THE PHONOLOGY OF CAPANAHUA AND ITS GRAMMATICAL BASIS

[inside front cover]

THE PHONOLOGY OF CAPANAHUA AND ITS GRAMMATICAL BASIS

SUMMER INSTITUTE OF LINGUISTICS PUBLICATIONS
IN
LINGUISTICS AND RELATED FIELDS

PUBLICATION NUMBER 20

THE PHONOLOGY OF CAPANAHUA AND ITS GRAMMATICAL BASIS

EUGENE EMIL LOOS

A Publication of the
SUMMER INSTITUTE OF LINGUISTICS
OF THE
UNIVERSITY OF OKLAHOMA
Norman

agosto, 1969 primera edición

Esta edición consta de 550 ejemplares

Impreso en México Printed in Mexico
5.5C 9-048

EDITOR'S NOTE

We welcome to this S.I.L. series a volume which uses the generative-transformational model as background. Eugene E. Loos, of the Peru branch of S.I.L., with experience as field worker, administrator and teacher, has given us perhaps the first study of a Panoan language cast in the form of a generative grammar, with its valuable emphasis upon the predication and refinement of linguistic universals. Structural studies in this series have relied largely upon Pike's tagmemic model (1955), e.g., Totontepec Mixe Phonotagmemics (#8), Chatino Syntax (#12), and Bolivian Grammars (#16). The editorial committee, however, which helps to select and approve volumes for publications, hardly represents a monolithic block of confirmed tagmemicists. Nor does the committee desire to project the myth that only tagmemic studies are acceptable for publications in this series. The Tzotzil Grammar (#18) rests upon the stratificational model, as modified by William E. Merrifield. Other generative studies of Middle and South American languages are in project for publication. It is our hope that this series will reflect faithfully the linguistic philosophy and practice of over 50 field workers, from several continents, who have completed graduate studies in linguistics and, as well, faithfully represent the linguistic studies of more than 2000 field workers whose intimate contact with indigenous peoples and fluency in their representative languages and dialects represent a prime source for structural studies from the worlds' languages.

Dow F. Robinson
Tlalpan, Mexico, D.F.
May 20, 1969

THE PHONOLOGY OF CAPANAHUA AND ITS GRAMMATICAL BASIS

by

EUGENE EMIL LOOS, B.A., B.D., M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas in Partial Fulfillment

of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS

Austin, Texas

August, 1967

ACKNOWLEDGEMENTS

I owe a special debt of gratitude to Professor Robert T. Harms whose many helful criticisms and hours of discussion have been invaluable in the preparation of this dissertation.

I also wish to express appreciation to the University of Texas for the generous fellowship that aided me during this past year.

June, 1967

CONTENTS

ABBREVIATIONS

A	adjective	Neg	negative
Adv	adverb	NP	noun phrase
Aj	adjective	Num	numeral
An	animate	PP	predicate phrase
AP	adjective phrase	Pro	pronoun
Asp	aspect	PS	phrase structure
Aux	auxiliary	Q	interrogative
Cond	condition	R	rewrite rule
CS	complex symbol	S	sentence
Decl	declarative mood marker	SC	sequential constraint (in MS component)
Excl	exclamation	SC	structural change (in T rules)
Exclm	exclamatory mood marker	SI	structural index
Imp	imperative	SS	segment structure
Md	mood indicator	T	transformational
MP	morphophonemic	V	verb stem
MS	morpheme structure	Vb	verb
N	noun	VP	verb phrase
Nar	narrative conjunction		
addr	addressee	cns	consonantal
cmp	compact	cnt	continuant

del	deletion
flt	flat
grv	grave
nas	nasal
obs	obstruent
per	peripheral
pit	pitch
pl	plural
pp	postpositional phrase
spkr	speaker
std	strident
strs	stress
syl	syllabic
voc	vocalic

I. INTRODUCTION

One of the goals of generative grammar is to formulate an evaluation measure that will determine the most highly valued formulation of a grammar. The evaluation measure is much more than counting features; it is the full set of notational conventions (Chomsky, 1965. 44). It is claimed that these conventions account for the ability of language learners to construct uniform grammars on the basis of exposure to disparate data.

> It is implied that a person learning a language will attempt to formulate generalizations that can easily be expressed (that is, with few symbols) in terms of the notations available in this theory, and that he will select grammars containing these generalizations over other grammars that are also compatible with the given data but that contain different sorts of generalization, different concepts of "natural class," and so on. These may be very strong claims, and need by no means be true on any a priori grounds. (Ibid., 45)

So far there have not been many attempts to test the evaluation measure on a full subsection of the grammar of a non-Indo-European language. It is therefore of theoretical interest to see if the conventions established will be equally applicable in a fairly comprehensive study of a language quite different from Indo-European.

A second thrust of current linguistics is to search for those data that support or refute conjectures about

language universals in order to make the strongest possible claims for the theory of language by restricting the possible forms of grammars. In this endeavor linguistic theory has probably not yet made claims as strong as those asserted for the evaluation measure:

> . . . we are very far from being able to present a system of formal and substantive linguistic universals that will be sufficiently rich and detailed to account for the facts of language learning. To advance linguistic theory in the direction of explanatory adequacy, we can attempt to refine the evaluation measure for grammars or to tighten the formal constraints on grammars so that it becomes more difficult to find a highly valued hypothesis compatible with primary linguistic data. . . . Thus the most crucial problem for linguistic theory seems to be to abstract statements and generalizations from particular descriptively adequate grammars and, wherever possible, to attribute them to the general theory of linguistic structure, thus enriching this theory and imposing more structure on the schema for grammatical description. Whenever this is done, an assertion about a particular language is replaced by a corresponding assertion, from which the first follows, about language in general. . . . In short, I am making the obvious comment that, wherever possible, general assumptions about the nature of language should be formulated from which particular features of the grammars of individual languages can be deduced. (Ibid., 46)

> In general, it should be expected that only descriptions concerned with deep structure will have serious import for proposals concerning linguistic universals. Since descriptions of this sort are few, any such proposals are hazardous, but are clearly no less interesting for being hazardous. (Ibid., 210)

The profusion of languages in South America makes it a rich hunting ground for the linguist. The names of more than 1500 South American Indian languages have been recorded, though as Alexander von Humboldt once observed, a more conservative figure would be more realistic:

> The historians of the Conquest, as well as the ecclesiastics who have described the progress of the Missions continually confound, like the ancients, geographical denominations with the names of races. They speak of Indians of Cumana and of the coast of Paria, as if the proximity of abode proved the identity of origin. They most commonly even give to tribes the names of their chiefs, or of the mountains or valleys they inhabit. This circumstance, by infinitely multiplying the number of tribes, gives an air of uncertainty to all that the monks relate respecting the heterogeneous elements of which the population of the Missions are composed. (1852, 298)

There were, nevertheless, a great many languages, and the picture has not been much clarified since Humboldt's day because, compared with the present state of language classification and description in North and Central America, linguistic work is just beginning in South America.

Among those which were early recognized as distinct language families was the Panoan group (De la Grasserie, 1888), which in a more recent grouping (McQuown, 1955) is listed as comprising some eighty different languages. Undoubtedly the problem of name reduplication has caused the Panoan number to swell beyond what would have been established if accurate linguistic data had been available, but the Panoans were still a significant block among the Indian languages of South America. Today there are perhaps only twenty or so Panoan languages in active use, scattered between the Amazon in Peru and Brazil into northern Bolivia. Most of these languages are spoken by small groups in various stages of acculturation.

The Capanahuas are among the more acculturated. Most of the younger adults are bilinguals (Spanish-Capanahua) with

an increasing tendency to prefer Spanish. They live on the Tapiche and Buncuya Rivers in eastern central Peru and number about 400 speakers of which about eighty speak a slightly divergent dialect not included in this study.

There have been, to my knowledge, no previous generative phonologies of a Panoan language. Such descriptions as exist are not concerned with the questions that are currently being asked in linguistics to construct an explanatory, unified theory of language. There are, then, two goals for a study of Capanahua phonology: first, to make available to linguistic science data analyzed from the point of view of this unified theory (though still in its nascient stages), and second, to test some conclusions suggested by those data, with respect to language universals.

II. PHRASE STRUCTURE RULES

The phrase structure and transformational rules provide the strings of formatives to which phonological features are assigned and also are the source of the labeled bracketing needed by the cyclical rules and word juncture rule. They are also the source of some syntactic features that are given phonological values by the MP rules, and they define the concatenation of sentences that hitherto has been called paragraph structure (Loos, 1963).

The following discussion presents only those base and transformational rules that are relevant to the factors mentioned above. The base rules necessary for a full grammar will be more complex by the addition of symbols but are not expected to be different in other respects. Rewrite rules will be prefixed with R and transformational rules with T, for identification.

R 1 Sent → # S # ($\left\{ \begin{matrix} \text{Excl} \\ (\text{NP})^{n} \end{matrix} \right\}$)

Sent: Sentence

S: sentence

Excl: exclamation

NP: noun phrase

A sentence may be followed by an exclamation or by a vocative noun phrase, but not by both. Only one exclamation follows a sentence, though several vocatives may (indicated by the n supercript on the NP parenthesis). Exclamations are words like 'anyhow,' 'then,' and 'well.'

rïboki ta ?ïn ka?ipiki, kokan
upriver I went, Uncle

?ïa ta ?ïn hawi, maman, papan
I have come, Mother, Father

kawï, ?ika?i sïn
go well (Well, go on!)

R 2 S → (# S #)^{n} Md (NP bïtan) NP PP (Nar # S #)

Md: mood indicator

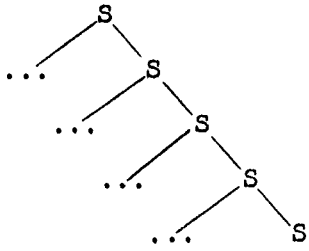
bïtan: conjunction 'with'

PP: predicate phrase

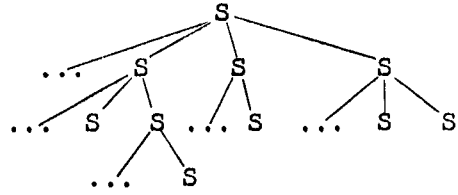
Nar: narrative conjunction

The Capanahua sentence can be expanded indefinitely by conjoined sentences, indicated by (Nar # S #) at the end of rule 2. In order to account for the possible expansion, the rule must either allow the conjoined S to be recursive in linear fashion as in rule 2, or it must permit recursion by a more general rule (# Nar S #)^{n}. The embedding of the conjoined S in rule 2 will produce only a right linear expansion, while n superscript embedding permits branching expansion. (The trees omit the Nar and # symbols.)

Right linear expansion



Branching expansion



The structural index of the T rules that embed conjoined sentences requires that the tense of each conjoined sentence be identical to that of the preceding sentence. The subject, object, and time NP deletions in conjoined sentences are also based on identity with the preceding sentence. The right linear expansion permits this to be done in a very simple manner, for the convention on cycling requires the cycle to begin on the most embedded S. In the branching tree, NP deletions will destroy information needed by later rules if a right-hand branch is not always cycled before a branch to its left, unless deletions are delayed until the information is no longer needed to establish NP identity. However, this is not a strong argument against the branching recursion, for delay of NP deletion turns out to be necessary for the right linear expansion as well (see T 13, T 14).

Stronger evidence for rejecting the branching tree is that Capanahua conjoined sentences show no evidence to support the underlying immediate constituent analysis of a branching recursion. The only permitted form of sentence conjoining is sequential: ". . . and then . . . and then . . . and then . . ." There are no sentence conjunctions

corresponding to 'or' or 'and' in English to produce sentences like the following examples.

1a. I will go and tell him about it.

2a. I will go and you will stay.

3a. I will go or stay.

4a. I will go or you will stay.

5a. I will go or you will go.

The first is expressed in Capanahua only by the sentence complement which requires an S source in the VP.

1b. I will go to tell him about it.

The other examples (2a-5a) are rendered by non-conjoined sentences.

2b. I will go. You will stay.

3b. I will go. Having not gone, I will stay.

4b. I will go. If I do not go, you will stay.

5b. I will go. If I do not go, you will go.

From this it can also be seen why Capanahua can have no sentences of the form

6. Jim and I will go and you and Fred will stay.

7. I will go and you will stay and Jim will work.

8. I will work and earn money and you will play and waste time.

There are only conjunctions of a non-branching type:

9. I went and then found turtle eggs and then fished.

The English conjunction '. . . and then . . .' parallels the Capanahua conjoined sentence in that it is strictly "linear." There seems to be no justification for assigning to an 'and

then' conjunction a structure such that underlying sentences are grouped into pairs of immediate constituents. It is sequential, as in the following:

10. Jim will go and then I will go and then you will stay and then Fred will stay.

To account for the restricted possible sentence conjunction in Capanahua, either the base rule which embeds the conjoined sentence must be limited to right linear expansions, or the embedding T rules that define the well-formedness of the conjoined strings must be formulated so that the surface string is linear, not branching into pairs. Since the right linear expansion restricts the form of recursion to the type compatible with the only conjunction possible, a very simple transformational rule will define the conjoined sentence as well-formed without the need of changing the tree structure in the process. The right linear expansion therefore is accepted for Capanahua.

Base rules increase in general linguistic interest according as they can be shown to be universal. Conversely, data are of interest when they constitute valid examples or counterexamples to linguistic universals. If there is a universal set of base rules, and if the types of possible sentence conjunction are determined in part by the tree structure in the base component, the Capanahua data suggest that the 'and then' conjunction should be distinguished from other types of conjunction.

The point at which the conjoined S can be embedded in the tree is not arbitrary. The recursion (Nar # S #) cannot be placed higher in the tree, that is, in rule 1, since it would then not be recursive, or if made recursive (Nar # Sent #), would produce branching trees. Hence it needs to be located in rule 2 or lower.

A second reason why the conjoined S should not be contained in Sent is that all sentences contained in another S must have the declarative mood indicator deleted. The conjoined sentence always has its declarative mood indicator deleted, but the declarative sentence contained in Sent does not. The simplest explanation for the deletion of the mood indicator in conjoined sentences, then, is that they are contained in an S as a result of the recursion in rule 2.

If the recursion for conjoined sentences is introduced lower in the tree by a rule PP → ... (Nar # S #), no serious consequences result. The conjoined S would be contained in PP instead of S, which would cause a slight complication in the transformational rules shifting VP to the end of the PP, but this could be circumvented by a modification in the structural index of the shifting rules. If the recursion is left as it is in rule 2, this minor complication is avoided. The simplicity criterion therefore requires that the recursion be in rule 2.

(NP bĩtan) is the conjoined subject noun phrase. In Capanahua, NP is not a recursive symbol, NP → NP(NP).

Lakoff and Peters (1966, VI-2) have argued that for English conjoined NP's it is necessary to have a rule expanding NP indefinitely: NP → NP (and NP)^{n}. The principal reason for this divergence from Chomsky's original formulation of conjoined NP's from conjoined S's:

Z X W and Z X W → Z X + and + X W

is that it can be shown that not all conjunctions have a S underlying them, since expansion of the conjunction gives contradictory sentences:

only very old men and very old women were in the room

cannot have the underlying sentences:

only very old men were in the room

only very old women were in the room.

Also, the condition that the formatives to be conjoined be constituents of the same type, i.e., of the same grammatical category, can be falsified:

The boy gave me a dime and the dog a bone.

The underscored are not each a constituent.

As a consequence, not all NP expansion for English takes place through an embedded S. Additional evidence is found in the fact that some verbs have the selectional restriction that the subject NP must be a plural or conjoined NP (e.g., 'meet together') so that plurality must be marked in the base component, which is not done by embedded S's.

Capanahua NP conjunction however, cannot be formed by a rule NP → NP (and NP)^{n} because NP conjunction is limited to two constituents, joined by a particle roughly equivalent to

'with' in English:

hawĭn ?aibo bĭtan yosibo ka?i 'his wife with old man goes'

There is no way of conjoining an unlimited number of subject NP's as in English. That is, there are no cases of

my wife and our children and our dog and our cat and I went

or:

my wife, our children, our dog, our cat, and I went.

If there are not too many NP's, the underlying sentences are preserved:

(a) ?ĭn ?aibo ka?i. nokĭn bakĭboribi ka?i. nokĭn ?očitiri?bi ka?i. nokĭn mišoribi ka?i.
'My wife went. Our children also went. Our dog also went. Our cat also went.'

If there are many NP's to be identified, they can be listed and summarized in a following sentence:

(b) ?iča honi ka?i. ?ĭn koka. ?ĭn papa. ?ĭn mama. ?ĭn boča. ?ĭn čoka. ?ĭn hočikabo. haabora?sibi ka?i. 'Many people went. My uncle. My father. My mother. My older brother. My sister. My younger brothers. Just all of those went.'

The listing of NP's in the fashion exemplified above is infrequently used, but can be elicited as a response to a question such as 'Who all went?' It is not productive enough to displace the full sentences of (a) or the frequently used generalized groupings such as 'they,' 'some people,' 'the

people in that house,' 'my mother with her folks,' etc. Yet it may be asked if this listing is not simply the Capanahua way of expressing the linguistic universal of NP conjunction. The answer is that it can be so regarded, but the consequences of such an analysis make it undesirable. First, in other languages NP conjunction is an embedding of NP's or of conjoined sentences within another sentence by embedding base rules, followed by transformational rules that delete the redundant parts of the embedded sentence and that conjoin the remainder to a formative in the matrix sentence. In Capanahua, however, the introductory and closing sentences to the listings are independent sentences, with mood-marking morphemes following the first constituent of the sentence (possible only in non-embedded sentences). But the listed NP's are outside of both the introductory and closing sentences. Then to avoid having a transformational rule that removes items from a sentence, it would be necessary to postulate the existence of a sentence (or some other structure) containing these sentences, in which the conjunction rule could operate. Such may be the case, but at present there is no independent evidence to support it.

Secondly, if such a conjunction rule exists in Capanahua, there is the question of why it should be pressed into service only for rare long sequences and not for shorter concatenations. Conjunction in other languages works equally well for two, three, five, or ten NP's. It therefore seems to be a weak account of the native speaker's linguistic

intuition to claim that he has acquired a rule but can make only limited use of it.

Only subject NP's may be conjoined by the conjunction 'with.' As with sentence conjunctions, there are no NP conjunctions corresponding to English 'and' and 'or.'

If the conjoined NP's were derived from underlying sentences, the analysis would require that ungrammatical sentences underlie the surface strings. It would claim that a sentence such as

yami bĭtan ta? ?ĭn hiwi rĭra?i 'With an axe I chop the tree.'

has the underlying sentences

yamin ta hiwi rĭra?iki 'The axe chops the tree.'

?ĭan ta ?ĭn hiwi rĭra?i 'I chop the tree.'

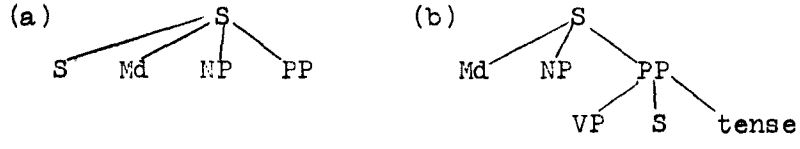
The normal cooccurrence restrictions of 'chop' are that it have an animate subject. This is violated by the underlying sentence in which the axe chops the tree.

An analysis which permits such ungrammatical sentences in the underlying strings has no way of preventing them from being well-formed surface strings when not embedded in conjunctions. Therefore, embedded sentences would seem not to represent the underlying structure of Capanahua conjoined NP's.

The rule which introduces the co-actor 'with' (bĭtan) in Capanahua can account for the fact that only two NP's can be conjoined by permitting an optional NP to be included in the sentence by the base rules. Rule 2 does this by the

inclusion of optional (NP bĭtan) before the subject NP. However, it is possible that 'with' in an instrumental sense should have a different source from the co-actor 'with' in order to account for the fact that the instrumental 'with' is normally in complementary distribution with the other form of the instrumental. If so, an underlying sentence could be the source of the co-actor 'with,' since the ungrammatical sentence would no longer be required. (Some means of restricting the co-actor 'with' to the conjoining of a maximum of two elements would be required.) Unlike clear cases of embedding, however, no sentential elements such as tense, verbs, adverbs, etc. survive from an underlying sentence as evidence that a sentence has in fact been embedded. The (NP bĭtan) solution is therefore tentatively retained.

The left-most (# S #) of rule 2 is the source of non-restrictive relative clauses. This S (or series of S's) must be contained in (a) the S dominating the other constituents of the sentence or (b) the PP.



In the surface strings, no non-restrictive relative clauses follow the AP or NP of an equational sentence formed by choosing $\left\{\begin{matrix}AP\\NP\end{matrix}\right\}$ in PP. It would be possible and desirable to account for this by locating the source S of the non-restrictive relative clause following the VP:

$$PP \rightarrow \left(\begin{Bmatrix} AP \\ PP \end{Bmatrix} \\ VP \ldots (S) \right) \quad Aux$$

Then the fact that no non-restrictive relative clauses occur with equational sentences is a simple consequence of the location of its source. However, the surface structure requires that the non-restrictive relative clauses be located at the beginning of the sentence, and if there are two sets of non-restrictive relative clauses, one must be shifted to the end of the sentence. The embedding of S in PP therefore requires two shift rules, one to shift them to the front and one to shift one to the end of the sentence.

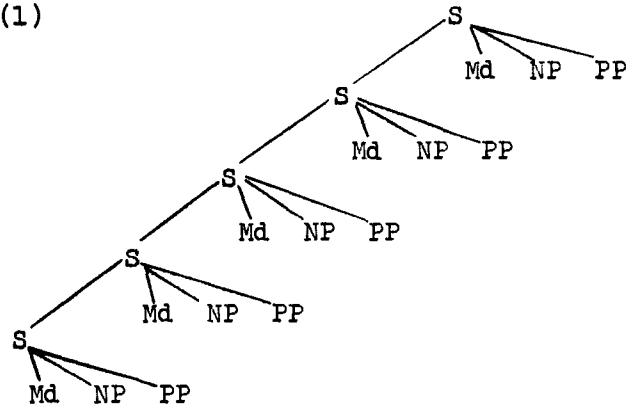
If the source S is located at the beginning of the sentence, only one shift rule will be needed. If one source S is placed at the beginning of the sentence and another at the end, no shift rules will be needed, but two embedding T rules will be needed for every type of non-restrictive relative clause--one to embed forward and another to embed backward--meaning the complexity of eight additional T rules in the grammar.

On the other hand, the T rule that specifies the particular verb endings for the different non-restrictive relative clauses can account for the absence of non-restrictive relative clauses in equational sentences simply by requiring, in the SI of the rule, that a verb be present in the matrix PP.

Thus simplicity considerations rule out having two independent sources, one at the beginning of the sentence and one at the end, but I have no strong arguments for locating the source in PP rather than preceding the Md symbol. To avoid the shift rule moving the clauses out of PP, therefore, the source S is provisionally placed before Md as in R 2.

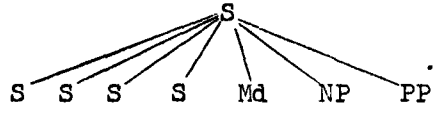
If the source of the non-restrictive relative clause is a rule of the form S → (S) Md NP PP, the relative clause expansions will always take the form of extended nesting (1):

(1)



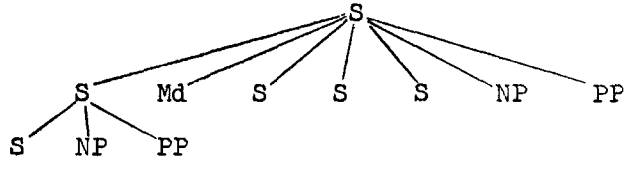
If the source is a rule $S \rightarrow (S)^n$ Md NP PP, the tree can take the form of the one above only if $n = 1$ for each repetition of the rule. If $n > 1$, a tree of form (2)

(2)



may be produced, but each S may in turn generate sub-trees. The surface string requires that the Md formative almost always immediately follow the first non-restrictive relative clause. To accomplish this shift it is necessary to have a

rule that will successively disembed the embedded S's of tree (1) until it has the form of tree (2). Tree (2) would also need a rule of this nature if the shift of the Md indicator were invariable, for it can produce, as mentioned, a tree (1) or a tree (2) with subtrees. However, some loquacious informants have given texts in which more than one non-restrictive relative clause precedes the Md. It is infrequently found so, but it occurs with sufficient regularity in some speakers that a tree of type (2) seems to be a better characterization of the underlying structure than (1), for "slips" due to a performance factor would be needed to explain the exceptions to the Md location with a tree of type (1). But if the rarity of deep embedding is attributed to performance, as seems probable because comprehension becomes more difficult as embedding increases, the rarity of sub-branching in type (2) is explained. Then the rule for Md location simply states that the Md is placed to the right of the first constituent of the S:



The "irregularity" is accounted for by the fact that the first S can in turn have an embedded S (non-restrictive relative clause).

Identification of the mood-carrying and topic-marking sentence is very simple in (1): the declarative mood marker

is simply deleted in every embedded S. The matrix sentence, not being embedded, remains as the mood-carrying S. The same operation will work for (2); the ease of identification of the mood-carrying sentence is therefore not a decisive factor in choosing the type of embedding rule.

The point at which the cycling rules begin in (1) is not difficult to determine, since the convention requires cycling the most embedded S first. In (2), however, either a special condition must be set requiring the cycle to begin at the extreme left, or all deletions affecting subsequent rules must be delayed until the very last cycle. Otherwise there is nothing to prevent the cycle from beginning on some other S than the initial one in the series. Then the deletions of subject or object NP's would block a transformational embedding rule operating on a string to its left. This is because NP identity is required between an S and its matrix S in (1) or between an S and the following S in (2). Tree (2) would therefore seem to require the greater complexity of deletion delay in the grammar, but it turns out that (1) also requires deletion delay because some embeddings effect an NP deletion in the matrix as well as in the embedded S, and by so doing require that deletions be delayed until the rules specifying transitive verb endings have operated. Thus, the matrix sentence below contains an embedded S that deletes the object of the matrix S:

$$\underset{S}{\lfloor}\#\ \underset{S}{[}\text{bakĭ}\S\text{o}\ \check{\text{s}}\text{iro?iton}\ \underset{S}{\rfloor}\ \#\ \text{baroman bakĭ}\S\text{o w̆irani}\ \underset{S}{]} \rightarrow$$

baby while-playing boy baby pushed

bakĭṣo širo?iton baroman wĭrankin 'while the baby was playing, the boy pushed him'

The second occurrence of 'baby' is deleted, and a late T rule changes the ending of the verb to -kin if the verb is transitive. However, if 'baby' is deleted, the information necessary for identifying the verb as transitive will be lost, and the intransitive marker will incorrectly remain.

Therefore, since a rule permitting branching ($S \rightarrow (\#\ S\ \#)^n \ldots$) accounts for the data better than a left linear expansion ($S \rightarrow (\#\ S\ \#) \ldots$) without additional complexity in the grammar, the branching type is taken as the source of the non-restrictive relative clauses.

R 3 $PP \rightarrow \left\{ \begin{matrix} \{AP\} \\ \{NP\} \\ VP \end{matrix} \quad \begin{matrix} ?i?k \\ (NP) \end{matrix} \quad \begin{matrix} \\ \\ (pp) \end{matrix} \right\}$ Aux

AP: adjective phrase

VP: verb phrase

?i?k: a form of the predicate 'to be'

pp: prepositional phrase

Aux: auxiliary

The choice of {AP, NP} will give equational sentences such as

(AP) ?ĭn baba ta ?ani ki 'My grandchild is big.'

(AP) ?ĭn baba ta siri ki 'My grandchild is good.'

(NP) ?ĭn baba ta bĭbo ki 'My grandchild is a boy.'

The predicative ?i?k is deleted when the tense is [-past, -fut], i.e., present tense.

The NP following VP is the instrumental noun phrase. Neither this NP nor the prepositional phrase (which is a locative construction) can occur in an equational sentence:

*?ĭn baba ta siri ṣobo?oki 'My child is good in the house.'

The instrumental NP following VP can also function as an adverbial, e.g., nontin 'by boat,' košin 'by strength, strongly,' ?iston 'by speed, quickly.' Since NP is not recursive, R 3 correctly predicts that there can not be more than one such NP in each sentence.

R 4 Md → { Decl (Q) / Imp / Exclm }

Decl: declarative mood marker
Q: question marker
Imp: imperative mood marker
Exclm: exclamatory mood marker

The sentence must carry one of the mood markers and with the declarative may optionally carry the interrogative marker.

The location of the Md symbol in R 2 is required by the Md shift rule (T 31) and the pronoun reduplication rule (T 33) which reduplicates a first or second person pronoun when it precedes the Md symbol.

The Md symbol must always follow the first constituent of the sentence, so that if there is a string of nonrestrictive relative clauses to its left, it will move left, and if there is no S to its left it will move right.

The mood markers will be realized as affixes on the

verb and will also determine whether phonological shape will be given to the person-features of [spkr] and [addr]. Both facts are accounted for by the selectional rules assigning features to the tense symbol.

Only the declarative marker will produce a well-formed embedded sentence, for the other mood markers will block an embedded sentence.

R 5 $VP \rightarrow \left(\left\{ \begin{matrix} NP \\ \#S\# \end{matrix} \right\} \right) Vb$ Vb: verb

In the verb phrase, the object must precede the verb for the morpheme lengthener rules (T 22 and MP 9) to apply properly, for T 21 says that the noun that precedes the verb is assigned a feature exempting it from the lengthening rule. The object noun is always short, a fact explained by its location preceding the verb.

The S preceding the verb (R 5) is the source of sentence complements taken by verbs of motion:

?ĭan ta <u>mani bĭna?i</u> ka?i 'I go <u>to look for bananas</u>.'

R 6 $Vb \rightarrow \left\{ \begin{matrix} AP \\ V \quad (NP) \end{matrix} \right\} \quad (\#\ S\ \#)\ (Adv)\ (Asp)$

Adv: adverb

Asp: aspect

V: verb stem

The AP in the verb will have the meaning 'to become' when the VP is intransitive, and the meaning 'to cause to become, to make to become' when the VP is transitive. In the latter case a transitive-marker is inserted by a T rule

(T 6 and MP 1):

hošin 'red' (Aj)

hošini 'become red' (Vb, intransitive)

hošinhakin 'to make it red' (Vb, transitive)

There are therefore two places in the tree where the rules permit an adjective to be placed: in the predicate phrase of an equational sentence (R 3) and in the verb (R 6). In the equational sentence the adjective has no verbal character and is the basis for noun modifier constructions. Thus the adjective ?ani 'big' in the sentence

?ĭn baba ta ?ani ki 'My grandchild (is) big.'

when embedded in a noun phrase becomes the attributive 'big'

?ĭn baba ?ani 'my big grandchild'

When placed in a verb, the adjective becomes a verbal.

