry A. and Jerrold J. Katz  
1964 The Structure of Language: Readings in the Philosophy of Language. Prentice-Hall, Inc.
- Halle, Morris  
1962 Phonology in a Generative Grammar. Word 18.  
1964 On the Bases of Phonology. In Fodor and Katz, 1964 (above).
- Hockett, Charles F.  
1942 A System of Descriptive Phonology. Language 18.3-21.  
1947 Problems in Morphemic Analysis. Language 23.321-43.
- Hoenigswald, Henry M.  
1960 Language Change and Linguistic Reconstruction. The University of Chicago Press.
- Koutsoudas, Andreas  
1966 Writing Transformational Grammars: An Introduction. McGraw-Hill.
- Lamb, Sydney M.  
1964 The Sememic Approach to Structural Semantics. American Anthropologist, Vol. 66, No. 3, Part 2.  
1966 Outline of Stratificational Grammar. Georgetown University Press.
- Lounsbury, Floyd G.  
1953 The Method of Descriptive Morphology. In: Oneida Verb Morphology, Yale University Publications in Anthropology No. 48.
- Matthews, G.H.  
1965 Hidatsa Syntax. Mouton & Co.
- Pike, Kenneth L.  
1947a Grammatical Prerequisites to Phonemic Analysis. Word 3:3.  
1947b Phonemics. U. of Michigan Press.



# COCAMA CLAUSE TYPES

Norma Faust

0. Introduction
1. Principal Independent Clause Nuclei
2. Subordinate Independent Clause Nuclei
3. Clause Margins
4. Internal Structure of Intra-Clause Tagmeme

A key to the abbreviations and symbols used in this paper. (Numbers in parenthesis refer to the section number in which the abbreviation or symbol first appears.)

Aff	Affirmation tagmeme (3)
App	Appositive tagmeme (3)
As	Aspect tagmeme (1.1)
Asso	Association tagmeme (3)
C	Cause tagmeme (3)
Con	Concern tagmeme (3)
Dir	Direction tagmeme (3)
I	Identification tagmeme (1.3)
IM	Intention-movement tagmeme (3)
Ins	Instrumental tagmeme (3)
L	Location tagmeme (3)
M	Manner tagmeme (3)
Mo	Mode tagmeme (1.1)
N	Non-imperative clause nucleus (3)
NA	Nominative absolute tagmeme (3)
Nc	Nucleus core (1.1)
Neg	Negation tagmeme (3)
O	Object tagmeme (1.1)
P <sub>1</sub>	Predicate <sub>1</sub> tagmeme (1.1)
P <sub>2</sub>	Predicate <sub>2</sub> tagmeme (1.2)
PC	Principal <sub>2</sub> clause (0)
PCA	Principal active clause type (0)
PCAN	Principal active clause nucleus (3)
PCI	Principal identification clause type (0)
PCIN	Principal identification clause nucleus (3)
PCS	Principal stative clause type (0)
PCSN	Principal stative clause nucleus (3)
Per	Perfective tagmeme (3)
Pur	Purpose tagmeme (3)
Q	Query tagmeme (3)
Qua	Quantity tagmeme (3)
R	Recipient tagmeme (1.1)

S	Subject tagmeme (1.1)
SC	Subordinate clause (0)
SCSN	Subordinate stative clause nucleus (3)
Sub	Subordinating tagmeme (2)
Sub <sub>1</sub>	Subordinating allotagma <sub>1</sub> (2)
Sub <sub>2</sub>	Subordinating allotagma <sub>2</sub> (2)
Sim	Similarity tagmeme (3)
Te	Tense tagmeme (1.1)
V	Vocative tagmeme (3)
-	Indicates phonological fusion between Cocama words (1.1)
∞	Alternates freely with (1.1)
...	Something precedes/something follows (narrative medial) (1.1)
'	Fusion of two levels (0)

0. The analysis of Cocama<sup>1</sup> syntax structure that is presented in this paper follows the theory of tagmemic analysis developed by Kenneth L. Pike in his Language in Relation to a Unified Theory of the Structure of Human Behavior.<sup>2</sup>

The theory presented in Pike's book involves the principle of the hierarchical structuring of all human behavior, including verbal behavior. This hierarchical structuring may be seen in the grammatical structure of the Cocama language. The minimum unit of this structure is the tagmeme, which is defined as a functional slot plus the filler of the slot. Both concepts, the slot type and its class of fillers, are necessary to the definition of a tagmeme in this theory. An example to show the slot-class correlation would be a subject slot filled by a substantive expression. The subject slot along with its filler, the substantive expression, manifests a subject tagmeme (see 4.2).

Tagmemes are here described in a "hierarchy of distributional matrices".<sup>3</sup> There are four major levels of distributional matrices in the syntax structure of Cocama. The lowest major level consists of stem and affix tagmemes included in the word matrix; words in turn constitute tagmemes included in the phrase matrix; phrases in turn constitute tagmemes included in the clause matrix; clauses in turn constitute tagmemes included in the sentence matrix. When the focus is on the word matrix the tagmemes included on this level are termed intra-word tagmemes. When the focus changes to the phrase matrix the tagmemes included on this level are termed intra-phrase tagmemes. Included tagmemes

in the clause matrix are termed intra-clause tagmemes, and included tagmemes in the sentence matrix are termed intra-sentence tagmemes.

The grammatical hierarchy of Cocama language structure may be seen in that intra-word tagmemes pyramid to intra-phrase tagmemes, which in turn pyramid to intra-clause tagmemes, which in turn pyramid to intra-sentence tagmemes.

This paper focuses on the clause matrix and in particular on the nuclear contrasts of the clause matrix. These nuclear contrasts differentiate the six major non-imperative clause types of the independent sentences of Cocama.

The remaining three major matrices--the word, the phrase, and the sentence--are discussed only briefly to show the relationship of the nucleus of the clause and of the clause matrix to the other matrices and to the tagmemes included in these matrices.

There are two main sentence types in Cocama: independent and dependent sentences. Dependent sentences are dependent on other sentences in narrative. Independent sentences are not. Only the latter are treated in this paper, since sequence-dependent sentences contain clauses of types described for independent sentences (plus a sequence indicator), whereas response-dependent sentences may contain no clause.

Independent sentences contain one obligatory principal predicating nucleus and may also contain one optional subordinate predicating nucleus. The independent sentence tagmeme is manifested by one of two variants. The variants are differentiated by the order of arrangement of their intra-sentence tagmemes and occur in free variation. These variants are symbolized: [ +PC' ±SC] ∅ [ ±SC +PC'], to be read, "obligatory principal clause tagmeme (PC) plus optional subordinate clause tagmeme (SC)". The symbol ' (in the formula PC') indicates that there is a fusion of levels in that PC' may occur either as the head tagmeme in an independent sentence matrix or may constitute the entire matrix. The principal clause may therefore constitute simultaneously an independent sentence matrix as well as a clause matrix. Henceforth in this paper fusion of two such levels will be indicated by the symbol '.

There are four main principal independent clause

types: (1) active (PCA), (2) stative (PCS), (3) identification (PCI), and (4) imperative which is subdivided into active imperative and stative imperative.

Active, stative, and identification clause types occur both in principal and subordinate varieties and are differentiated from each other by distinctions within the internal structure of their nuclear intra-clause tagmemes.

The imperative clause types are differentiated from the non-imperative by distinctions within the internal structure of their nuclear tagmemes--which we will not discuss here--as well as corresponding differences within the marginal intra-clause tagmemes. The imperative clause occurs as a sub-division of the class of principal clauses and not of the class of subordinate clauses. The imperative clause will not be treated further in this paper, which focuses on the nuclear contrasts of non-imperative clauses. Examples, however, of the imperative clause are: *cáni úri* 'come' (*cáni* 'imperative particle', *úri* 'stative verb, to come'), *úmi kái ái* 'look at it!' (*úmi* 'active verb, to see', *kái* 'imperative particle', *ái* 'third person singular substantive').

1. Principal Independent Clause Nuclei. The Cocama clause consists of an obligatory nucleus and an optional margin. The nucleus may be viewed as a matrix within the clause matrix and consists of an obligatory core plus optional modifying tagmemes. The nucleus of the principal non-imperative clause types will be discussed in Sections 1.1, 1.2, and 1.3; the nuclei of the subordinate clause types will be discussed in Section 2; and the margin, which is identical for both the principal and subordinate clause types, will be discussed in Section 3.

1.1. The core of the principal active clause nucleus consists of three obligatory intra-clause tagmemes: a subject tagmeme (S) manifested by a substantive expression, a predicate tagmeme ( $P_1$ ) manifested by an active verb expression, and an object tagmeme (O) manifested by a substantive expression. The subject tagmeme precedes the predicate tagmeme. The object tagmeme may either precede or follow the subject-predicate sequence, and may be formularized as follows:  $[+S +P_1 +O] \oslash [+O +S +P_1]$ . The above formula represents minimum non-conditioned variants of the principal active clause nucleus. Other rare variants occur with the object tagmeme as optional; the minus options are

conditioned in that they occur only in narrative medial position. They do not occur narrative initially. They are formularized as follows: [...+S +P<sub>1</sub> ±O...] ∞ [...±O +S +P<sub>1</sub>...].

The above are the minimal formulas for the principal active clause nucleus, as well as the minimal formulas for the principal active clause as a whole. The marginal tagmemes are optional and are included in the maximum formula for the clause (see 3).

The following examples illustrate minimum forms of the principal active clause nucleus:

+S		+P <sub>1</sub>		+O	
ipirawūra-púra		čikuaráta		púnu	'The dolphins follow us.'
dolphins		follow		us	
+O	+S	+P <sub>1</sub>			
yanamáta	ca-	kúpü			'I cultivate the grass.'
grass	I	cultivate			
+S	+P <sub>1</sub>				
umánu	ayúka				'The fever hit (her).'
fever	hits				

The theoretical maximum formula for the principal active clause nucleus is: +NC' ±As ±Te ±Mo ±R, to be read, "obligatory nucleus core (Nc) plus optional aspect tagmeme (As) plus optional tense tagmeme (Te) plus optional mode tagmeme (Mo) plus optional recipient tagmeme (R)".<sup>4</sup> No occurrence of maximum structure has been found in the present data. The following examples, however, illustrate all the tagmemes symbolized in the maximum formula:

┌-----Nucleus-----┐					
┌-----Nc'-----┐					
+S	+P <sub>1</sub>	+O	±As	±Te	
úri	debiška	t- íma-	úkuá-	cúri	'He was in debt to
he	owes	my brother	habitual	far	my brother.'
			aspect	past	
				tense	

+O	+S	+P <sub>1</sub>		±Mo	
ta-kú my garden	t- I	iára want	yaúki- make	mía potential mode	'I would like to make my garden.'
+S	+P <sub>1</sub>	+O		±Te	±R
ái he	yúmi gives	y- it		úi immediate past tense	inú- cu them to
					'He gave it to them.'

1.2. The core of the stative clause nucleus consists of two obligatory intra-clause tagmemes: a subject tagmeme (S) manifested by a substantive expression and a predicate tagmeme (P<sub>2</sub>) manifested by a stative verb<sub>1</sub>, stative verb<sub>2</sub>, or stative verb<sub>3</sub> expression (see 4.1).

The subject tagmeme may either precede or follow the predicate tagmeme, and may be formularized as follows: [+S +P<sub>2</sub>] ∞ [+P<sub>2</sub> +S]. The above formula represents minimum non-conditioned variants of the principal stative clause nucleus, as well as the minimal formula for the principal stative clause as a whole.

The following examples illustrate minimum forms of the principal stative clause nucleus:

+S	+P <sub>2</sub>		
t- I	úcu go (stative verb <sub>2</sub> )		'I go.'
+P <sub>2</sub>		+S	
yumíra is angry (stative verb <sub>1</sub> )	t- ía my heart		'I am angry.'
+P <sub>2</sub>		+S	
tümúndü there is no (stative verb <sub>3</sub> )	kuríki money		'There is no money.'

The theoretical maximum formula for the principal stative clause nucleus is [+P<sub>2</sub> +S ±As ±Te ±Mo ±R] ∞ [+S +P<sub>2</sub> ±As ±Te ±Mo ±R] ∞ [+S<sup>2</sup> +P<sub>2</sub> ±R ±As ±Te ±Mo].<sup>5</sup>



clause as a whole.

The following examples illustrate minimum forms of the principal identification clause:

+I	+S		
wüpü-nán	ca-mümirakiramími		'I have only one little son.'
one only	my little son		

+S	+I		
úcu	inu- tía		'I am their aunt.'
I	their aunt		

The theoretical maximum formula for the principal identification clause nucleus is [+S ±As ±Te ±Mo +I] ∅ [+I +S ±Te] ∅ [+I +S ±R]<sup>6</sup>.

The following examples illustrate the expanded nucleus:

+I	+S	±Te	
yaúpü	ya-pur-	úi	'He was there.'
there	he	immediate past tense	

+S	±Mo	+I	
ya-putiakuarar-úra		ya-túa tíni	'Its chest is white.'
its chest	subjunctive mode	it white	white.'

+S	±Te	+I	
učimatan-	cúri	yúwa	'That which was taken out was a thorn.'
that-which- was-taken- out	far past tense	thorn	

+S	±As	+I	
na-kumic-úcu	ya-kuára		'Your speech will be inside of it (tape recorder).'
your speech	aspect of going	it in	

2. Subordinate Independent Clause Nuclei. Subordinate clauses are differentiated from principal clauses by virtue of their optional occurrence in independent sentences and the occurrence of a subordinating tagmeme within the nucleus of the clause. In the present data tense and aspect tagmemes are of rare occurrence in the subordinate clause. Mode and recipient tagmemes have not been found, although it is highly probably that they may potentially occur as optional tagmemes of the subordinate clause in the Cocama language structure.

The tense tagmeme occurs following the subordinating tagmeme in the subordinate stative clause. It has not as yet been found in the subordinate active or subordinate identification clause. The aspect tagmeme occurs following the predicate<sub>2</sub> tagmeme and preceding the subordinating tagmeme in the subordinate stative clause; it occurs following the subordinating tagmeme and the object in the active clause; it has not as yet been found in the subordinate identification clause.

The following formula shows the structure of the subordinate stative clause as found in the data to date:  
 [+P<sub>2</sub> +S +Sub ±Te] ∞ [+S +P<sub>2</sub> ±As +Sub].

The following examples illustrate the subordinate stative clause:

+P <sub>2</sub>	+S	+Sub	±Te	
caipurári drunk	ya-puránu- they	íkua- because	íkua yester- day tense	'Because they were drunk yesterday.'
+S	+P <sub>2</sub>	+Sub		
ta-kuríki my money	ümütú- there is	ra if		'If I had the money.'
+S	+P <sub>2</sub>	±As	+Sub	
pünu we	iri- return	úcu- going aspect	púka when	'When we return.'

The following formula shows the structure of the subordinate identification clause as illustrated by the one

example that has been found in the present data:

+S +I +Sub

úñü üram- íkya 'Because you are good.'  
 you good because

In the subordinate active clause the subordinating tagmeme follows the predicate<sub>1</sub> tagmeme. Following the predicate<sub>1</sub> tagmeme the subordinating<sub>1</sub> allotagma occurs preceding the object, and the subordinating<sub>2</sub> allotagma occurs following the object. The subordinate active clause is formularized as follows: [+0 +S +P<sub>1</sub> +Sub<sub>2</sub>] ∞ [+S +P<sub>1</sub> +0 +Sub<sub>2</sub>] ∞ [+S +P<sub>1</sub> +Sub<sub>1</sub> +0].

The following examples illustrate the subordinate active clause:

+0 +S +P<sub>1</sub> +Sub<sub>2</sub>

ašún umanuarinán ucún c- ürucu wíri 'Although I  
 one-about-ready-to-die I took although took one who  
 was about  
 ready to die.'

+S +P<sub>1</sub> +0 +Sub<sub>2</sub>

ca-wihu-púra ukúra ca- wíri 'Although my old man  
 my old-man kept back me although kept me back.'

+S +P<sub>1</sub> +Sub<sub>1</sub> +0

t- amastáka ráši úra 'If I can do it.'  
 I can do if it

3. Clause Margins. Various marginal intra-clause tagmemes of modification may be attached to the nucleus of each of the six clause types described in detail in this paper (principal active, principal stative, principal identification, subordinate active, subordinate stative, and subordinate identification). Certain of these marginal tagmemes occur preceding or following the nucleus. The marginal intra-clause tagmemes that occur only preceding the nucleus of the clause include affirmation, negation, perfective, and query. Those that occur either preceding or following the nucleus include vocative, nominative absolute, intention movement, location, direction, association,

instrumental, cause, concern, purpose, quantity, time, manner, similarity, and appositive. The affirmation is always first; otherwise there is no fixed order. (There is some possibility, however, that degree of emphasis may affect the order of two intra-clause marginal tagmemes preceding the nucleus, or of two following the nucleus; the one closer to the nucleus may prove to be the one emphasized.)

In the data studied to date there are no further structural restrictions on the co-occurrence of the marginal tagmemes in the various principal and subordinate clause types.

There are, of course, some semantic restrictions such that the intention-movement marginal tagmeme occurs chiefly with the stative verb<sub>2</sub> *úcu* 'to go' as manifesting the predicate<sub>2</sub> tagmeme and with other words that express movement. Examples of sentences containing this tagmeme are:

+S +P<sub>2</sub> ±Intention Movement ±Location ±Appositive

y-	úcu	kamatatára		ahánga	orejanáka
he	goes	work	to	here	Orellana in

'He goes to work here in Orellana.'

±Intention Movement +S +P<sub>1</sub> +O ±Te ±Mo

mucanakatára	c-	ürúra	y-	ui-	mía
to-treat-	I	bring	her	immedi-	poten-
with-medicine				ate	tial
				past	mode
				tense	

'I brought her to be treated.'

The non-imperative clause formula is given below. Tagmemes listed in the formula as following the nucleus are from the group that may either precede or follow it (a fact which is indicated by the arrow); those before the nucleus are from the group which cannot follow the nucleus; free variation of order of these marginal tagmemes in reference to one another is as stated above, but not symbolized: ±affirmation (Aff) ±negation (Neg) ±query [+slot<sub>1</sub> +query] (Q) ±perfective (Per) +Nucleus (N) ±←similarity (Sim) ±←manner (M) ±←quantity (Qua) ±←purpose (Pur) ±←concern (Con) ±←cause (C) ±←instrumental (Ins)

±← association (Asso) ±←direction (Dir) ±←location (L)  
 ±←intention movement (IM) ±←nominative absolute (NA)  
 ±←appositive (App) ±←vocative (V).

Of these, it should be further noted that the query tagmeme is manifested by a clitic word stem which is phonologically dependent on the manifested form of the preceding tagmeme other than the affirmation. When no marginal tagmemes precede the nucleus, the query tagmeme occurs within the nucleus following the first nuclear tagmeme, and its manifested form is phonologically dependent upon the manifested form of the first nuclear tagmeme. When the affirmation tagmeme occurs in the clause (which, if present, always occurs as the first tagmeme) the query clitic can only follow the manifested form of the second occurring tagmeme, whatever it may be, and is phonologically dependent upon it.

The formula listed above is the theoretically maximum formula for the non-imperative clauses of the independent sentences. No more than three marginal tagmemes, however, have been found in any one clause. The following examples illustrate all of the tagmemes symbolized in the above formula:

	┌---PCSN---┐						
±L	±Q	+S	+P <sub>2</sub>				
maka- where	típa query	n- you	úcu go			'Where are you going?'	
	┌-----PCAN-----┐						
+O	±Q	+S	+P <sub>1</sub>	±Te	±T		
mari- thing	típa query	n- you	igauki-úi do	immediate past tense	ikún today	'What did you do today?'	
	┌-----PCIN'-----┐						
+S	±Q	+I					
mari- thing	típa query	ahánga here					'What is here?'

┌-----SCSN-----┐

±M		+S	+P <sub>2</sub>	+Sub	±Asso	±App
wúpü	yatíri	rana-caipura-	cún	ta-múki	kuyara-múki	
one	together	they are	drunk	for me with	garden with	owner

'In order for them all to get drunk with me, the owner of the garden.'

┌-----PCIN-----┐

±Aff	+I		±Q	+S	±R	
áha	aicüwanán	cačín-	dip-	ái	ya-	papa- cúpü
yes	little	loved	query	he	his father	recipient indicator

'Yes, is he just a little loved by his father?' (rhetorical question)

┌---PCSN---┐

±M	±App	+S	+P <sub>2</sub>	±NA	
čún	elvira	y-	umánu	aikuán yía	cačín
almost	Elvira	she	die	sick heart	hurt

'Elvira almost died from a sick heart.'

┌----PCSN-----┐

±Per	±Neg	+S	+P <sub>2</sub>	
ái	tíma-purái	yakikuára	cači	'(Her) head does not hurt anymore.'
now	not	head	hurts	

┌----PCSN-----┐

+S	+P <sub>2</sub>	±Te	±IM	±Asso
t-	uri-	cúri	kamatatára	t- ima- múki
I	come	far	work to	my brother with
		past		
		tense		

'I came to work with my brother.'



	┌---PCAN-----┐				
±Con	+S	+P <sub>1</sub>	+O	±As	
rafael-ári Rafael about	t- I	ikuáta advise	n- you	úcu will	'I will tell you about Rafael.'

	┌-----PCAN-----┐					
±T	+S	+P <sub>1</sub>	+O	±Te	±Qua	
ikún today	amuakána others	ürúra bring	ipirapur- fish	úi immediate past tense	číta many	

'Today the others brought many fish.'

4. Internal Structure of Intra-Clause Tagmemes. In this section some data on the internal structuring of the intra-clause tagmemes are to be presented. Each of the intra-clause tagmemes constitutes a phrase matrix. The breakdown of the phrase matrix will not be given in detail except where it is pertinent to the clause structure. The first part of this section (4.1-8) deals with the nuclear tagmemes; the remainder (Sections 4.9-27) deals with the marginal tagmemes.

4.1. The predicate tagmemes are manifested by verb expressions. They are distinct from other tagmemes in that the words manifesting the head tagmemes of the predicate phrase are made up of verb stems accompanied by verbal suffixes. There are two distinct predicate tagmemes, predicate<sub>1</sub> and predicate<sub>2</sub> as mentioned in 1.1 and 1.2.

The predicate<sub>1</sub> intra-clause tagmeme is manifested by an active verb expression. Within this expression the active verb word occurs as manifesting an obligatory intra-phrase head tagmeme. (Note that intra-clause and intra-phrase tagmemes are on different structural levels. When, however, the phrase is comprised of a single word, there is a fusion of levels--see the introduction--so that the one word simultaneously manifests both the intra-phrase head and the total intra-clause predicate.) The stative verb<sub>1</sub> word, stative verb<sub>2</sub> word, and certain particles and clitics occur as manifesting optional intra-phrase tagmemes which modify directly the intra-phrase head--in which case there

is no fusion of levels since the head verb itself does not manifest the total intra-clause predicate tagmeme. The meaning of the predicate<sub>1</sub> tagmeme is action directed toward a goal. The following examples illustrate the predicate<sub>1</sub> phrase:

+Head ±Modifying particle

tikíta	cápa	'to tie (it) soon'
tie	soon	

±Negative particle +Head

tíma	čipíta	'not to blame
not	blame	(someone)'

±Modifying stative verb<sub>2</sub> +Head

ucupurána	iku	'to know (it)
immediately	knows	immediately'

±Modifying stative verb<sub>1</sub> +Head

úpa	kúpü	'to finish culti-
finish	cultivate	vating (it)'

The predicate<sub>2</sub> intra-clause tagmeme is manifested by a stative verb expression. Within this tagmeme the stative verb word occurs as an obligatory intra-phrase head tagmeme and the stative verb<sub>1</sub>, stative verb<sub>2</sub>, and certain particles and clitics occur as optional intra-phrase tagmemes which modify this head.

Stative verbs are sub-divided into three classes: (1) stative verb<sub>1</sub> which manifests the predicate<sub>2</sub> and the manner intra-clause tagmemes; (2) stative verb<sub>2</sub> which manifests the predicate<sub>2</sub>, manner, and aspect intra-clause tagmemes; and (3) stative verb<sub>3</sub> which manifests the predicate<sub>2</sub> tagmeme. The stative verb<sub>3</sub> class consists of only two members, ümütü 'there is' and tümündü 'there is no'.

The meaning of the predicate<sub>2</sub> tagmeme is state of action, being, or quality of a subject. For examples of the three classes of stative verb within the stative clause nucleus see Section 1.2.

The following examples illustrate the predicate<sub>2</sub> phrase:

±Modifying stative verb<sub>2</sub> +Head stative verb<sub>2</sub> ±Modifying  
clitic

úri	úri	kat
come	come	again

'to come repeatedly'

±Modifying similarity particle +Head stative verb<sub>1</sub>

yíá	kumic-
like that	talks

'(he) talks like that'

±Modifying negative particle +Head stative verb<sub>1</sub>

tíma	amask-	'not to be
not	able	able'

±Modifying stative verb<sub>1</sub> +Head stative verb<sub>1</sub>

yupúni	üratayar-	'begin to
begin	to put in order	organize'

4.2. The subject intra-clause tagmeme is manifested by a substantive expression. The head intra-phrase tagmeme of the subject phrase matrix is manifested by a substantive class of nouns and pronouns which, with some restrictions, occur with noun suffixes.