?ĭn baba ta ?ani?iki 'My grandchild grows.'

The rules permit the verbal AP in Vb to be transitive

?ĭan ?ĭn baba ?anihakin 'I raised (caused to become big) my grandchild.'

The AP is thus distinguished from the NP in R 3 in that NP cannot be a verbal and from the verb stem in that V cannot occur in an equational sentence, which explains why verbs do not occur as noun modifiers:

*?ĭn baba ta bo ki '*My child is carry.'

*?ĭn baba bo '*my carry grandchild'

The indirect object (NP in R 6) is found only when a verb stem also occurs; it is not found with AP in Vb. The S embedded in the Vb is the source of causative constructions

such as

haa ta ?ĭn hiwi rĭrama?i 'I cause him to chop the tree.'

In causative constructions one verb is deleted, giving a surface string which is always "transitive," though one or both underlying sentences may be intransitive:

haa ta ?ĭn noyaki?nai 'I fly him.' i.e., 'I cause him to fly by flying with him' from the underlying sentences 'I fly' 'He flies.' with the second embedded in the first. As a result of the deletion of the verb stem of the matrix verb without deletion of the Adv and Asp suffixes of the matrix, the resultant verb will have affixes from both verbs (covered in T 8 and T 9).

e.g. $[\ [\ V\ Adv\ ]_{V}\ (caus)\ Adv\ Asp\ Imp\ ]_{V}$

mapĭt ka?in ma šon tan wĭ
ascend quickly cause for me over there imp
'Make it go up quickly over there for me!'

If the source of the causative sentence is located in the matrix verb preceding the Adv series of affixes, the surface string is seen to be a result of the deletion of the matrix verb. Otherwise a special shift rule will be necessary to place the affixes in their proper sequence.

R 7 AP → {Aj, N, pp} Aj: adjective
N: noun

The nouns which can be dominated by AP are a restricted class of numerals and colors which have been marked

as nouns because they may fill the head slot of a noun phrase

hawĭn hoṣo 'his whiteness'

hawĭn rabĭ 'his second one'.

The prepositional phrase dominated by AP is a locative. Thus, while the locative occurs following a VP, its complementary distribution with the adjective requires that it be located in the AP also. Like adjectives, the prepositional phrase may serve as a noun modifier:

?ĭn baba rĭboki?a? 'my upriver grandchild'

R 8 Adv → (Adv1) (Adv2) ((NP)Adv3) (Adv4)

R 9 Asp → (Asp1) (Asp2) (Asp3) (Asp4)

The identity of the goal of Adv3 is specified by the optional NP which is erased under identity conditions with another NP in the sentence. Otherwise the goal NP will be shifted to precede the PP.

The Asp series of adverbial suffixes is differentiated from the Adv series by the causative sentence which introduces a causative suffix after the Adv series. Though theoretically all eight series of affixes in Asp and Adv may occur in any one verb, what is assumed to be a performance limitation restricts the actual sequence to a maximum of six.

The Adv and Asp suffixes will be listed in the lexicon but are given here to explain their nature:

Adv1 pa 'downward motion,' ?i?ni 'upward motion,' ka?in 'rapid motion.'

Adv2 bo?an 'invertedly, turned over,' ki?ran 'bent

over, motion in a curve,' kawan 'sideways, evasive motion,' bona 'while going along (away from speaker),' ?ona 'while coming (towards speaker),' ba?ina 'all day long,' šina 'all night,' wan 'action whose effects have ceased,' mĩno 'for the first time,' ba?in 'completely, before going,' rĩnĩn 'very abruptly,' kana 'iteratively,' bĩ?man 'before coming,' ri 'hurriedly.'

Adv3 ṣon 'for the benefit of someone,' ?an 'to the detriment of someone,' piška 'please' (used only in imperative sentences).

Adv4 koin 'truly'

Asp1 kaskima 'in vain,' kaɸi 'want to, wish to,' ko (intensifier for negative).

Asp2 yama (negative) 'no,' yora (positive intensifier) 'yes indeed,' ta? 'without delay.'

Asp3 ta?ni 'more or less, somewhat.'

Asp4 tan 'at a distance away.'

R 10 V → CS

R 11 Aux → {(NP) tense, Nom} Nom: nominalizer

The NP in R 11 is a time noun phrase. There are no tense markers in nominalized sentences; and the complement nominalizer -ti, which corresponds sometimes to the English nominalizer -er and sometimes to to + infinitive, is formed by a nominalizing T rule that requires Nom (see T 10, T 11, T 12).

[S ?iston kati]S kĭĭni 'he wants to go fast'
quickly go - wants

[S mai pokinti]S ?onanyamakin 'he doesn't know how
earth dig - know-neg to dig'

[S ha?an kabori bači bĭnati]S ... 'that with which
he turtle egg searches - one searches for
turtle eggs'

There are a few nouns that are marked in the lexicon with cooccurrence restrictions requiring that they be preceded by an S in NP: [+S ___, - ___,]. The T rule (not included in this study) that forms the NP with these nouns as "head" nouns also requires Nom. The nouns are:

ti?i 'the amount that one ____'
piti?i 'the amount one eats'

mis 'the one who habitually ____'
pimis 'a big eater'

nika 'the one who by nature ____'
pinika 'one who by nature eats'

wan 'the one who excessively ____'
pi?anwan 'a glutton'

R 12 Tense → CS

R 12a [+tense] → [±past]

R 12b $\begin{bmatrix}+\text{tense}\\-\text{past}\end{bmatrix}$ → [± future]

R 12c [+past] → [-future]

R 12d $\begin{bmatrix} +past \\ +future \end{bmatrix} \rightarrow [\pm remote, \pm distant]$

The subcategorization of the tense symbol will give four past and four future tenses and one present.

$\begin{bmatrix} +past \\ +remote \\ +distant \end{bmatrix}$ remote past tense (six months or more ago)

$\begin{bmatrix} +past \\ +remote \\ -distant \end{bmatrix}$ remote past (one to six months ago)

$\begin{bmatrix} +past \\ -remote \\ +distant \end{bmatrix}$ recent past (from one day to one month ago)

$\begin{bmatrix} +past \\ -remote \\ -distant \end{bmatrix}$ recent past (from the present to one day ago)

$\begin{bmatrix} -past \\ -future \end{bmatrix}$ present tense

$\begin{bmatrix} +future \\ -remote \\ -distant \end{bmatrix}$ immediate future (within a couple of hours)

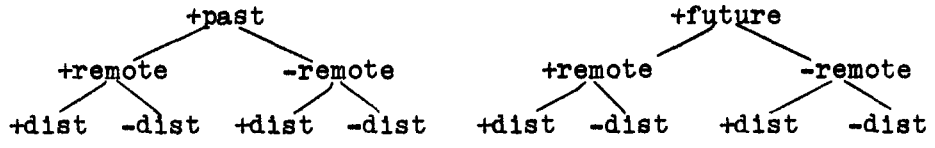
$\begin{bmatrix} +future \\ -remote \\ +distant \end{bmatrix}$ near future (tonight or tomorrow morning)

$\begin{bmatrix} +future \\ +remote \\ -distant \end{bmatrix}$ distant future (tomorrow or later)

$\begin{bmatrix} +future \\ +remote \\ +distant \end{bmatrix}$ remote future (indefinite)

R 12e $\left\{ \begin{bmatrix} +tense \\ +future \\ +remote \\ -distant \end{bmatrix}, \begin{bmatrix} +past \\ -remote \\ +distant \end{bmatrix} \right\} \rightarrow [+B]$

The time designated by each tense is approximate and varies somewhat between individuals. But as the translations show, the tenses sharing the same features in past and future are not separated from the present tense by the same amount of time, even though the past and future are logically symmetrical:



There are no lexical entries corresponding to these time divisions, so that it is not possible to subcategorize the tense symbol in terms of a lexical entry. No time words except one (nɨatɨma 'long ago') make a distinction between past and future tense. Thus, the lexical entry ba?kiš means 'tomorrow' when the future tense is used, and 'yesterday' when the past tense is used, and ?okɨaba?kiš means 'the day after tomorrow' with the future and 'the day before yesterday' with the past. One lexical entry therefore will not have the same features of remoteness and distance for both past and future: ba?kiš will be [-remote, +distant] for past and [+remote, -distant] for the future.

Furthermore, if the tense were to be subcategorized with respect to the time noun, it would have to be subcategorized not only with respect to the head noun, but with respect to the whole time noun phrase. That is, the noun nɨtɨ 'day' will have its own past tense features, but the

NP will have the features [-remote, +distant] (assuming past tense) when nĭtĭ is preceded by wĭɸa 'other' (wĭɸa nĭtĭ 'the other day') and the features [-remote, -distant] when preceded by nĭa 'this' (nĭa nĭtĭ 'this day'). Therefore the subcategorization of the tense symbol would have to depend on a semantic reading of the whole time noun phrase.

I leave this complex problem unsolved and allow the CS subclassification rules to develop the necessary tense features. These tense features will later be spelled out as phonological segments.

[+B] is an arbitrary label for a feature assigned to some of the tenses. It is needed by later rules to place the plural markers after [+B] tenses but before the others (see T 41 and T 48) and needed also in the phonological component to lengthen verbs (MP 6) and to assign high pitch (Stress Rule 3).

R 13 pp → NP p p: postposition (locative suffix)

Since the suffixes are postposed, the "prepositional phrase" is actually a postpositional phrase. The locative suffixes will be listed in the lexicon: e.g., ?o? 'inside of,' ki 'toward,' naman 'under,' kačian 'on top of,' patas 'beside,' ?okĭ 'behind,' etc.

ṣobo?o 'inside the house'

hiwi naman 'under the tree'

hiwi pataṣ 'beside the tree'

hiwi ?olĭ 'behind the tree'

hiwi čičo 'inside the tree'

hiwi bočiki 'above the tree'

hiwi kačian 'on top of the tree'

hiwi ki 'toward the tree'

hiwi nĭkĭ 'on this side of the tree'

R 14 NP → (# S #) N (Neg) Neg: negative

Only one S source has been posited for noun phrase expansions by noun and adjective modifiers. Modifying nouns and modifying adjectives expand to the right. Possessives expand only to the left. In order to account for expansion in both directions, it is possible to posit two sources thus, NP → (S) N (S) (Neg) rather than one source with a shift rule for examples such as these (head noun is underlined):

e.g., hawĭn <u>nonti</u> 'his canoe'

hawĭn <u>nonti</u> tiroma 'his worthless canoe'

hawĭn <u>nonti</u> tiroma piška 'his small worthless canoe'

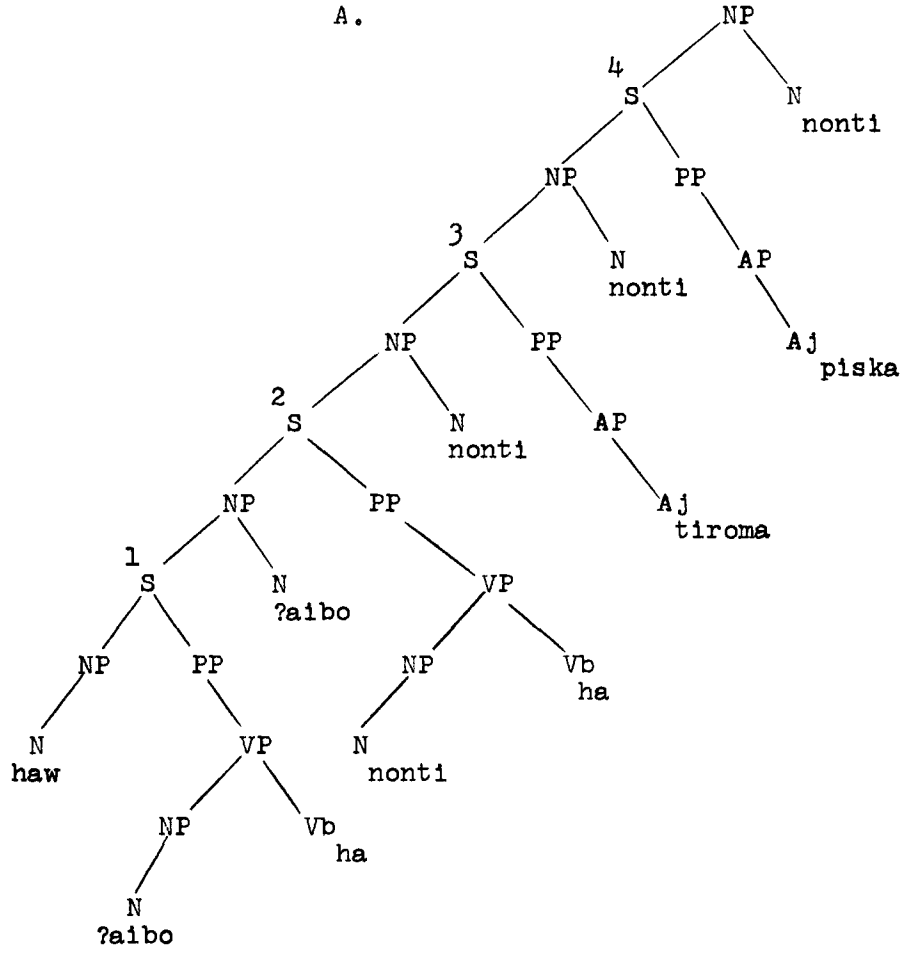
bakĭn <u>nonti</u> 'a child's canoe'

hawĭn bakĭn <u>nonti</u> 'his child's canoe'

hawĭn ?aibar bakĭn <u>nonti</u> 'his wife's child's canoe'

hawĭn ?aiban bakĭn <u>nonti</u> tiroma piška 'his wife's child's small worthless canoe'

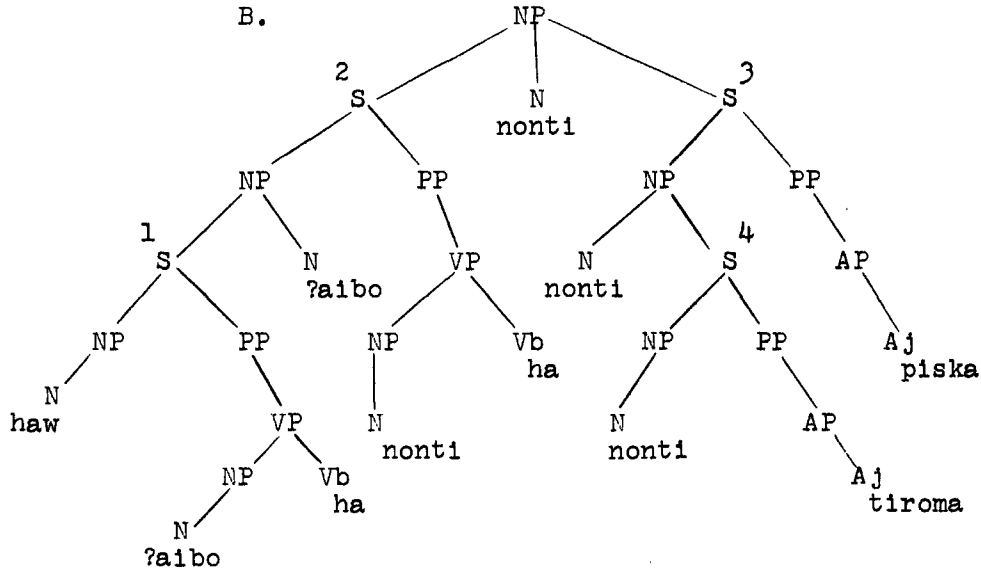
In order to account for the surface structure of a phrase such as hawĭn ?aiban nonti tiroma piska 'his wife's small worthless canoe' with one S source, a tree of type A below will be needed (# signs and Aux omitted):



The numerals indicate the underlying sentences

1. he has a wife
2. S-wife has a canoe
3. S-canoe is worthless
4. S-canoe is small

A double source for the modifiers, one S on each side of the head noun, will produce a tree such as B below:



1. he has a wife
2. S-wife has a canoe
3. canoe-S is small
4. canoe is worthless

Apart from the semantic considerations involved such as whether the projection rules operating on underlying phrase markers should state (as in B) 'she has a canoe, the canoe is small' rather than (A) 'she has a canoe, S-canoe is small, etc' the two approaches can be compared in terms of simplicity. First, either way, the cyclical P rules will have to assign stress and pitch to the NP with modifiers both preceding and following the head noun. Simplicity of cyclical P rules is therefore not a consideration, unless

the structure of A after the shift is different from that of B.

Second, A will require a shift rule to move some nouns and adjectives to the right of the head noun, but must not shift the possessive nouns. If the shift takes place after the T rules that embed an S by removing # junctures and deleting nouns, at least one such embedding T rule is needed, beside the deletion of Verb for possessives.

B will require two embedding T rules--one embedding left and one embedding right for the S's on each side of the head noun--and in addition will require some means of insuring that modifying nouns and adjectives are inserted in an S on the proper side of the head noun. Restricting the structural index of the transformational embedding rules might accomplish this, adjective phrases being embedded to the left and everything but adjective phrases to the right. But since A requires no special SI restrictions of this sort, it seems to be somewhat preferable if the necessary shift rule or rules are not too complex.

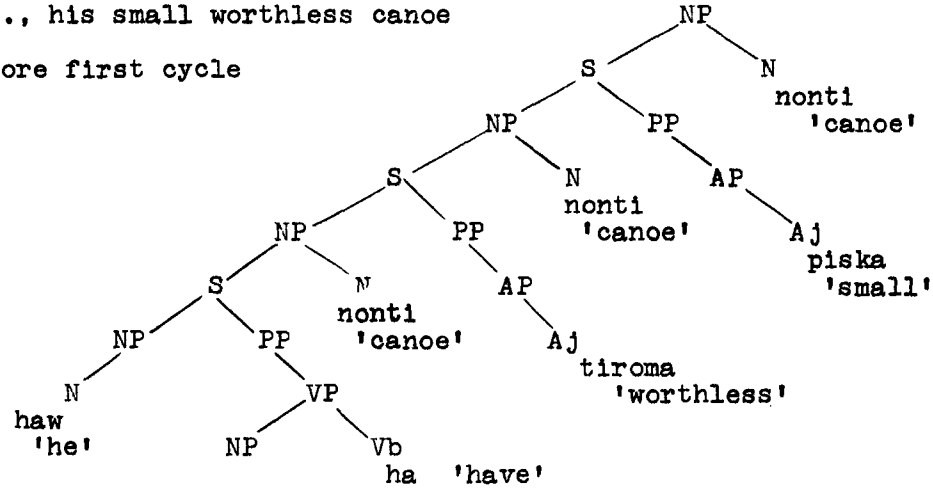
In fact, the shift of all and only the proper adjectives and nouns in A can be accounted for by one simple rule which shifts an AP from preceding a head noun to following it:

$$X \underset{S}{[} Y [[Z]]] N W \qquad 2 \rightarrow \emptyset$$

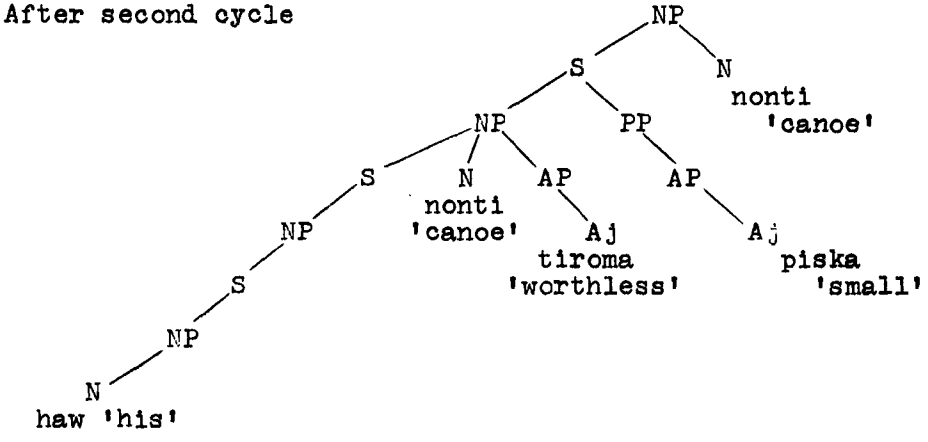
$$\underset{1}{} \quad PP \underbrace{AP\ AP}_{2} \underbrace{PP\ S}_{3} \underset{4}{} \qquad 3 \rightarrow 3+2$$

e.g., his small worthless canoe

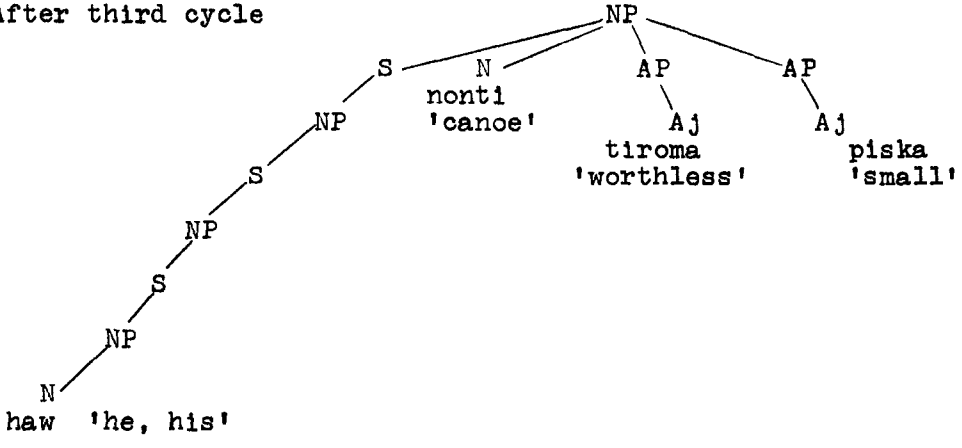
Before first cycle



After second cycle



After third cycle



The rule will not shift an N in NP, which is correct, for noun modifiers in NP are built up on the left, unlike adjectives:

?ino ?aibo bakĩ (tiger-woman-child) 'a female tiger cub'

The rule will not put all the modifying nouns and adjectives in their proper surface order. Nevertheless, it is clear that a single S source is adequate, given the shift rule.

Negation in Capanahua is expressed by negative morphemes in a noun phrase or in a verb, or both. Verbs are negated by including the suffix yama in the Asp series in the verb:

haa ta kayama?i?ki 'he does not go'
he Md go not

Noun phrases are negated by the negative suffix ma attached to the final formative of the NP:

haama ta ka?i?ki 'the wrong one is going',
he-not Md go 'it is not he that goes'

R 13' (Provisional modification of R 13)

NP → (# S #) (N) N (Neg)

There are some noun-noun relationships that are not accounted for in the modifier treatment, which I am calling compound nouns for lack of a better term and representing by (N)N in this rule. Compound nouns seem to have no underlying sentence, for they either reflect a genitive relationship

(for which there are no prepositions or other particles corresponding to English 'of' and 'for') or are paraphrased in idiosyncratic ways so that defense of one or another underlying S to explain the surface structure is very weak. R 13' is therefore a tentative formulation for lack of regularities on which to base an argument for a different description. The first N has a genitive or modifying relationship to the second:

?atapa ṣobo 'chicken house' a house for chickens
?očiti ta?ĭ 'dog track' a track made by a dog
?awa ba?i 'tapir trail' a trail used by tapirs
?i?sa pĭ?i 'bird feather' the feather of a bird
mai kini 'earth hole' a hole in the earth
hiwi kini 'tree hole' a hole in a tree
makan roĭ 'stone axe' an axe made of stone
bawa rono 'parrot snake' a snake that resembles a parrot
bawa ṣĭta 'parrot tooth' an ant with pincers shaped like a parrot's beak
kaši nawa 'bat people' people of the bat tribe
?oṣnĭ nĭtĭ 'moon light' light from the moon
yami ṣĭta 'iron tooth' axe-head
bina ṣobo 'bee house' bee-hive
yošĭr ba?i 'demon trail' rainbow
hiwi tašo 'tree stem' a slender stick
yapa bĭro 'fish eye' the eye of a fish
bĭro yošin 'eye spirits' spirits inhabiting eyes

paama bimi 'paama fruit' fruit of the paama tree

čaṣo ?ino 'deer tiger' a tiger with a deer's color

Personal pronouns will be marked with the subcategorization features [-N__, -__N, -S__], and thus cannot be inserted in the CS of a compound noun or of a noun with an attributive. Only animate nouns may precede Adv3 suffixes in the verb, and some verbs require animate subject nouns. No time nouns may be inserted in the CS of a subject NP noun. The selectional rules must therefore mark the CS of the noun with features that guarantee well-formedness.

Demonstratives are pronouns which may be inserted in the CS of a noun modified by S but not in the CS of a compound noun. When they are preceded by a modifier, a T rule will shift them to precede the modifier.

e.g. non hisipi ?oa we-saw-that →

?oa non hisipi 'that which we saw'

This treatment of the demonstratives has eliminated the need of complex rewrite rules to place the demonstratives preceding the modifiers, yet permitting the demonstratives to occur without the noun as shown by these two rules which, to be collapsed, require at least one additional node between NP and N.

NP → Dem (S) (Neg)

NP → (Dem) (S) N (Neg)

R 14 N → CS

R 14a [+N] → [±count] count: count nouns

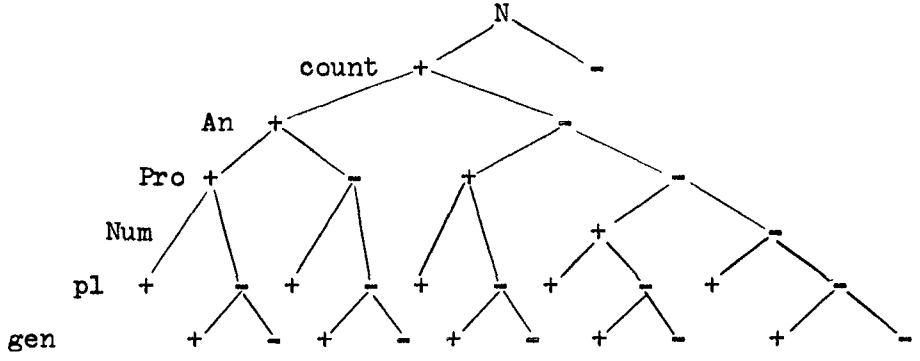
R 14b [+count] → [±An, ±Pro]

R 14c $\begin{bmatrix}+\text{count}\\-\text{An}\\-\text{Pro}\end{bmatrix}$ → [±Num]

R 14d [+count] → [±pl]

R 14e $\begin{bmatrix}+\text{count}\\-\text{pl}\end{bmatrix}$ → [±gen]

An: animate
Pro: pronoun
Num: numeral
gen: generic



Subclassification rules 14a through 14c define inherent features of the noun necessary for application of the lexical rule. Rules 14d and 14e define those features of the noun which are not lost under NP deletion, but are preserved by assignment as features to the verb if the head noun of the NP is deleted.

[-count] nouns cannot be pluralized and cannot be modified by a numeral (e.g., hĭnĭ 'water'). Verbs carry cooccurrence restriction features relevant to the animateness of the subject NP. Numerals can be pluralized like other nouns, can be immediately dominated by NP (hence, function as the 'head' of a noun phrase), and can be modified by adjectives and possessives. [-count] nouns cannot be assigned to a CS containing a [+generic] feature.

Lexical redundancy rules will develop the appropriate features for the lexical entries.

The [+pl] and [+gen] features will be spelled out as suffixes to the NP by the T component. [-pl] and [-gen] will remain unmarked in the surface structure.

All pronouns are specified for person by features of speaker and addressee and are so marked in the lexicon. A lexical redundancy rule will define all [-Pro] nouns and all demonstratives as [-spkr, -addr], i.e., third person. The pronouns are:

[+pro, -S__]

[+spkr] ʔĩ 'I'

[-spkr, +addr] mi 'you'

[-spkr, -addr] ha 'he'

[+Pro, +S__] (Demonstratives)

[+location of spkr] nĩat 'here, this'

[-location of spkr, +location of addr] toat 'that, there'

[-location of spkr, -location of addr] ʔoat 'that, there'

R 15 Neg → ma (Neg)

Negation in nouns is recursive. The number of recursions seems to be limited only by performance (memory). Each negative negates everything preceding it in the NP, so that every added negation switches the value of the total previous negation.

e.g. haa 'he'

haama 'not he'

haamama not-not-he (he indeed: no one else)

haamamama not-not-not-he (someone else)

etc.

R 16 $X \rightarrow CS$

R 17 $[+V] \rightarrow CS \; / \begin{cases} \alpha\,[\ldots\_\_\,]_{VP} \\ [\beta\ldots\_\_\,]_{VP} \end{cases}$ where α is a N and β is a N

Verbs are subclassified according to human, animate, inanimate subject NP's, transitivity (object NP), and indirect object NP (covered by the subcategorization convention since V and the indirect object NP are dominated by the same node).

The indirect object and time NP's are marked by suffixes. It seems to make no real difference to the description whether their suffixes are assigned by base rules or by transformational rules, apart from claims made for the significance of nodes representing a case system, as in Fillmore's recent proposal (1967). In any treatment, a selectional feature must be assigned to the CS of the head noun of the NP to determine the type of noun that will be inserted by the lexical rule. Thus, only "time" words may be dominated by Aux. e.g., ba?kiš 'tomorrow,' bari 'sun,' 'year' but not rono 'snake.'

This can easily be done if the subcategorization rules by convention assign a feature of the node immediately higher

than NP, as well as [+N] and [+NP], to the CS of the noun.

Prideaux has shown that the conventions formulated by Chomsky (1965) for subcategorization need to be expanded to include the NP category dominating N as well as N in the subcategorization: "Next higher node convention: If X is a lexical category and Y is the least major category of X with respect to a given derivation, then [+X] → [+Y]." (Prideaux, 1966. 115) My position goes just a step farther than Prideaux's in order to account for the fact that in Capanahua the next higher node over NP determines the function of the NP and thereby imposes restrictions on the choice of head noun of the NP.

Then the time NP's head noun will by convention carry a feature [+Aux], making it unnecessary to assign a [+time] feature to it by a special rule. Likewise the instrumental NP need not be marked [-human] and the benefactive [+animate], for a lexical redundancy rule will assign the proper selectional feature to the nouns.