Within the subject intra-clause tagmeme a substantive occurs as manifesting an obligatory intra-phrase head tagmeme. The head intra-phrase tagmeme may be accompanied by optional modifying intra-phrase tagmemes manifested by substantive words, location expressions, stative verb<sub>1</sub> words, and certain modifying particles and clitics.

The subject tagmeme indicates: (1) the subject of the action directed toward a goal in an active clause, (2) the subject described as being in a state of action, being or quality in a stative clause, and (3) the subject identified in an identification clause.

The following examples illustrate the subject phrase:

±Modifying noun +Head noun ±Modifying location expression

ahán	medíko	puérto prádo-katín
this	doctor	Puerto Prado at

'the doctor from Puerto Prado'

±Modifying pronoun ±Modifying noun +Head noun

ahán	ikracún	čipitára	'this child's
this	child	bewitcher	bewitcher'

±Modifying noun +Head noun

wainakána	yapuračitáta	'the women's
women	dance maker	dance maker'

±Modifying noun +Head noun

úpi	áwa	'everyone'
all	man	

±Modifying stative verb<sub>1</sub> +Head noun

ipípu	yawára	'the dog
inside	dog	inside'

+Head noun ±Modifying clitic

kuyaramburá-	ya	'the owner of
owner-of-the-garden	reportative	the garden
		(it is said)'

4.3. The object intra-clause tagmeme is manifested by a substantive expression (see 4.2), or by the hesitation particle, ápu.<sup>7</sup> The object tagmeme indicates the goal toward which the activity is directed in an active clause.

The following examples illustrate the object phrase:

±Modifying quantity noun +Head noun

mucaprika	waránga	čupáti	'three thousand
three	thousand	palm leaves	palm leaves'

±Modifying negative particle +Head noun

tíma	mári	'nothing'
no	thing	





( +P<sub>2</sub> +S) ±R

(cáku kuaráči) kupüwaraní- cu '(The sun is hot)  
(hot sun) cultivators recipient for the cultivators.'

Identification clause (unique example).

( ±Aff +I ±Q +S) ±R

(dha aicüwanán cačín-dip- dí) ya- papa- cüpü  
(yes little only loved query he his father recipient  
indicator

'(Yes, is he just a little loved) by his father.'

4.6. The aspect intra-clause tagmeme is manifested by a stative verb<sub>2</sub> word. The membership of the stative verb<sub>2</sub> class of words includes: (1) úcu 'to go'; (2) úri 'to come'; (3) úkuá 'to go about, habitual'; (4) yúti 'to be, temporal'; and (5) iára 'to desire'.

The aspect tagmeme indicates aspectual modification of the clause nucleus core. The following examples illustrate the aspect phrase:

( +S +P<sub>1</sub> +O) ±As

(pünu yumíta y-) úkuá '(We teach her)  
(we teach her) habitual habitually.'

( +S +P<sub>1</sub> +O) ±As

(ta- tikíta cápa r-) úcu '(I) am going (to  
(I tie soon it) to go tie it right away).'

( ±L +S +P<sub>2</sub>) ±As (±Te)

(ca-mümirakuñáka c- uwatat-) uri- (úí)  
(my daughter at I visit to come (immediate  
past  
tense)

'(I) came (to visit my daughter).'

( +S) ±As ( +I)

(na- kumic-) úcu (ya-kuára) '(Your speech) will  
(your speech) to go (it in) (be inside of it).'

(+S +P<sub>2</sub>) ±As ( ±Te)

(r- umánu-) iara- (cúri) ' (He) wanted (to  
(he die) desire (far die).'  
past  
tense)

( ±Per +S +P<sub>1</sub> +O) ±As (±Te)

(ai-púra ya- pápa ayukápa ini-) yuti- (úi)  
(now his father hits us to be (immediate  
past tense)

'(His father has already hit us.)'

4.7. The tense intra-clause tagmeme is manifested by a tense clitic. The membership of the tense class of clitics includes: (1) cúri *ø* curípü 'far past tense', (2) ikuá 'yesterday past tense', (3) úi 'immediate past tense', and (4) á 'future tense'. The following examples illustrate the tense phrase:

(±Aff ±IM +S +P<sub>1</sub> +O) ±Te

(áha mucanakatára c- ürúra y-) úi ' (Yes, I  
(yes for treatment I bring her) immediate brought her to  
past tense be treated.)'

( +S +P<sub>1</sub> +O) ±Te

(ínu úyu ya-) cúri ' (They ate it.)'  
(they eat it far  
past  
tense)

( +S) ±Te ( +I)

(ca-kiwira-) cúri (inu- pápa) ' (My brother) was  
(my brother far (their father)  
past  
tense (their father).'

( +O +S +P<sub>1</sub>) ±Te

(úpi mári inu- tá-) ikuá ' (They received  
(everything they receive yesterday  
past  
tense everything) yes-  
terday.'

( ±M	+S		+P <sub>2</sub> )	±Te
(tutukári	ca-mümirakuñapuraná	ndi	yuriti kat-	á
(washing	my daughter		stays again	future
				tense

'(My daughter) will (remain to wash).'

4.8. The mode intra-clause tagmeme is manifested by a mode clitic. The membership of the modal class of clitics includes:<sup>9</sup> *úra* ∅ *úránu* 'subjunctive mode', and *mía* 'potential mode'. The following examples illustrate the mode phrase:

( ±Asso	+S	+P <sub>1</sub>	+O	±Te)	±Mo
(č- ia-	múki	c- üruráka	y-	ui-)	úra
(my heart with	I	bring	her	immediate	subjunctive
				past	mode
				tense)	

'(With deep feeling I brought her.)'

( ±Per	±Q	+S	+P <sub>2</sub> )	±Mo	
(úcu-pura-táku	y-	umanu-)	mía		'(He) might (al-
(already perhaps	he	die)	potential		ready be dead).'
			mode		

4.9. Sections 4.1 through 4.8 have dealt with the nuclear intra-clause tagmemes; Section 4.9 through 4.27 deal with the marginal intra-clause tagmemes.

The affirmation intra-clause tagmeme is manifested solely by the affirmation particle *áha* 'yes', with no modifiers permitted:

±Aff ( ±IM	+S	+P <sub>1</sub>	+O	±Te)	
áha	(mucanakatára	c- ürura	y-	úi)	'Yes, (I
yes	(for treatment	I	bring	her	brought her
				immediate	to be
				past	treated).'
				tense)	

4.10. The negation intra-clause tagmeme is manifested by a negation expression. The negation expression consists of the obligatory negation particle *tíma* plus an optional modifying clitic. The following examples illustrate the

negation phrase:

±Neg ( +S +P<sub>1</sub> +O +Sub)

tíma (na- čipíta ca-pur- íkuá) '(Because you did)  
not (you pay me because) not (pay me).'

±Neg ( ±Per +S +P<sub>2</sub> ±L)

tíma-pur- (ái ra-cüt- úri ikiáka) '(He does) not  
not emphatic (now he wants come here) (want to come  
clitic here any more).'

4.11. The query intra-clause tagmeme is manifested by an unmodified question or pondering clitic. The membership of the query class of clitics includes: (1) típa~dípa, (2) táku~dáku, (3) nda, (4) ráku, (5) ta, (6) ra, and (7) ku. The variants as symbolized in numbers 1 and 2 are phonologically conditioned; dípa and dáku occur following the phoneme /n/, típa and táku occur elsewhere. The following examples illustrate the query phrase:

(±Dir) ±Q ( +S +P<sub>2</sub>)

(maka- cui-) típa (n- úri) '(Where do you  
(where from) query (you come) come from)?'

( ±L ) ±Q ( +S +P<sub>2</sub>)

(maká-) nda (n- úcu) '(Where are you  
(where) query (you go) going)?'

( +S) ±Q ( +P<sub>2</sub>)

(ra- cukuará-) ra (cáči) '(Does his body  
(his body) query (hurt) hurt)?'

( ±T) ±Q ( +S +P<sub>2</sub> ±Te)

(maña-puká-) ta (ini-úri uri- kat- á) '(When will we  
(when) query (we come come again fu- come again)?'  
ture  
tense

( +I) ±Q ( +S)

(mári yawá-) ku (na- ritáma) '(What is your  
(thing like) query (your city) city like)?'

( ±Per)	±Q	(+S +P <sub>2</sub>	±Mo)	
(ucu-pura-) (already)	táku pondering	(y- umanu- he die	mía) potential mode)	'(He might al- ready be dead.)' or '(Might he al- ready be dead)?'
( +S)	±Q	( +I)		
(mari-) (thing)	ráku query	(ahan-guára) (this in)		'(What is in this)?'

4.12. The perfective intra-clause tagmeme is manifested by a perfective expression, which consists of an obligatory perfective particle plus an optional modifying clitic. The membership of the perfective particle class includes úcu, ái, and áwü. Ucu and ái have been found with the emphatic clitic. Awü has not been found with a modifying clitic. The following examples illustrate the perfective phrase:

±Per	( +S +P <sub>2</sub> )			
úcu already	(y- umánu) (she die)			'(She is) al- ready (dead).'
( ±Ins)	±Per	(+S +P <sub>2</sub> )		
(ya- pú) (it by means of)	ái already	(y- úra) (he well)		'(By means of it he is) now (well).'
±Per	( ±M +S +P <sub>2</sub>	±Te)		
ai- already	púra emphatic clitic	(úra y- úra well he well far past tense)		'(He is) now (very well).'
±Per	( +S +P <sub>1</sub>	+O ±Mo)		
áwü already	(inu- mucanáka they treat	ya- her	mía) potential mode)	'(They may) al- ready (have treated her).'

4.13. The similarity intra-clause tagmeme is manifested by a similarity expression, which consists of an obligatory substantive expression or stative verb<sub>1</sub> plus an obligatory similarity particle plus an optional modifying

clitic. The similarity class of particles includes two members: *yá* and *indíru*. The following examples illustrate the similarity phrase:

+Stative verb<sub>1</sub> +Similarity particle

<i>úra</i>	<i>yá</i>	'very well'
good	like	

+Substantive expression +Similarity particle

<i>tíma ámuá medikúnu</i>	<i>yá</i>	'not like
not other doctors	like	other doctors'

+Pronoun +Similarity particle ±Modifying clitic

<i>r-</i>	<i>íá</i>	<i>nán</i>	'just like it'
it	like	restrictive	

+Similarity particle +Substantive expression

<i>indíru</i>	<i>áwa itikán</i>	'like a man
like	man one-thrown-down	thrown down'

4.14. The manner intra-clause tagmeme is manifested by an obligatory stative verb<sub>1</sub> expression, stative verb<sub>2</sub> expression, manner particle, or stative clause nucleus core, plus an optional modifying clitic:

+Stative verb<sub>1</sub> ±Modifying clitic

<i>úra-</i>	<i>na</i>	'just good'
good	restrictive	

+Stative verb<sub>2</sub>

<i>ucári</i>	'going'
--------------	---------

+Manner particle

<i>yanikáka</i>	'fortunately'
-----------------	---------------

+Stative clause nucleus core

+S +P<sub>2</sub>

<i>tá</i>	<i>ucári</i>	'my going'
I	go	

4.15. The time intra-clause tagmeme is manifested by a time expression. The time expression may consist of an optional modifier plus an obligatory time particle; or it may consist of an obligatory stative verb<sub>1</sub> expression, active verb expression, substantive expression, or similarity expression, plus an obligatory time particle or clitic, plus an optional modifying clitic:

+Stative verb<sub>1</sub> +Time clitic

kanat-	ái		'when it is light'
is light	when		

±Modifying noun +Time particle

wúpü	ipíca		'the whole night'
one	night		

+Similarity +Time ±Modifying  
expression clitic clitic

r- íá	pukátu-	nán	'when it was just like
it like	when	restrictive	that'

+Substantive +Time clitic

amua	púca		'another time'
other	when		

+Active verb +Time particle

úyu	cakapíri		'after eating'
to eat	after		

4.16. The quantity intra-clause tagmeme is manifested by a substantive expression:

+Substantive

úpi	'all'
-----	-------

číta	'many'
------	--------

4.17. The purpose intra-clause tagmeme is manifested by a purpose expression, which consists of an obligatory substantive expression plus the obligatory purpose clitic ra ∞ rán:

+Substantive expression +Purpose clitic

t-úka cá	rán	'for my roof'
my roof	for	

4.18. The concern intra-clause tagmeme is manifested by a concern expression which consists of an obligatory stative verb<sub>1</sub> expression, optional object plus obligatory active verb, or substantive expression, plus an obligatory concern clitic:

+Stative verb<sub>1</sub> +Concern clitic

kuatjaranyar-	ári	'concerning writing'
writing	about	

±O +P<sub>1</sub> +Concern clitic

úri yumuyari-	ári	'concerning helping him'
him help	about	

+Substantive expression +Concern clitic

úpi mári	ári	'concerning everything'
all thing	about	

4.19. The cause intra-clause tagmeme is manifested by a cause expression, which consists of an obligatory substantive expression plus the obligatory cause clitic íkuá plus an optional modifying clitic:

+Substantive expression +Cause clitic

ra- mím-	íkuá	'because of his pet'
his pet	because	

+Substantive +Cause clitic ±Modifying clitic

r-	íkuá	purára	'because of him'
him	because	emphatic	

4.20. The instrumental intra-clause tagmeme is manifested by an instrumental expression, which consists of an obligatory substantive expression, hesitation particle, or stative clause nucleus core, plus the obligatory instrumental clitic pu oo púpü plus an optional modifying clitic:

+Substantive expression    +Instrumental clitic    ±Modifying clitic

ya-pua-                    pupü-                    nán                    'by means of  
his hand                    by means of            restrictive            only his hands '

+Hesitation particle    +Instrumental clitic

apú-                                    pu                                    'by means of  
it                                        by means of                        it'

+Stative clause nucleus core +Instrumental clitic

+S

+P<sub>2</sub>

y-    fa                    cačí                                    pu                                    'by means of  
her heart                    pains                                    by means of                        a pained heart'

4.21. The association intra-clause tagmeme is manifested by an association expression, which consists of an obligatory substantive expression plus an obligatory association clitic:

+Substantive expression    +Association clitic

ca-    mímira-                                    múki                                    'with my son'  
my    son                                        with

ya-    mirikuá                                    ndi                                    'with all his  
his family                                    with all                                    family'

4.22. The direction intra-clause tagmeme is manifested by a direction expression which consists of an obligatory stative verb<sub>1</sub> expression, substantive expression, or location expression, plus an obligatory direction clitic plus an optional modifying clitic:

+Stative verb<sub>1</sub>    +Direction clitic

amu-                                    suíka                                    'from far away'  
far away                                    from

+Location particle    +Direction clitic

máka                                    cúi                                    'from where'  
where                                        from

+Substantive expression	+Direction clitic	$\pm$ Modifying clitic	
iruáka kuarači- four days	cúi from	nán restrictive	'after four days'

4.23. The location intra-clause tagmeme is manifested by a location expression which consists of an obligatory substantive expression plus an obligatory location clitic plus an optional modifying clitic, or which consists of a location particle:

+Substantive expression	+Location clitic	$\pm$ Modifying clitic	
kuín iwír- this tree	arí on	ya reportative	'on this tree, it is said'

+Modifying particle

yaúpü	'there'
-------	---------

+Substantive expression +Location clitic

c- uká- my house	ka to, toward	'to my house'
---------------------	------------------	---------------

4.24. The intention-movement intra-clause tagmeme is manifested by an intention-movement expression, which consists of an optional substantive plus an obligatory verb expression plus an obligatory intention-movement suffix:

+Stative verb<sub>1</sub> +IM suffix

ukiri- sleep	tára in order to	'in order to sleep'
-----------------	---------------------	---------------------

+Active verb +IM suffix

kamata- work	tára in order to	'in order to (go) to work'
-----------------	---------------------	-------------------------------

$\pm$ Substantive +Active verb  $\pm$ IM suffix

señóra lady	yumita- teach	tára in order to	'in order to teach the lady'
----------------	------------------	---------------------	---------------------------------

±Substantive +Stative verb<sub>2</sub> +IM suffix

t-	ucu	tára	'in order for me to
I	go	in order to	go'

4.25. The nominative absolute intra-clause tagmeme is manifested by a nominative absolute expression, which consists of an obligatory active clause nucleus core, stative clause nucleus, or identification clause nucleus core, plus the obligatory nominalizing clitic n:

+Active clause nucleus core +nominalizing clitic

+S	+P <sub>1</sub>		'(He died) by
bruhúnu	umanutá	n	a witch doc-
witch doctors	kill		tor killing.'

+Stative clause nucleus +nominalizing clitic

+P <sub>2</sub>	+S	±As	
ukuatári	íni-	ucú-	n
walking	we	go	
			'when we
			went walking'

+Identification nucleus core +nominalizing clitic

+I	+S		
aangatína	kuaračí	n	'when the
here	sun		sun is here'

4.26. The appositive intra-clause tagmeme is manifested by a substantive expression, locative expression, similarity expression, intention-movement expression, instrumental expression, and associative expression. Although other manifestations of the appositive tagmeme have not as yet been found, it seems highly probable that any expression which manifests any intra-clause tagmeme may potentially manifest the appositive tagmeme as well. See examples in Sec. 3.

4.27. The vocative intra-clause tagmeme is manifested by a sub-class of substantive words determined by their membership in the vocative tagmeme. The vocative tagmeme indicates vocative or direct address. See Sec. 3 for an example.

## FOOTNOTES

<sup>1</sup>Cocama, according to Tessman in the Handbook of South American Indians, 3.60, 687-89 (1948), is a language spoken by some 10,000 people living largely in Peru along the Amazon and some of its tributaries. McQuown in "Indigenous Languages of Latin America" American Anthropologist, 57. 522 (1955), lists Cocama as a member of the Upper Amazon Tupian Family.

For a description of the phonemes of Cocama see Faust and Pike, "The Cocama Sound System", Publicações Avulsas do Museu Nacional, Série Lingüística Especial, No. 1, Rio de Janeiro, 1959, pp. 10-55. [The present paper was written in 1958-59 in connection with a study program at the University of Michigan and has, unfortunately, remained unpublished until now. Its contribution to our knowledge of this geographically peripheral Tupian language compensates for a possible lack of relevancy to current theoretical questions.]

<sup>2</sup>Janua Linguarum, Series Major XXIV, Mouton & Co., The Hague, 1967.

<sup>3</sup>The concept of tagmemes-in-matrices used in this paper follows R. E. Longacre's modification of Pike's tagmemic theory. This concept is introduced and developed in Grammar Discovery Procedures, Janua Linguarum, Series Minor XXXIII, Mouton & Co., The Hague, 1964.

<sup>4</sup>One exception to this formula that has been found in the present data is the occurrence of the far past tense clitic cúri within the nucleus core following the subject or predicate tagmeme.

<sup>5</sup>One exception to this formula that has been found in the present data is the occurrence of the far past tense clitic cúri within the nucleus core following the subject tagmeme.

<sup>6</sup>Only one example has been found to date of the occurrence of the recipient tagmeme within the identification clause.

<sup>7</sup>The hesitation particle may occur as manifesting a modifying tagmeme in any phrase matrix and indicates hesitation of the speaker. It is not shown specifically in our regular formulas, although it can occasionally occur in slots where the optional presence of particles is indicated.

<sup>8</sup>One informant accepted only the variant cu in this

position; however other informants have given both variants in what appears to be free variation.

<sup>9</sup>A few others may have to be added to this class. There is a small number of clitics occurring too infrequently in our corpus to allow us to determine whether they belong in the modal or in the tense class.



# INTERNAL CLASSIFICATION OF THE TUPI-GUARANI LINGUISTIC FAMILY

Miriam Lemle

0. Introduction
1. Sources of Data
2. Phonemic Systems
3. Reconstructed Phonemic System of Proto-Tupi-Guarani
4. Vocabulary List
5. Classification

0. Introduction.<sup>1</sup> It is evident from the most superficial inspection of word lists that there is a remarkable similarity among the nearly forty dialects that constitute the Tupi-Guarani family (geographically the most widely spread of any in South America, with representatives from northern Brazil--the Emerillon on the borders of the territory of Amapá and French Guiana--to the Guarani of Argentina in the south, and from the east coast, where Tupinambá was spoken at the time of the discovery of Brazil, to Peru in the west, where today the Kokama still live). Yet hypotheses of possible subgroups within the family need to be tested. Until a few years ago this would have been impossible because of the lack of descriptive data for most of the languages. Now such material is available thanks to the phonological analyses and the filling out of the "Formulário dos Vocabulários Padrões para Estudos Comparativos Preliminares nas Línguas Indígenas Brasileiras" by linguists of the Summer Institute of Linguistics who are working under contract with the Museu Nacional. This material, now available for comparative purposes, is more reliable than previous material collected without the methodological orientation of modern linguistics.

The present study presents a tentative reconstruction of the phonological system and a restricted portion of the lexicon of Proto-Tupi-Guarani insofar as this has been made possible by the data contained in the copies of the "Formulário Padrão" (which we consider to be reliably filled out) and by data from other sources accessible to us. Based on the reconstruction, a classification is proposed, and then dialects which were not used in establishing the classification are assigned to specific branches.

1. Sources of Data. Some of the vocabulary lists for Tupi-Guarani languages filed in the Linguistics Depart-

ment of the Museu Nacional are incompletely filled out or are not phonemicized. For making the reconstruction it was necessary to select the more complete ones and those that had been phonemicized. The following were selected:

Asurini, collected by Carl H. Harrison, 1960, Posto Trocará, Tocantins River, Pará.

Guajajara, David and Margaret Bendor-Samuel, 1960, Posto Gonçalves Dias, Pindaré-Mirim, Maranhão.

Parintintin, Helen Pease and LaVera Betts, 1962, Canavial, Ipixuna River, Amazonas.

Kamayurá, Carl H. Harrison, 1964, Posto Leonardo Villas Boas, Mato Grosso.

Urubu, James Y. Kakumasu, 1963, Posto Pedro Dantas, Canindé, Maranhão.

Guarani, Edna Aaron, 1966, Posto Laranjeiras do Sul, Paraná.

Kokama, Norma Faust, Ucayali, Marañon, and Amazonas Rivers, Peru.

Other sources used from published material were as follows:

Guarayo. Frei Alfredo Hoeller, O.F.M. Mis. Apco., Guarayo-Deutsches Wörterbuch. Guarayos, Dep. S. Cruz de la Sierra, Bolivia. Stuttgart, 1932.

Sirionó. P. Fr. Anselmo Schermair E., Vocabulario Siriono-Castellano. Innsbrucker Beiträge zur Kulturwissenschaft. Innsbruck, 1958. And Perry N. Priest, San Pedro, Departamento Beni, Bolivia, 1960.

Tupinambá. Vocabulário na Língua Brasílica. Second edition, revised and coordinated with Ms. fg. 3144 of the Biblioteca Nacional de Lisboa, by Carlos Drummond. Bulletin no. 137 and 164 of the Serie Etnografia e Tupi Guarani, Faculdade de Filosofia Ciências e Letras of the Universidade de São Paulo, São Paulo, 1953. The phonological analysis is by Aryon dall'Igna Rodrigues, Phonologie der Tupinambá Sprache, doctoral thesis submitted to the University of Hamburg, 1958. Although this is the only one of the languages that is already extinct, it is perhaps the one whose phonemicization is most certain, and we feel that the three centuries that separate it from the other languages do not

invalidate its inclusion in the comparison. Furthermore, there are special reasons for including it to establish its position in the group.

## 2. Phonemic Systems.<sup>2</sup>

Asurini:	p t k ʔ y w m n ŋ h r i e i a o
Guajajara:	p t k ʔ c y w m n ŋ h r i e i ə a u o
Parintintin:	p t k ʔ č y w m n ŋ h r b i e i a u o ĩ ě ĩ ǎ ũ ǒ
Kamayurá:	p t k ʔ č y w m n ŋ h r i e i a u o ĩ ě ĩ ǎ ũ ǒ
Urubu:	p t k ʔ s š y w m n ŋ h r i e i a u o ĩ ě ǎ ũ ǒ
Kokama:	p t k c č s š y w m n h r i ũ a i u
Guarani:	p t k ʔ č y w m n ŋ h r b i e i a u o ĩ ě ĩ ǎ ũ ǒ
Guarayo:	p t k ʔ c č y w m n ŋ h r b i e i a u o ĩ ě ĩ ǎ ũ ǒ
Sirionó <sup>+</sup>	p t k c s š y m n ŋ r b i e i a u o ĩ ě ĩ ǎ ũ ǒ
Tupinambá:	p t k ʔ s š y w m n ŋ r b i e i a u o ĩ ě ĩ ǎ ũ ǒ

### 3. Reconstructed Phonemic System of Proto-Tupi-Guarani.

\*p Before back rounded \*V.

p (As, Gj, Pt, Ub, Ko, Gn, Gy, Tb), h (Km), h, # (Si).  
61, 82, 183, 186.