R 18 $[+\text{tense}] \rightarrow CS \;/\; \alpha \ldots (\beta\hat{}\,VP)\ldots\_\_$

where α = Md, β = N

i.e., the features of the mood marker and the features of the subject NP head noun are assigned to the tense symbol.

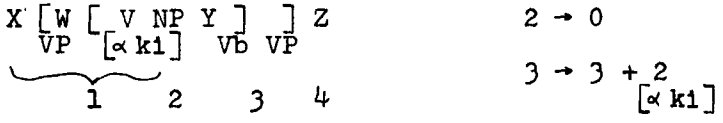
The concord of the tense with the person features of the subject NP, shown in the surface structure by suffixes following the tense morphemes or by the shape of the tense markers and mood markers, is effected by rules which

spell out the shapes of the respective morphemes from the feature assigned to the tense symbol. The person features are not assigned to the tense symbol if there is no verb phrase in the sentence.

The phonological shape of the tense and person morphemes is conditioned in part by the mood features assigned to the tense symbol. The Md symbol is now needed only to determine pronoun reduplication and to position the clitic ta?, marker of the declarative mood. Otherwise Md will be deleted, and the features assigned to the tense symbol will block the sentence if it is not well formed, i.e., the tense symbol will not be given a phonological shape if the proper features are not present in the tense symbol. Thus, the symbol will remain unconverted if the imperative features of the Md are assigned to the tense symbol in a sentence containing no verb phrase, since a verb phrase is necessary for the imperative.

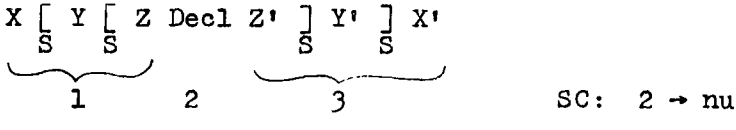
III. TRANSFORMATIONAL RULES

T 1 Indirect object shift



i.e., the indirect object NP is shifted from inside the VP brackets to outside the VP brackets. Its location inside the brackets was required by the fact that an indirect object NP cannot follow a verb containing an adjective. It can only follow a verb containing a verb stem. The surface structure requires that the indirect object be external to the Vb, the VP shift rule (T 26) requires that the indirect object not be included in the VP bracketing following the Vb, and the noun lengthening rules (T 21, T 22) require that the indirect object follow the verb. It must therefore follow the VP bracketing.

T 2 Md marker deletion



i.e., the declarative mood marker is deleted in all embedded sentences, and thus remains only in the matrix S. By

ordering T 2 early, the SI of S-embedding rules is simplified in many cases, since Md will not need to be mentioned. The deletion is not unrecoverable because only Decl may be deleted, and any other Md in an embedded S will block the S.

Noun Phrase forming rules

T 3 Possessive and attributive formation

$$X \left[_{\bar{N}P} \# \left[_{\bar{S}} \left[Y\ N\ (AP) \right]_{\bar{N}P} Z\ X' \begin{bmatrix} +\text{tense} \\ -\text{past} \\ -\text{fut} \end{bmatrix} \right]_{\bar{S}} \# N \right]_{\bar{N}P} Z'$$

SI: 1 2 3 4 5 6 7 8 9 10 11

Possessive:

Cond: 6 = 10

7 = $\begin{bmatrix} +V \\ ha \end{bmatrix}$ 'to have'

SC: 2, 6, 7, 8, 9 → 0

4 → 4 [+pos]

5 → [-next rule]

Attributive:

Cond: 6 = {NP, AP}

7 = ?i?k

4 = 10

SC: 2, 4, 6, 7, 8, 9 → 0

10 → 10 + 6

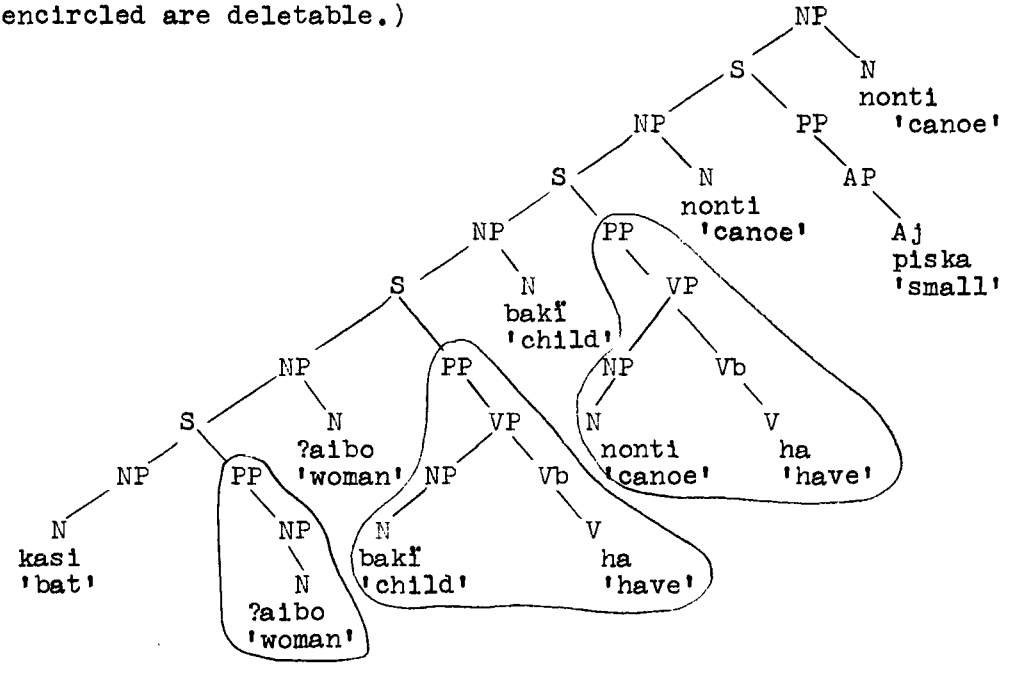
i.e., both the possessive- and attributive-forming rules require present tense and delete the tense and the verb or ?i?k. Both also require that none of the optional PP formatives be present in the embedded S and each permits deletion of only a particular constituent under identity conditions with the head noun.

The possessive assigns a [+pos] feature to the possessing noun, which is used by a later rule (T 36) to spell out the proper form of the pronoun.

The reflexivizing rule (T 7) can either precede or follow this rule, for the deletions will prevent the one from fitting the structural description of the other. But if the deletions of one rule were to be delayed, the other rule would incorrectly be able to operate on the same string, with the result that a noun both possessive and reflexive could emerge in the noun phrase. The deletions must therefore be effected immediately in these rules, though as will be shown, in other rules they must be delayed.

The variable Y preceding N is necessary in order to permit modifier build-up from successive embeddings of S in NP.

e.g., 'the Cashibo woman's child's small canoe' (Items encircled are deletable.)



The optional (AP) permits carry-over of the AP from a

previous embedding. The AP shift (T 4) effected by the attributive rule must precede the morpheme lengthener rules (T 21 and T 22) so that if there is an AP, it is lengthened instead of the noun when the lengthening rules operate on the matrix sentence. But if an AP is thus lengthened, it is no longer subject to the AP shift rule (guaranteed by the [-next rule] feature, example 5 below).

The morpheme lengthening rules will operate on the underlying S before the embedding deletions take place. Therefore the possessives will have long forms of the noun because they are separated from the verb in the underlying sentences, but the attributive nouns will be short, since there are no verbs in their sentences and the lengthener rules only operate when a verb is present. This accounts for the fact that possessives have a different form from the same word used as an attributive or compound, yet the added portion is the same (MP 9) as that of the noun when lengthened in any other syntactic construction.

e.g. 1. ʔawa bakĭ 'a tapir calf'

2. ʔawapan bakĭ 'a tapir's calf'

3. ʔamĭn taʔĭ 'a capybara track'

4. ʔamĭman taʔĭ 'a capybara's track'

5. ʔamĭn hošinin taʔĭ 'a red capybara's track'

As was discussed in the PS treatment of the noun phrase, the nouns and adjectives which can be dominated by AP follow the head noun and must be shifted. The shift rule thus makes it possible to state the cooccurrence restrictions

only once for those entries which can follow the noun, by stating them for the AP in an equational sentence. The ungrammaticalness of noun phrases like

6. *bakĭ [aibo] (NP) child-woman 'girl'

7. *kimiša [tiroma] (AP) nonti three-worthless-canoe

8. *[hošini] (AP) pi?a red-bow

9. *[hošoni] (AP) ?atapa white-chicken

is explained by the fact that in 6 the NP should not have been shifted, and in the others the AP should have been shifted. (There are some instances when the AP may be shifted back to the left side of the noun but we will not go into these stylistic variations.)

T 4 AP modifier shift

X [[Y [Z] (AP)] (NP) N] (NP) W

SI:	1	2	3	4	5
SC:	1	2	0	4+3	5

i.e., the AP preceding a noun is shifted to follow it.

If the AP has been lengthened by the lengthener rules while in the possessive sentence, it can no longer be shifted; T 22 therefore assigns a feature [-T 4]. All other AP's will be successively shifted to follow the head noun. (The head noun is identified by underscoring, below.)

10. hawĭn <u>?aibo</u> piška rabĭ his-wife-small-two
'his two little wives'

11. hawĭn ?atapa hoṣoni piška his-chicken-white-small 'his little white chicken'

12. čašo hošinin bakĭ piška deer-red-child-small 'the red deer's small fawn'

While the shift rule will put the adjectives on the proper side of the head noun, it will not put them in the correct sequence.

T 5 The relative clause

$$X\ [_{\overline{NP}}\ \#\ [_{\overline{S}}\ Y\ [_{\overline{NP}}\ N\ ]_{\overline{NP}}\ X'\ ]_{\overline{S}}\ \#\ N\ ]_{\overline{NP}}\ Z'$$

SI: 1 2 3 4 5 6 7 8

Cond: 4 = 7

SC: 1 0 3 0 5 0 7 8

The relative clause differs from the possessive and attributive rules by requiring identity of the whole NP to the head noun, not just head noun identity. It also has no restriction on the verb or on the tense, and there is freedom to choose an NP either preceding or following the verb.

e.g. $[_{S}\ [_{\overline{NP}}\ [_{\overline{NP}}$?atapa $]_{\overline{NP}}$ [ino $]_{\overline{NP}}$ [nĭtĭ $]_{\overline{NP}}$ pi?ipi$]_{S}$ # N'

chicken tiger day ate Noun

Depending on identity with the head noun, the rule will produce three possible relative clauses from the example above.

13. ?atapa nĭtĭ pi?ipi ?ino (?ino = N')

'the tiger that ate the chicken in the daytime'

14. ?atapa ?ino pi?ipi nĭtĭ (nĭtĭ = N')

'the day that a tiger ate the chicken'

15. ?ino nĭtĭ pi?ipi ?atapa (?atapa = N')

'the chicken that the tiger ate in the daytime'

An NP with an attributive or a possessive will not satisfy the identity condition and cannot be deleted. To prevent an attributive noun alone from meetin the identity condition, an additional condition is needed, such that 'X does not contain an S dominating the noun (4)' or some such condition. Thus, # [[[N]] ?atapa] ?ino pi?ipi # N' (N-chicken-tiger-ate) will not produce 16 or 17:

16. *kimiša ?atapa ?ino pi?ipi (kimisa = N')

*'the three that the tiger ate chicken'

17. *?ino pi?ipi min ?atapa (min ?atapa ≠ N')

*'the your chicken that the tiger ate'

Identity of the entire NP prevents a head noun from being deleted, leaving possessives or modifiers.

18. *[min] ?ino pi?ipi ?atapa (?atapa = N')

*'the chicken that your the tiger ate'

Thus the conditions imposed on the formation of the relative clause meet the well-formedness requirements, generating grammatical strings and excluding the ungrammatical ones.

There are no relative particles corresponding to 'who,' 'that,' and 'which' of English. No wh-attraction rules are needed, nor shift rules to place the clauses before or after the head noun, for they can only precede the noun.

The possessive and attributive rule must precede the relative clause rule, for the relative clause would otherwise prevent the formation of attributives and possessives, since the SI of the latter two fits that of the former.

The noun-phrase-forming rules should precede the non-restrictive relative clause rules (non res rel cl) and the causative rules, because the non res rel cl and causatives require NP identity between an embedded S and the matrix S. Since the NP's of the embedded S will have been formed before the rules cycle on the matrix S, formation of the matrix S NP's before the embedding T rules operate is necessary for NP identity.

T 6 The transitivizer

$$X\ [\ NP\ \begin{Bmatrix} V \\ AP \end{Bmatrix}\ \underbrace{Y\ ]}_{VP}\ Z$$

SI: 1 2 3

SC: 1 2+[trn] 3

i.e., a feature of transitivity [+trn] is adjoined to the verb or AP if it is contained in a transitive VP.

The feature will be given phonological shape by a later rule (MP 1) according to the type of verb or adjective that carries the feature.

e.g., (transitivizer underscored)

19. winti wisohawi 'make the paddle black'
    paddle black imp

20. kapa kayatankin 'he pounced on the squirrel'
    squirrel jump

21. kĩnti botonkin 'he took the pot down'
pot descend

T 7 The reflexive verb

X NP [NP {A, V} Y]$_{VP}$ Z

SI: 1 2 3 4 5 Cond: 2 = 3

SC: 1 2 0 4 5
[+R]

i.e., if the subject and object NP's are identical, the object NP is erased and the verb stem or adjective is assigned a reflexive feature [+R].

The rules which assign transitivity concord markers to verbs (T 23, T 24) and the morpheme lengthening rules (T 21, T 22) which lengthen nouns not immediately followed by a verb require that the object NP of reflexive verbs be deleted before they operate, for they will treat the reflexive as *an* intransitive verb. If a later ad hoc rule for deleting the reflexive object NP is to be avoided, the object NP must clearly be deleted in this rule. The transitive-marking rule T 6 must therefore precede T 7.

A later rule (MP 3) will spell out the phonological shape of the [+R] feature according to the phonological features present in the verb.

22. ta haa haa ṣa?tĩ [-past, -fut] 'he cuts he' →
Md he he cut

ta haa ša?tĩ?t [-past, -fut] → ta haa ṣa?tĩti?ki
'he cuts himself'

The causatives

T 8 Causative (ma)

$$X\ NP\ \underset{PP}{[}NP\ V\ \#\ \underset{S}{[}NP\ \underset{VP}{[}(NP)\ \begin{Bmatrix}AP\\V\end{Bmatrix}\ (Adv)\underset{VP}{]}\ tense\underset{S}{]}\ \#\ Y\ tense\underset{PP}{]}\ Z$$

1 2 3 4 5 6 7 8 9 10 11 12 13 14

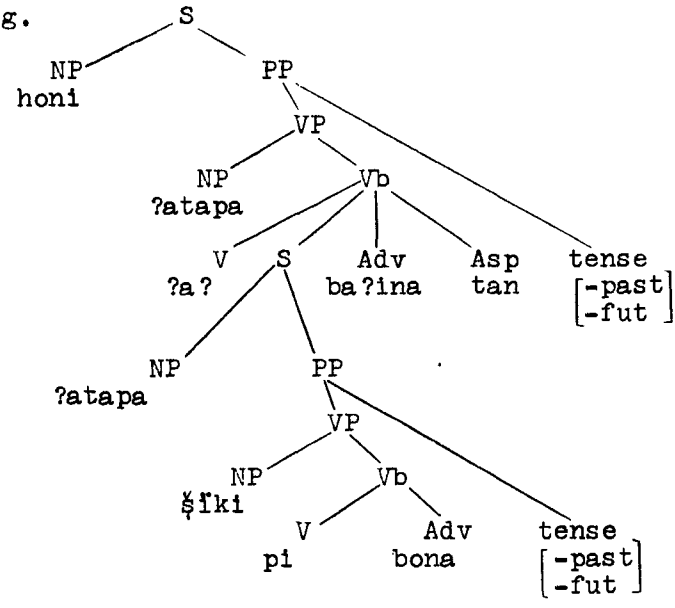
Cond: 3 = 6
10 = 13
4 = $\begin{bmatrix}?a?\\+V\end{bmatrix}$

SC: 5,6,7,8,9,10,11 → 0
3 → 7+3
4 → 8+9+ma

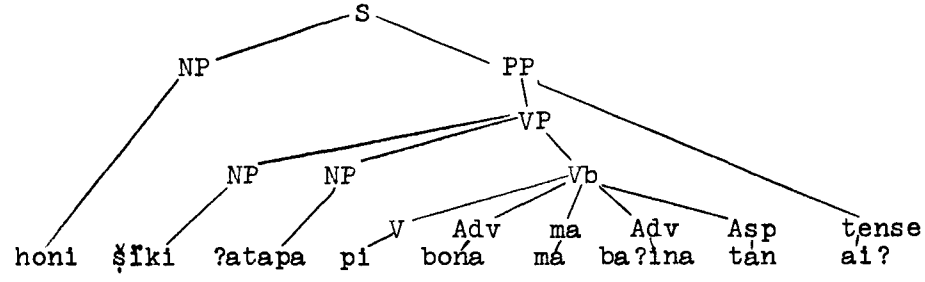
i.e., if there is a sentence embedded in a verb, and the object NP of the matrix sentence is identical to the subject NP of the embedded S, and the tenses are identical, and the verb of the matrix S is ?a? 'to make, to do, to say, to command, to cause,' the embedded S is deleted, a copy of the embedded S's object NP (if there is one) is adjoined to the matrix object NP, and a copy of the embedded verb stem and adverbial suffixes replaces the matrix verb.

There remains then only one verb, but it may have two series of Adv suffixes: the series immediately following the verb has its source in the embedded verb, and the series following ma comes from the matrix verb.

e.g.



becomes modified to



man corn chicken eat along all-day away pres.

'the man went along all day causing his chicken to eat while going'

There are no cooccurrence restrictions on any constituent of PP, VP, or Vb that require the causative S to be embedded in the verb by the base rules. The causative S could as well be contained in VP or PP, then be shifted by the embedding T rule to the verb where it now stands. The only advantage the present treatment seems to offer is that

the source S is located where its verb affixes will remain.

The fact that the embedded S may have an Adv series of affixes explains why two such sets may occur in the same verb in the surface structure. The identity of subject NP of the embedded S with the object NP of the matrix explains why this NP, in the surface string, is both agent and goal of the action expressed by the verb, and the fact that the object NP of the embedded S will be assigned a [-lengthen] feature before it is shifted explains why it is not lengthened in the surface structure.

No causative constructions with the subject NP identical to the object NP are grammatical: *'I caused myself to . . .' The ordering of the reflexive T 7 before T 8 and T 9 will prevent T 8 and T 9 from operating on reflexives, by deleting the object NP needed for NP identity in the latter two rules. The ordering thus makes it unnecessary to impose a special condition in T 8 and T 9 that the subject is not identical to the object of the matrix sentence.

The meaning of the sentences that undergo T 8 is 'to cause someone to do something.'

e.g. 23. honin ṣ́ĭki pikin 'the man eats corn'

24. honin ṣ́ĭki ʔatapa pimakin 'the man causes the chicken to eat corn'

25. ʔaiban bakĭ ʔoṣ́amakin 'the woman caused the baby to sleep'

26. ʔočitinin yawa kamakin 'the dog caused the pigs to go'

27. ?očitinin yawa kaka?inmakin 'the dog caused the pigs to go rapidly'

T 9 Causative conjunction (ki?n)

X NP $[_{PP}$ (NP) {V,AP} # $[_{S}$ NP $[_{VP}$ (NP) {V,AP} (Adv)$]_{VP}$ tense$]_{S}$ # Y tense$]_{PP}$ Z

1 2 3 4 5 6 7 8 9 10 11 12 13 14

Cond: 3 = 7
4 = 8
10 = 14

SC: 3,5,6,7,8,9,10,11 → 0
4 → 7+6+8+9+ki?n

i.e., if the object NP's are present they must be identical, the verbs must be identical, and the tenses also identical. Like T 8, T 9 deletes the embedded S and replaces the matrix verb by the constituent verb and Adv series of affixes. A causative indicator ki?n is attached instead of the ma of T 8 and a copy of both the subject and object NP's of the constituent S are adjoined to the verb, but in reverse order.

The meaning of the embedded S will be 'to cause someone to do something by doing it with him.' Thus, ?očiti ni?ki?nkin means 'to walk the dog.' The actor
dog walk
participates in the walking, whereas T 8 produces a sentence ?očiti ni?makin 'to cause the dog to walk' without the actor getting involved in the walking activity. The causation of T 9 involves getting the object to do something, with no emphasis on the involvement of the subject in the activity, though in the glosses I have had to use 'with' to make reasonable English in many cases. Evidence to support the

treatment of the subject of the underlying sentence as object in the surface string is that a transformation requiring object NP identity (T 13) operates under identity with the underlying subject, now object, if the matrix verb of T 13 is a T 9 causative.

Like T 8, there are no cooccurrence restrictions: the causative ki?n can be used with any verb. But if the verb is intransitive in the underlying string, it will be transitive after the transforation, since the underlying subject is an object NP in the surface string; and if the underlying sentence is transitive, the transformed structure will be "di-transitive" with two object NP's.

Both the underlying subject and object NP's will be short forms in the surface string. I account for this by the switch in order of the subject and object in the transformation. In the underlying sentence, the object precedes the verb so is marked [-lengthen]. After the transformation the subject precedes the verb and is marked [-lengthen] on the next cycle by the lengthener rules. The subject NP of the matrix S will in any case be lengthened, for it will always have an NP between itself and the verb. This means that the lengthener rules must follow the causative rules, for if they preceded them the wrong subject NP would receive the [+lengthen] feature. The relative order of T 8 and T 9, however, is inconsequential.

28. piroto noya?i? 'the pilot flies'

29. honi noya?i? 'the man flies'

30. pirotonĭn honi noyaki?nkin 'the pilot flies the man' (i.e., flies with the man)

31. ?očiti niči 'the dog walks'

32. ?i?bo niči 'the owner walks'

33. ?i?ban ?očiti ni?ki?nkin 'the owner walks the dog'

34. ?očiti bĭnati 'the dog mourns'

35. ?i?bo bĭnati 'the owner mourns'

36. ?očitinin ?i?bo bĭnaki?nkin 'the dog mourns with the owner'

37. ?i?ban ?očiti bĭnaki?nkin 'the owner mourns with the dog'

38. ?i?ban pikin 'the owner eats'

39. ?očitinin pikin 'the dog eats'

40. ?i?ban ?očiti piki?nkin 'the owner eats with the dog'

41. ?očitinin ?i?bo piki?nkin 'the dog eats with the owner'

42. ?i?ban šĭki pikin 'the owner eats corn'

43. ?i?ban šĭki ?očiti piki?nkin 'the owner eats corn with the dog'

The lengthening of the matrix's subject NP when the verb is intransitive-causative, as in 30 above, has been easily explained by the fact that the lengthening rules reapply when the rules are recycled on the matrix sentence. It constitutes evidence that all of the cyclical transformation rules must be cycled afresh on each new domain, moving from

the most to least embedded S. Ross and Lakoff (1967)[1] have recently challenged the notion of the cyclical rule, suggesting that the rules should be applied individually, first rule 1 through all the sentences from most to least embedded, then rule 2 in the same way, etc. After the final rule, the T component would have finished its course with each rule having applied successively on all the sentences, but in sequence so that after rule 3, rule 2 never applied again. If this approach were used, forms that should be lengthened would be short, and vice versa, in T 9. But without an embedded S, either cyclical or non-cyclical rules would produce the same result.

44. ʔiʔsa noyaʔiʔ 'bird flies'

ʔiʔsa is short because it stands before the verb.

[ʔiʔsa noya [hawĭn bakĭ noyaʔiʔ] ʔiʔ]
S S S S
bird fly her child fly tense

If the rules are applied as Ross suggests, the sentence above would have neither 'bird' nor 'child' lengthened. After the embedding, both would be short in the surface structure:

45. *ʔiʔsa hawĭn bakĭ noyakiʔnai 'the bird flies with her young' (i.e., teaches it to fly)

ʔiʔsa, however, should be long:

46. ʔiʔsapan hawĭn bakĭ noyakiʔnai

If the verb is transitive, e.g.

[ʔiʔsa ṣĭki pi [hawĭn bakĭ ṣĭki piʔiʔ] ʔiʔ]
S S S S
bird corn eat her child corn eat pres tense

'the bird eats corn with her child'

and the rules are applied non-cyclically, only 'corn' will not be lengthened in either cycle. The transformed sentence will then find 'bird' and 'her child' both long (incorrectly):

47. *?i?sapan ṣ̌ĭki hawĭn bakĭn piki?nai

Applied cyclically, the correct forms appear:

48. ?i?sapan ṣ̌ĭki hawĭn bakĭ piki?nai

If the lengthener rules were applied only after the embedding again the wrong forms would be lengthened:

49. *?i?sapan ṣ̌ĭkin hawĭn bakĭ piki?nai

Like T 8, the underlying embedded S of T 9 does not necessarily have to be found in the verb, but the results of applying the transformational rules in non-cyclical order would be the same as described above if the embedded S were dominated by VP, PP, or even S. To explain the location of the verb affixes as in T 8, the source S is located in the Vb; and since the two causatives are in complementary distribution, the same source S is used.

The glosses given for the examples show that there is more than just a causative sense to the embedded S. One might suspect that there is simply a conjunction of S's or of NP's underlying the surface structure, rather than a causative construction.

If the causative conjunction were the result of the sentential conjunction of two underlying sentences such as 'the bird flies,' 'her child flies,' the effect of applying the T rules cyclically or sequentially would be the same as

has just been described: the sequential application would produce ill-formed sentences.

If the conjunction were NP conjunction (i.e., without an S source), some non-ad hoc way to account for the lengthenings still would need to be found. I can find none. A more serious objection to a conjunction analysis is that the type of conjunction that can be shown to exist (explained in base rule R 2) does not affect transitivity. The non-restrictive relative clauses and narrative sentences have verb endings that require different allomorphs if the sentence each is related to is transitive, i.e., if a noun is standing between the subject NP and the verb. The conjoined subject NP's do not change the transitivity of the sentence. But the NP which I have argued here as being the subject NP in an underlying S does change an intransitive sentence to a transitive sentence, for it now serves as object NP. The transitive allomorphs of the verb endings are required when the underlying sentence's subject NP precedes the verb.

Nor can the same NP be merely a second object NP inside the VP bracketing from a vule VP → (NP) (NP) Vb, with the explanation that it is short because it is dominated by VP (i.e., VP domination produces short forms), for the subject NP of intransitive verbs is not dominated by VP, yet is short in form. It would also not account for the paraphrases (by native speakers) of T 8 and T 9 with sentences in which the subject NP of one is that of the matrix S, and that of the other belongs to the constituent S.

Complements

T 10 Sentence complement (1)

$$X\ VP_{VP}[\#[_{S}\ NP[_{VP}\ Y\ (NP)\ Vb]_{VP}\ Nom]_{S}\ \#\ Vb]_{VP}\ Z$$

$$1\ 2\quad 3\quad 4\quad 5\quad 6\quad 7\quad 8\ 9\quad 10$$

Cond: $2 = 4$ $\qquad$ SC: $7 \rightarrow \begin{bmatrix}\alpha J\\ \beta K\end{bmatrix}$

$9 = \begin{bmatrix}+cmpl\\ \alpha J\\ \beta K\end{bmatrix}$ $\qquad$ $3, 4, 8 \rightarrow 0$

i.e., the sentence complement which has its source in the S in VP (R 5) requires that the subject NP's be identical and that Nom and not a tense be dominated by Aux. The features of the complement permitted by the verb are then assigned to the Nom symbol and will later be spelled out in phonological shape.

The "J" feature is characteristic of verbs of motion and a few others, and will assign an i suffix to all verbs. The "K" feature is carried by a group of verbs difficult to classify. "K" assigns a -ti suffix to the verb and will be covered in the the next rule. Many verbs and all adjectives will have a [-complement] feature and will thereby block the sentence.

The morpheme lengthener rules (T 21, T 22), which must follow the complement rule, require that the subject NP of the underlying sentence in the complement be deleted, not merely marked for deletion and deleted later; the lengthener rules lengthen any noun not immediately followed by a verb or a suffix. In 51 below, where 50b is made a

complement, ʔĭa would be lengthened to ʔĭan of the deletions were not carried out before the lengthener rules apply.

50a. ʔĭa kaʔi 'I go' (short)

50b. ʔĭa ʔoṣaʔi 'I sleep' (short)

51. ʔĭa ⌊ʔĭa ʔoṣaʔi⌋ kaʔi → ʔĭa ʔoṣaʔi kaʔi
'I go to sleep' (short)

52. ʔĭa [ʔĭa hiwi rĭraʔi] kaʔi → ʔĭan hiwi rĭraʔi kaʔi 'I go to chop a tree' (long)

The lengthened subject NP ʔĭan in 52 is accounted for by the fact that the undeleted NP hiwi 'tree' intervenes between the subject NP and the nearest verb. In 51 nothing intervenes after the deletion, so the NP ʔĭa remains short. The complement rule with its deletion of subject NP must therefore precede the lengthening rules.

The complements must precede the non res rel cl's because the latter must refer to subject and object NP's in the following sentence, and the complement rules define the subject or object NP of the following sentence by embedding the complement sentence and deleting the subject NP of the constituent sentence.

T 11 Sentence complement (ti)

$$X\ \underset{VP}{NP}\ [\ [\ \#\ [\ \underset{S}{NP}\ [\underset{VP}{Y}\ (NP)\ [\underset{Vb}{W}]\ ]\ \underset{VP}{Nom}]\ \#\ [\underset{NP}{N}]\ \underset{VP}{Vb}]\ Z$$

1 2 3 4 5 6 7 8 9 10 11 12

Cond: 2 = 4 SC: 3, 4, 7, 8, 9 → 0

$$11 = \begin{bmatrix} +cmpl \\ \alpha\ J \\ \beta\ K \end{bmatrix} \qquad 10 \rightarrow 7+8\begin{bmatrix} \alpha\ J \\ \beta\ K \end{bmatrix}$$

$$10 = \begin{bmatrix} +pro \\ -spkr \\ -addr \end{bmatrix}$$

i.e., if the subject NP of the complement and matrix are identical and a complement-taking verb stem is included in the Vb of the matrix, and the third person pronoun fills the head noun slot of the complement-including NP, then the complement's subject NP is deleted, the pronoun is replaced by the verb stem and its affixes, and the complement features are assigned to the Nom symbol. Nom will later be spelled out as the suffix -ti.