Before \*w.

p (Tb), k (As, Ub, Gn, Gy), k, k#, # (Si), h (Km).  
9, 70, 119, 157, 173, 191, 206

In other environments.

p (As, Gj, Pt, Km, Ub, Ko, Gn, Gy, Tb), h, # (Si).  
1, 6, 8, 11, 13, 14, 15, 17, 21, 24, 36, 39, 42, 49,  
53, 65, 66, 79, 88, 90, 96, 98, 102, 108, 110, 112,  
121, 139, 140, 144, 152, 155, 160, 162, 164, 165, 168,  
172, 175, 177, 181, 184, 185, 194, 196, 198, 202, 203,  
205, 206, 212, 213, 219.

\*t Before \*i and \*ĩ.

t (Pt, Ko, Tb), č (As, Km, Gn, Gy), c (Gj), č, ʒ (Si),  
ʒ (Ub). 8, 30, 51, 56, 64, 73, 95, 128, 138, 145,  
162.

In other environments.

t (As, Gj, Pt, Km, Ub, Ko, Gn, Gy, Si, Tb). 10, 11,  
29, 50, 53, 55, 80, 83, 87, 90, 91, 95, 96, 97, 100,  
132, 134, 136, 137, 139, 161, 167, 184, 199, 205, 207,  
213.

\*k In final position.

k (Gj, Km, Ko, Tb), k, # (Ub), ŋ, # (As, Pt), # (Gn, Gy,  
Si). 12, 13, 23, 41, 45, 47, 52, 55, 63, 67, 80, 97,  
120, 180, 194, 208.

Before \*w.

k (As, Gj, Pt, Km, Ko, Gn, Gy, Si, Tb), # (Ub). 32,  
56, 73, 170, 195, 199. See also sequence \*kw.

In other environments.

k (As, Gj, Pt, Km, Ko, Gn, Gy, Si, Tb), k, ʒ (Ub). 29,  
33, 39, 43, 44, 46, 61, 64, 66, 68, 75, 76, 80, 85, 97,

107, 112, 116, 124, 125, 127, 129, 135, 150, 163, 171, 176, 178, 182, 186, 187, 188, 192, 197, 200, 216.

Note the development in Ub of the sequence kw in items 76, 124, and 186 under the influence of a preceding u.

- \*ʔ ʔ (As, Gj, Pt, Km, Ub, Gn, Tb), ʔ, # (Gy), # (Ko, Si). 11, 16, 21, 25, 34, 35, 39, 40, 45, 52, 58, 59, 60, 69, 71, 77, 82, 86, 87, 88, 89, 93, 94, 99, 102, 105, 108, 116, 120, 125, 130, 132, 152, 154, 162, 172, 183, 193, 197, 204, 210, 214, 220.
- \*c h, # (As, Gj, Pt), h, y, # (Km), s, h, # (Ub), c, č (Ko, Gy), č, # (Gn), s, š (Si, Tb). 13, 31, 62, 63, 83, 84, 85, 93, 95, 98, 100, 106, 111, 113, 115, 118, 122, 130, 140, 144, 147, 168, 175, 195, 200, 203, 210.
- \*m In final position.  
m (As, Gj, Pt, Km, Ub, Ko, Tb), ẽ# (Gn, Gy, Si). 62, 82, 96, 113, 129, 132, 163.  
Before \*w.  
m (Tb), ŋ (Gn, Gy). 2, 46.  
In other environments.  
m (all languages). 2, 3, 5, 12, 18, 26, 31, 46, 54, 55, 57, 67, 69, 75, 78, 81, 109, 114, 120, 123, 127, 131, 133, 136, 149, 167, 180, 203.
- \*n In final position.  
n (As, Gj, Km, Ko, Tb), n, ẽ# (Pt, Ub), r, ẽ# (Gy), ẽ# (Gn, Si). 50, 54, 78, 79, 123, 139, 157, 174, 175, 220.  
In other environments.  
n (all languages). 20, 24, 48, 55, 108, 120, 127, 131, 143, 149, 151, 185, 189, 218.
- \*ŋ ŋ (As, Gj, Pt, Km, Tb), n (Ko), ẽ# (Ub, Gn, Gy, Si). 30, 33, 69, 86, 95, 145, 150, 192, 211, 212, 213.
- \*y In final position.  
y (As, Gj, Pt, Km, Gn, Gy, Tb), y, #i (Ub, Si), i in monosyllabic words, #i in polysyllabic (Ko). 18, 28, 31, 44, 48, 57, 58, 72, 75, 111, 119, 133, 146, 168, 176, 179, 182, 193.

In other environments.

y (As, Gj, Pt, Km, Ub, Ko, Gn, Gy), ñ, ċ (Si). 7, 20, 27, 31, 37, 45, 78, 83, 86, 107, 115, 117, 124, 135, 143, 148, 154, 155, 156, 169, 188, 196, 200, 205. See also \*py, 208.

\*w As second member of consonant cluster.

w (all languages). 2, 9, 32, 46, 56, 70, 73, 119, 150, 157, 170, 173, 191, 195, 199, 206. See also sequence \*kuw, 187.

In other environments.

w (As, Gj, Km, Ub, Gn, Ko, Gy, Tb), ηw (Pt), g, k (Si). 43, 68, 148, 158, 179, 190, 211, 220.

\*b In final position.

b (Tb), b, # (Pt), p (Km), w, m (As), # after u, w elsewhere (Gj), # (Ub, Ko, Gn, Gy, Si). 7, 16, 19, 26, 34, 68, 71, 89, 92, 100, 110, 114, 126, 137, 151, 153, 166, 171, 178, 187, 191, 198, 202.

In other environments.

w (As, Gj, Km, Ub, Ko), b (Pt, Gn, Gy, Si, Tb). 15, 23, 28, 44, 47, 67, 90, 100, 103, 128, 132, 145, 146, 159, 180, 201, 207, 209, 210, 216, 217.

\*r In final position.

r (Gj, Ub, Ko, Tb), r, # (Pt, Gy), n, r, t (As), t (Km), # (Gn, Si). (Not only should conditions of the alternation n~r~t be reviewed for Asurini, but also its occurrence in the other languages should be investigated. The tentative nature of the grammatical data which we had did not permit a definitive evaluation of the actual situation.) 9, 11, 15, 32, 35, 38, 49, 53, 73, 74, 76, 90, 103, 109, 132, 141, 148, 150, 165, 169, 173, 188, 199, 203, 215.

In other environments.

r (all languages). 27, 31, 46, 60, 63, 81, 87, 93, 99, 104, 107, 114, 142, 156, 158, 159, 164, 165, 182, 185, 191, 195, 212, 221.

\*a Preceding a final nasal consonant.

a (Pt, Km, Ko, Tb), a, ǎ# (Ub), ǎ# (Gn, Gy, Si), ï, o (As), e (Gj). 33, 54, 62, 78, 82, 93, 132, 150, 157, 163, 192, 211, 212, 213. Note also that in Guajajara e occurs in some forms as a reflex of \*a preceding a final syllable that contains a nasal vowel or that is closed by a nasal consonant: 54, 95, 129, 131, 182.

Preceding a final y.

a (As, Gj, Pt, Km, Ub, Gn, Gy, Tb), # (Ko), #, o, e (Si). 31, 119, 176, 179.

In other environments.

a (all languages). 8, 9, 10, 11, 15, 16, 19, 20, 21, 29, 31, 32, 33, 34, 35, 36, 38, 39, 43, 46, 47, 51, 52, 55, 56, 65, 68, 71, 73, 74, 77, 83, 87, 91, 93, 94, 95, 99, 100, 101, 102, 105, 107, 110, 112, 115, 120, 122, 124, 125, 126, 128, 134, 137, 140, 144, 145, 147, 148, 149, 152, 154, 155, 156, 158, 159, 161, 162, 164, 166, 169, 170, 177, 178, 181, 183, 185, 186, 187, 191, 194, 195, 197, 199, 202, 203, 208. But note # reflex in some Sirionó items preceding a final -?V syllable. 116, 193.

\*ï Before a final nasal consonant.

ï (As, Gj, Pt, Km, Ub, Ko, Tb), ǐ# (Gn, Gy), ǐ#, ǐ# (Si). 96, 113, 129

In other environments.

Generally ï (As, Gj, Pt, Km, Ub, Ko, Gn, Gy, Tb), i (Si). 4, 8, 12, 13, 15, 16, 20, 21, 53, 63, 67, 80, 87, 98, 102, 110, 114, 115, 118, 126, 132, 133, 136, 137, 139, 140, 144, 146, 159, 160, 171, 177, 184, 188, 189, 190, 194, 200, 201, 203, 204, 207, 209, 213, 221. But note unexplained irregularities in 1, 22, 25, 26, 28, 38, 44, 47, 64, 83, 85, 89, 90, 93, 108, 117, 145, 158, 167, 168, 195, 197.

\*e Before a final nasal consonant.

e (As, Gj, Km, Tb), ẽ# (Ub, Gn, Gy, Si), I# (Pt), ü (Ko). 69, 86, 123, 220.

After \*k or \*y.

e (As, Gj, Km, Ub, Gn, Gy, Si, Tb), i (Pt), ĩ (Ko). 31, 76, 86, 200, 205.

In other environments.

Generally e (As, Gj, Pt, Km, Gn, Gy, Si, Tb), e, ĩ (Ub), ü, i (Ko). 1, 2, 3, 14, 17, 22, 23, 36, 42, 46, 48, 50, 53, 65, 66, 69, 77, 81, 84, 85, 96, 107, 108, 111, 112, 134, 141, 142, 143, 147, 150, 155, 165, 167, 173, 189, 191, 196, 205, 210, 217, 218, 219, 220. But note irregularities in 26, 58, 151, 204, 208, 216.

\*i Before a final nasal consonant.

i (As, Gj, Pt, Km, Ko, Tb), ĩ# (Ub, Gn, Gy, Si). 30, 79, 95, 145.

In other environments.

Generally i (all languages). 11, 39, 55, 56, 60, 73, 81, 84, 88, 100, 116, 120, 126, 128, 149, 152, 161, 162, 164, 165, 172, 175, 210, 212. But note nasalization of final i following ʔ in some Tb forms. 39, 88.

\*o In final syllable and in penultimate syllable preceding a final syllable with \*o.

o (Gj, Pt, Km, Ub, Gn, Gy, Si, Tb), a (As), u, ua (Ko). 17, 40, 41, 44, 45, 49, 57, 75, 81, 92, 106, 109, 120, 121, 155, 180, 181.

In other environments.

o (As, Pt, Km, Gn, Gy, Tb), u (Gj, Ub, Ko), u, o (Si). 12, 31, 67, 78, 88, 90, 93, 111, 142, 162, 168, 203, 205, 209, 210, 216.

\*u Preceding a final nasal consonant.

u (Gj, Pt, Km, Ko, Tb), u, ü# (Ub), ũ# (Gn, Gy), o (As), ø# (Si). 50, 139, 174, 175.

In other environments.

u (Gj, Pt, Km, Ub, Ko, Gn, Gy, Tb), u, o, i (Si), o (As). 7, 24, 25, 27, 29, 59, 61, 66, 68, 82, 89, 97, 100, 103, 104, 110, 122, 124, 127, 130, 135, 144, 151, 152, 153, 156, 165, 169, 178, 179, 183, 184, 186, 188, 190, 196, 207, 215.

\*ã Following \*Cw.

ã (Gn, Gy, Si, Tb), a (As, Km). 2, 46, 70, 206.

In other environments.

ã (Pt, Km, Ub, Gn, Gy, Si, Tb), a (Ko), o (As), ə (Gj). 24, 72, 98, 108, 135, 167, 182, 185. But note Sirionó ẽ preceding a final y. 72, 182.

\*ĩ ĩ (Pt, Km, Gn, Gy, Tb), i (As, Gj), ě, ĭ (Si), # (Ub, Ko). 58, 133, 146, 193. Note that examples are few and all precede a final \*y.

\*ě ě (Gn, Gy, Si, Tb), ě, e (Pt, Km, Ub), e (As, Gj), i (Ko). 6, 98, 154, 155, 206, 214, 219.

\*ĩ ĩ (Ub, Gn, Gy, Si, Tb), ĭ, i (Pt, Km), i (As, Ko, Gj). 8, 43, 51, 64, 138, 205.

\*ö ö (Pt, Gn, Tb), ö, o (Km), ö, ẽ (Si), ö, o, ũ, u (Ub), o (As, Gj, Gy). 5, 18, 48, 75, 87, 99, 131, 221.

\*ü ü (Gn, Gy, Tb), ũ, ö (Pt), ö (Km, Si), u (Gj, Ko), o (As). 37, 60, 102, 105, 112, 136, 172.

\*kw kw (As, Gj, Pt, Km, Gn, Gy, Si, Tb), w (Ub), ku (Ko). 32, 56, 73, 170, 195, 199.

\*kuw kuw (Tb), kw (all other languages). 68, 187.

\*Cu?V and \*Ci?V are generally reflected as CwV and CyV in Ko. 152, 162, 183. But note exception in 82.

\*V?V in which the vowels are identical is reflected as a single vowel in Ko. 11, 40, 220.

4. Vocabulary List. The reconstructed forms of 221 vocabulary items together with their reflexes in the languages studied are presented below. The gaps that exist are due either to the fact that the item is missing in the original vocabulary list or that the item is not a cognate.

1. \*pipe 'to, in' pipe (As, Pt, Gy), pupe (Gj, Km, Tb), ie (Si).

2. \*memwă 'accident' meŋwă (Gn, Gy), etomemwă (Tb).

3. \*aemee 'sharp' hoyme (As), hayme (Gj, Pt), ayme (Km, Ub, Gn, Gy), aimin (Ko), himey (Si), aemee (Tb).
4. \*i 'water' i (As, Ub, Gy, Tb), ?i (Gj, Km), ihî (Pt), iî (Gn), i (Si).
5. \*amõ 'some' omo (Gj), amuta (Ub), amõwe (Gn), amo (Gy), amõ (Tb).
6. \*pě 'align, braid' pě (Pt, Tb), ě (Si).
7. \*yub 'yellow' iyuahî (Gj), yup (Km), i'un (Ko), yu (Gn), žu (Si), yub (Tb). Cf. 188.
8. \*apitî 'tie' ?apitî (Pt), apitî (Tb).
9. \*pwar 'tie' hwat (Km), pukwar (Ub), kwa (Gn, Gy, Si), pwar (Tb).
10. \*ata 'walk' ata (As, Gj, Tb), wata (Ub, Gn, Gy, Tb), kwata (Si).
11. \*tapi?ir 'tapir' tapi?ira (As), tapi?ir (Gj, Pt, Ub, Tb), tapi?it (Km), tapira (Ko), tapi?i (Gn).
12. \*momik 'tighten' mumik (Gj), momî (Gy), mumi (Si), momik (Tb).
13. \*picik 'squeeze' pihi (As), pii (Gn), picî (Gy), isi (Si), pisik (Tb).
14. \*pe 'that' eope (As), pero (Pt), pe (Gy).
15. \*ibirapar 'bow' iwirapar (As), wirapar (Gj), ibirapar (Pt, Tb), iwirapat (Km), irapar (Ub), wirapa (Gn), rapar (Gy), kiraá, raá ngira (Si).
16. \*?ib 'tree' iw.a (As), iba (Pt), ?ib (Tb).  
\*?ib ?a 'fruit' ib.a (Si). Cf. 94.
17. \*pepo 'wing' pepa (As), pepo (Gj, Pt, Km, Gn, Gy, Tb), pîpo (ub), pûpu (Ko), o (Si).
18. \*amõy 'grandfather' amõy (Pt, Tb), tamũy (Ub), aměy (Si).
19. \*kab 'fat' kaw (As, Gj), ikab (Pt), ikap (Km), iša (Ub), ikawa (Ko), kab (Tb).
20. \*yani 'lard' yani (Gn, Gy, Tb).

21. \*piʔa 'stomach, liver' piʔa (As, Gj, Pt, Km, Ub, Gn, Gy, Tb), pia (Ko), ia (Si).
22. \*ie 'stomach' herie (Gj), ie (Gn, Gy), r-ie (Si), iʔe (Tb).
23. \*ebek 'stomach' eweŋ (As), rebeŋ (Pt), cūwūka (Ko).
24. \*nupā 'hit' nopo (As), nupā (Pt, Km, Gy, Tb), inupa (Ko), noā (Si).
25. \*iʔu 'drink' iʔo (As), iʔu (Gj, Pt, Km, Gn, Tb), uʔu (Ub), yuti (Ko), iu (Gy, Si).
26. \*eimab 'animal' heomaw (As), herəmau (Gj), h-ima (Ub, Gn), eimba (Gy), imba (Si), eimbab (Tb).
27. \*yuru 'mouth' yoro (As), yuru (Gj, Pt, Km, Ub, Ko, Gn, Gy, Tb), čuru (Si).
28. \*bebiy 'float' wewoy (As), bebuy (Pt, Tb), wewiy (Km), wewi (Ub), bebiy (Gy).
29. \*katu 'good' kato (As), katu (Pt, Gj, Km, Ub, Gn, Gi, Tb), katupiri (Gy).
30. \*tiŋ 'white' čiŋ (As), ciŋ (Gj, Km), tiŋahi (Pt), tinin (Ko), čī (Gn, Gy), šī (Si), tiŋ (Tb).
31. \*yemocaray 'play' yemoaray (As), yemuaray (Gj), yiŋwaray (Pt), yemoray (Gn), yemoocaray (Gy), cemesa'ri ~ čumusari (Si), yemosaray (Tb).
32. \*kwar 'hole' kwar (Gy, Tb), kwa (Si).
33. \*akaŋ 'head' akiŋ (As), əkəŋ (Gj), akaŋ (Pt, Km, Tb), akā (Ub, Gn, Gy), yakī (Ko), ākī (Si).
34. \*ʔab 'hair' ʔaw (As, Gj), ʔab (Pt, Tb), ʔap (Km), ha (Ub), ca (Ko), ʔa (Gn), a (Gy, Si).
35. \*ʔar 'fall' ʔan (As), ʔar (Gj, Ub, Tb), ʔa (Pt, Gn), ʔat (Km), u'ari (Ko), a (Gy, Si).
36. \*pe, ape 'path' pe (As, Gj, Ub), pehe (Pt), tape (Km), pū (Ko), -ape (Gn), ape (Tb).
37. \*yū 'field' yū (Pt, Tb).
38. \*iar 'canoe' ihar (As, Pt, Gy), iar (Gj), iat (Km),

- yarusu (Ub), i'ara (Ko), iia (Gn), i'ar (Gy, Tb).
39. \*kapiʔi 'grass' kapiʔi (Gj, Ub, Gn, Gy), kapiʔi (Tb).
40. \*oʔo 'meat' aʔaa (As), hoʔo (Gj), aʔo (Km), oʔo (Ub, Gn, Tb), u (Ko), oo (Gy, Si).
41. \*ok 'house' aŋ (As), oŋa (Pt), ok (Km, Ub, Tb), uka (Ko), o (Gn), oi ~ oig (Gy).
42. \*pe 'bark' pekwer (Gj, Gy), ñipe (Pt), ipe (Km), e, eke (Si), ape (Tb).
43. \*kawĩ 'corn drink' kaŋwĩ (Pt), kawĩ (Tb).
44. \*ibikoy 'dig' iwikay (As), wikiy (Gj), ibikoy (Pt, Tb), čiuki (Ko), ibikoi (Si).
45. \*yoʔok 'dig' yoʔok (Km, Tb), yoʔo (Gn), yoo (Gy), čoo 'remove' (Si).
46. \*karamemwã 'basket' karamewã (Gn, Gy), karamemwã (Tb).
47. \*ibak 'sky' iwaŋ (As), iwak (Gj, Km), ibaŋ (Pt), iwa (?) (Ub), iba (Gn), iba (Si), ibak (Tb).
48. \*enöy 'call' enöy (Pt, Tb).
49. \*por 'full' por (Gj, Tb), po (Gy), o (Si).
50. \*etun 'smell' eton (As), etun (Gj, Pt, Km, Tb), etũ (Ub, Gn, Gy), etö (Si).
51. \*atĩ 'horn' ači (As), atĩ (Pt), ačĩ (Km, Gy), ašĩ (Si).
52. \*ʔak 'horn' ʔak (Gj, Tb), iʔak (Km, Ub), iaka (Ko), akwa (Gn).
53. \*piter 'suck' piten (As), piter (Gj, Ub, Tb), pite (Pt, Gn), pitet (Km), siite (Si).
54. \*aman 'rain' amin (As), əmən (Gj), aman (Pt, Km, Ub, Tb), amana (Ko), amar (Gy), amã (Si).
55. \*tanimuk 'ashes' tanimuk (Gj, Tb), tanimuŋ (Pt), tatimuk (Ub), tanimuka (Ko), tanimu (Gn, Gy).
56. \*kwati 'coati' kwači (Gy), kwati (Tb).

57. \*moy 'snake' may (As), moy (Gj, Pt, Km, Ub, Gn, Gy, Tb), mui (Ko), mey (Si).
58. \*eʔɨy 'scratch' eʔɨy (Pt, Tb), ĩ (Si).
59. \*ʔu 'eat' ʔo (As), ʔu (Gj, Pt, Km, Ub, Gn, Tb), ũyu (Ko), u (Gy, Si).
60. \*ʔirũ 'companion, brother' ʔirũ (Pt, Tb), irõ 'accompany' (Si).
61. \*puku 'long' poko (As), puku (Gj, Pt, Ub, Gn, Gy, Tb), huku (Km), ipukun (Ko), hoko (Si).
62. \*cam 'string' topahom (As), ihəm (Gj), tupaham (Pt, Km), tĭpaham (Ub), ičã (Gn), cã (Gy), sã (Si), sam (Tb).
63. \*cirik 'flow' irik (Gj), čiri (Gn), ciri (Gy), siri 'move' (Si), sĭrik (Tb).
64. \*kitĩ 'cut' kiči (As), kici (Gj), kiti (Pt), kiči (Km), kičĩ (Gn, Gy), kišĩ (Si), kitĩ (Tb).
65. \*ape 'back' ape (As), ʔape (Km).
66. \*kupe 'back' kupe (Gj, Gn, Gy, Tb), kupepĭter (Pt), šupe (Ub), yatukupi (Ko).
67. \*mobik, mobibik 'sew' miwĭk (Gj), miwĭ (Ub), o mo bibi (Gn), mo bibi (Gy), mobibik (Tb).
68. \*akakuwab 'grow' akakuwab (Tb).
69. \*meʔeŋ 'give' meʔeŋ (Km, Tb), meʔě (Ub, Gn), meě (Gy, Si).
70. \*pwã 'finger' ikwakiŋa (As), uã-ě 'fingernail' (Si), pwã (Tb). Cf. 206.
71. \*ʔab 'lie' ʔam (As), ʔaw (Gj, Km), ʔau (Pt).
72. \*ãy 'tooth' oy (As), iy (Gj), ãy (Pt, Km, Ub, Gy, Tb), ai (Ko), ěy (Si).
73. \*kwatiar 'draw' kwačia (Gy), kwatiar (Tb).
74. \*ar 'day' ar (As), ʔar (Gj, Tb), ʔat (Km), ara (Gn), ar(i) 'sun' (Gy), arõ (Si).
75. \*mokõy 'two' mokoy (As, Gj), mokõy (Pt, Km, Ub, Gn,

- Si, Tb), mu'kui (Ko).
76. \*ker 'sleep' ker (As, Gj, Tb), ki (Pt), ket (Km), ukwer (Ub), ukiri (Ko), ke (Gn, Si), k'ye (Gy).
77. \*aʔe 'he' aʔe (As, Gj, Km, Ub, Tb), ai (Ko), haʔe (Gn), ae (Gy, Si).
78. \*moayan 'push' muyən (Gj), omoayă (Pt), yumuyani (Ko), moayă (Gn), moayăi (Gy), maňă (Si), moayan (Tb).
79. \*pin 'rub' pin (As, Km, Tb), pĩ (Gn, Gy), ĩ (Si).
80. \*kitik 'rub' kitik (As), kitik (Gj, Ub, Tb), kiti (Pt, Gn, Gy), giti ~ kiti (Si).
81. \*emireko 'wife' (h)emireko (Gj, Km), (h)emirekohë (Pt), mirikua (Ko), emireko (Gn, Gy, Tb).
82. \*puʔam 'stand' poʔom (As), puʔəm (Gj), puʔam (Ub, Tb), ipamawa (Ko), puă (Gn, Gy), uă (Si).
83. \*yacitata 'star' yahita'ta (As, Gj), yaitataʔi (Pt), yaitataʔi (Km), yahirata (Ub), yačitata (Gn), časi-tata (Si), yasi-tata (Tb).
84. \*(i)ce 'I' iyee (As), ihe (Gj), iye (Km), ihë (Ub), ücü (Ko), čee (Gn), če (Gy), se (Si), (i)še (Tb).
85. \*kice 'knife' kihe (As), takihe (Gj), kie (Km), kise (Ub, Tb), kičikira (Ko), kiče (Gn), kice (Gy), kise (Si).
86. \*yeʔeŋ 'speak' yeʔeŋ (As, Gj, Km, Tb), yiʔĩ (?) (Pt), yeʔë (Ub), yeë (Gy), ñeë ~ čeë (Si).
87. \*ʔitarö 'full' ʔitarö (Pt, Tb).
88. \*poʔi 'thin' ipuʔiahiaʔi (Gj), poʔi (Km, Gn, Gy), puʔi (Ub), oi (Si), poʔĩ (Tb).
89. \*uʔib 'arrow' oʔiw (As), uʔiw (Gj), iʔip (Km), 'ua (Ko), uʔi (Gn), uhu (Gy), uu (Si), uʔub (Tb).
90. \*potir, ibotir 'flower' iwotira (As), putir (Gj, Ub), ibatiri (Pt), ipotit (Km), potir (Gy, Tb), uti (Si).
91. \*tata 'fire' tata (As, Pt, Km, Ub, Ko, Gn, Gy, Si, Tb), hata (Gj).
92. \*ob 'leaf' op (Km), o (Gn, Gy, Si), ob (Tb).