53. ʔĭan rĭrati ʔonankin 'I know how to chop'
I chop know

54. ʔĭan nonati ʔonanyamakin 'I do not know how
I swim know not to swim'

55. ʔĭan nonabaʔinati ʔaʔtipanyamakin 'I am not able
I swim all day able not to swim all day'

56. ʔĭan ʔatapa pimati ʔaṣĭkin 'I learned how to
feed chickens'

57. ʔĭan ṣoboaʔti kĭĭni 'I want to build a house'

58. ʔĭan nonati kĭĭni 'I want to swim'

59. maṣopan mai pokinti pišini 'the possum needed
to dig a hole'

The "nominalized" ti complement does not carry a noun label but no longer has Vb bracketing. Therefore the subject noun phrase (4 of the SI) will be lengthened, being separated from the nearest verb by an NP.

60. ʔĭan [nonati]$_{\overline{NP}}$ ʔonankin 'I know how to swim'
I swim-to know

The transitivity concord markers (e.g., kin in 60) show that the verb is transitive and thus require that the verb be preceded by an object NP. i.e., nonati above must carry an NP bracketing. Selectional restrictions on the verbs that can take the -ti complement also require that the CS of the verb be transitive, i.e., preceded by NP in VP. These are the reasons for requiring that the complement be bracketed by NP. The head noun of the NP, a unique member of its class, is deleted. There are no extra-position rules cr "it" pronouns like there are in English to strengthen the argument for using a deletable pronoun to account for the NP bracketing on the complement, but the facts stated above would require additional complexity of selectional restrictions and transitive-concord rules if the -ti complement were derived from an S source in VP as the -i complement is.

T 12 Noun complement

	X	[	Y	#	[S	NP	[	(NP)	[Vb	W Vb]	Nom]	#	N] NP	Z
SI:	1			2		3		4		5	6	7	8	9
SC:	1			0		3 [+comp]		4		0	0	0	4+5 [+comp, +K]	9

Optional SC: 3 → 0

Cond for T 12: $3 = \begin{bmatrix} +pro \\ -spkr \\ -addr \end{bmatrix} = 8$

i.e., a noun complement is formed by a sentence modifying a unique pronoun. The pronoun is replaced by the verb stem

and its affixes and a nominalizing feature [+K] is assigned to the Nom symbol. This will be spelled out as a suffix ti. The subject NP, also a third person pronoun, is usually deleted, but for paraphrasing and for clarrification when a new form is nominalized, it may be retained.

As in T 11, the structural description of the resultant form must not be a verb, for nouns preceding it are lengthened.

61. ?ĭan ɸa?oti hiskin 'I saw a chair'
I sit see

62. haan tooa?ti mĭrakin 'he found a gun'
he shoot find

The transitive concord markers on verbs require that the noun complement have an NP bracketing. This is automatic since the complement is developed inside an NP.

The noun head of the NP is always erased. It must therefore be a distinguished representative of its class, to avoid deletion of unrecoverable information. Pronouns are never modified by a sentence or adjective unless they are demonstratives. Of the demonstratives, only hawa 'what' may occur to the right of its modifiers. I therefore regard it as the best choice for the deletable pronoun. The w of hawa changes to ? when the [+comp] feature is present.

63. ha?an ɸa?oti (to sit) 'chair'

64. tooa?ti (to make a noise) 'gun'

65. nonti (to float) 'canoe'

66. winti (to paddle) 'a paddle'

67. šaʔtɨ̄ti (to cut) 'shears'

68. sawɨ̄ti (to clothe) 'dress'

69. kɨ̄oti (to sing) 'radio'

70. ʔoroti (to cut grass) 'lawn mower'

Non-restrictive relative clauses

T 13 Simultaneous action (-ton)

X # [Y NP [Z Vb X']VP]S [Y' tense]S # (Md) (NPbɨ̄tan) NP

1 2 3 4 5 6 7 8 9

[X" NP Vb]VP Z" tense W

10 11 12 13

Cond: 4 = 11

6 = [-past, -fut]

SC: 11 → 11 [+del]

6 → 6+ton

8, 9, 10, 11, 12 not part of an NP 2, 7 → 0

i.e., if a sentence produced by the (# S #)^n part of R 2 precedes a series of sentence constituents such that the subject NP of the first is identical to the object NP of the second, and the tense of the first is [-past, -future] (present), then the object NP of the second is marked for deletion and an affix ton is adjoined to the tense symbol.

The non-restrictive relative clause sentences produced by the base will be in a string dominated by the same S. The convention of applying the rules to the most embedded S first, therefore, will not specify which of the string of

S's is to be processed first. The SI of the rule will specify the beginning point, however, for the single # juncture after the embedded S requires that the right-most S be processed first. Since each of the other S's in the string will be embedded on the basis of subject or object NP identity, the deletion of NP's by the non-restrictive relative clauses must be delayed until the information required for the embedding of the other S's is no longer needed.

If the non-restrictive relative clauses were hierarchically nested, the rules would be cycled from the most embedded (left-most) to the lest embedded, but as the present rule will show, the NP deletions would still need to be delayed because the deletion of the object NP will erase the information needed on the next cycle if the next-highest is a narrative sentence or an ?a? non-restrictive relative clause, both of which require the object NP in their SI. Deletion delay is therefore required whether the non-restrictive relative clauses are hierarchically nested or sequentially arranged.

The reason for the constraint that the second sentence is not part of a noun phrase (or some such condition) is that the non-restrictive relative clause may precede the relative clause of an NP. But the non res rel cl must be subordinated only to strings dominated by S, not by NP. That is, if a relative clause follows, it will be included

in NP 8 or 9 of the SI. The theory requires that labeled bracketing must be present if it is included in the SI of the rule, but nothing is indicated by the absence of bracketing in the SI. The SI therefore cannot exclude the NP bracketing except by a special condition, as stated.

The condition that the tense be present tense makes the action of the clause simultaneous to that of the following verb. For this reason I have translated the -ton clause by 'while ...' in order to distinguish it from the other clauses.

71. ʔawa mapɨ̈taiton ʔɨ̈an tooaʔkin
tapir ascend-while I shot

'I shot the tapir while he climbed the river bank'

72. ʔɨ̈n boča piʔiton čokan wɨ̈rankin
my brother eat-while sister shoved

'Sister shoved my brother while he was eating'

73. ʔatapa kɨ̈otaiton maṣopan biʔkin
chicken crow-while possum caught

'the possum caught rooster while he crowed'

T 14 Preceding action

X # [Y NP [Z NP {AP,V} X'] Y' tense] # (Md)(NPbɨ̈tan) NP VP Z'
S VP VP S

1 2 3 4 5 6 7 8 9 10 11 12 13

Cond: 6 = 11
8 = [+past]
10, 11, 12 not part of an NP

SC: 2, 9 → 0
6 → 6
[+del]

i.e., if the tense is past and the object NP is identical to the subject NP of the following sentence (with the same

domination conditions as T 13), the object NP is assigned a deletion feature.

The deletion feature which enables deletion to be delayed is required independently by a later rule that will erase identical object NP's in consecutive non-restrictive relative clauses. Immediate deletion would destroy the basis of NP identity.

There are no passives in Capanahua, but T 14 produces a structure that can be translated by a passive or by a non-restrictive relative clause with past tense.

74. binan ʔaʔkaʔ ʔĭa habati
bee stung I fled
'having been stung by a bee, I fled'

75. ʂĭniban rĭʔtĭni ʔiso yamaskai
forebears killed monkey is not
'the monkeys which were killed off by our ancestors, no longer exist'

76. pirotonĭn ʔiwĭpi bakĭ koinkoini
pilot brought child happy
'the child who was brought by the pilot is happy'

T 15 Preceding action (aš)

X # [S Y NP VP Z tense] # (Md)(NPbĭtan) NP [VP X' (NP) V' Vb] Y' tense Z'

1 2 3 4 5 6 7 8 9 10 11 12 13 14

Cond a: 9, 10, 11, 12, 13 not part of an NP

Cond b: 4 = 10

7 = [+past]

SC: 2, 8 → 0

4 → 4 [+del]

7 → 7+aš

i.e., if the tense is [+past] in the non-restrictive relative clause, and the subject NP's of the two sentences are identical, the first subject NP is assigned the deletion feature.

The condition (a) that 9, 10, 11, 12, 13 are not part of an NP is a general condition for all rules T 15 through T 19 and therefore need not be repeated for each rule. It has therefore been separated from the conditions of (b) for the purposes of the simplicity metric, but the rules will still be read so that all the conditions apply, not (a) or (b). Condition (a) does not actually accomplish what it intends in terms of formal theory but will have to suffice for lack of a more precise means of specifying the condition.

The transformed sentence may be translated 'having . . .' 'when he had,' 'if he had.'

77. mani pi?aṣ ?iso ka?i
banana eat monkey go

'having eaten the banana, the monkey goes'

78. kayama?aṣ ta ?ĭn piti pišinti?i
go not Md I food lack

'If I don't go, I will lack food'

79. taama bananiš ?ĭa hoyamaska?ipiki
peanuts plant I came not more

'having planted the peanuts I did not return'

80. šabaka?inaṣ howĭ noka?i
shine suddenly lamp extinguished

'the lamp flashed then went out'

A later rule will change the shape of the suffix adjoined to 7 if the following sentence is transitive:

81. kaṣon ta ʔĩn mani pitiʔi
go Md I banana eat

'if I go, I will eat a banana'

In T 15 and the following non res rel cl's, the time adverbial that may occur in the clause is restricted by the tense condition. Thus, a time NP nĩa nĩtĩ 'this day, today' cannot be used if the tense is specified as past. Otherwise there seem to be no restrictions on the time adverbial, though if the adverbial is identical to that of the next clause, it will be deleted by a later rule (T 24).

T 16 Subsequent action (noṣon)

SI: Same as T 15	SC:	4 → 4 [+del]
Cond a: Same as T 15		
Cond b: 7 = [+fut]		7 → 7 + noṣon
4 = 10		2, 8 → 0

i.e., if the subject NP's are identical and the tense of the first S is future, mark the first subject NP for deletion and adjoin noṣon to the tense of the first S.

The meaning of the transformed structure will be 'in order to,' 'so that.'

82. banĩnoṣon ta ʔĩn kaʔi
remain Md I go

'I am going in order to stay'

83. yapa miškinoṣon ta ʔĩn taʔno kĩĩnai
fish hook Md I grub want

'I want grubs in order to catch fish'

84. kayaša?nošon haan čopa bïšoa?i
go [fut] he clothes fix

'he is fixing his clothes in order to leave someday'

85. ?ošanošon rakawï
sleep lie down [imp]

(in order to sleep, lie down)

'lie down and go to sleep'

86. ?oa hisnošon hiwin haa ?inati
that see tree he climb

'he climbed the tree so that he could see that'

87. hano rakanoša?nošon čašon mai taškakin
there lie down [fut] deer earth scratch

'the deer pawed the ground in order to lie down there'

T 17 Subsequent action (non)

SI: Same as T 15

Cond a: Same as T 15 SC: 2, 8 → 0

Cond b: 7 = [+fut] 7 → 7 + non

Cond c: 4 ≠ 10

Cond d: 4 = 12 SC: e or f

e: $4 \rightarrow \underset{[+del]}{4}$

f: $12 \rightarrow \underset{[+del]}{12}$

i.e., if the conditions are the same as for T 16, except that the subject NP's are different, no subject NP deletions

take place, and a suffix non is adjoined to the tense of the first sentence. Furthermore, if the object NP of the second sentence is identical to the subject of the first, either of two changes may be effected: (e) mark the first subject NP for deletion, or (f) mark the second S's object NP for deletion.

88. han banĩnon kawĩ
he stay go [imp] (so that he will stay, go)

'go on and leave him'

89. ?ĩn ?oṣanon nĩtĩwĩ
I sleep be quiet[imp]

'so that I can sleep, stop talking'

90. rono mawanon haan ɸakakin
snake die he hit

'he hit the snake so that it would die'

91. mawanon haan rono ɸakakin
die he snake hit

'he hit the snake so that it would die'

92. han ?onannon papan yo?ikin
he know father told

'so that he would know, father told him'

93. ?onannon papan haa yo?ikin
know father he told

'so that he would know, father told him'

94. bimi hayaṣa?non banakin
fruit be[fut] plant

'so that there will be fruit he planted it'

95. hayaṣa?non bimi banakin
be [fut] fruit plant (Same as 94)

T 18 Simultaneous action (ya)

SI: Same as T 15

Cond a: Same as T 15 SC: 7 → 7 + ya

Cond b: 7 = $\begin{bmatrix}-\text{past}\\-\text{fut}\end{bmatrix}$ 2, 8 → 0

4 ≠ 10

i.e., if the tense is present tense in the first S and the subject NP's are different between the first S and the second, neither of the subject NP's is marked for deletion, but a ya suffix is adjoined to the tense symbol.

This sentence can be translated 'if . . .' or 'when . . .' Its similarity to T 15 is shown by the fact that the same SI is used, but with the ya suffix only present tense may be used.

96. haa ʔisinaiya ʔĩa kaʔi
he sick I go

'if he gets sick, I will leave'

97. ʔoi hawiya ʔĩa tĩĩyamaʔi
rain comes I work not

'if it rains, I will not work'

98. nai ranca nokotaiya ta ʔĩn mia kĩnatiʔi
sky launch arrive Md I you call

'I will call you when the plane comes'

99. bari ʂanaʔiya wai mapoawĩ
sun hot clearing burn

'set fire to the clearing when the sun is hot'

100. patoronĩn mia sinaaʔiya ʔĩa honĩtipiki
employer you scold I hide

'I hid when the employer was scolding you'

T 19 Simultaneous action (kin)

SI: Same as T 15

Cond a: Same as T 15

$7 = \begin{bmatrix} -\text{past} \\ -\text{fut} \end{bmatrix}$

SC $7 \rightarrow 7 + \text{kin}$

$4 \rightarrow 4\ [+\text{del}]$

$2, 8 \rightarrow 0$

i.e., if the first tense is present tense and the subject NP's are identical, mark the first one for deletion and adjoin the suffix kin to the tense symbol.

This rule is the "same subject" counterpart of T 18 and will be translated by 'while' in most instances.

101. namakin haan yošin hiskin
dream he spirit saw

'he saw a spirit while dreaming'

102. mai pokinkin nokĩn kori mĩrakin
earth dig we gold found

'we found gold while digging in the ground'

103. kcšin winakin haa winoka?ini
strong paddle he pass rapidly

'by paddling furiously, he passed quickly by'

104. ?ošakin haan hawĩn bakĩ bištašķin
sleep she her child squash

'while she was asleep, she lay upon her offspring'

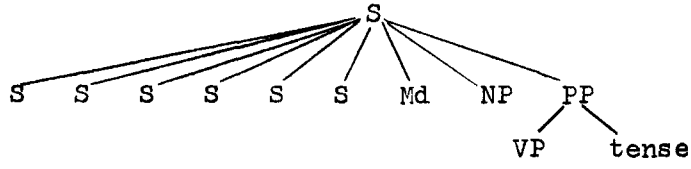
105. kakin ?ĩan haa yo?ikin
go I him told

'while going, I told him'

106. kakin ?ĩa rašķoba?ini
go I slip

'while going, I slipped down (the hill)'

The order of the rules T 13 through T 19 with respect to one another is arbitrary. The delay of deletion enables any of them to apply in any order to the linear string of S's, and the cycle could begin at any point in the string. Thus, given the tree



the fact that the string of S's are all equally embedded presents no problem to the operation of the rules. The S's are cycled one by one until they have all been cycled. The fact that the S's are processed right-most first because of the SI of the rule is a consequence of arbitrarily choosing to include only one # symbol on the right of the S, which does not affect the need to delay deletion.

The non-restrictive relative clauses, as I have labeled them, are probably not all true non-restrictive relative clauses as we think of them in English. Another term may be more appropriate to some of them, depending on whether one focuses on the condition of tense or the condition of NP identity specified by the rule. In any case, the similarities of SI and SC warrant their grouping together and make it possible to capture what seems to me to be a linguistically significant generalization: that they all have the same source and are embedded under conditions set on tense and NP identity.

T 20 The narrative (conjoined) sentence

X	tense	Nar	#	[S	Y	tense	]S	#	Z
1	2	3	4		5	6		7	8
1	2	3	&		5	i [+Nar]		&	8

Cond: 2 = 6

i.e., a conjoined sentence is well-formed if the tense is identical to that of the matrix sentence. The narrative symbol is deleted and i with a narrative feature replaces the tense symbol. The sentence boundaries are changed to pause markers (&).

The narrative feature has now been spelled out as a morpheme and the tense features will not be given phonological features. Thus all conjoined sentences will lack overt tense and person markers and will always share the tense of the matrix sentence. Depending on the tense of the matrix sentence, the tense of the narrative suffix can be present, past, or future, though by nature the conjoined sentences in legends, narrations, tales, etc. will usually be a past tense.

T 21 Lengthener exception

X	[Y Z]NP	{(Vb, sf) / W ?1 ?k}	Z'
1	2	3	
1	2 [-next rule]	3	

Cond: 2 = (N, V, A, sf)

i.e., if a noun, verb, adjective, or suffix stands in a noun phrase immediately before a verb or suffix outside the NP, or if it is followed by ?i?k 'to be' it is assigned a feature (heretofore referred to as a [-lengthen] feature) that marks it as an exception to the noun lengthener rule.

The forms so marked will not be phonologically lengthened. Thus, an intransitive verb will have a subject NP with a short form, and a transitive verb will have a subject NP with a lengthened form.

107. maṣ̌o ka?i 'the possum goes'
possum goes

108. maṣ̌opan ṣ̌ĭki pikin 'the possum eats corn'
possum corn eats

T 22 Lengthener Assignment

$$W \underbrace{[\ X\ Y\ ]}_{NP}\ Z$$
$$1 \quad 2 \quad 3$$

Cond: $2 = \{N, V, A\}$ SC: $2 \rightarrow 2\ [+M\ -T\ 4]$

i.e., the final lexical entry in a noun phrase is marked with the lengthening feature [+M] which will be used by the phonological rules to lengthen the phonological form of the morpheme.

It is necessary to specify the last lexical entry of the NP because marking the whole NP as [+M] would lengthen all of the words dominated by NP. The use of a word juncture will not serve to eliminate the condition by identifying the

morpheme to be lengthened because the NP may have many word junctures internally and will also have internal NP bracketings, some of which will have nouns that are lengthened from underlying sentences (i.e., possessives) and others not.

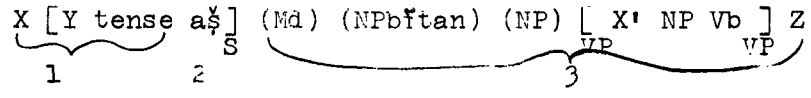
The lengthener rules show why in my earlier treatment (1963) subject, instrumental, possessive, vocative, and time NP's all seemed to use the same suffixes, for if a time noun occurred without the time suffix, it then took the same endings as subject, instrumental, etc.

109. čašon ta bariya?an hawĭn ta?ĭn mai wiša?iki
deer Md by day his foot earth paws

kokan 'The deer is pawing the earth in the daytime, Uncle.'
uncle

Transitive allomorphs

T 23 (šon)



SC: 2 → šon

i.e., if the VP following the non res rel cl terminating in aš is transitive, change aš to šon.

T 23 rule should precede the VP shift rule (T 25) in order to facilitate specification of the VP without permitting variables between aš and the VP, wherein one of the variables might be a VP, preventing application only to the VP immediately following the non res rel cl.

110. ?oro?aṣ̌ ka?i
cultivating go

'having cultivated, he went' (intr)

111. ?oroṣ̌on ?oat hiskin
cultivating that see

'having cultivated, he saw that'

112. ?oroṣ̌on Jose kĭnakin
cultivating Joe called

'having cultivated, he called Joe'

T 24 (kin)

X [Y NP Z i] W
VP [+Nar]

1 2 3

SC: 2 → kin

i.e., if the verb is transitive, the narrative suffix is changed to kin.

113. wasi ?orokin

grass cultivated 'he cultivated the grass'

114. mani banakin

banana planted 'he planted the bananas'

115. haa hiskin

that saw 'he saw that'

T 25 Miscellaneous NP deletions

X (NP) [Y (W) Z tense] (&) (NP) [X' (W) Y' tense] Z'
PP PP PP PP

1 2 3 4 5 6 7 8 9 10

General condition: 7, 8, 9 are not part of an NP

SC 1: Cond: 2 = 6 $\quad 6 \rightarrow \underset{[+del]}{6}$

SC 2: Cond a: 4 = 8 = NP = [+Aux] $\quad 8 \rightarrow \underset{[+del]}{8}$

Cond b: 4 = 8 = NP = [+VP] $\quad$ "

Cond c: 4 = 8 = pp $\quad$ "

i.e., if the subject NP's, time NP's, object NP's, or prepositional phrases of two consecutive sentences (either non-restrictive relative clauses or conjoined sentences) are identical, delete the second.

This rule is a merger of four structural changes that would otherwise be included in the non-restrictive relative clause rules and the conjoined sentence rule. Extracting them from those rules has not only reduced the number of symbols needed in the rules, but points up the fact that it is the redundant repetition of these NP's that the rule eliminates. The rule works in conjoined sentences in the same way as in the non res rel clauses.

Selectional restrictions require that a feature of the node dominating the NP be assigned to the head noun of the NP (R 18). To develop the proper suffixes of the time, indirect object, and associative (not shown in the base rules) NP's, another rule exploits the feature so assigned to the noun CS and spells out the shape of the morpheme related to the feature. The morphemes could be spelled out early in the T rules but it is convenient to spell them out after the deletions, so that a noun phrase will not be deleted leaving a morpheme (suffix) unattached to an NP.

Final cycle rules

The transformational rules are cyclical in operation so that they are reapplied in the order given on each successively larger domain, beginning with the most embedded sentence. However, there are some rules which must apply only on the final cycle because the structural changes (i.e., deletions or adjunctions that change the structural index of the string) which they effect would otherwise prevent the proper application of other rules in the following cycle. These rules have been designated as final cycle rules.

There are also some rules (e.g., T 28, T 29) that are final cycle rules only because their structural indices cannot be met by embedded sentences. Although such rules are not final cycle rules for the same reason as those mentioned in the paragraph above, they fall naturally into the same section of the transformational component.

Finally, there are some rules (such as those which spell out the phonological shape of morphemes) that need not be classed as final cycle rules, for they could be ordered almost anywhere in the transformational component. These have been placed late in the transformational cycle because of their relatedness to the phonological component, but no strong reason has been found to prevent them from being placed earlier.

T 26 Verb phrase shift

X	VP	Y	{Nom / tense}	Z	SC: 2 → 0
1	2	3	4	5	4 → 2 + 4

The location of the verb phrase in the base was required by the lengthener rules. The subject NP must not be lengthened before intransitive verbs but lengthened before transitive verbs. Noun phrases that follow the verb in the base rules are always lengthened (e.g., vocative).

A shift rule would have been required even without the lengthener rules, for the verb and tense must be final in the sentence. If the time NP were to follow the tense in the base, a rule to shift it back would be necessary. If it precedes the tense, a rule to shift VP is required.

116. ʔĭa ta ʔĭn rakĭt ʔaʔki baʔkiš [past] →
I Md I fear of him yesterday

117. ʔĭa ta ʔĭn ʔaʔki baʔkiš rakĭtipiki

'I was afraid of him yesterday'

T 27 Deletion of NP

X	[+del]	Y	SC: 2 → 0
1	2	3	

i.e., delete all deletable elements.

After T 27, many strings will be pared down to a sequence of verbs.

118. wĭsti ṣontakonĭn ta šĭki bananiški & banaʔipi
one girl Md corn planted planted

hokoni & ?ani?i & hoa?i & howĭska?iton bi?kin &
sprouted grew flowered ripened took

ǿinkikin & rĭnĭkin & rĭnĭṣon, bataakin & bataaṣon
gathered ground grinding sweetened sweetening

kĭntiki nanĭkin & nanĭṣon kobinhakin
in a pot put putting cooked

'A certain girl planted corn. After she had planted it, it sprouted, then grew and flowered. When it was ripe she took it and gathered it. And then she ground it. Having ground it. she sweetened it (by chewing it). Having sweetened it, she then put it in a pot. Having placed it in a pot she cooked it.'

T 28 Topic marker assignment

$$X \begin{Bmatrix} ta \\ haskaska?iya \end{Bmatrix} NP \underbrace{PP \quad Y}_{4}$$

1 2 3

SC: 3 → 3
[+topic]

i.e., a topic feature is assigned to a subject NP that is preceded by the mood marker or by the idiomatic non res rel cl construction haskaska?iya 'that was the way it was when.'

All NP's having the same referent as the topic will be pronominalized by the next rule. If the NP following haskaska?iya has the same referent as that following Md, there will be one series of pronominalized NP's. But if the NP's marked by 3 in the SI have different referents, the NP's following haskaska?iya will be pronominalized by both applications, since Md must always precede haskaska?iya in the string. (The rules for producing haskaska?iya have been

omitted in the base.) There will then be one series preceding and two series following haskaska?iya.

Topic assignment may precede or follow the deletion rules. If it precedes, and the key NP is deleted, the pronominalization rule (T 29) will not work. If it follows the deletion, it will not assign a topic if the NP is missing. The pronominalization rule must follow the deletion in either case to prevent pronominalization of NP's without an antecedent. To keep the topic-related rules together, T 28 precedes the pronominalization rule, and both follow the deletion rule (T 27).

T 29 Topic pronominalization

X NP Y NP Z

1 2 3 4 5

Cond: 2 = [+topic]

2i = 4i

SC: 4 → 4 [+pro]

i.e., pronominalize all NP's having the same referent as the topic.

119. honi ?ani ta honiški & hawiton ?ĩn bakĩn
man big Md came coming my son

hiskin & hawĩn čopa waṣati & hawĩn nonti
saw his clothes torn his canoe

?anita?ma?i
small be

'a big man came, and my son saw him come
His (topic) clothes were town. His (topic)
canoe was very small.

T 30 Pronominalization

$$Z''\ [W\ [X]_{\bar{N}\ NP}\ ]\ Y'\ [\ [\ Y\ ]_{NP}\ Z\ ]_{NP}\ Z'\ [W'\ [X']_{\bar{N}\ NP}\ ]\ Z''$$
$$\underbrace{\qquad}_{1}\ \ 2\ \ 3\ \ 4\ \ 5\ \ \underbrace{\qquad}_{6}\ \ 7\ \ 8$$

Cond a: The only S node over 7 is that dominating the string 1 - 8

$4_i = 7_i$

SC: $4 \rightarrow \underset{[+pro]}{4}$

Cond b: The only node over 2 is that dominating the string 1 - 8

$2_i = 4_i$

SC: $4 \rightarrow \underset{[+pro]}{4}$

i.e., an NP may be pronominalized if its referent is identical to that of the head noun of an NP constituent of the sentence. The NP constituent may be on either side of the pronominalized element.

120. [[bakĭn] [maman]] [[maman] [bakĭ]]
NP NP N N NP NP NP N NP
child's mother mother's child

mĭrakin →
found

121. [[hawĭn] maman] [[hawĭn] bakĭ] mĭrakin
her mother her child found

'the mother found her child'

The non-restrictive relative clauses cannot have subject or object NP's pronominalized by identity to another NP in the matrix, or vice versa:

122. *haa ka?aṣ̌ hono nokoti
he going peccary arrived

*'when he had gone the peccary arrived'

123. *hono ka?aṣ̌ haa nokoti
peccary going he arrived

*'when the peccary had gone he arrived'

124. *ka?aṣ̌ haa nokoti
going he arrived

*'having gone he arrived'

125. *haṣ̌oman ?a?ka? haa ka?i?
ant stung he go

*'he went after an ant stung him'

126. *han ?ona?a?na čoka bĭwa?i
she sew while sister sings

*'sister sings while sewing'

127. *čcka ?ona?a?na han bĭwa?i
sister sew while she sings

*'sister sings while sewing'

One way to account for the fact that pronominalization occurs in possessives but not in the non-restrictive relative clauses is to put a special condition on the SI of the rule: 4 = [+poss]. The rule would then operate only on formatives to which the [+poss] feature has been assigned. The [+poss] feature is needed independently, so the only complication would be the condition stated in the rule.

Another solution is to order the pronominalization rule after the deletion rule. The absence of pronominalization then becomes a simple consequence of the fact that the

basis for the NP identity is gone. This seems to be more reasonable than an ad hoc condition specifying that the NP must be [+poss].

Without the [+poss] condition, the pronominalization rule cannot precede the deletion rule, for it would in some cases pronominalize an NP under identity with another that was to be deleted, producing the ungrammatical sentences 129 and 131 from 128 and 130.

128. ʔatapanĩn ʂ̬ĩki piʔiton Joaman ʔatapa biʔkin →
chicken corn eat while John chicken grabbed

129. *haan ʂ̬ĩki piʔiton Joaman biʔkin
he corn eat while John grabbed

*'John grabbed (the chicken) while it was eating corn'

130. ʔatapanĩn ʂ̬ĩki piʔiya Joanman sĩki biʔkin →
chicken corn eat while John corn grabbed

131. *ʔatapanĩn haa piʔiya Joanman biʔkin

*'while the chicken was eating the corn, John grabbed it (the corn)'

The pronominalization then, must follow the deletion rules or be restricted by a special condition.

There is the question of how far this rule should range, for as is expected, a noun separated from a noun phrase by a long series of conjoined sentences will not intelligibly pronominalize the noun phrase unless it is the sentence topic. Perhaps the solution to this is to specify a condition that Y' or Z' in the rule cannot contain

the & juncture of a conjoined sentence. This seems to give correct results, but it is in an area of indeterminacy for the present.

T 31 Md shift

SI 1 # X S S Md X

1 2 3 4

SI 2 # Md {Vb / NP(sf)} X

1 2 3 4

SC: 2 → 3

3 → 2

i.e., if the Md indicator is preceded by more than one sentence, shift it to the left until only one S precedes it. If it is preceded by the juncture, shift it to the right by permuting it with either the verb or the nearest NP with its suffix.

In the first case the rule will apply iteratively: # S S S S Md NP PP X becomes # S Md S S S NP PP X.

132. rono hišiš ta rakĭtaṣ biški?ika?inaṣ ?ĭa
snake see Md afraid run fast I

pakĭtipiki
fell

'I saw a snake, feared, ran fast and fell.'

In the second case, the rule will shift the subject NP or one of the NP's from the PP to the left. There is much flexibility in the surface order of the NP's, which is outside the scope of this study. But any stylistic NP

shift would precede the Md shift.