93. \*roʔi, roʔica 'cold' roihĩ, roʔioho (As), ruic (Gj), irotĩ (Pt), iroʔica (Km), risã (Ub), iroʔiçã (Gn), roi, roica (Gy), rui (Si), roʔi, roʔisa (Tb).
94. \*ʔa 'fruit' ʔa (Gj, Km, Gn, Tb), a (Gy, Si). Cf. 16.
95. \*tatatiŋ, catatiŋ 'smoke' tata-čĩŋ (As), hətəcin (Gj), tatatiŋ (Pt, Tb), tatačĩŋ (Km), tataši (Ub), tatatini (Ko), tatačĩ (Gn, Gy, Si).
96. \*petim 'tobacco' petim (As, Km, Tb), petipiar (Gj), pĩtim (Ub), pũtima (Ko), petĩ (Gn).
97. \*kutuk 'pierce' kotoŋ (As), kutuk (Gj, Km, Ub, Tb), kutu (Pt, Gn, Gy), kĩtika (Ko).
98. \*picãpẽ 'claw' piãpe (Gj), piãpẽ (Ub, Gn), picapi (Ko), piãpẽ (Tb).
99. \*ʔarõ 'like' ʔarõ (Pt, Tb), arõ (Si).
100. \*tubicab 'large' tuwiyap (Km), tiha (Ub), tubiça (Gn, Gy), tubișab (Tb).
101. \*aba 'man' awaceʔi (As), awa (Gj), sawaʔe (Ub), 'awa (Ko), aba (Gn, Gy, Si, Tb).
102. \*ipaʔũ 'island' ipaʔõ (Pt), ipaʔũ (Tb).
103. \*bubur 'swell' bubu (Pt), wuwut (Km).
104. \*ruru 'swell' iruru (Ko), i-ruru (Gn), ruru (Gy, Si).
105. \*paʔũ 'space' paʔõ (Pt), paʔũ (Tb).
106. \*co 'go' oho (Gj, Pt, Ub), ucu (Ko), o-o (Gn), co (Gy), so (Si, Tb).
107. \*yakare 'alligator' ya'kare (As), yakare (Gj, Km, Ub, Gn, Gy, Tb), yãkare (Pt), yakari (Ko), nikare (Si).
108. \*enipĩʔã 'knee' renipĩʔã (Pt), cünipa (Ko), renapĩʔã (Gn), enĩpia (Gy), eniã (Si), enipĩʔã (Tb).
109. \*momor 'throw' maman (As), momor (Gj, Tb), momo (Pt, Gy, Si), momot (Km), omor (Ub), yumuc (Ko), o-momo (Gn).
110. \*ipab, iupab 'lake' ipaw (Gj), iwpap (Km), ipa (Ub), i-upa (Gn), upab (Tb).

111. \*yocey, ey, c-ey, yac-ay 'wash' hey (As), yuhey (Gj), pi<sup>h</sup>ey (Km), yocey (Gy), yosey (Tb).
112. \*ape-kũ, kũ 'tongue' ko (As), apeku (Gj), kũ (Pt, Gy), kō (Km), kumira (Ko), apekũ (Gn, Tb), ekō (Si).
113. \*cīm 'smooth' him (As, Gj, Ub), ihmāhī (Pt), yim (Km), icīman (Ko), čīī (Gn), cī (Gy), sim (Tb), ī, šī 'rough' (Si).
114. \*...mirib 'far' momiri (Gn), amomiri (Gy), wimirib (Tb).
115. \*yaci 'moon' yahī (As, Gj, Pt, Ub), yai (Km), yaci (Ko, Gy), yači (Gn), časi (Si), yasi (Tb).
116. \*kaʔi 'monkey' 'kaʔi (As), kaʔi (Gj, Pt, Km, Ub, Gn, Tb), kai (Ko, Gy), ki (Si).
117. \*yi 'ax' yi (As, Km, Ko, Gy, Tb), itayi (Gj), yīkwaʔri (Pt), yi (?) (Ub).
118. \*ci 'mother' hikee (As), hi (Gj), ihē (Pt), i (Km), čī (Gn), ci (Gy), si (Si), sī (Tb).
119. \*pway 'command' kway (Gy), kwa (Si), pway (Tb).
120. \*maniʔok 'manioc' maniʔaŋ (As), maniʔok (Gj, Ub, Tb), maniʔoŋ (Pt), maniʔo (Gn), manio (Gy).
121. \*po 'hand' pa (As), po (Gj, Pt, Ub, Gn, Gy, Tb), pua (Ko), o (Si).
122. \*acu 'left hand' (nepo ne)ahur (Gj), aču (Gn), (po)acu (Gy), asu (Tb).
123. \*men 'husband' man (As, Gj, Tb), mūna (Ko), mē (Gn, Si), mēr (Gy).
124. \*yuka 'kill' yoka (As), yuka (Gj, Pt, Km, Ko, Gn, Gy, Tb), yukwa (Ub), ika / ika (Si).
125. \*kaʔa 'woods' 'kaʔaa (As), kaʔa (Gj, Pt, Km, Ub, Gn, Tb), kaa (Gy), kʔaa (Si).
126. \*aib, aib (?) 'bad' ai (Km), ahi (Ub), ai (Gy), ay (Si), aib (Tb).
127. \*kunumī 'boy' kunuʔum (Km), kurumī (Ub), kunumi (Gy), kunumī (Tb).

128. \*abati 'corn' awači (As), awaci (Gj, Km), abati (Pt, Tb), awaši (Ub), aw (Ko), abači (Gn, Gy), ibaši / abači (Si).
129. \*akim 'wet' 'akim (As), ə'kim (Gj), akim (Pt, Ub, Tb), ʔakim (Km), akĩ (Gn, Gy).
130. \*cuʔu 'bite' oʔo (As), huʔu (Pt), uʔu (Km), suʔu (Ub), čuʔu (Gn), cuʔu (Gy), suʔu (Tb).
131. \*manǝ 'die' mǝno (Gj), manǝ (Pt, Gn, Si, Tb), mano (Km, Ub, Gy), uma'nwari (Ko).
132. \*ibitir, ibiʔam (+ -usu) 'hill' iwitir (As, Gj), ibiʔamuhu (Pt), iwiʔam (Ub), iwata (Ko), ibiti (Gn), ibitir (Gy, Tb), ibi (Si), ibiʔam 'ravine' (Tb).
133. \*mǝy 'move' mǝy (Pt, Tb).
134. \*eta, c-eta 'many' eta (As, Gj, Ub, Gy, Si, Tb), čita (Ko), heta (Gn).
135. \*kuyǎ 'woman' koyo (As), kuyǝ (Gj), kuyǎ (Pt, Km, Ub, Gn, Tb), 'kuyǎ (Gy).
136. \*mitũ 'variety of bird' mitũ (Pt, Tb), mitǝ (Si).
137. \*itab 'swim' itaw (Gj), ita (Pt, Gy), ʔitap (Km), ʔita (Gn), ita (Si), itab (Tb).
138. \*tĩ 'nose, beak' či (As), ci (Gj), tĩ (Pt, Tb), čĩ (Km, Gn, Gy), ti (Ko), šĩ (Si).
139. \*pitun 'night' ipiton (As), pitun (Gj, Ub, Tb), ipitun (Pt), ʔipitun (Km), ipitunin (Ko), pitũ (Gn, Gy), itǝ (Si).
140. \*pica 'night' ipica (Ko), isa (Si), pisare (Tb).
141. \*er 'name' er (As, Gj, Pt, Ub, Tb), et (Km), čira (Ko), e / ere (Si).
142. \*ore 'we (excl.)' oree (As), ure (Gj, Si), ore (Pt, Km, Gn, Gy, Tb).
143. \*yane 'we (incl.)' yanee (As), yane (Gj, Pt, Ub, Gn, Gy, Tb), yene (Km), ini (Ko), ñane, nane (Si).
144. \*picacu 'new' piahu (Gj, Pt, Ub), piau (Km, Gn), ipicacun (Ko), piacu (Gy), pisasu (Tb), iasu (Si).

145. \*ibatiŋ 'cloud' iwaciŋ (Gj), iwaciŋ (Km), ibĩčĩ (Gy), ibeisĩ (Si), ibatiŋ (Tb).
146. \*ibĩy 'hollow' ibĩy (Pt, Tb).
147. \*eca 'eye' eha (As, Gj, Ub), eakwar (Pt), ea (Km), čicakuara (Ko), eča (Gn), eca (Gy), esa (Si, Tb).
148. \*yawar 'jaguar' yawar (As, Gj, Gy, Tb), yãʔŋwãr (Pt), yawat (Km), yaŋwate (Ub), yaware (Ko), yawa (Gn), ñaka (Si).
149. \*nami 'ear' nami (As, Gj, Pt, Km, Ub, Ko, Gn, Gy, Tb).
150. \*kaŋ, kaŋ-wer 'bone' kiŋ (As), kəŋ (Gj), kaŋ (Pt, Km, Tb), šaŋwer ~ kaŋwer (Ub), kanwara (Ko), kã gwe (Gn), kã (Gy), kěy / kaŋ (Si).
151. \*enub 'hear' enom (As), enu (Gj, Pt, Ub, Gn, Gy), cunu (Ko), anu (Si), enub (Tb).
152. \*upiʔa 'egg' opiʔa (As), upiʔa (Gj, Pt, Km, Gn, Tb), piʔa (Ub), cupya (Ko), upia (Gy), [uθia]/ečia (Si).
153. \*ub 'father' ow (As), u (Gj, Ub, Gn, Gy, Si), ub (Pt, Tb), up (Km).
154. \*yaʔě 'pan' yaʔee (As), yaʔe (Km), yaě (Gy), ya'ʔě (Tb).
155. \*yaěpepo 'pan' yapepo (Gj, Ub, Gn), yaĩpepo (Pt), yaěpepo (Tb).
156. \*ayuru 'parrot' ayoro (As), ayuru (Gj, Km, Gy, Tb), ayuruʔi (Pt), čuru (Si).
157. \*pwan 'pass' kwã (Gy), kwa (Si), pwan (Tb).
158. \*wira 'bird' wira (As, Km, Ub, Ko, Gn, Gy, Tb), wirəmiri (Gj), ŋwiraʔi (Pt), gira (Si).
159. \*ibira 'stick' iwira(ʔi) (As, Km), iwira (Gj), ibira(ʔi) (Gn), ibira (Gy, Tb), ira (Si).
160. \*pi 'foot' pi (As, Gj, Pt, Km, Ub, Gn, Gy, Tb), pita (Ko), i (Si).
161. \*ita 'rock' ita (As, Gj, Km, Ub, Gn, Gy, Si, Tb), itaki (Pt), itaki (Ko).