133. mani ta howĭti?ki
banana Md ripen 'the bananas are ripening'

134. mani ta ?ĭn hisi
banana Md I see 'I see bananas'

135. ba?kiš ta koka kanoši?iki
tomorrow Md Uncle go [fut] 'Uncle will go tomorrow'

If all the NP's have been deleted (or shifted to follow the verb), the verb is placed to the left of the Md indicator. The tense symbol remains unshifted, and a tense-carrier verb is later inserted before the tense.

136. bana ta ?ĭn haipiki
plant Md I did 'I planted'

137. rĭra ta haiki
chop Md is 'he is chopping'

The Md shift must follow the deletion rule (T 27) since it permutes only with undeleted formatives. It must also follow the topic marker rule (T 28), for the topic marker rule must precede the stylistic and emphatic-marking shift rules which must precede the Md shift.

T 32 Tense carrier

$$\underbrace{X\ Vb\ Md\ (NP)}_{1}\ \underset{2}{tense}\ \underset{3}{Y}$$

SC: 2 → ha + 2

i.e., if the verb has been separated from the tense by a Md shift, a tense-carrier pro-verb ha is adjoined to the tense. (See examples 136 and 137 above.)

T 33 Pronoun reduplication

$$S \quad \begin{bmatrix} +N \\ +pro \\ -S\_\_ \\ +spkr \\ +addr \end{bmatrix} \quad Md \quad Y$$

1 2 3 4

Cond: 3 = [+Decl] SC: 3 → 3 + 2 [-emp]

i.e., if a first or second person pronoun precedes a declarative or interrogative mood marker, a copy of the pronoun is adjoined to the mood marker and marked [-emphatic] which will determine its phonological shape.

138. mia ta? min ?ani ki
you Md you big Md 'you are big'

139. mian ta? min mani pi?i
you Md you banana eat 'you eat bananas'

140. mia min kin
you you Md 'Is it you?'

141. ?ĭa ta? ?ĭn ki
I Md I Md 'it is I'

142. ?ĭa ta ĭn ka?i
I Md I go 'I am going'

143. ?ĭan ta ?ĭn ʂĭki pi?i
I Md I corn eat 'I am eating corn'

Third person pronouns are not reduplicated:

144. haa ta? ki 'it is he'

145. haa ta ka?iki 'he is going'

Pronoun development

T 34

$$\begin{bmatrix} N \\ Pro \\ -S\_\_ \\ +pl \\ \langle ?i \rangle_1 \\ \langle mi \rangle_2 \end{bmatrix}_1 \quad 1 \rightarrow \begin{Bmatrix} \langle no \rangle_1 \\ \langle ma \rangle_2 \end{Bmatrix}$$

i.e., the singular first and second person pronouns, ?i and mi, respectively become no and ma in the plural.

The rule requires that one element be chosen from each of the braces on the left to be assigned the phonological features indicated by the corresponding set of angles.

T 35

$$\begin{bmatrix} +N \\ +Pro \\ -S\_\_ \\ \begin{Bmatrix} +addr \\ +spkr \end{Bmatrix} \\ -pl \\ +poss \end{bmatrix}_1 \quad 1 \rightarrow \underset{[+emp]}{1}$$

i.e., the first and second person pronouns are assigned the [-emphatic] feature when they are used as possessives.

The possessive forms of first and second singular persons are not unique, but they must always have the same phonological features that characterize the unemphatic forms.

T 36

$$\begin{bmatrix} +N \\ +Pro \\ -S\_\_ \\ +poss \\ \begin{Bmatrix} \langle -spkr, -addr \rangle \\ -pl \\ +pl \end{Bmatrix} \end{bmatrix}_1 \quad 1 \rightarrow \underset{[+emp]}{1} + \langle w \rangle$$

i.e., all plural possessives are marked [+emp], and the possessive third person singular is marked [+emp] and a w is added it it.

The form of the third person is now haw. Rule T 34 must precede T 35 and T 36, but the order of T 35 to T 36 is arbitrary.

T 37 $\begin{bmatrix} +N \\ +Pro \\ -S\_\_ \\ -emp \end{bmatrix}_1 \quad 1 \rightarrow 1 + n$

i.e., an n is added to all [-emphatic] pronouns.

The transformational rules have marked some pronouns as [+emp]. By convention all those not marked "+" are "-" for that feature, giving the series:

sg.	1st	ʔĩn	pl.	1st	non
	2nd	min		2nd	man
	3rd	han		3rd	han

T 38 $X\left[\begin{matrix} +N \\ +Pro \\ +emp \\ \left\{\begin{matrix} \langle +sg \rangle_1 \\ \langle +pl \\ +spkr \rangle_2 \\ \left\langle \left\{\begin{bmatrix} +pl \\ +addr \end{bmatrix} \begin{bmatrix} -spkr \\ -addr \\ +pl \end{bmatrix}\right\} \right\rangle_3 \end{matrix}\right\} \end{matrix}\right] \langle \# \rangle_1 \; Y$

$2 \rightarrow 2 + \left\{\begin{matrix} \langle a \rangle_1 \\ \langle kĩ \rangle_2 \\ \langle to \rangle_3 \end{matrix}\right\}$

1 2 3

i.e., the emphatic forms of the pronouns are assigned an ending. If no affix is adjoined to the singular forms, they

receive an a ending: ?ĩa 'I,' mia 'you,' haa 'he,' hawa 'his, hers.' The plural forms receive a terminal whether or not an affix follows: nokĩ 'we,' mato 'you,' hato 'they,' matoki 'of you.'

T 39 $\left\{\begin{bmatrix} +N \\ \left(\left[\begin{matrix}-Pro\end{matrix}\right]\right) \\ \left\{\begin{matrix}[-emp \\ +Pro\end{matrix}\right\} \\ \left(\left[\begin{matrix}-S\_\_\end{matrix}\right]\right) \\ -spkr \\ -addr \\ +pl \end{bmatrix} \\ [+gen] \end{matrix}\right\}$ 1 → 1 + bo

i.e., adjoin a bo suffix to a noun or a [-emp] third person pronoun if it is [+plural] or to any form marked [+generic].

bakĩbo 'children' (pl) hiwibo 'tree' (gen)

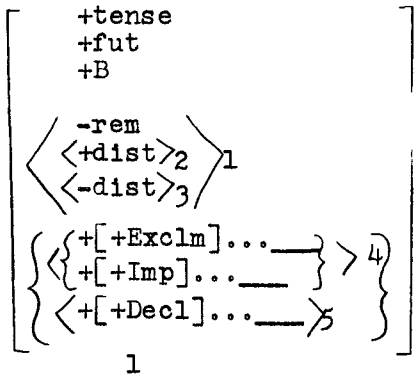
nononbo 'duck-like habo 'they'

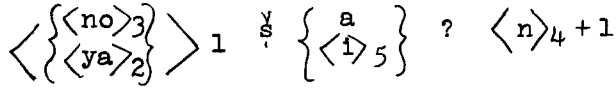
miabo 'your kind' (sg. gen)

Tense, plural, person, and mood developing rules

T 40 $\begin{bmatrix} +tense \\ +past \\ +rem \\ -dist \end{bmatrix}$ 1 SC: 1 → ipi + 1 ka?ipiki 'went'

T 41



SC: 1 → 

kanošaʔnwĭ 'go soon'

i.e., a [+B] future tense is spelled out basically as šaʔ which becomes šiʔ if the sentence is declarative and šaʔn if it is imperative or exclamatory. If the tense is [-remote, +distant] the segments no precede šaʔ, and if it is [-remote, -distant] ya precede šaʔ.

There are some instances when a fourth mood not included in this description because of its very limited use (to put a curse or hex on someone) utilizes only the basic form of the [+B] tense, šaʔ.

T 42 [+tense, +fut, +rem, +dist] 1

SC: 1 → tiʔ+1

katiʔi 'will go'

i.e., the most remote future tense is spelled out is tiʔ

T 43 $\begin{bmatrix} +\text{tense} \\ +\begin{bmatrix} -\text{spkr} \\ -\text{addr} \end{bmatrix} \ldots\_\_\_ \\ +[+\text{Decl}] \ldots\_\_\_ \\ +\text{pl} \end{bmatrix}$

1

SC: 1 → kan+1

banakani 'they plant'

bana?ipikanaʂki 'they planted'

i.e., in a declarative sentence the plural morpheme assigned (by a rule not given) to the verb by the deletion of a third person subject NP is spelled out as kan. The ordering of T 43 with respect to other tense-defining rules determines that the plural kan follows the tenses of T 40, T 41, and T 42 but precedes other tenses.

T 44 $\begin{bmatrix} +\text{tense} \\ +\text{pl} \\ +[+\text{Imp}]\ldots\_\_ \\ \begin{Bmatrix} \langle +[+\text{spkr}]\ldots\_\_ \rangle_1 \\ \langle +[+\text{addr}]\ldots\_\_ \rangle_2 \end{Bmatrix} \end{bmatrix}$

1

1 → $\begin{Bmatrix} \langle \text{na} \rangle_2 \\ \langle \text{non} \rangle_1 \end{Bmatrix} + 1$

banawĭ 'plant it'

bananonwĭ 'let's plant it'

i.e., the plural in imperative sentences is spelled out as na for second person and non for first person.

T 45 $\begin{bmatrix} +\text{tense} \\ +\text{past} \\ \left\{ \begin{matrix} \langle -\text{rem} \\ -\text{dist} \rangle_1 \\ +\text{rem} \end{matrix} \right. \left\langle \begin{Bmatrix} \langle -\text{dist} \rangle_3 \\ \langle +\text{dist} \rangle_4 \end{Bmatrix} \right\rangle_2 \left. \vphantom{\begin{matrix} a \\ b \\ c \end{matrix}} \right\} \end{bmatrix}$

1

1 → $\left\{ \left\langle \begin{Bmatrix} \langle \text{a} \rangle_1 \\ \left\langle \begin{Bmatrix} \langle ?\text{o} \rangle_3 \\ \langle \text{ni} \rangle_4 \end{Bmatrix} \right\rangle_2 \end{Bmatrix} \right\} + 1$

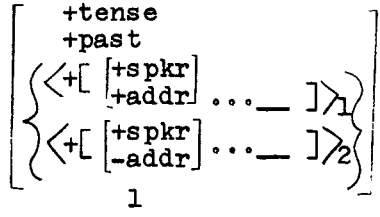
ka?a 'went'

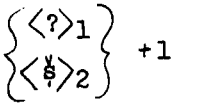
ka?o 'went'

kani 'went'

i.e., the past tenses are a [-rem, -dist], ?o [+rem, -dist], and ni [+rem, +dist].

T 46



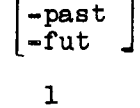


ka?a?ki 'I went'

ka?aski 'he went'

i.e., with past tenses, ? marks first and second person and ɣ marks third person.

T 47

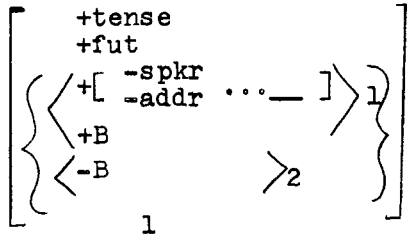


1 → ai? +1

mapĭtai 'goes up'

i.e., the present tense is marked by the morpheme ai?.

T 48



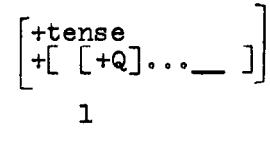
1 → i +1

kati?i 'will go'

i.e., a [-B] future tense and a [+B] future tense third person add i to the string. The pluralizer kan will now intervene between the tense and i:

katikani 'they will go'

T 49



1 → 1 + n

ka?in 'go?'

i.e., the interrogative is marked by a final n.

T 50

$$Z\underbrace{\left[_{S}\ Y\left[_{S}\ X\left[\begin{array}{l}+\text{tense}\\ +[[+\text{Decl}]\ldots\_\_]\\ -[[+\text{Q}]\ldots\_\_]\end{array}\right]\right]_{S}\ X'\right]_{S}}_{}\ Y'$$

SI: 1 2 3

SC: 1 0 3

i.e., an embedded declarative non-interrogative sentence has the tense symbol deleted. Any other mood in an embedded S will be blocked.

T 51

$$ai?\left[\begin{array}{l}+\text{tense}\\ +[[+\text{Decl}]\ldots\_\_]\\ -\text{past}\\ -\text{fut}\\ \left\{\begin{array}{l}+[[+\text{spkr}]\ldots\_\_]\\ +[[+\text{addr}]\ldots\_\_]\end{array}\right\}\end{array}\right]\ \#$$

SI: 1 2 3

SC: 1 0 3

[+Decl]

mapĭtai 'I go up'

i.e., the tense symbol is deleted in declarative S's with first or second person present tense, and a declarative feature is assigned the present tense morpheme.

T 52

$$\left[\begin{array}{l}+\text{tense}\\ -\text{past}\\ +[\text{VP}\ldots\_\_]\\ \left\{\begin{array}{l}(<+[[+\text{Imp}]\ldots\_\_]>_1)\\ (<+[[+\text{Excl}]\ldots\_\_]>_2)\end{array}\right\}\end{array}\right]\ \#$$

1 2

$$1 \rightarrow \left\{\begin{array}{l}<\text{wĭ}>_1\\ <\text{mis}>_2\end{array}\right\}$$

kawĭ 'go'

kamis 'it might go!'

i.e., the imperative marker is wĭ and the exclamation marker is mis. If the # boundary is not present, the undeveloped tense symbol will block the sentence. A VP must be in the

sentence for either of these morphemes to be developed.

T 53 $\begin{bmatrix} +\text{tense} \\ +[[+\text{Decl}]\ldots\_\_\ ] \\ \quad 1 \end{bmatrix}$ 1 → ki

ka?iki 'goes'

i.e., elsewhere, the tense symbol becomes ki in declarative sentences.

T 54 Word juncture

X Y Z

1 2 3

Cond: 2 = {N, A, V} SC: 2 → # 2

i.e., insert a word juncture before each lexical category: noun, adjective, and verb stem.

All the affixes and clitics follow the formative to which they are attached. It is therefore possible to insert the juncture correctly in every case by placing it only before each formative. This rule cannot precede the morpheme-forming rules. But since word junctures play no part in the transformational rules, there is no need to introduce them until the transformational rules have been exhausted. Word junctures cannot be inserted following the formative, for to do so would separate affixes and clitics from their head nouns and verb stems.

The rule will place a juncture at the beginning of a sentence but none at the end. At the end one is not needed, since the sentence juncture there will count as one. At the beginning of the sentence it may be desirable to

assume that of two adjacent junctures, the lower is deleted or merged with the higher. In Capanahua it happens that the rule which inserts a glottal catch before r and b in utterance initial position is the only one which would be affected by the sentence-initial word juncture and could be handled by being regarded as a sentence-level rule. The word juncture would then be ignored by the rule. However, it would seem to be more general to assume that the assimilation or deletion obtains in all cases of juncture concatenation so that it is unnecessary in any language to specify # ≠ or ≠ #, or even + # + in a rule, if only the highest is relevant and the level of the rule is less than the lower of the two junctures. (By convention, the sentence juncture # is now designated as ≠.)

144. ≠ ʔawap ta #ʔaniki ≠
tapir Md big Md 'the tapir is big'

145. ≠ baʔkiš ta #pakĭkaʔinipikanašĭki≠
yesterday Md fall suddenly past pl 3rd Md
'Yesterday they fell quickly.'

IV. THE PHONOLOGICAL FEATURES

The theory of generative phonology claims that the phonological component of a grammar will encompass a lexicon, morpheme structure rules, blank-filling redundancy rules, and morphophonemic rules. These sub-components are interrelated and the most highly valued grammar will be the one which makes the strongest claims for that language in terms of all four sub-components, that is, correctly predicts the output consistent with the data available by exploiting all the regularities of the sub-components of the phonology in the most economical fashion.

The simplicity metric is the means of assigning an objective, empirical value to a formulation. The words economy and economical will be used to refer to the most valued formulation attained by applying the simplicity metric. General and generality will be used in a generic sense to refer to regularities expressed by a rule or rules, and generalization will refer to specific instances of regularities expressed by a rule. Thus, the fact that Capanahua w, y, and ? function alike in the nasalization rules is a generalization captured by the economy achieved in grouping them into a natural class. So economy and generality go

hand in hand. The most general statement in the grammar is the simplest, and the simplest is the most general.

The use of a simplicity metric to evaluate competing formulations of the grammar of a language requires that the generalizations to be claimed by the description be established first. The different treatments expressing these generalizations are then weighed according to the simplicity metric to determine which is the most economical statement, and the simplicity metric itself, if properly formulated, should lead one to discover other generalizations that are or should be covered in the description.

However, the assumption that economy measured by counting feature specifications can be used as a simplicity metric carries with it some problems. The fact that in our theory the phonological component comprises several subcomponents permits economy to be achieved in one suboomponent at the expense of another, but our notational system does not enable the economy of one subcomponent to be weighed against that of the other.[2] Ideally a change in the feature matrix will effect similar results in both lexicon and MP rules, but if an improvement in one produces adverse results in the other, some way of giving priority to one is needed.

Halle is reported to have proposed that "economy in the phonological rules is to take precedence over economy in the morpheme structure rules and dictionary." (McCawley, 1965. 70) As Harms has pointed out, however,

there can be cases where the position of Halle may be undesirable (Southern Paiute)(Harms, 1966), but the opposite extreme of giving precedence to the lexicon's economy is also undesirable:

> The answer appears to be in determining some factor enabling a meaningful comparison of economy in the various phonological components of the grammar. A reliable answer--and an effective evaluation procedure--can be found only after our present meager knowledge of individual languages in these respects has been substantially bolstered.[3]

Requiring that generalizations be captured by the MP rules does not usually present much of a problem. If two different sets of rules will both correctly predict the facts (i.e., both capture the generalizations), and one set is simpler, it should be evident that it is simpler because either the rule formulations are more efficient, hence, more general, or the natural classes in the feature inventory of the one lend themselves to more general statements.

In the MS component, however, generalizations are harder to identify, so that the more economical formulation may simply be economical by virtue of having done a poorer job of capturing the generalizations. A statement parallel to Halle's above to the effect that economy of MP rules takes precedence over economy of the lexicon would presuppose that MS generalizations are identifiable and included within the scope of competing descriptions, but such is not always the case. In this respect the theory needs refining.[4]

The distinctive features

There are fifteen consonants and four vowels in the inventory of phonemes necessary to specify the lexical entries in Capanahua. Both the consonants and the vowels form asymmetrical patterns when charted articulatorily:

Articulatory Chart of the Consonant and Vowel Segments

	Bilabial	Alveolar	Alveopalatal	Palatal	Velar	Pharyngeal	Laryngeal
Stops	p	t			k		?
Continuants	b						
		s	š ṣ			h	
Affricates		ȼ	č				
Flap		r					
Nasal	m	n					
Semiconsonant	w			y			

	Front	Central	Back
High	i		ɨ
Mid			o
Low		a	

o ranges from mid close to high open, unconditioned.

Consonants

The consonants are the unaspirated stops p, t, k, ?; the bilabial voiced fricative b; the affricated stops ȼ and č; the grooved fricatives s (alveolar), š (alveopalatal), and ṣ (retroflexed alveopalatal); the flap r; the nasals m

and n; and the glides w and y. Since h occurs only in word initial position before vowels and vowels never occur initially unpreceded by h or another consonant, h can be regarded as a predictable redundancy. The vowels are i, a, close o, and high back unrounded ɨ.

For the consonants of Capanahua, the simplest solution for the lexicon requires the specification of 68 features in the inventory for non-vowel phonemes. This is achieved by marking b, r, m, n as liquids and w, y, ? as glides. The next best solution, marking only b and r as liquids, requires 71 features.

Solution I. Parsimonious arrangement of non-vowel phonemes: 68 features

	p	t	k	ø	c	s	š	ʂ	b	r	m	n	y	w	?
cns	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
voc	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-
std	-	-	-	+	+	+	+	+							
nas									-	-	+	+			
cnt				-	-	+	+	+					+	+	-
cmp	-	-	+	-	+	-	+	+							
grv	+	-				-	+	+	-	+	-	-	+		

Solution II. 71 features

	p	t	k	ø	c	m	n	s	ś	ʂ	b	r	y	w	?
cns	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
voc	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-
cnt	-	-	-	-	-	+	+	+	+	+			+	+	-
std	-	-	-	+	+	-	-	+	+	+					
cmp	-	-	+	-	+			-	+	+					
grv	+	-				+	-				+	-	-	+	
flt									-	+					

b can be regarded as a resonant because it patterns with r differently from the obstruents and glides. Obstruents form consonant clusters with nasals and continuant stridents, but b and r do not. The obstruents and nasals may occur morpheme final, but b and r do not. Only b and r occur with a glottal onset in initial position, and the nasalization of vowels is weakened before b and r.

It is possible to treat w and y as vowels in the underlying forms and allow a rule to change their vocalic feature to non-vocalic early in the P rules. Since w and y occur only preceding vowels, and there would be no sequences of four vowels such that the two central ones would be o and i, the rule would change o and i to the respective glide in the case that a vowel or juncture preceded and another vowel followed: $\begin{Bmatrix} o \\ i \end{Bmatrix}_1 \rightarrow \begin{Bmatrix} w \\ y \end{Bmatrix}_1 \; / \; \begin{Bmatrix} V \\ \# \end{Bmatrix} \; \_\_V$

Sequences such as boo 'hair,' ʂoo 'green,' ʂoó 'wind,' koo 'sap,' ?oa 'that,' ʂoa 'fat,' poa 'potato,' koa 'spear,'

?oi 'rain,' koi 'chin' would be unaffected. Forms beginning with a vowel cluster would undergo the rule:

oina	wina	'to paddle'	oĭot	wĭot	'to hang on'
iaoiš	yawis	'armadillo'	oa?o	wa?o	'mud'
?aoin	?awin	'wife'	oai	wai	'garden'
baoin	bawin	'catfish'	oaon	waon	'crybaby'
?ioi	?iwi	'stingray'	iĭnĭ	yĭnĭ	'grasshopper'
?oio	?oyo	'to suck'	iaoa	yawa	'boar'
baoa	bawa	'parrot'	iama	yama	'none'
?ioĭ	?iwĭ	'heavy'	iaia	yaya	'mother-in-law'
			iobĭ	yobĭ	'shaman'

Apart from other considerations this treatment would permit some savings in the lexicon since ? could then be specified by just the two features [-cns, -voc], a saving of one feature for every occurrence of ?.

However, h would then have to be specified in the lexicon because vowel-initial words that should otherwise have an h redundantly specified would turn up with a w instead of h:

oina	hoina	'to breathe' (minimal pair with 'to paddle' above)
oyo	hoyo	'mute'
oa	hoa	'flower'
oi	hoi	'language'
iwi	hiwi	'tree'
iin	hiin	'ant'
owĭ	howĭ	'light'

oit hoit 'to hang suspended'

oo hoo 'yes'

There will therefore be no savings effected in the lexicon because ? will again need three features, the vowels will require four each (the same as w and y), and in addition, h will have to be specified with three features for each occurrence. The solution eliminating w and y from the systematic phoneme inventory therefore achieves no economy of feature specification. In addition, the treatment of h as a redundancy is more economical because though h occurs only initially, it occurs before all vowels and has a higher frequency in the lexicon than initial y and w combined.

It is clear, of course, that only one of the glides can be predicted in. It would be circular to try to predict first h on the basis of w and y, then y and w on the basis of h. Otherwise the first to be predicted in would cover all vowel-initial forms, leaving no case where the second could be inserted.

Finally, it should be observed that instead of h, ? can be considered to be predictable in initial position. Its frequency of occurrence initially is about the same as h. But while the savings effected in cutting out the features needed to specify ? initially would then be the same as for h, there would be no concomitant reduction of features necessary to specify h, y, w elsewhere, since ? would not be completely removed from the inventory. There would still need to be an extra feature each for y and w as shown below,

in addition to the four needed for h.

	h	w	y	ʔ
cns	-	-	-	-
voc	-	-	-	-
cnt	+	+	+	-
cmp	+	-	-	
grv		+	-	

Apart from economy, the very limited role of h in the language with its limited initial occurrence and frequent deletion in the surface string makes it seem that to consider h predictable is a more valued generalization than predicting w, y, or ʔ. The simplicity metric supports this intuitive judgement by showing that the treatment of h as redundant is the most economical.

A sampling of the lexicon indicates that for every 100 entries, 52.5 per cent of the phonemes are consonants, glides, or liquids and 47.5 per cent are vowels. The phonemes have the following frequencies of occurrence.

consonant	per cent of phoneme occurrences in the lexicon
p	3.57
t	5.58
k	6.02
ɸ	.86
č	2.29
s	2.47
š	1.43
ʂ	2.14

consonant	per cent (continued)
m	4.00
n	7.30
b	3.86
r	2.00
w	2.29
y	2.00
?	4.55
(h	2.14)
	52.50

vowels	per cent
a	17.90
i	10.72
ɨ	8.87
o	10.01
	47.50

Apart from the economy effected by morpheme structure rules, therefore, the distinctive features chosen will achieve maximal simplicity in the lexicon if the stops and nasals are marked by the lowest possible number of features at the expense of the number marking the strident group, for the stridents have a much lower frequency of occurrence. This is what Solution I does. Apart from economy achieved by exploiting redundancies, for every 100 phoneme occurrences in the lexicon Solution I will specify an average of 214.8 consonantal features.

Solution II specifies the lexical entries with an average of 226.4 consonantal features for every 100 phonemes.

The irregularity of the features needed to distinguish the sibilants leads one to suspect that one of them, perhaps š or ṣ, comes from an underlying cluster. There are no sequences sr, sr, sn, sr in the lexicon, so that either š or ṣ might be derived from sn or sr. But treating

š or ṣ as a cluster would create a unique cluster in initial position with a consequent loss of the redundancy exploitable by the MS rules, since the prediction could no longer be made that if the first segment is a consonant, the second is a vowel. The cluster treatment would also produce the unique instances of triple clusters in medial position. It would require one less feature to specify the occurrences of š or ṣ, but the economy would be offset by the need to specify the other consonant in the cluster. Assuming that ṣ were considered to be a cluster of sr, every specification of s, š, and ṣ would require the following number of features:

	s	š		ṣ		total
unitary	12.35	8.58		12.84		33.77
				s	+ r	
cluster	12.35	7.15	Sol.I	12.35	8.56	40.41
	"	"	Sol.II	12.35	6.42	38.27

The cluster treatment would allow the MS rules to effect a savings of two features [+cnt, +std] predictable for the first member of every sr cluster, but this would be more than offset by the loss of the redundancy predicting that the second member of a cluster is always [-voc]. The cluster treatment is therefore the less economical.

Solution I regards ṣ as having an underlying compact form [x] and makes a neater chart. It is suggested by historical evidence but offers no advantages in the grammar by putting x in a natural class with some other segments. Solution II uses the flatness feature to distinguish š from ṣ.

In corresponding more closely to the phonetic reality, Solution II obviates the need for a late phonetic rule converting x to ʂ.

Vowels

Two vowel solutions, A and B, were tested and rejected because they turned out uneconomical and unable to compete with solutions C and D in capturing the generalizations in the morpheme structure rules. A, suggested by comparative evidence, interpreted ɨ as having an underlying form e. B distinguished i and ɨ from a and o by a diffuseness feature:

A	i	e	a	o
grv	-	-	+	+
cmp	-	+	+	-

B	i	ɨ	a	o
dif	+	+	-	-
grv	-	+	-	+

Solutions C and D competed closely both in capturing significant generalizations and in economy.

Solution C:	a	i	ɨ	o
syl	+	+	+	+
cmp	+	-	-	-
grv		-	+	+
flt			-	+

13 features

Solution D:	i	o	ɨ	a
syl	+	+	+	+
per	+	+	-	-
dif			+	-
flt	-	+		

12 features

In Solution C ɨ and o are distinguished by graveness from i and from one another by the flatness feature. Solution D utilizes peripheralness to separate i and ɨ from o and a. The ɨ is differentiated from i by the

diffuseness feature and o from a by the flatness feature. Solutions C and D correspond closely to the phonetic features so that late feature-changing rules are unnecessary.

Comparing the vowel solutions as to economy, D requires 190.32 features per 100 phoneme occurrences in the lexicon, against 190.98 for Solution C. Apart from the MS rules, therefore, Solution I (with D for the vowels) would give the greatest economy in the lexicon. Nevertheless Solution IIC will be the one used in the following rules because overall it is the best solution.

V. MORPHEME STRUCTURE RULES

If all the lexical entries of a grammar were clearly monomorphemic the formulation of the morpheme structure rules would be a matter of formulating rules that specify the phonological redundancies directly for the morphemes, following which the P component begins its task of phonological interpretation.

In fact, however, not all entries are clearly monomorphemic, so that what is going to be permitted as a lexical entry must be determined before the redundancy rules can be formulated. There are some forms which seem best regarded as single entries to explain cooccurrence restrictions, but which are static results of rules that have become lost. For example, nĭi?no 'weasel' comes from a rule that has united hĭnĭ 'water' and ?ino 'tiger.' The first syllable of hĭnĭ has been lost, and the glottal and i of ?ino have been metathesized, producing a ?n cluster. In this case, the language abounds in ?n clusters already, but this is the only form found in the language with an ĭi vowel cluster.

Similarly, the check of a 1500 word sample from the dictionary shows only one word with an sb sequence, risbi

'rope.' The derivation is similar to that of nĭi?no except that the full form of the first word of the construction is not recoverable: ris- (meaning unknown), biči 'bark, hide.' The final syllable of bici has been lost by known historical (but now only partially operative) rules. The result of accepting risbi as a single lexical entry is to introduce a unique case of a sibilant followed by a liquid.

The effect of such constructions, if they are regarded as base forms, is to fill in gaps in the permitted CV sequences and thereby either increase or reduce the economy achievable by the redundancy rules. The ĭi sequence reduces the economy, for with it the grammar loses the rules that specify that any [-cmp, +grv] vowel (o) is [+flt] when immediately followed by a [-grv, -cmp] vowel (i). The sb sequence would increase the attainable economy (if there were enough occurrences of it) because b would become the only liquid that may follow a sibilant so that the features of graveness that distinguish it from r become redundant there.

The paucity of words having phoneme sequences that increase economy in this way, in contrast to the more numerous cases that decrease economy, means that economy in the lexicon attainable by lexical redundancy rules is seriously impaired by the admission of complex entries. Since the redundancy rules not only effect economy but also

make a claim about the native speaker's intuition concerning which sequences of sounds are permitted in the lexical entries of Capanahua, the oddity of the rare sequences constitutes an additional reason for wanting to regard forms such as nĭi?no as two entries, not one.

That the lexicon should contain other, more productive derivational rules is shown by forms such as tĭṣ̌pa 'creek,' composed of tĭ 'throat' and pa 'to open' with an š inserted between the morphemes. Neither tĭ nor pa can be inserted by the lexical rule into a complex symbol (CS) because they must be combined first with some other form. Yet the combination must have the structural index of a single entry (Noun or Verb) for the P rules to operate correctly. If rules generating these forms are permitted in the lexicon, and the phonological redundancy rules are formulated only for the simple forms, i.e., excluding all those produced by such derivational processes, there will be no sequences sp, šp, ṣ̌p or ṣ̌b within morpheme boundaries, and the redundancy rules will be able to exploit the lacunae in the pattern to achieve greater economy. That is, since there is no contrast of p-t after a sibilant, t is redundantly [-grv] there.

Both the extremely limited rules that produce forms like nĭi?no and the more productive ones producing forms like tĭṣ̌pa have been considered to be operative in the lexicon for the computation of phoneme frequencies previously given, and for evaluating the solutions for the consonant

and vowel feature matrices, and are assumed also for the following MS formulations. Thus nĩi?no is considered to be two separate morphemes in the formulation of the following morpheme structure rules.

Constraints on morpheme structure (MS) rules

The rules of the MS component are ordered rules. Ordering will be necessary in the case that two equally desirable generalizations are such that one requires information supplied by the other. Thus, if the quality of vowels is predictable before certain consonants, and the consonantal feature of segments occurring final is also predictable, the consonantal feature must be specified first.

McCawley (1966.2) argues for a feature of syllabicity to replace vocalicity because it enables a description to make generalizations that the Jakobsonian set of features does not. He shows that syllabicity gives the advantage of covering in one rule the changes i → y and r̥ → r (as in Sanskrit) by a loss of syllabicity in both cases. In the Jakobsonian system this would be a change of consonantality in one case and of vocalicity in the other. He also shows that he can claim a generalization in the SC rules of a language where there are no vowel clusters, which the feature of vocalicity does not permit, namely that the syllabicity of a segment is predictable on the basis of the preceding segment: [ ] → [-syl] / [+syl]__ . The Jakobsonian feature system permits the prediction of either the vocalicity of

glides or the consonantality of liquids in such a case, but no feature of obstruents would be predictable: either the rule $[-\text{cns}] \rightarrow [-\text{voc}] \,/\, \begin{bmatrix}-\text{cns}\\+\text{voc}\end{bmatrix}\_\_$ or $[+\text{voc}] \rightarrow [+\text{cns}] \,/\, \begin{bmatrix}-\text{cns}\\+\text{voc}\end{bmatrix}\_\_$.

However, McCawley's formulation of the rule to predict syllabicity on the basis of the preceding consonant involves a problem. It must either require the condition that the rule begins at the left-most extreme of the phonological string and tests by moving to the right until it finds a set of segments that meets the structural description, or it violates the distinctiveness criterion. It is interesting to note that a finite automaton can be constructed that reads either from the left or the right, and SC rules are of the type that are formulated for such a device. They can be assumed to begin operating on the left-most extreme of any morpheme. But if the rules are applied randomly over a stretch of segments for which the syllabicity feature has been redundantly unspecified, they will sometimes require the + or - specification in a segment for which the feature is unspecified.

The question then is, should such left-tc-right operation be accepted as part of the theory of SC rules? It is evident that if it is, the generality achieved will be greater than if it is not. On the other hand, there are some cases where a feature of the second of two segments must be specified before another one in the first segment is, so the rules cannot be viewed as applying exhaustively to

each segment until it is fully specified, then moving on to the next. The device (i.e., the application of the rules) must therefore be viewed as moving once across each stretch of segments (each morpheme) for each rule, then beginning again on the left of the same stretch until the rules are exhausted. This is what I propose to be the interpretation of SC rule application. If it is accepted, it means that in a language with a CVCV pattern such that all morphemes begin with a consonant, all of the syllabic features of a given string of segments are predictable, corresponding to the fact that one need only know the position of a segment in a morpheme to know if it is + or - syllabic.

The rule for the redundant specification of syllabicity in Capanahua could therefore be stated as follows:

$$[\ ] \rightarrow [+syl] \ / \ [-syl]\ [-syl]\ \_\_\_$$

There are more consonant clusters than vowel clusters. Therefore the rule exploits the more frequent set to specify a segment as + syllabic when preceded by two non-syllabics. This convention enables only one rule to specify a feature of two or three consecutive vowels in a language like the Kolokuma dialect of Ijo where, if geminate vowels are treated as phonemically long vowels, each vowel must be the opposite in diffuseness from the preceding vowel (data from Williamson, 1965.14): $[+syl] \rightarrow [-\alpha dif] \ / \ \begin{bmatrix} +syl \\ \alpha dif \end{bmatrix} \_\_\_$

The same rule which predicts the diffuseness of the second vowel applies to the third vowel, using the diffuseness

specified for the second vowel.

e.g. suei 'thirty'

buou 'board (a boat)'

In Halle's system, obstruence was not regarded as a feature. Halle (1959, 1961a, 1962a, 1964) and Jakobson, Fant and Halle (1961) class the nasal consonants as stops because the articulators produce contact and impede the passage of air between the glottis and the lips. However, McCawley (1966) treats the Finnish nasals as resonants.

Each segment presents an array of features, some distinctive and some redundant, some of which can be used in one way in one language and in other ways in another language. The difference, in a generative approach, is shown by the way each language treats the features in phonological rules. Since the nasals make use of both the oral and nasal cavities, it does not seem surprising that they should be open to a wider interpretation than some other segments. By regarding the change of features produced by the oral closure as significant they can be treated as stops, as in Halle (1959), or by regarding the added features of nasal murmur produced by the resonating nasal cavity with unimpeded air passage as syllabics for German (morpho-phonemic rules) or as resonant non-syllabics for Finnish (McCawley, 1966). In Capanahua they are treated as continuant obstruents, using the features characteristic of the oral closure to define obstruence and the nasal feature of the supplementary nasal resonator to permit a continuant feature to be distinctive.

The MP and SC rules of Capanahua require that the nasals be classed with the continuant obstruents in the underlying forms, and the SC rules also require that in clusters in which both segments are consonantal the first must be a continuant. The MP rules delete all non-continuant consonants and all resonants in the environment $\_\_\left\{\begin{matrix}\#\\C\end{matrix}\right\}$. The features in nasals that are required by the grammar are thus found to correspond to actual acoustic features, so that even though the Capanahua treatment of the nasal as continuant obstruents may be an innovation in a generative framework, it has always been a potential arrangement of the nasal features.

The previous solutions can therefore be modified as follows:

	p	t	k	ɸ	č	m	n	s	š	ş	b	r	y	w	ʔ	a	ɨ	i	o
syl	(-	-	-	-	-	-	-	-	-	-	-	-	-)	-	-	-	+	+	+
cns	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-
obs	+	+	+	+	+	+	+	+	+	+	-	-	(-	-	-	-	-	-	-)

The defect in this arrangement is that there is no way in which true consonants can be made a natural class with the glides without including the liquids. The justification for accepting this consequence in the rules below is that the advantages gained offset the disadvantages, in terms of economy, in the MP rules.

Solution II-C, as modified above, is used in the following morpheme structure rules.

MS 1 $[\ ] \rightarrow \begin{bmatrix}-\text{cns}\\+\text{syl}\end{bmatrix} / \left\{ \begin{matrix}\#\\ [\ \{\begin{matrix}+\text{cns}\\-\text{syl}\end{matrix}\}\ ]\end{matrix} \right\} \quad [\ \{\begin{matrix}+\text{cns}\\-\text{syl}\end{matrix}\}\ ]\ \_\_$

i.e., after an initial consonant, and after a consonant cluster, the next segment must be a vowel. (There are no morpheme-initial consonant clusters.)

There are no triple consonant clusters. This rule takes advantage of the fact that there are more consonant clusters than vowel clusters to predict the vowel features on the basis of the consonants. In the examples given, features specified by the rules are placed in parentheses.

	'hair'			'moon'					'butterfly'									
	b	o	o	?	o	š	n	ĭ	p	k	o	?	k	o	š	t	o	b	i
cns	+	(-)	-	-	(-)	+	+	(-)	+	+	(-)	-	+	(-)	+	+	(-)	+	-
syl	(-)	(+)	+	-	(+)	(-)	-	(+)	(-)	(-)	(+)	-	(-)	(+)	(-)	(-)	(+)	(-)	+

MS 2 [+cns] → [-syl]

i.e., all consonantal segments are redundantly specified as [-syllabic].

MS 1 must precede MS 2 because the second segment, if a vowel, has been unspecified for the consonantal feature until MS 1 has operated. Examples: see MS 1.

MS 3 [+cns] → [+obs] / [+cns] ___

i.e., in any consonant cluster, the second member is always a true consonant if the first also is. There are no cases of a liquid followed by an obstruent or of a consonant followed by a liquid or a glide. MS 3 must follow MS 1 for the same reason as MS 2.

MS 4 $[+cns] \rightarrow \begin{bmatrix} +obs \\ +cnt \\ \langle -\alpha std \rangle \end{bmatrix} / \_\_ \begin{bmatrix} +cns \\ +obs \\ \langle \alpha std \rangle \\ \langle \beta cnt \rangle \end{bmatrix}$

Cond: if α is "-", β is "+".[5]

i.e., if the first member of a CC is consonantal it is also [+obstruent] and [+continuant]; only the nasals and sibilants precede another true consonant. If the second member is strident ([+cnt] or [-cnt]), the first will be a non-strident (i.e., a nasal). If the second is a [-strident] continuant (a nasal), the first will be a strident.

Consonant clusters are thus composed of a consonant preceded by a sibilant or nasal. If the second member is a non-strident occlusive (p, t, k), the first member is unpredictably a nasal or sibilant, but if the second is a strident occlusive or a continuant, the stridency of the first is predictable.

If the first C is a nasal, it will be identical in compactness and graveness to a following stop, but it is unnecessary to formulate a rule for this redundancy here because an MP rule will assimilate all nasals to the following consonant. No distinction between m and n is made in the base forms before another consonant. Examples of the economy achieved by MS 4:

	'to split'				'moon'					'lump'					'bowl'				
	k	a	š	k	ĭ	?	o	š	n	ĭ	p	t	o	n	k	o	k	ĭ	n	ǒ	a	p
cns	+		+	+				+	+		+	+		+	+		+		+	+		+
syl	(-)		(-)	(-)				(-)	(-)		(-)	(-)		(-)	(-)		(-)		(-)	(-)		(-)
obs	+		(+)	(+)				(+)	(+)		+	+		(+)	(+)		+		(+)	(+)		+
cnt	-		(+)	-				(+)	+		-	-		(+)	-		-		(+)	-		-
std	-		+	-				(+)	-		-	-		-	-		-		(-)	+		-
cmp	+		+	+				+			-	-			+		+			+		-
grv									-		+	-		(-)					(-)			+
flt			+					+														

MS 3 and MS 4 operate on the same cluster, so MS 3 must precede MS 4 to specify obstruency in the second segment.

$$\text{MS 5} \quad \begin{bmatrix}\text{-syl}\\ \text{-cnt}\end{bmatrix} \rightarrow \begin{bmatrix}\text{-std}\\ \text{-cmp}\\ \text{+grv}\end{bmatrix} \;/\; \underline{\quad}\;)_{N}$$

i.e., the only non-continuant segment terminating nouns is p. Verbs and adjectives, however, are not limited, for they may terminate in any true consonant except ø. In nouns, therefore, there are predictable features of a final stop.

MS 2 must precede MS 5 to supply the syllabic feature for the condition of MS 5.

	?	o	š	n	ĭ	p	'moon'
cns	-	(-)	+	+	(-)	+	
syl	-	(+)	(-)	(-)	(+)	(-)	
obs			(+)	(+)		(+)	
cnt	-		(+)	+		-	
std			(+)	-		(-)	
cmp			+			(-)	
grv				-		(+)	
flt			+				

$$\text{MS 6} \quad \begin{bmatrix}\text{-cns}\\ \text{-syl}\end{bmatrix} \rightarrow \left\{\begin{matrix}[\text{-cnt}] \;/\; \underline{\quad}\;[\text{-syl}]\\ [\text{+cnt}] \;/\; [\text{-syl}]\;\underline{\quad}\end{matrix}\right\}$$

i.e., if the first element of a cluster is a glide, it will be a glottal stop. (No cases of w or y are followed by a consonant.) But the second element may be a consonant, glide,

or liquid. If the second is a glide, it is w or y.

MS 6 must also follow MS 2 because it needs the syllabicity feature provided by MS 2.

	'grub'					'bird'					'owner'					'to mix'				
	t	a	?	n	o	?	i	?	s	a	p	?	i	?	b	o	s	i	?	w	a	n
cns	+	-	+			-	-	+			+	-	-	+			+	-	-			+
syl	(-)	-	(-)			-	-	(-)			(-)	-	-	(-)			(-)	-	(-)			(-)
obs	+		+					+			+			-			+		-			+
cnt	-	(-)				-	(-)	+				-	(-)				+	(-)	(+)			+
std	-		-					+			(+)						+					
cmp	-		-					-			(-)						-					
grv	-		-								(+)			+					+			(-)

MS 7 [+cns] → [+obs] / ___ +

i.e., no liquids are morpheme-final; true consonants, glides, and vowels may be..

	i	s	i	s	'ant'
cnt	+		+		
syl	(-)		(-)		
obs	+		(+)		
cnt	+		+		
std	+		+		
cmp	-		-		

MS 8 $\left\{\begin{bmatrix}+\text{syl}\end{bmatrix}\right.$ $\left.\begin{bmatrix}+\text{cns}\\-\text{obs}\end{bmatrix}\right\}$ → $\begin{bmatrix}+\text{cnt}\\-\text{std}\end{bmatrix}$

i.e., all vowels and resonants are redundantly [+continuant] and [-strident].

MS 8 must follow MS 7 because the obstruency feature predicted by MS 7 is required in MS 8.

MS 9 $\begin{bmatrix}-\text{cns}\\-\text{syl}\end{bmatrix}$ → [-std]

i.e., all glides are redundantly [-strident]. All segments are now specified for continuancy and stridency.

MS 10 [-cns] → [-obs]

i.e., all vowels and glides are [-obstruent].

MS 11 [ ] → [-nas]

$$\begin{bmatrix}+\text{obs}\\+\text{cnt}\\-\text{std}\end{bmatrix} \rightarrow [+\text{nas}]$$

i.e., m and n are [+nasal]; all other segments are [-nasal].

MS 11 must precede MS 12.

MS 12 [+nas] → [-grv] / ___+

i.e., a final nasal will be n.

	w	a	r	a	n
	'squash'				
cns	-		+		+
syl	-		(-)		(-)
obs			-		(+)
cnt					+
omp					
grv	+		-		(-)
nas	(-)		(-)		(+)

MS 13 Strident harmony

$$\begin{matrix}+\text{std}\\ \langle\beta\text{cnt}\rangle\end{matrix} \rightarrow \begin{bmatrix}\alpha\text{cmp}\\ \langle\gamma\text{flt}\rangle\end{bmatrix} \;/\; \begin{bmatrix}+\text{std}\\ \alpha\text{cmp}\\ \langle\beta\text{cnt}\rangle\\ \langle\gamma\text{flt}\rangle\end{bmatrix} \quad [+\text{syl}]\ \_\_$$

i.e., strident consonants separated by only one vowel agree in compactness, and if they are alike in the continuant feature they agree in flatness also.

For maximum economy of feature specification, MS 9 must precede MS 13 to permit the omission of [+cns, +obs] in the condition and environment of MS 13. There are no known morphemes with more than two stridents, but if some are eventually found, the left-to-right application will

permit MS 13 to apply to the second and third stridents without violating the distinctiveness criterion.

	'whisper'				'tray'				'ant'				'whistle'				'fish'			
	b	a	s̆	ṯ	s̆		ṣ	a	ṣ	o	i	s	i	s	s	ṯ	s	o	ɸ	a	ɸ	a
cns	+		+		+		+		+		+		+		+		+		+	
syl	(-)		(-)		(-)		(-)		(-)		(-)		(-)		(-)		(-)		(-)	
obs	-		+		+		+		+		+		+		+		+		+	
cnt			+		+		+		+		+		+		+		-		-	
std			+		+		+		+		+		+		+		+		+	
cmp			+		(+)		+		(+)		-		(-)		-		(-)		-		(-)	
grv	+																					
flt			+		(+)		+		(+)													

	'grandmother'				'breast'			
	č	i	č	i	č	o	č	o
cns	+		+		+		+	
syl	(-)		(-)		(-)		(-)	
obs	+		+		+		+	
cnt	-		-		-		-	
std	+		+		+		+	
cmp	+		(+)		+		(+)	
grv								

	'bosom'				'deer'				'ashes'					'buttocks'					
	ṣ̆	o	č	i	č	a	ṣ̆	o	ɸ	i	s		t ṯ	č	i	š́		k	i
cns	+		+		+		+		+		+	+		+		+	+		
syl	(-)		(-)		(-)		(-)		(-)		(-)	(-)		(-)		(-)	(-)		
obs	+		+		+		+		+		(+)	(+)		+		(+)	(+)		
cnt	+		-		-		+		-		(+)	-		-		(+)	-		
std	+		+		+		+		+		(+)	-		+		(+)	-		
cmp	+		(+)		+		(+)		-		(-)	-		+		(+)	+		
grv											(-)	-							
flt	+						+									-			

MS 14 $\begin{bmatrix}+\text{obs}\\-\text{cnt}\\-\text{std}\\-\text{cmp}\end{bmatrix}$ → [-grv] / [+std] ___

i.e., sibilants are not followed by p; t is redundantly [-grv] after s, š, ṣ̆.

Since there are no cases of affricates preceding p, it is not necessary to distinguish continuant from

non-continuant stridents here.

MS 3 and MS 10 must precede MS 14 to specify the features required in the condition and environment of MS 14.

	w	a	s	t	ĭ	'herb'
cns	-	(-)	+	+	(-)	
syl	-	(+)	(-)	(-)	(+)	
obs	(-	-)	(+)	(+)	(-)	
cnt	+	(+)	(+)	-	(-)	
std	(-	-)	(+)	-	(-)	
cmp		+	-	-	-	
grv	+		-	(-)	+	
flt					-	
nas	(-	-	-	-	-)	

MS 15 $\begin{bmatrix}+\text{std}\\-\text{cnt}\end{bmatrix} \rightarrow [+\text{cmp}] \;/\; \_\_\_ +$

i.e., a morpheme-final affricate is [+compact], c. The affricate ǵ does not occur final.

	b	i	ǵ	'to take'	n	i	ǵ	'to walk'
cns	+	(-)	+		+	(-)	+	
syl	(-	+	-)		(-	+	-)	
obs	(-)	(-)	(+)		+	(-)	(+)	
cnt	(+	+)	-		+	(+)	-	
std	(-	-)	+		-	(-)	+	
cmp		-	(+)			-	(+)	
grv		-			-	-		
flt								
nas	(-	-	-)		(+	-	-)	

MS 16 $\begin{bmatrix}+\text{syl}\\+\text{cmp}\end{bmatrix} \rightarrow [+\text{grv}]$

i.e., a is redundantly [+grave].

MS 17 $\begin{bmatrix}+\text{syl}\\ \alpha\text{cmp}\\ \alpha\text{grv}\end{bmatrix} \rightarrow [-\text{flt}]$

i.e., a and i are redundantly [-flat].

MS 16 and MS 17 must precede the vowel feature specification rules to permit maximum economy in the latter.

MS 18 $\begin{bmatrix}-\text{cns}\\-\text{syl}\\+\text{cnt}\end{bmatrix} \rightarrow [-\alpha\text{grv}] \;/\; \_\_\_ \begin{bmatrix}+\text{syl}\\ \alpha\text{flt}\\ \alpha\text{grv}\end{bmatrix}$

i.e., y before o and w before i are redundantly opposite in graveness from the following vowel.

There are no occurrences of yi or wo. Vowel Solution D can predict a feature of the vowel on the basis of the glide, or vice versa, with the same savings and the same cost in terms of features required to specify the rule in either case. Solution C can only predict the graveness of the glide on the basis of the vowel.

	k	f̈	y	o	'to finish'	w	i	n	a	'to paddle'
cns	+	(-)	-			-	(-)	+	-	
syl	(-	+)	-	+		-	(+)	(-)	+	
obs	+	(-	-	-)		(-	-)	+	(-)	
cnt	-	(+)	+	(+)		+	(+)	+	(+)	
std	-	(-	-	-)		(-	-)	-	(-)	
cmp	+	-	(-)	-		-		+		
grv		+	(-)	+		(+)	-	-		
flt		-		+						
nas	(-	-	-	-)		(-	-	+	-)	

MS 19 (*it)

$\begin{bmatrix}+\text{syl}\\-\text{cmp}\end{bmatrix} \rightarrow [+\text{grv}] \;/\; \_\_\_ \begin{bmatrix}+\text{obs}\\-\text{cnt}\\-\text{std}\\-\text{cmp}\\-\text{grv}\end{bmatrix}$

i.e., i does not precede t; o and f̈ are redundantly [+grv] before t.

	p	o	t	o	'dust'	n	f̈	t	f̈	'day'
cns	+	(-)	+	-		+	(-)	+	-	
syl	(-)	(+)	(-)	+		(-)	(+)	(-)	+	
obs	+	(-)	+	(-)		+	(-)	+	(-)	
std	-	(-)	-	(-)		-	(-)	-	(-)	
cnt	-	(+)	-	(+)		-	(+)	-	(+)	
cmp	-	-	-	-			+	-	+	
grv	+	(+)	-	+		-	(+)	-	+	
flt		+		+			-		-	
nas	(-	-	-	-)		(+	-	-	-)	

MS 20 $\begin{bmatrix}+cnt\\+std\\+cmp\end{bmatrix} \rightarrow [+flt] \;/\; \begin{bmatrix}+syl\\-cmp\\-flt\\+grv\end{bmatrix}$

i.e., ǩ does not precede or follow ĭ; ǧ is redundantly [+flat].

	k	ĭ	ǧ	a	'lip'
cns	+	(-)	+	-	
syl	(-)	(+)	(-)	+	
obs	+	(-)	+	(-)	
cnt	-	(+)	+	(+)	
std	-	(-)	+	(-)	
cmp	+	-	+	+	
grv		+			
flt		-	(+)		
nas	(-	-	-	-)	

MS 21 $[+syl] \rightarrow [+grv] \;/\; \#\ ([-syl])\ \_\_([-syl])([-syl])\ \begin{bmatrix}+syl\\-cmp\\+grv\\-flt\end{bmatrix}$

i.e., a ĭ, and o are redundantly [+grv] preceding ĭ in a second syllable.

There are no clusters iĭ, ĭi, or iCĭ. The possibility of exploiting the graveness feature of all three vowels a, o, ĭ in the environment before and following ĭ (without an intervening consonant) is excluded by the fact that to do so would first predict the graveness of the first ĭ on the basis of the second, then vice versa, in the sequence ĭĭ. Therefore the only possibilities are to predict a graveness feature of o before and following ĭ and when followed by ĭ with an intervening consonant or to restrict the environment more and gain the economy of predicting a feature for both ĭ and o. The latter turns out to be the more productive, for there are relatively few

occurrences of ĩo, and the only vowel clusters are in the first two syllables of a morpheme. The cost of adding #([-syl]) to the rule therefore gained the economy of predicting graveness for every ĩ in the first syllable before another ĩ in the second.

The graveness feature of a is a segment structure redundancy and therefore cannot be counted in the simplicity metric here, but to exclude a would have cost the rule several features without gaining any additional economy.

The total economy of the rule, then, is to predict the graveness feature of o and ĩ in the first syllable when followed by ĩ in the next syllable.

	ʂ	o	k	ĩ	'toucan'	t	ĩ	k	ĩ	'piece'
cns	+	(-)	+	-		+	(-)	+	-	
syl	(-	+)	(-	+)		(-	+)	-	+	
obs	+	(-)	+	(-)		+	(-)	+	(-)	
cnt	+	(+)	-	(+)		-	(+)	-	(+)	
std	+	(-)	-	(-)		-	(-)	-	(-)	
cmp	+	-	+	-		-	-	+	-	
grv		(+)	+	+		-	(+)		+	
flt	+	+		-			-		-	
nas	(-	-	-	-)		(-	-	-	-)	

MS 22 $[-\text{cmp}] \rightarrow [+\text{grv}] / [-\text{syl}] \left[\begin{Bmatrix}+\text{cmp}\\-\text{grv}\end{Bmatrix}\right] [-\text{syl}] \_\_ \left[\begin{Bmatrix}-\text{std}\\-\text{cnt}\end{Bmatrix}\right]$

i.e., ĩ and o are redundantly [+grv] in the environment $C\begin{Bmatrix}i\\a\end{Bmatrix}C\_C$, in which the final consonant is a non-continuant or non-strident.

Since there are no triple consonant clusters, the environment bar must represent a vowel, as must also the second segment. Syllabicity is therefore redundant in the bracket on the left of the arrow and in the second segment of the environment.

	m	a	p	o	p	'clay'	m	a	p	ĭ	t	'ascend'
cns	+	(-)	+	-	+		+	(-)	+	-	+	
syl	(-)	(+)	(-)	+	(-)		(-)	(+)	(-)	+	(-)	
obs	+	(-)	+	(-)	(+)		+	(-)	+	(-)	+	
cnt	+	(+)	-	(+)	-		+	(+)	-	(+)	-	
std	-	(-)	-	(-)	-		-	(-)	-	(-)	-	
cmp		+	-	-	-			+		-		
flt				+						-		
grv	+		+	+	+				+	+	-	
nas	(+)	-	-	-	-)		(+	-	-	-	-)	

MS 23 $[+syl] \rightarrow [+grv] / \begin{bmatrix} +std \\ +cnt \\ +cmp \\ +flt \end{bmatrix} \_\_$

i.e., ĭ, a, and o are redundantly [+grv] following ṣ̌.

There are two exceptions to this rule, ṣ̌aṣ̌i 'gourd' and ṣ̌i? 'future tense,' which will have to be marked with a feature [-23]. As in previous rules, the rule here applies to a redundantly. Only the feature predicted for o and ĭ can be counted as a savings in the lexicon.

	ṣ̌	a	ṣ̌	o	'tray'
cns	+	(-)	+	-	
syl	(-	+)	(-)	+	
obs	+	(-)	+	(-)	
cnt	+	(+)	+	(+)	
std	+	(-)	+	(-)	
cmp	+	+	+	-	
grv		(+)		(+)	
flt	+		+	+	
nas	(-	-	-	-)	

The economy achieved in the lexicon by the MS rules above is calculated by multiplying the features predicted as redundant in each rule by the number of segments to which they can apply in the lexicon. The number of segments to which a rule applies is given in terms of a percentage of each 100 phonemes in the lexicon, which provides the following percentages for each rule. Rules which are mergers

of two rules have been computed on the basis of each rule comprised by the formulation, separately.

MS 1	CC	2.5	MS 13		.1
MS 2	#C	20.	MS 14		.1
MS 3		.9	MS 17	y	.2
MS 4	s	.6		w	.2
	n	.3	MS 18		.1
MS 5		.66	MS 19		.6
MS 6	?C	1.6	MS 20		4.
	?w	.1	MS 21		.66
MS 7		3.7	MS 22		1.
MS 11		1.8			

Alpha variables in the rules have been counted as one feature each. The reason this seems justified is that alpha variables are instances of a collapsing of rules so that one alpha replaces one or two +'s or -'s. The variable is thereby rewarded, since in the computation of the redundancy specifications, the rules underlying each alpha variable were each figured individually. That is, a rule

$$\text{(MS 18)} \quad \begin{bmatrix} -\text{cns} \\ -\text{syl} \\ +\text{cnt} \end{bmatrix} \rightarrow [-\ \text{grv}] \ / \ \underline{\quad} \ \begin{bmatrix} +\text{syl} \\ \alpha\text{flt} \\ \alpha\text{grv} \end{bmatrix}$$

represents two rules, one with w and one with y. The redundancies predicted by each underlying rule are computed independently so that the alpha values are explited twice here. If a rule were the merger of three or more rules, the alpha would thus give a yield in redundancy specification in proportion to the number of rules represented. There seems

to be no other way of computing the redundancies, for if in the above case yo had a higher frequency than wi, and the computation were done on the basis of the highest or lowest, the exact redundancy specification would be lost.

The count of alpha as one feature then, seems justified if one compares the uncollapsed rules with the collapsed one (a difference of four features in the rule above) and if one considers the redundancy output of the rule. Generality is given a numerical value in terms of high savings at low cost.

Diachronic change is usually in the direction of greater generality. Thus if in a rule in some language a segment is marked $\begin{bmatrix}-\text{cns}\\-\text{voc}\end{bmatrix}$ and the rule is then extended to apply to resonants as well as glides, the segment will be designated $\begin{bmatrix}\alpha\,\text{cns}\\\alpha\,\text{voc}\end{bmatrix}$. It would be desirable to reflect the greater generality of the alpha variables in the rule by having the alpha variable count less than a "+" or "-" feature, but this would lead to absurd consequences in a case where the specification of an a vowel, $\begin{bmatrix}+\text{cmp}\\+\text{grv}\end{bmatrix}$, would be more highly valued when formulated as $\begin{bmatrix}\alpha\,\text{cmp}\\\alpha\,\text{grv}\end{bmatrix}$, a formulation made possible simply because by an accident of the language no high front vowel happens to contrast with a in that rule. A unit value for alpha variables avoids this, but the problem of getting the simplicity metric to reflect greater generality in historical change in this case is unsolved.

The preceding morpheme structure rules consist of interspersed segment structure rules and sequential constraint rules. In both the case of II-C and I-C, when SS and SC rules were not interspersed the cost of the rules was much higher than when interspersed. When the SS rules all followed the SC rules, I-C required about 20 more features than II-C. But when the rules were interspersed, the difference between I-C and II-C diminished. The following chart shows the results achieved by interspersing the rules:

	II-C	II-D	I-C	I-D
gross number of features /100 phonemes	417.38	416.72	405.78	405.12
savings	66.34	62.21	62.94	59.21
net	351.04	354.51	342.84	395.91
SC cost	95		95	
SS cost	45		49	
total MS cost	140		144	

The effect of interspersing the rules was to diminish the differences of savings and costs between I and II, since both could then exploit many of the same features.