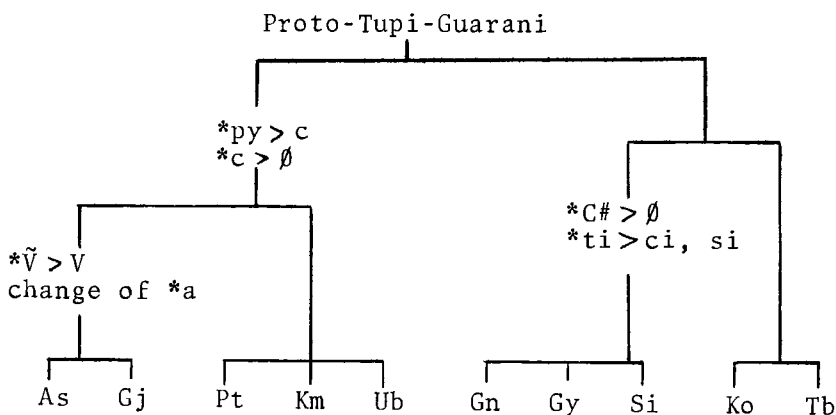
162. \*poti<sup>?</sup>a 'chest' poti<sup>?</sup>a (Pt, Tb), poci<sup>?</sup>a (Km), piši<sup>?</sup>a (Ub), putyakwara (Ko), poči<sup>?</sup>a (Gn), počia (Gy), isia (Si).
163. \*kam 'breast' kom (As), kəm (Gj), kami (Ub), kā (Gy, Si), kam (Tb).
164. \*pira 'fish' ipira (As, Km, Ko), pira (Gj, Pt, Ub, Gn, Gy, Tb), hira ~ sira (Si).
165. \*pir, piruer 'skin' pirer (As, Gj, Ub, Gy), ahampir (Pt), pīret (Km), piruara (Ko), pire (Gn), i (Si), pir, pīrwer (Tb).
166. \*ab, c-ab, c-a-wer 'feather' aw (As), hawer (Gj), ab (Pt, Tb), ap (Km), ca (Ko), a (Gy, Si).
167. \*etimã 'leg' etimoci<sup>?</sup>a (As), etimə (Gj), etimakaŋ (Km), timã (Ub), etimã (Gn, Gy, Tb), etimã (Si).
168. \*pociy 'heavy' pohoy (As), puhiy (Gj), pohiy (Pt), powiy (Km), puhī (Ub), i'puci (Ko), poiy (Gn), pociy (Gy), usi (Si), posiy (Tb).
169. \*ayur 'neck' yor (As), ayu<sup>?</sup>iw (Gj), yur (Pt), ayut (Km), ayu<sup>?</sup>u (Gn), ayu (Gy), aču / ečurui (Si), ayur (Tb).
170. \*akwa 'person' akwaime<sup>?</sup>e (Pt), akwama<sup>?</sup>e (Km).
171. \*kib 'louse' kiw (As, Gj), kib (Pt, Tb), kip (Km), iši ~ ki (Ub), kiwa (Ko), ki (Gn, Gy), ki (Si).
172. \*pi<sup>?</sup>ũ 'variety of gnat' pi<sup>?</sup>ũ (Pt, Tb).
173. \*pwer 'past tense' -ke (Si), -pwer (Tb).
174. \*un, c-un 'black' cunin (Ko), ũ (Gn), cū, hū (Gy), ō (Si), un (Tb).
175. \*picun 'black, dark' pihon (As), pihun (Gj, Ub), picun (Km), išō (Si), pišun (Tb).
176. \*kay 'burn' kay (As, Gj, Pt, Km, Gn, Gy, Tb), kway (Ub), uki (Ko), key (Si).
177. \*api 'burn' api (As, Gj, Km, Gn, Tb).
178. \*akub 'hot' akom (As), aku (Gj, Ub, Gn, Gy, Si), hakub ahi (Pt), akup (Km), caku (Ko), akub (Tb).

179. \*uway 'tail' way (As, Km), uay (Gj), wayã (Pt), uway (Ub, Gn, Gy, Tb), cui (Ko), okoy (Si).
180. \*mobok, bok 'split' mowok (Km, Ub), mobo (Gn, Gy), moṃbo, bo (Si), mobok (Tb).
181. \*apo 'root' iwapa (As), apo (Gj, Km, Ub, Gn, Gy, Tb), hapo (Pt), capwa (Ko), ao (Si).
182. \*karãy 'scrape' kariy (As), kərəy (Gj), karãy (Pt, Gn, Gy, Tb), kari (Ko), karëy (Si).
183. \*apu?a 'round' iapo?a (As), iapu?a (Gj), i?ahu?a (Km), yupu?a (Ub), ya'pwan (Ko), iyapu?a (Gn), apua (Gy), hua (Si), apu?a (Tb).
184. \*pitu 'breathe' pitoekic (As), pituhem (Gj), ipitu (Pt), yepitu (Km), pituë (Gn, Gy, Tb), itu 'breath' (Si).
185. \*paranã 'river' parano (As), parana (Km, Ub, Ko), paranã (Gn, Tb).
186. \*puka 'laugh' poka (As), puka (Gj, Gn, Gy, Tb), huka (Km), pukwa (Ub), apukari (Ko), ika (Si).
187. \*kuwaab 'know' kwaham (As), kwaw (Gj), kwaha (Pt), kwahap (Km), kwa (Ub), ikwa (Ko), kwaa (Gn, Gy), kwa (Si), kuwab (Tb).
188. \*yukir 'salt' yokiri 'green' (As), yukir (Gj, Pt, Ub, Gy, Tb), yokit (Km), yuki (Gn). Cf. 7.
189. \*eni 'saliva' eni (As, Gj, Pt, Km, Gy(?), Tb), eni (Si).
190. \*uwi 'blood' owi (As), iwikwer (Gj), ṃwĩhi (Pt), wi (Km), uwi (Ub, Gn, Gy, Tb), cui (Ko), uki (Si).
191. \*pwerab 'heal' kwera (Gy), kera (Si)(?), pwerab (Tb).
192. \*kaṅ 'dry' kã (Si), kaṅ (Tb).
193. \*a?iy 'seed' a?iy (As), i?iy (Gj), a?iy (Pt, Km, Gn, Tb), a?i (Ub)(?), cai (Ko), aiy (Gy), ëy (Si).
194. \*apik 'sit' apiṅ (As), apik (Gj, Km, Tb), api (Pt)(?), mapik (Ub), yapika (Ko), api (Gn), wapi (Gy).
195. \*kwaraci 'sun' kwarahi (As, Gj), warahi (ub), kuarači (Ko), kwarai (Gn), kwarasi (Tb).

196. \*peyu 'blow' peyo (As), peyu (Pt, Ub, Gn, Gy, Tb).
197. \*kiʔa 'dirty' šiʔa ~ kiʔa (Ub), kiʔa (Gn, Tb), kia (Gy), kia (Si).
198. \*ipib 'dirty' ipiw (Pt), ipibuhu (Km).
199. \*takwar 'bamboo' takwar (Gy, Tb), takwa (Si).
200. \*cikiye 'fear' kiiye (As), kiye (Gj, Km, Ub, Gn), kihäyi (Pt), cikiye (Gy), sikiče (Si), sikiye (Tb).
201. \*ibi 'land' iwī (As, Gj, Km, Ub), ibi (Pt, Gn, Gy, Tb), ibi (Si).
202. \*pab 'all' paw (Gj), pabëy (Pt), upa (Ub), pabë (Gn), opa (Gy), ha (Si), (o)pab (Tb).
203. \*mocapir 'three' mohapī (Pt), moʔapit (Km), mapir (Ub), mucapirika (Ko), moapī (Gn), mocapī (Gy), mosapir (Tb).
204. \*iʔe 'intestines' ieʔi (As), ie (Gj, Gn), iapoyub (Pt), i.ʔiepo (Km), ie (Si), iʔe (Tb).
205. \*oyepeteĩ 'one' oyepe (As, Tb), pitei (Gj), oyipeyĩ (Pt), oyepete (Km), pete'ĩ (Ub, Gn), wüpu (Ko), yepeĩ (Gy).
206. \*pwã-pë 'fingernail' koá-pe (As), hwa-pe (Km), pë kuã (Gy), oãë (Si), pwã pë (Tb). Cf. 70.
207. \*ibitu 'wind' iwīto (As), iwītu (Gj, Km, Ub, Ko), ibitu (Pt, Gn, Gy, Tb).
208. \*epyak 'see' ečaŋ (As), ečak (Gj, Km), usak (Ub), esa (Gn), epia (Gy), ea (Si), epyak (Tb).
209. \*obi 'green' huiahi (Gj), hobiahi (Pt), čowi (Km), howi (Ub), owi (Gn), obi (Gy, Tb), ubi (Si).
210. \*ceboʔi 'worm' ewoʔi (As, Km), eboʔi (Pt), ceboʔi (Gy).
211. \*waŋ 'red' i.ŋwaŋ (Pt), i.waŋ (Km), waŋ(Tb).
212. \*piraŋ 'red' piroŋ (As), pireŋ (Gj), pirã (Gy), irã (Si), piraŋ (Tb).

213. \*pītaŋ 'red' pītā (? Km, Gn), pītanin (Ko), itā (Si), pītaŋ (Tb).
214. \*(?)ē 'pour' ?ē (Pt), ē (Tb, Si).
215. \*ur 'come' on, ot (As), hu (Pt), ?ut (Km), u (Gn, Gy, Si), ur (Tb).
216. \*eko, ekobe 'live' iko (Gj), kowe (Km), šuwe (Ub), ikobe (Gn, Gy, Si), eko, ekobe (Tb).
217. \*bebe 'fly' wewe (As, Km), ewew (Gj), bebe (Pt, Gn, Si, Tb), uwū (Ko).
218. \*ne, ene 'you (sing.)' enee (As), ene (Km, Tb), ünü (Ko), ne (Gj, Pt, Ub, Gy, Tb, Si), nee (Gn).
219. \*pe...ē 'you (pl.)' pehee (As), pehe (Pt, Ub), pehē (Km), üpü (? Ko), peč (Gn, Gy), hē (Si), pe?ē (Tb).
220. \*we?en 'vomit' we?en (As, Ub, Tb), u'wünü (Ko), weč (Gy).
221. \*irō 'mad' irō (Pt, Tb), irō (Si).

5. Classification. Based on retained items and innovations common to the phonemic systems of the languages studied as compared to the system reconstructed for the proto-language, the following tree diagram representing the genetic classification of this group of languages is obtained:



The following statements may be made on the basis of a rapid inspection of certain Tupi-Guarani languages which, because of lack of data, were not used in the reconstruction.

Tapirapé (data from Yonne Leite) has much in common with Asurini, including the loss of contrast between \*u and \*o, the merging of certain other vowels with \*a, and the phonetic manifestation of /y/ as [y] in syllable-final position and as [ç] in other environments. However, its inclusion in the same section of the diagram as Asurini is difficult, since vowel nasalization remains as a phonemic feature of Tapirapé but not of Asurini and Guajajara.

Kayabi (data from Helga Weiss and Rose Dobson) and Emerillon (data from Vaughn Collins) have a zero reflex of \*c and phonemic vowel nasalization. Thus they belong to the Parintintin-Kamayura-Urubu branch. Kayabi appears to be closest to Parintintin.

Kaiwá (data from Loraine Bridgeman) and Xetá da Serra dos Dourados (data from Aryon Rodrigues) are similar to Guarani, including the characteristically Guarani distribution of zero and /c/ reflexes of \*c.

#### FOOTNOTES

<sup>1</sup>The present study was made possible through extended consultation with Prof. Aryon Rodrigues, author of a classification of the Tupi linguistic stock and of various studies of Tupinambá. Without his counsel and help in solving a number of problems we would not have been able to progress in the study. His assistance is here gratefully acknowledged.

<sup>2</sup>In order to standardize the use of phonemic symbols, we made certain substitutions for symbols used by the original authors. In most cases the substitution was on a one to one basis, thus not implying a different phonemicization. In cases in which we found complementary distribution, we took the liberty of substituting one symbol for various others. We have also reinterpreted /k<sup>w</sup>/ and /ŋ<sup>w</sup>/ as sequences, /kw/ and /ŋw/, rather than single phonemes.

Substitutions made in the various languages are as follows:

As: k<sup>w</sup> > kw; č > y. Gj: k<sup>w</sup> > kw; ŋ<sup>w</sup> > ŋw; z > y. Pt: k<sup>w</sup> > kw; j, ŷi, ŷñ > y; g<sup>w</sup> > w; g > ŋ; v > b; ŷi > ŷy; ŷñ > ŷy.  
 Km: k<sup>w</sup> > kw; m<sup>w</sup> > m; h<sup>w</sup> > h. Ub: k<sup>w</sup> > kw; ŋ > ŋ; ng<sup>w</sup> > ŋw.  
 Ko: iy > yv; uy > wv; Gn: k<sup>w</sup> > kw; j, ñ, ŷi > y; g<sup>w</sup>, ŋ<sup>w</sup>, ŷw > w; v > b; ~ on last nasal vowel only; ŷi > ŷy; Gy: k<sup>w</sup> > kw; z > c; ch > č; mb > m; nd > n; ng > ŋ; v > b; gu > w; y, ŷi, ŷv > y; ñv > y<sup>v</sup>.