Vowel Solution D was more economical that C (by a faint margin) in specifying the lexicon without redundancy rules. After redundancies, D was shown to be less highly valued, and the computation of SS and SC cost was no longer continued because it was apparent that if D were pressed to capture some of the minor generalizations possible with C, the cost in features of the SC rules would render it very

undesirable.

If lexical specification is more highly valued than MS rule economy, I-C is to be preferred. If total MS rule economy is more highly valued, II-C is preferred.

I-C specifies the lexicon more economically not because it captures more generalizations in the MS rules but because its symmetrical feature distribution in the feature matrix makes consonants of high frequency (p, t, k) very economically specified. But the economy of specifying the lexicon by I-C does not reflect the native speaker's intuition about the sequences possible in his language in the same way that II-C does. For example, II-C in MS 3 predicts the obstruent quality of the second segment of a consonant cluster. I-C cannot do this because n is classed as [-obstruent]. The redundancy yield of the simple rule required to do this in II-C is too small (one feature per consonant cluster) to give II-C much advantage in lexical specification, but since it does make a statement about the character of the second member of all consonant clusters with much more yield than the number of features required by the rule, it is making what I feel to be a linguistically significant generalization. MS 7 predicts an obstruent feature for all final consonants. This is another generalization that I-C cannot capture. Given the consonantality and obstruency of the final segment, I-C can predict the nasality and graveness of final n, but it can make no statement about final p, t, k, s, š, ṣ̌, č. The economy achieved by II-C in this rule

reflects the difference between it and I-C, but again it is not enough to give II-C the advantage in specifying the total lexicon.

If I-C and II-C are regarded as different theories, total lexical economy will not be the first criterion for deciding in favor of one or the other. Rather, the one which best captures the desired generalizations is to be preferred. II-C seems to do this in a way that I-C does not. A suggested means of using the simplicity metric to support such a decision in close cases is as follows.

Other things being equal (a binary arrangement of distinctive features with MP rules equally well handled and the specification of the lexicon and MS rules not disproportionately different in some sense) we can require that both descriptions capture equivalent generalizations. If one cannot do so, reject it. If both can, compare the cost in features and the economy achieved. Choose the one that captures the generalizations the most economically, or if they are identical on that score, choose the one that predicts the greatest number of redundancies. This amounts to putting economy of MS rules above the lexical feature count, provided the generalizations are otherwise captured.

An example of the application of this rule of thumb is MS 23. Vowel Solution D can predict a feature for ĩ or for o following ʂ, but not for both. C predicts a feature for both and thus captures a generalization that C does not.

Solution D can be made to state a generalization similar to MS 22, but the number of features predicted are fewer than the number of features needed to state the rule. If pressed to capture the generalization, D will be shown to be less economical in the SC rules. C, then, is preferred because the natural classes formed by the feature arrangement permit it to make more general statements.

VI. MORPHOPHONEMIC RULES

Although no status is ascribed to a phonemic level between the input to the rules and the output, and there is not a one-to-one correlation between the input of this component of a generative phonology and the morphophonemic level of taxonomic phonemics, I have retained the label morphophonemic (MP) to designate this component. The term phonological (i.e., P rules) refers in general to the total phonological component comprising the morpheme structure rules and the MP rules.

The feature inventory after the morpheme structure rules have operated:

	p	t	k	ø	č	m	n	s	š	ṣ	b	r	y	w	?	h	a	ɨ	i	o
cns	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
syl	(-	-	-	-	-	-	-	-	-	-	-	-)	-	-	-	-	+	+	+	+
obs	+	+	+	+	+	+	+	+	+	-	-	(-	-	-	-	-	-	-	-	-)
cnt	-	-	-	-	-	+	+	+	+	+(+	+)	+	+	-	+	(+	+	+	+	+)
std	-	-	-	+	+	-	-	+	+	+(-	-	-	-	-	-	-	-	-	-	-)
cmp	-	-	+	-	+	(-	-)	-	+	+(-	-)	-	-(-)		+	+	-	-	-	
grv	+	-(+)		(-	-)	+	-	(-	-	-)+	-	-	+(-)		(-)	(+)	+	-	-	+
flt	(-	-	-	-	-)	(-	-)	-	-	+(-	-	-	+	-	-)	(-)	-	(-)	+	
nas	(-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-)

The following junctures are needed for Capanahua MP rules:

	+	#	≠
morpheme juncture	+	+	+
word juncture	-	+	+
sentence juncture	-	-	+

Following Harms (1967.78), the junctures are marked so that the lowest in rank is the morpheme juncture. The word level, being [+ morpheme] as well as [+word], counts for both a word and morpheme juncture. Likewise the sentence juncture counts for all three. The feature marking permits a rule to specify only the highest juncture needed for the rule, and it will by its feature specification also count for all lower junctures.

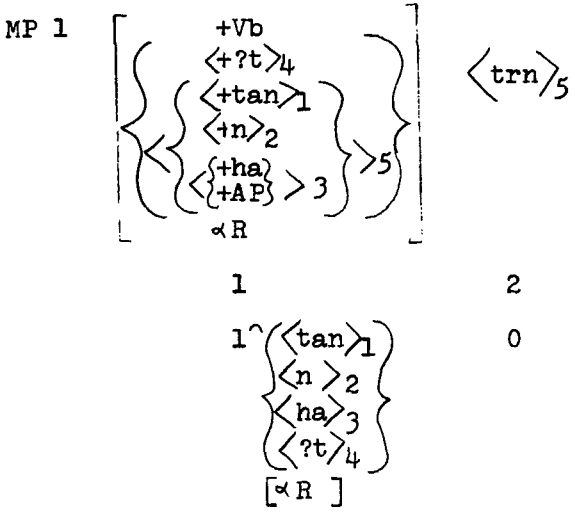
The rules are ranked according to the domain relevant to the rule. Morpheme-level rules operate only within the domain of a morpheme. It is therefore unnecessary to mark the morpheme boundaries which flank the extremes of the domain of a morpheme-level rule, since by definition a morpheme-level rule will not operate across any boundary for which the feature of morpheme juncture is marked "plus."

The transitive-marking morpheme

The transitivizer rules are properly late transformational rules, but are here given as background for the rules to follow. There is a class of "middle" verbs and

adjectives that may be inserted by the lexical rule into the CS of either a transitive or intransitive verb, but they require that a transitivizer morpheme be attached to them when they are transitive and an intransitivizer morpheme when they are intransitive.. These verbs and adjectives are marked in the lexicon by the syntactic features $[+[NP\_\_]_{VP}]$ whereas transitive verbs are marked $[+[NP\_\_]_{VP}, -[\_\_]_{VP}]$ and can be assigned only to the CS of a transitive verb. Intransitive verbs are marked $[-[NP\_\_]_{VP}, +[\_\_]_{VP}]$ and are assigned only to the CS of intransitive verbs.

The middle verbs must also carry a feature that will enable the rules to spell out the proper shape of the transitivizer or intransitivizer marker according to idiosyncratic features of the verb. There are three forms of the transitivizer: tan, n, a, each produced by the corresponding feature, and one form of the intransitivizer: ?t.



i.e., if a verb carries a [+?t] feature and no trn symbol follows the verb, ?t is adjoined. If a middle verb carrying one of the features [+tan], [+n], or [+ha], or an adjective phrase is adjacent to the transitive marker trn assigned by T 6, the segments corresponding to the features are adjoined to the verb. The transitive marker trn is deleted.

?t	nono?t	'to float'		
	rono?t	'to hang'		
	mo?i?t	'to awaken'		
	ɸa?o?t	'to sit'		
	raka?t	'to lie down'		
	wĭo?t	'to hang to'		
	haɸa?t	'to be bothered'		
	boto?t	'to descend'		
	kopi?t	'to cost'		
	rabi?t	'to turn over'		
tan	kaya	'to jump'	kayatan	'to pounce on'
	non	'to float'	nontan	'to float it'
	ron	'to hang'	rontan	'to hang it'
n	mo?i	'to awaken'	mo?in	'to wake him up'
	ɸa?o	'to sit'	ɸa?on	'to set'
	raka	'to lie down'	rakan	'to lay'
	wĭo	'to hang to'	wĭon	'to hang it to'
	nič	'to stand'	ničn	'to stand it up'
	haɸa	'to be bothered'	haɸan	'to bother'
	boto	'to descend'	toton	'to bring down'

<u>ha</u>	širo	'to play'	široha	'to tease'
	kĭsa	'to tell a lie'	kĭsaha	'to do falsely'
	kopi	'to cost'	kopiha	'to pay'
	rabi	'to be with'	rabiha	'to mix'
	kobin	'to seethe'	kobinha	'to cook'
	hošin	'red'	hošinha	'to make it red'
	?ani	'big'	?aniha	'to raise'
	sirip	'good'	siripha	'to do it well'
	nĭnkĭp	'long'	nĭnkĭpha	'to lengthen'

MP 2 $\begin{bmatrix} +\text{nas} \\ \langle +\text{M} \rangle_1 \\ \langle +\text{R} \rangle_2 \end{bmatrix} \rightarrow [+\text{grv}] \; / \; \_\_ + \begin{bmatrix} \langle -\text{M} \rangle_1 \\ \langle -\text{R} \rangle_2 \end{bmatrix}$

i.e., a morpheme-final nasal becomes [+grv] if the morpheme contains a feature [+M] or [+R] assigned by the T rules to lengthen the morpheme ([+M]) or to introduce a reflexive marker morpheme ([+R]), and a morpheme that does not contain an identical feature ([+M] or [+R]) follows.

All final nasals will have been specified as [-grv] by the MS component in order to predict an <u>n</u> in this position if any vowel-initial suffix is added to the stem. There seems to be no phonological feature in the R and M morphemes that can be used to explain the shift in graveness of the nasal.

MP 3 $\left[\left\langle \begin{Bmatrix} \langle +k \rangle_1 \\ +nas \end{Bmatrix} \right\rangle_2 \quad [-R] \right.$
$\left. +R \right]$

SI: 1 2

SC: $1 \rightarrow 1 + \langle \langle k \rangle_1 \; ĭ \rangle_2 \; ? \; t$

i.e., if a feature [+R] (found only in verbs) and a [+k] feature are found in the same segment, and a [-R] segment follows, kĭ?t is added to the [+R] segment. If [+R] is found in the same segment as a nasal feature in the same environment as above, ĭ?t is added. Otherwise ?t is added.

The reflexive feature R has been assigned to the verb by a T rule which deletes an object NP identical to the subject NP of a verb. The allomorphic shapes of the reflexive-marking morpheme are determined in part by the phonological shape of the verb.

[+k] and [+nas] are disjunct in this rule. [+k] is expanded first because by convention the topmost set in a brace constitutes the first application of the rule. It turns out, in this rule, that it would make no difference whether k or m were expanded first, since they are here in complementary distribution, but the expansion convention will be shown to be more important in later rules.

The shapes of the morpheme spelled out by this rule are thus kĭ?t, ĭ?t, and ?t.

hoš̤ohakĭ?t 'to turn white'

hošinhakĭ?t 'to turn red'

hamakhakĭ?t 'to step softly'

taramĭ?t 'to roll oneself'

ʂotomĭ?t 'to push oneself'

ra?omĭ?t 'to cure oneself'

ʂatĭ?t 'to cut oneself'

nĭʂa?t 'to tie oneself'

øĭka?t 'to pull oneself out'

tĭ?ka?t 'to break oneself'

MP 4 $\begin{bmatrix}+\text{syl}\\+\text{cmp}\end{bmatrix} \rightarrow$ null /

$$\left\{\begin{array}{l}\\\\\end{array}\right.$$

$$+ \underset{\begin{bmatrix}+\text{syl}\\-\text{grv}\end{bmatrix}}{\overset{\text{i}}{\_\_\_}} \underset{\begin{bmatrix}-\text{cns}\\-\text{syl}\\-\text{cnt}\end{bmatrix}}{\overset{?}{}} + \text{k}$$

$$\left\{\begin{array}{l}\#\left\{\begin{bmatrix}+\text{obs}\\-\text{syl}\\-\text{cns}\end{bmatrix}\right\}[+\text{syl}]\left(\begin{bmatrix}-\text{cns}\\-\text{syl}\\-\text{cnt}\end{bmatrix}\right)[+\text{cns}]\\\qquad\qquad[+\text{syl}]\end{array}\right\}\langle+\rangle_3\_\_\_\langle+\begin{bmatrix}-\text{cns}\\-\text{syl}\\-\text{cnt}\end{bmatrix}\rangle_1$$

(cont. below)

$$\left\{\begin{array}{l}\\\\\end{array}\right.$$

$$\begin{bmatrix}-\text{cmp}\\-\text{grv}\\-\text{std}\\\langle+\text{obs}\\-\text{cnt}\rangle_1\\\langle+\text{syl}\rangle_2\end{bmatrix} \quad \langle\begin{bmatrix}-\text{cns}\\-\text{syl}\\-\text{cnt}\end{bmatrix}\rangle_2 + \left\{\begin{array}{l}\\\\\end{array}\right.$$

i.e., the vowel a is deleted if it is followed by i?+k. Also, the a vowel is deleted if it follows a morpheme of shape CV?C or CVC and precedes ?t or i? and a + juncture, and finally, a is deleted if it follows a vowel and precedes

the latter part of the environment above.

before i?+k:

katanai?ki → katani?ki 'he goes'

?oš̤aka?inai?ki → ?oš̤aka?ini?ki 'he falls asleep'

mapɪ̈tai?ki → mapɪ̈ti?ki 'he ascends'

pokinai?ki → pokini?ki 'he digs'

wɪ̈ranai?ki → wɪ̈rani?ki 'he pushes it'

hamakai?ki → hamaki?ki 'he steps'

ninkakai?ki → ninkaki?ki 'he hears'

nonaai?ki → nonai?ki 'he swims'

?onanai?ki → ?onani?ki 'he knows'

By rule 4, all transitive verb stems ending in a will have the a deleted if the reflexive marker follows, with the exception of forms that do not have a consonantal segment beginning the second syllable: (e.g., nia 'to place' and nɪ̈?a 'to tie up' are of the type excluded from the rule).

tɪ̈?ka?t → tɪ̈?k?t 'to break oneself'

nɪ̈š̤a?t → nɪ̈š̤?t 'to tie oneself'

wɪ̈š̤a?t → wɪ̈š̤?t 'to scratch oneself'

pɪ̈ka?t → pɪ̈k?t 'to untie oneself'

tɪ̈sa?t → tɪ̈s?t 'to come apart'

øɪ̈ka?t → øɪ̈k?t 'to pull oneself out'

š̤oka?t → š̤ok?t 'to skin oneself'

pa?øa?t → pa?ø?t 'to slap oneself'

tana?t → tan?t 'to test oneself'

Monosyllabic stems of shape CVC cause the a of the present tense to drop:

bičai? → bič̌i? 'I grab'

ničai? → ničǐ? 'I walk'

hisai? → hisi? 'I see'

Following a vowel and preceding i?+, a is dropped:

šǐ?poai? → šǐ?poi? 'I close it'

tǐ?kaai? → tǐ?kai? 'I break it'

šǎ?tǐai? → šǎ?tǐi? 'I cut it'

po?poai? → po?poi? 'I shut it'

boai? → boi? 'I carry'

piai? → pii? 'I eat'

niaai? → niai? 'I place it'

But the vowel will not be deleted if a closed syllable preceding the present tense is not a monosyllabic verb stem:

wǐranai? 'I pushed it'

kayatanai? 'I went and jumped'

sǐnǐnai? 'I reach it'

kaisai? 'I choose it'

mapǐtai? 'I ascend'

hamakai? 'I stepped on it'

Since middle verbs ending in a cannot be reflexivized unless they are first transitivized, they will never occur with any reflexive marker other than kǐ?t. The ordering of the rules that spell out the transitive markers prior to those that spell out the reflexive marker is

therefore essential.

The first syllable of the stem must begin with a consonantal or a glide, but not a resonant. This generalization is only awkwardly captured by the feature system I have chosen and is the price paid for the advantage of using the syllabic instead of the vocalic feature. $\left\{\begin{bmatrix}+\text{obs}\end{bmatrix}\right.$ $\left.\begin{bmatrix}-\text{syl}\\-\text{cns}\end{bmatrix}\right\}$ Otherwise, [-vocalic] would capture the generalization, using the Jakobson-Halle feature system. However, the gain in economy elsewhere compensates for the loss of generality here.

Another possibility would be to allow the rule to specify only [-syl] for the first segment, then mark all resonant-initial stems as exceptions to the rule.

rĩra	'to chop'	bana	'to plant'
bĩna	'to search for'	boya	'to plug'

These would either be marked in the lexicon or marked by a redundancy rule predicting in the feature, for only the second line of the rule. The other lines must still operate, i.e., rĩraai? → rĩrai? 'I chop', so that the generalization captured by collapsing the rules would be lost.

MP 5 $\quad \underset{[-\text{syl}]}{\text{C}} \quad \underset{[+\text{syl}]}{\text{V}} \quad \underbrace{\left(\begin{bmatrix}-\text{cns}\\-\text{voc}\\-\text{cnt}\end{bmatrix}\right)^{?} \quad \underset{[+\text{cns}]}{\text{C}}}_{3} \quad \underbrace{\left(\begin{bmatrix}-\text{cns}\\-\text{voc}\\-\text{cnt}\end{bmatrix}\right)^{?} \quad \underset{[+\text{cns}]}{\text{C}}}_{4} +$

SI: $\quad 1 \quad 2 \quad 3 \quad 4$

SC: $\quad 3 \rightarrow 3 + 2$

i.e., a morpheme with no vocalic segment is separated from a preceding consonantal segment by the insertion of a redupli-cated vowel before the morpheme.

By this rule, a form resulting from the transitivizer of rule 1 receives a vowel:

nič̣n → ničin 'to stand it up'

Also, forms which have been assigned a reflexive marker ?t by rule 3 receive an epenthetic vowel, if the verb stems are consonant-final:

rĭok?t → rĭoko?t 'to scatter itself'

ra?payok?t → ra?payoko?t 'to decay bodily'

yokak?t → yokaka?t 'to inquire for oneself'

yo?yok?t → yo?yoko?t 'to go to waste'

kais?t → kaisi?t 'to choose oneself'

Similarly, forms which have undergone the vowel deletion of rule 4 are assigned a vowel before the intransi-tivizer:

nĭṣ̌?t → nĭṣ̌ĭ?t 'to tie oneself'

tĭ?k?t → tĭ?kĭ?t 'to break itself'

wĭṣ̌?t → wĭṣ̌ĭ?t 'to scratch oneself'

pĭk?t → pĭkĭ?t 'to untie oneself'

tĭs?t → tĭsĭ?t 'to come apart'

ɸĭk?t → ɸĭkĭ?t 'to pull oneself out'

ṣ̌ok?t → ṣ̌oko?t 'to peel oneself'

paɸ?t → paɸa?t 'to slap oneself'

tan?t → tana?t 'to test oneself'

The third person marker ş will, when following the plural marker kan, form a cluster that is resolved by this same rule:

ka+ipi+kan+ş+ki (go-past-pl-3rd person-declarative)

ka?ipikanaşki 'they went'

bana+ipi+kan+ş+ki (plant- past-pl-3rd-decl)

bana?ipikanaşki 'they planted'

Middle verbs that take the intransitive-marking morpheme also undergo the rule:

non?t → nono?t 'to float'

ron?t → rono?t 'to hang'

The ordering of rule 4 preceding rule 5 is crucial to the ability of 5 to specify the vowels for all intransitivized forms as well as for the other sources of consonant clusters.

The effect of rule 5 has thus been to prevent a morpheme consisting of only a consonant as its phonological shape from forming a cluster with a preceding consonant-final form.

The final + boundary in 5 prevents a vowel from being inserted when a CC cluster is formed at morpheme junctures when the second morpheme consists of more than a single consonant. Thus, sequences such as hamak+tan 'step-over a distance' will not be affected by 5.

There are two exceptions to rule 5: toba- 'to break' and čoka- 'to wash', which become tobi?t and čoki?t, respectively. This is a residue not yet explainable.

MP 6 $\text{null} \rightarrow \begin{bmatrix} n \\ +nas \\ +grv \end{bmatrix} \begin{bmatrix} o \\ +syl \\ +flt \end{bmatrix} / \left[\begin{bmatrix} C \\ -syl \end{bmatrix} \begin{bmatrix} V \\ +syl \end{bmatrix} \left(\begin{bmatrix} (C) \\ -syl \end{bmatrix} \right) \right]_V \_\_ \begin{bmatrix} CVC+ \\ +B \end{bmatrix}$

i.e., verb stems consisting of only one syllable have another syllable adjoined to the stem if the stem immediately precedes the [+B] morpheme feature which is also used by the T rules to spell out the phonological shapes of tense morphemes.

biˇ+ṣi?	biˇnoṣi?	'will get'
?a?+ṣi?	?a?noṣi?	'will do'
ka+ṣi?	kanoṣi?	'will go'
pi+ṣi?	pinoṣi?	'will eat'
niˇ+ṣi?	niˇnoṣi?	'will walk'
his+ṣi?	hisnoṣi?	'will see'

But if the [+B] morpheme is bisyllabic, the rule does not apply:

niˇ+ipi	niˇ?ipi	'did walk'

Lengthening rules

The syntactic component assigns a morpheme-lengthening feature M to nouns, adjectives, and verbs. The lengthener takes a different shape on demonstrative pronouns and verbs than it does on nouns, pronouns, and adjectives. A separate rule for verb and demonstrative lengthening is therefore necessary.

MP 7 $\left[\left\{ \begin{matrix} [+Pro \\ +S\_\_] \\ +V \\ +M \end{matrix} \right\} \right] \left\{ \begin{matrix} [-M] \\ \# \end{matrix} \right\}$

SI: 1 2 SC: 1 → 1 + to?on

mawataito?on 'the one who died'

?oato?on 'that one'

nĭato?on 'this one'

MP 8 $\begin{bmatrix} o \\ +syl \\ +flt \\ +suffix \\ +M \end{bmatrix}$ [-M]

SI: 1 2 SC: 1 → [-flt, +cmp]

The homophonous morphemes bo 'plural' and bo 'generic' become ba when the lengthener feature M is present. Since there are other cases where the same sequence of segments does not undergo a change before a nasal, e.g., šobo+n 'in the house,' it seems best to assume that a morpheme feature triggers the rule. There are also words that end in -ba (baba 'grandchild') so that positing an underlying form ba which changes to bo before # is ruled out. An aw sequence in the verb haw 'to come' also shows an a-o change, but different from that above, aw becoming o before consonants or word juncture and remaining aw before vowels, so that positing an aw sequence in the underlying form of bo is likewise ruled out. (In the examples, the final n is produced eventually by the [+M] feature.)

generic: ?aibo 'woman' ?aiban

plural: honibo 'people (man-pl)' honiban

The remainder of the lengthener forms are determined by the number of syllables in the morpheme to be lengthened and by the phonological feature of the final segment of the

morpheme.

The rules require the condition that no morpheme juncture be included in their domain. Since by convention, morpheme junctures are not written unless their presence is relevant to the rule, their absence cannot be indicated by omitting them. This makes it necessary to identify this as a morpheme level rule: i.e., the domain of the rule is the morpheme. Unless the environment specification of a morpheme-level rule specifies where a juncture may occur, the rule is limited in application to a stretch of segments bounded by morpheme (or higher) junctures and containing no junctures.

MP 9 (Morpheme level)

$$\underbrace{(\ CVC_0\langle V\rangle\ )\ (\ C\ )\ \left\langle \begin{Bmatrix} CV \\ C \end{Bmatrix}_1 \right\rangle}_{1} \quad \underbrace{\begin{Bmatrix} \# \\ [-M] \end{Bmatrix}}_{2}$$

Cond: $1 = [+M]$ SC: $1 \rightarrow 1 + \left\langle \begin{Bmatrix} n \\ a \end{Bmatrix}_1 \right\rangle n$

i.e., a morpheme which has been assigned the [+M] feature is lengthened if it precedes a word juncture or [-M] segment. The application of the rule requires adherence to the non-iterative rule convention and the rule expansion convention.

The non-iterative rule convention stipulates that a rule cannot apply to its own output, hence, cannot apply to the same string more than once. If rules could reapply, alpha-switching rules would in some cases endlessly reverse the feature values of the output of each previous application.

There have not yet been any reasons adduced for requiring that only alpha-switching rules are non-iterative, and the effectiveness of other rules formulated with the non-iterative rule convention seems to argue against there being any such ad hoc requirement for alpha rules.

The expansion convention states that an expandable rule must apply first to the longest stretch permitted by the longest reading of the rule. In this way, when angles, which stand for numbered parentheses in pairs (or more) as in rule 11 are found, first all the possible expansions are applied using the angles, then all the possible expansions without angles but with optional C, and finally, all possible expansions permitted by the upper and lower limits of the first parenthesis. The upper limit, in this case, is unspecified.

The possible expansions in their order of expansion (omitting some sub-expansions permitted by C_o, which are not relevant to the rule):

		non-occurring	accepted	pre-empted
1.	(CVCV)(CVCV)(C) CV	x		
2.	(CVCV)(CVCV) CV	x		
3.	(CVCV)(C) CV		x	
4.	(CVCV) CV		x	
5.	(CVCV)(C) C	x		
6.	(CVCV) C		x	
7.	(CVC)(CVC)(CVC) C	x		
8.	(CVC)(CVC)(CVC)		x	
9.	(CVC)(CVC)(C)	x		

10.	(CVC)(CVC)			(6)
11.	(CVC)(C)	x		
12.	(CVC)		x	
13.	(CV)(CV)(CV)(CV)(C)	x		
14.	(CV)(CV)(CV)(CV)		x	
15.	(CV)(CV)(CV)(C)		x	
16.	(CV)(CV) C			(6)
17.	(CV) C			(12)
18.	(CV)(CV)		x	
19.	CV		x	

Having utilized 〈 〉 to formulate the largest expansions, this set of angles is not reapplied in forming smaller expansions. Thus, a sequence of CVCCV cannot be produced by choosing CVC from the first parenthesis and CV from the latter pair of angles. That set of angles, being optional, is in the longer expansions only.

The interesting support the rule gives to the convention is that, if the above proscribed interpretation were used, the rule would put a nĭn lengthener on forms that should only receive n. On the other hand, the present interpretation permits cclapsing into one rule a fairly complex set of conditions.

Examples for the expansions listed above:

2	ko?koṣtobi	ko?koṣtobin	'butterfly'
3	bĭmanan	bĭmanannĭn	'face'
4	ṣontako	ṣontakonĭn	'girl'
	čičika	čičikanĭn	'knife'

	tininti	tinintinín	'governor'
	?očiti	?očitinín	'dog'
5	hisis	hisisan	'ant'
	maraš	marašan	'poison'
	po?pos	po?posan	'mud'
	ka?moš	ka?mošan	'snake'
	naíš	naíšan	'center'
	níšníš	níšníšan	'brownbird'
	?awap	?awapan	'tapir'
	?amín	?amíman	'capybara'
6	honi	honin	'man'
	?ino	?inon	'tiger'
	šobo	šobon	'house'
	šítí	šítín	'buzzard'
	wasi	wasin	'grass'
	bara	baran	'bullet'
	boo	boon	'hair'
7	?ín	?ínn	'I'
	min	minn	'you (sg)'
	bo	ban	'they, people'
	bi	bin	'very'
	ko	kon	'place'

It has up to now been assumed by linguists in generative phonology that the lowest level rule in the hierarchy is the word level rule for the MP component. Chomsky and Halle have been reported to feel that all morpheme level rules must be in the morpheme structure

component: all other rules, pre-cyclical and otherwise are of a higher rank.[6]

There is no evidence to show that Capanahua morpheme-level rules should follow a higher rank rule, but there is evidence that they must precede some of the higher level rules: they must precede the vowel harmony rules, for example, that operate on vowels inserted by the lengthener rules.

The order of the lengthening rules with respect to the vowel-deletion and vowel insertion rules, and the rules spelling out the shapes of the R and transitivizer morphemes is arbitrary. The latter apply only to verbs, and the lengthener rules (with the exception of the one spelling the shape of the lengthener on verbs) apply to non-verbal elements.

The convention that + junctures are not represented in the rule unless their presence is relevant to the rule is founded on the hypothesis that no morpheme-level MP rules (i.e., rules for which the domain is restricted to the string bounded by a pair of morpheme junctures and which contain no such junctures within the domain) are necessary.

Since the morpheme level rule here constitutes a counter-example to what has been assumed to be, or possibly be, a language universal, it is of interest to ascertain how widely such rules can be found.

Notice the following Capanahua surface-string words, with their reconstructed forms for Proto-Pano:

Proto-Pano	Capanahua	
*marin	mari	'rodent'
*manin	mani	'banana'
*biin	bii	'mosquito'
*?awara	?awa	'tapir'
*kapĭtĭ	kapĭ	'alligator'
*?amĭnĭ	?amĭ̆	'capybara'
*mĭkĭnĭ	mĭkĭ̆	'hand'
*yosibo	yosibo	'old man'
*bĭtonko	bĭtonko	'forehead'
*maraš̆ĭ	maraš̆	'poison'
*ka?moš̆ĭ	ka?moš̆	'snake'

Comparative evidence indicates that the grammar of Proto-Pano included a noun shortening rule that operated on all nouns standing immediately preceding a verb. These nouns underwent rules that shortened them by deleting the final vowel or consonant somewhat in the following way.

First, assume that a T rule assigned a shortening feature to the noun standing before the verb. This feature made the noun subject to the shortening rules which, presumably, were pre-cyclical rules.

SI: X N V Y

1 2 3

SC: 2 → [+T 1]

1. [+T 1] → null / CVCV(C) ___#

i.e., the first shortener rule deleted the final segment of a two or three syllable noun containing the shortening

feature. (Without the non-iterative rule convention, a special apply-once-only restriction would have to be stated for this rule; otherwise the rule would continue to apply and remove much more than the final segment.)

*kapĭtĭ	'alligator'	*maraṣ̌ĭ	'poison'
kapĭt		maraṣ̌	
*mĭkĭnĭ	'hand'	*marin	'rodent'
mĭkĭn		mari	

2. V → [+strs] / ___ C #

i.e., final closed syllables were stressed.

kapĭ́t maráṣ̌ mĭkĭ́n mari

3. V → [+nas] / ___ [+nas] #

i.e., vowels were nasalized preceding final nasals. (I am not sure of the fact that only before final nasals were vowels nasalized, but at least these were. The rule may well have been without the juncture.)

kapĭ́t maráṣ̌ mĭkĭ̃́n mari

4. $\begin{bmatrix} +\text{cns} \\ -\text{voc} \\ -\text{std} \end{bmatrix}$ → null / ___ #

i.e., final non-strident obstruents were deleted.

kapĭ́ maráṣ̌ mĭkĭ̃́ mari

Chacobo[7] and Pacaguara (in Bolivia) and Marinahua and probably also Yamiaka, Azairi and Atsahuaca[8] (in Peru) still retain the original syntactic rules almost intact and Chacobo and Marinahua still retain the phonological rules, with the exception of Rule 3 which Chacobo has lost.

Most of the other Panoan languages now show the shortened forms (found in rule 4) in their surface string as subject of intransitive verbs or object of transitive verbs. The rules have undergone a modification so that instead of a deletion feature being assigned to nouns standing immediately before a verb, a [-lengthen] feature is assigned them. The base forms are now those found in rule 2. (Capanahua has further restructured the base forms so that almost all nouns which terminated in any occlusive originally now terminate only in p.) The lengthener rules now apply to all forms which have not been assigned the [-lengthen] feature. Each language seems to differ somewhat in the rules that develop the long forms, but from the information available it is apparent that they all tend towards simplification by regularizing in a systematic way the endings now used to lengthen nouns, similar to the Capanahua pattern.

Thus, the Proto-Panoan noun shortening rules have now been replaced by noun lengthening rules. But the Proto-Panoan shortening rules seem to have been morpheme-level rules as are the present Capanahua lengthening rules, for the only forms that have been retained unmodified are the monosyllabic morphemes, or compounds in which the final morpheme is monosyllabic, e.g., min 'you,' sonta-ko (young one) 'adolescent girl.' (Most of these are now fused forms without internal morpheme junctures in the lexicon.) Stems of two syllables were not reduced to one syllable, nor were monosyllables shortened. The shortening rule must have

applied only to CVCVC, and longer, morphemes.

The fused forms now undergo the lengthening rules as single morphemes. Compound nouns with internal junctures still permit only the final juncture-free stretch to undergo the rule: bĭ+tonko (face-lump) 'forehead' becomes bĭtonkon (i.e., as a bi-syllabic word), not bĭtonkonĭn as would be expected of a tri-syllabic morpheme; and yoš̆an+bo (old-person) 'old woman' is lengthened as a monosyllabic morpheme, yoš̆anban.

There is therefore evidence that we might expect morpheme-level rules to be found in perhaps twenty Panoan languages.

Therefore the current convention in generative phonology that there are no morpheme-level rules needs to be modified to allow them. It also seems possible to view MP 11 as being a transformational rule in the syntactic component and not properly a part of the phonological component, so that the convention excluding morpheme-level rules would be unaffected. This would require the assumption that the phonological features are assigned before the transformational rules apply, whereas this description assumes that they apply after the transformational rules, but shifting the location of the rule just to preserve a theory about the phonological component would mean the theory is only weakly motivated.

MP 10 $$\begin{matrix} a \\ \begin{bmatrix} +syl \\ +cmp \end{bmatrix} \end{matrix} \rightarrow \begin{matrix} ĭ \\ [-cmp] \end{matrix} / \begin{matrix} w,s,s,s \\ \begin{bmatrix} -syl \\ +cnt \\ -nas \end{bmatrix} \end{matrix} + \underline{\quad}$$

i.e., The a vowel inserted by the vowel lengthener rules, the pronominal a introduced by T rules, and the a of the past tense and past participle morpheme become ĭ when preceded by a strident or w.

maraš̭+an	maraš̭ĭn	'poison'
hisi+an	hisisĭn	'ant'
po?pos+an	po?posĭn	'mud'
ka?moš̭+an	ka?moš̭ĭn	'snake'
naĭš̭+an	naĭš̭ĭn	'center'
haw+aš̭	hawĭš̭	'having come'
haw+a+ki	hawĭki	'I came'
haw+a+n	haw+in	'he'
his+aš̭	hisĭš̭	'having seen'
bič+aš̭	bičĭš̭	'having taken'
bič+an+kin	bičĭnkin	'took it away from someone'

It happens that b, r, y, and h do not occur morpheme-final, so the rule has taken advantage of their faulty distribution to specify the environment as [-nas]. If for some reason one wished to exclude the possibility of b, r, and y here, the environment would be:

(modified 10) $$\begin{bmatrix} -syl \\ +cnt \\ \{+std\} \\ \{+flt\} \end{bmatrix} + \underline{\quad}$$

The rule then says that in this environment the strident continuants and flat continuants function alike.

Since only two morphemes terminate in w, it might be desirable to use alternate treatments.

It would be possible to mark the two w-final morphemes with a [+10] feature. The rule would than say that 'come,' and 'he, she, it' are exceptions in that they obligatorily satisfy the environmental specification of rule 10, even though they do not terminate in a strident:

$$\ldots \begin{Bmatrix} [+\text{std}] \\ [+10] \end{Bmatrix} \_\_$$

This approach is simple but presents the difficulty of evaluating [+10] in the simplicity metric, since [+10] is a morpheme feature, not a phonological feature.

Another possibility is to regard w as having different underlying features in this instance, which might be suspected on the basis of comparative evidence, for the related language Marinahua shows many p^{w} correspondences to w.

If w in this case were to be regarded as a strident p^{w} or f^{w} in the underlying form, the modified rule 10 could be formulated in a more economical way:

$$10' \qquad \ldots \begin{matrix} \text{f}^{\text{w}} \text{ or } \text{p}^{\text{w}} \\ [+\text{std}] \end{matrix} + \_\_$$

In the feature inventory there would be four continuant stridents in this solution, specified either in the lexicon or developed by P rules prior to 10', with the features:

	f^w	s	š	ṣ
std	+	+	+	+
cnt	+	+	+	+
cmp	-	-	+	+
flt	+	-	-	+

If w were regarded in these words as f^w in the underlying form, the fact that it occurs final would be perfectly regular, for all the sibilants occur final. This treatment would require that s be specified with one more feature for each occurrence in the lexicon, but would make possible an additional savings in the redundancy rules whereby the syllabicity of final segments is exploited to predict their obstruency and consonantality:

$$[\text{ syl}] \rightarrow \begin{bmatrix} -\alpha\text{obs} \\ -\alpha\text{cns} \end{bmatrix} \quad / \_\_\_ \#$$

Since the occurrence of s per 100 phonemes is 1.43 in the lexicon, and the occurrence of a segment in final position is much higher, a savings would be achieved and it would be unnecessary to introduce a feature-changing rule to convert w to f^w in the P rules prior to 10:

$$\overset{w}{\begin{bmatrix} -\text{cns} \\ -\text{syl} \\ +\text{cnt} \\ +\text{grv} \end{bmatrix}} \rightarrow \overset{f^w}{\begin{bmatrix} +\text{cns} \\ +\text{std} \\ +\text{flt} \end{bmatrix}} \quad / \_\_\_ +$$

This f^w treatment would also simplify the blank-filling rules, for all the continuant stridents would now be already specified for flatness, eliminating the need to extend the environment of the flatness feature-specification rule to cover s.

However, it would be necessary to have a rule changing f^w to w before the nasalization rules operate. This rule would require the specification of 6 features:

$$10'' \quad \begin{bmatrix} +std \\ +cnt \\ -cmp \\ +flt \end{bmatrix} \rightarrow \begin{bmatrix} -std \\ -cns \end{bmatrix}$$

The f^w solution then, would gain economy in the MS rules and lexicon, but would require four more features than rule 10 in the MP rules. The [-nas] solution is preferred to the f^w solution by the simplicity metric if we grant that economy in the MP rules is to be valued above economy in the lexicon. But the simplicity metric will not choose the w solution (modified rule 10) over the morpheme feature solution, for there is no way of evaluating a morpheme feature against three phonological features. Rule 10, using purely phonological features, is predictive, but the morpheme feature treatment is not. Rule 10 predicts that if another w-final word is found, it should produce the same results. It is testable. The morpheme feature makes no predictable claim and is thereby a weaker solution, to be exploited when a generalization is not statable in purely phonological terms.

$$\text{MP 11} \quad \overset{a}{\begin{bmatrix} +syl \\ +cmp \end{bmatrix}} \rightarrow \overset{o}{\begin{bmatrix} +flt \\ -cmp \end{bmatrix}} \; / \; \underset{\overset{w}{\begin{bmatrix} -cns \\ -syl \\ +cnt \\ +grv \\ -cmp \end{bmatrix}}}{\_\_} \; \left\{ \begin{matrix} [-syl] \\ \# \end{matrix} \right\}$$

This change of vowel quality is found only in the

one verb stem haw 'come,' which is the only morpheme that can produce a wC cluster.

haw+i → hawi 'come'

haw+?o+ṣ+ki → how+?o+ṣ+ki → ho?oṣki 'he came'

haw+ni+ṣ → . . . honiṣki 'he came (long ago)'

haw+yama → . . . hoyama 'not come'

An alternate solution here as in 10 would be to assign a morpheme feature to the verb marking haw as obligatorily undergoing the minor rule 11b.

$$\text{11b} \quad \begin{bmatrix} a \\ +\text{syl} \\ +\text{cmp} \\ +\text{llb} \end{bmatrix} \rightarrow \begin{bmatrix} o \\ -\text{cmp} \\ +\text{flt} \end{bmatrix} \; / \; \_\_ \; [+\text{cns}] \; \left\{ \begin{bmatrix} \# \\ -\text{syl} \end{bmatrix} \right\}$$

It would no longer be necessary to specify the features of w, for the [+llb] feature occurs in no other word.

Again, it seems to me to be undesirable to use morpheme features to explain phonological feature changes that can otherwise be explained by a phonological rule using only phonological features. The fact that only one morpheme undergoes the rule can in this case be as well explained by the fact that only this morpheme has this particular sequence of features, as by a distinctive feature. Also, the predictive quality of the phonological formulation is lost in the morpheme feature solution.

$$\text{MP 12} \quad \begin{bmatrix} ɨ \\ +\text{syl} \\ -\text{cmp} \\ +\text{grv} \\ -\text{flt} \end{bmatrix} \rightarrow \begin{bmatrix} i \\ -\text{grv} \end{bmatrix} \; / \; \begin{bmatrix} i \\ +\text{syl} \\ -\text{grv} \end{bmatrix} \; \begin{bmatrix} C \\ -\text{syl} \end{bmatrix} \; \_\_$$

i.e., ɨ becomes i when preceded by i in the preceding syllable.

The basic forms have begun as a, have been shifted in some cases to ɨ, and now some of the latter are shifted to i just in case the preceding vowel is i. It would have been possible to formulate 10 so that both surface forms are produced in the same rule:

$$\begin{bmatrix}+syl\\+cmp\end{bmatrix} \rightarrow \begin{bmatrix}-cmp\\ \langle -grv \rangle\end{bmatrix} \; / \; \begin{bmatrix}+syl\\ \langle -grv \rangle\end{bmatrix} \quad \begin{bmatrix}-syl\\+cnt\\-nas\end{bmatrix} \_\_$$

However, rule 12 would still be needed to shift some occurrences of ɨ to i when no strident intervenes, and the combined formulation of a → $\begin{Bmatrix}i\\ɨ\end{Bmatrix}$ of rule 10 would therefore be less economical.

Rule 12 must follow 10, but the ordering of 11 is arbitrary.

hisisɨn	hisisin	'ant'
hisɨšᵩ	hisišᵩ	'having see'
hisɨki	hisiki	'saw'
bičɨnkin	bicinkin	'took it away from someone'
ničɨšᵩ	ničišᵩ	'having walked'
tinitinɨn	tinintinin	'governor'
ʔocitinɨn	ʔocitinin	'dog'
yawišɨn	yawišin	'armadillo'
pipipipičɨšᵩ	pipipipičišᵩ	'having eaten continually'

MP 13 (morpheme level)

$$\begin{bmatrix}+syl\\-grv\end{bmatrix} \rightarrow \overset{ɨ}{[+grv]} \; / \; \# \overset{C}{[-syl]} \overset{a}{\begin{bmatrix}+syl\\+cmp\end{bmatrix}} \overset{C}{[-syl]} \_\_ \begin{bmatrix}-syl\\-cnt\end{bmatrix}$$

i.e., an i vowel in the second syllable of a morpheme becomes

ĭ when the preceding vowel is a and i is followed by an occlusive. MP 13 may precede MP 11.

Though this rule operated in early Panoan on a larger number of forms, it is now required in Capanahua only by the form rabĭ 'two' and the related homophonous rabĭ 'to mix'.

rabĭ 'two' must have a final consonant in the base form to account for the final stress and pitch. When the lengthener rules apply, the final consonant appears, č. By the normal rule, the lengthened form should be rabĭčĭn. In fact, it is rabĭčin, with an irregular i that has to be accounted for.

The final consonant of the reconstructed form is *t: *rabita. Before the deletion of final consonants took place, a rule (t → č / i __) operated, changing the strident feature of t following i. The resultant forms are now preserved in Capanahua base forms (nič 'to walk,' bič 'to grab,' biči 'hide,' etc.) and the absence in the base forms of an it sequence is exploited by the redundancy rules in the lexicon. The fact that these words reconstruct as *it sequences suggests that perhaps the t → č rule was first an MP rule and was eventually incorporated into the Capanahua MS rules. If it had been an MS rule in early Panoan, there should be no ič sequences reconstructable as it. However, early incorporation of this rule into Capanahua MS rules caused the original final t to be preserved as č.

Without the form rabit 'two', Capanahua MS rules would require a rule exploiting the fact that morphemes of

the form CaCiC are not found: in one such rule ĭ would be redundantly [+grv] in the environment CaC_C where the final C is not a continuant. This is essentially rule MP 13. If this rule is incorporated into the MS rules, rabič 'two' will have to be entered in the lexicon in something closer to its surface form rabĭč and will in addition have to carry a special morpheme feature to trigger an ad hoc rule forming the irregular noun-ending (lengthened form) rabĭč-in 'two.' But if rule MP 13 is incorporated into the MP rules, the lengthened ending turns out to be perfectly regular, and the absence (in the surface structure) of nouns and verbs of the form CaCiC is still explained.

The decision to formulate rule 13 on the basis of purely phonological features rather than a morpheme feature must rest on the fact that a linguistically significant generalization is captured by so doing, and it involves a modification of Halle's dictum that economy of MP rules is valued above lexicon and MS rule economy: a generalization in the MP rules is valued above a generalization in the MS rules, since total economy does not always reflect a capturing of specific generalizations.

There is no good way of comparing the economy achieved by the two possible solutions, for the economy of an ad hoc morpheme feature cannot be compared with the phonological feature count of rule 13. For this reason Halle's valuation of MP economy over MS economy is insufficient as a basis to choose between the two competing

solutions. Nor will it do to propose that morpheme features should always be rated as less economical than a purely phonological one, for one could always list the phonological features of the morpheme in such a case. And if there were several different formatives, each with a different phonological matrix offering no phonological generalization, but all sharing the same morpheme feature, the morpheme feature would seem to catch the generalization and therefore be rated as more economical than a simple listing of phonological feature matrices.

Rule 13 seems to me to capture a generalization because it shows the grammar to be perfectly regular on phonological grounds, but a morpheme feature does just the opposite. Rule 13 also shows why no CaCiC form appears in the surface string, but a morpheme feature does not, since the MS rules made possible by the morpheme feature solution can exploit a feature of a, o, or ĭ in the second syllable, or of i, o, or ĭ in the first syllable, or of one of the consonants (e.g., [+cnt] in the final C), depending on which one provides the best savings in the lexicon.

Yet rule 13 does permit savings in the lexicon if rabič is marked with an "exception" feature, exempting it from the redundancy rule that otherwise capitalizes on the lack of CaCiC forms. Such an "exception" morpheme feature in the lexicon should be more highly valued than the otherwise necessary morpheme feature in the MP rules, because it means only that a rule already existing does not apply.

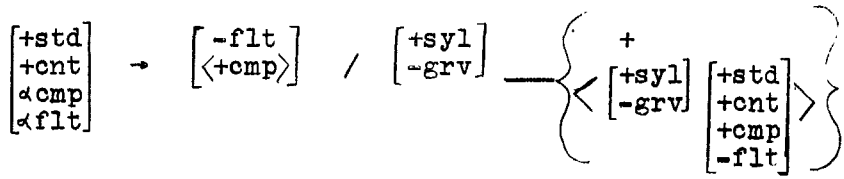
The other morpheme feature in the MP rules would necessitate the formulation of a rule which relies on that feature to specify an output, and is therefore incomparable with other formulations.

It is of interest to note that MP 13, like the redundancy rule (t → č / i__), was also a rule in early Panoan, though apparently not as early, for some of the languages (i.e., the Mayoruna group) which preserve the more archaic forms lack it. It is possible that those languages which separated early lost the rule, but the scattered distribution of six of those languages (Mayoruna Sauvage, Poynahua, Nokaman, Yambinawa, Chacobo 2, Atsahuaca) over the Panoan territory from northern Peru to Bolivia argues against it. It would seem rather that it is a later innovation. In any case, the rule is now at the point where only one recorded lexical entry prevents it from being entered in the lexical redundancy rules instead of in the MP rules. The accident that rabič is the only recorded word in which the (t → č) rule and MP 13 intersected in final position underlies the retention of 13 and preserves a case that shows how a rule can move from the MP component to the MS component.

rabĭ 'two' compared with wĭsti 'one':

base form	rabič	wĭstič
Rule 9	rabičan	wĭstičan
Rule 10	rabičĭn	wĭstičĭn
Rule 12	rabičin	wĭstičin
Rule 13	rabĭčin	wĭstičin

MP 14 Strident, flatness, and graveness shift



i.e., if an s or š precedes a morpheme border and is preceded by i, š becomes š (s is unchanged). If s or š is preceded and followed by i and š follows the second i, both s and š become š.

The expansion convention requires that the items in the braces be taken top line first. This gives the first change, which produces the allomorphic shapes of the third person marker and the non-restrictive relative clause closure. The second part of the rule effects a sibilant harmony similar to that in the MS rules.

The conjunctive ordering convention permits expansions from lower items in braces to apply to the output of expansions from higher items. In this way 14b, formed by taking the items in angles, applies to the output of 14a formed by taking the morpheme juncture.

base form	ka+ipi+š+ki	bič+a+š+ki	bič+aš	his+aš	nič+aš
rule 10		bičǐški	bičǐš	hisǐš	ničǐš
rule 12		bičiški	bičiš	hisiš	ničiš
rule 14a	kaipiški	bičiški	bičiš	hisiš	ničiš
rule 14b				hišiš	
	'he went'	'he took'	'having taken'	'having seen'	'having walked'

base form	pi+pi+pi+pi+č+aš+	kaič+aš	his+a	kais+aš
rule 10	pipipipičǐš	kaičǐš	hisǐ	kaisǐš
rule 12	pipipipičiš	kaičiš	hisi	kaisiš
rule 14a	pipipipičiš	kaičiš	hisi	kaisiš
rule 14b				kaišiš
	'having eaten'	'having multiplied'	'he saw'	'having chosen'

MP 15 $\begin{bmatrix}+\text{std}\\-\text{cnt}\end{bmatrix}\ \begin{Bmatrix}\#\\C\end{Bmatrix}$ → $\begin{bmatrix}?\\-\text{cns}\\-\text{syl}\\-\text{cnt}\end{bmatrix}$

1 2 1

i.e., the strident stop č is replaced by a glottal stop if it is final or followed by a consonant, liquid, or glide.

There are no morpheme-final occurrences of ø, so that it is unnecessary to specify č as [+cmp] because the cC cluster is formed only over morpheme junctures. The ordering of 15 before 16 is desirable in order to economize the specification of the stops in 16, where stridency no longer need be specified if rule 15 precedes 16.

bič- 'to take'	niš- 'to walk'
bi?makin 'caused to take'	ni?nawǐ 'walk (pl)'
bi?wǐ 'take it'	ni?wǐ 'walk (sg)'
bi?nawǐ 'take it (pl)'	ni?ki?nwǐ 'walk with him'
bi?ba?inkin 'took it away'	ni?tanwǐ 'walk over there'
bi?tanwǐ 'go take it'	ni?bonawǐ 'walk along'
bi? ta haiki 'he is taking it'	ni?yasa?nwǐ 'walk (fut)'
	ni? ta haiki 'he is walking'

MP 16 Consonant loss

$$\begin{bmatrix} +\text{obs} \\ -\text{cnt} \end{bmatrix}^{p,\ t,\ k} \rightarrow \text{null} \ / \ \_\_\_ \left\{ \begin{matrix} \# \\ [-\text{syl}] \end{matrix} \right\}$$

i.e., the first consonant is dropped from a cluster if it is a stop and the second is a consonant, liquid, or glide.

Rule 15 must precede 16 to avoid deletion of č in this environment. 16 must precede 17 to permit nasalization of the vowel preceding the consonant deleted by 16.

nanĭtba?ini	nanĭba?ini	'get right aboard'
nanĭtmawĭ	nanĭmawĭ	'make him get aboard'
hamakwĭ	hamawĭ	'step on it'
nanĭt?oš̆ki	nanĭ?oš̆ki	'he got aboard'
hamak?oš̆ki	hama?oš̆ki	'he stepped on it'
?i?sap#	?i?sa	'bird'
hamak?ona	hama?ona	'coming stepping'
?awap?on	?awa?on	'about a tapir'
hamakhakin	hamahakin	'to step down on it'

Continuant consonants are not affected:

hisis#	hisis	'ant'
waran#	waran	'squash'
waran+?on	waran?on	'about squash'
?ĭn+?on	?ĭn?on	'about me'

MP 17 Glottal gain

$$\text{null} \rightarrow \begin{bmatrix} -\text{cns} \\ -\text{voc} \\ -\text{cnt} \end{bmatrix}^{?} \ / \ \begin{matrix} \overset{b,\ r}{\begin{bmatrix} +\text{cns} \\ +\text{voc} \end{bmatrix}} \\ \_\_\_ \\ \overset{V}{[+\text{syl}]} + \_\_\_ \overset{V}{[+\text{syl}]} \end{matrix}$$

i.e., initial liquids at sentence juncture acquire a glottal onset. But after a juncture of lower rank the glottal onset does not appear. A glottal is also inserted between vowels at a morpheme juncture.

Initial:

?banawĭ 'plant'

?bimi banawĭ 'plant fruit'

?ra?mabi rĭrawĭ 'chop now'

Medial:

bana+i	bana?i 'planting'
pi+i	pi?i 'eating'
hoš̆oha+i	hoš̆oha?i 'making it white'
bana+ipi+ki	bana?ipiki 'planted it (yesterday)'
bana+a+?+ki	bana?a?ki 'planted it (today)'
bana?ona+i	bana?ona?i 'coming along planting'

The introduction of h (by the redundancy rules), and glottal here, has eliminated the need for ad hoc junctures or morpheme features in a number of morphemes. The surface structure requires that some morphemes have initial glottal stops that delete and others that do not.

a. non-deletable ?: remote past

bana+?oš̆ki	bana?oš̆ki 'he planted'
mapĭt+?oš̆ki	mapĭ?oš̆ki 'he ascended'
hamak+?oš̆ki	hama?oš̆ki 'he stepped'

b. deletable ?: ?a recent past

bana+?a+š̆+ki	bana?aš̆ki 'he planted'

mapït+?a+š+ki mapïtaški 'he ascended'

hamak+?a+š+ki hamakaški 'he stepped'

c. There are other cases where no glottal can appear, but consonant loss still occurs:

bïnï+ha+kin bïnïakin 'she married him'

sirip+ha+kin siriakin 'did it well'

hamak+ha+kin hamaakin 'stepped down'

All three cases are seen as regular if the "deletable" glottals are lacking in the base forms but inserted by the glottal insertion rule, which precedes the h-deletion rule. The third group has an initial h in the underlying form, which causes the consonant loss just as the glottal of group (a) does. Since the h deletion rule is needed independently to cover other cases of h in word-medial position (e.g., reduplication hošoošo 'very white'), the formulation of the glottal insertion rule obviates the need for an ad hoc glottal deletion rule to cover the (b) cases and special junctures to produce the forms in the (c) group.

MP 18 Nasalization

$$\begin{bmatrix} V,w,y,?,h \\ -cns \end{bmatrix} \rightarrow [+nas] \; / \; \_\_\_ \; [-cns]_0 \; [+nas]$$

i.e., a vowel or glide preceding a nasal consonant is nasalized if no consonant intervenes.

warãn 'squash'

põỹãn 'arm'

bãw̃ĩn 'catfish'

cĩ?ĩn 'by fire'

cipõnki 'downriver'

bõõn 'hair'

wĭrãnwĭ 'push it'

wĭrãnyasã̃ʔnwĭ 'push it sometime'

bĭmi 'fruit'

Rule 18 nasalizes glottal stops as well as vowels and continuant glides. This does not mean that there is a feature of nasality perceptible during the occlusion of the glottal stop, but indicates that the velum remains open during the glottal closure. Since the glottal closure automatically means a complete absence of acoustic output during the interval of the closure, the nasal feature finds no acoustic realization.

The rule would have to be complicated by the addition of several optional glottal stops and vowels to work properly if the nasalization of glottal stops were ruled out, to produce forms such as cĭʔĭʔõn 'about the fire.'

MP 19 Nasal loss

$$\begin{bmatrix} n, m \\ +nas \end{bmatrix} \rightarrow \text{null} \; / \left\{ \begin{array}{l} \_\_\_ \left\{ \begin{bmatrix} \# \\ +cns \\ +cnt \\ -nas \end{bmatrix} \right\} \\ \begin{matrix} C \\ [-syl] \end{matrix} \_\_\_ \begin{matrix} C \\ [-syl] \end{matrix} \end{array} \right\}$$

i.e., nasal consonants are deleted before a continuant strident consonant, glide, liquid, or word boundary. A nasal is also deleted in case it occurs as central member of a triple consonant cluster, of which the first is a glide (all other clusters have already been reduced). The nasalization on

the preceding vowel remains in all instances.

warã 'squash'

põỹã 'arm'

bãw̃ĩ 'catfish'

cĩ?ĩ 'by fire'

cipõnki 'downriver'

bõõ 'hair'

wĩrãyasã?wĩ 'push it sometime'

wĩrãwĩ 'push it'

Before n, m, p, k, t, c, ∅: no deletion

?õnãn+na+wĩ 'learn (pl)'

?õnãn+ma+wĩ 'teach'

?õnãn+pã 'I learn'

?õnãn+kĩ 'learned'

?õnãn+tã+wĩ 'go learn it'

kĩnča 'bowl'

kãn∅ĩ 'banana'

MP 20 Second nasalization rule

V,w,y,?,h

$$[-\text{cns}] \rightarrow [+\text{nas}] \;/\; [-\text{cns}] \begin{bmatrix} -\text{cns} \\ +\text{nas} \end{bmatrix}_0 \_\_$$

i.e., a vowel or glide is nasalized if it follows a non-consonantal element nasalized by the now-deleted nasal consonant.

This rule thus nasalizes in the opposite direction from that of rule 18, and the present ordering of 19 before 20 is essential to the ability of 20 to cover all the

possible cases. The nasalization spreads progressively from a nasalized vowel or glide over the next expanse of vowels or glides until a consonantal segment stops it. Rules 18, 19, and 20 account for the fact that regressive nasalization is conditioned only by the presence of a nasal consonant, and progressive nasalization occurs only when a nasal consonant has been deleted. No progressive nasalization can occur if the nasal consonant has not been immediately followed by a glide, and hence deleted, because rule 20 will not nasalize from a nasal consonant, nor over a segment with a [+cns] feature.

	'push it over'	'push it over sometime'	'make it red sometime'	'plant it'
underlying form	wĭranwĭ	wĭranyaʂa?nwĭ	hoš̆inhayaʂa?nwĭ	banawĭ
rule 17	wĭrãnwĭ	wĭrãnyaʂã?nwĭ	hošĭ̃nhayaʂã?nwĭ	bãnawĭ
rule 18	wĭrãwĭ	wĭrãyaʂã?wĭ	hošĭ̃hayaʂã?wĭ	bãnawĭ
rule 19	wĭrãw̃ĭ	wĭrãỹãʂã?w̃ĭ	hoš̆ĭ̃hãỹaʂã?w̃ĭ	bãnawĭ

MP 21 Nasal assimilation

$$\begin{bmatrix}+\text{cns}\\+\text{nas}\end{bmatrix} \rightarrow \begin{bmatrix}\alpha\,\text{cmp}\\\beta\,\text{grv}\end{bmatrix} \;/\; \_\_ \begin{bmatrix}+\text{cns}\\\alpha\,\text{cmp}\\\beta\,\text{grv}\end{bmatrix}$$

m, n → m,n, / ___ p,t,k,ø,ŏ,m,n

i.e., nasal consonants undergo homorganic assimilation to the following consonant.

After rule 19 there remain nasal consonants only before interrupted or nasal consonants, so that specification of the features $\begin{bmatrix}-\text{voc}\\-\text{cnt}\\+\text{nas}\end{bmatrix}$ is unnecessary for rule 21.

basic form	18-19-20	21	
čiponki	čipõnki	čipõ ki	'downriver'
poyan	põỹã	põỹã	'arm'
min katanti?in	mĩ katãntĩ?ĩ	mĩ katãntĩ?ĩ	'will you go away'
?onannawĩ	?õnãnnawĩ	?õnãnnawĩ	'learn (pl)'
?onanmawĩ	?õnãnmawĩ	?õnãmmawĩ	'teach'
?onanpan	?õnãnpã	?õnãmpã	'I will learn'
?onankin	?õnãnkĩ	?õnã kĩ	'learned'
kĩnčap	kĩnčap	kĩnčap	'bowl'
kanøin	kãnøĩ	kãnøĩ	'banana'

MP 22 Glide loss

?,w,y,h

$$\begin{bmatrix}-\text{ons}\\-\text{voc}\end{bmatrix} \rightarrow \text{null} \;/\; [+\text{std}]\ \_\_$$

i.e., a glide is dropped if it is the second member of a cluster.

Since nasals and stops have already been deleted preceding glides only the strident continuants remain there, and specification of more than [+std] is unnecessary.

his+wĩ	hisĩ	'see'
kais+wĩ	kaisĩ	'choose'
baš̆ĩš̆+wĩ	baš̆ĩš̆ĩ	'whisper'
hisis+ya	hisisa	'with an ant'
his+ya+šã?+wĩ	hisašã?w̃ĩ	'see sometime'
maraš+ya	maraša	'with poison'
?õnã+?õ+?+ki	?õnã?õ?ki	'I knew (last month)'

his+?o+?+ki hiso?ki 'I saw (last month).

MP 23 ?C cluster reduction

$$\begin{bmatrix}-\text{cns}\\-\text{voc}\\-\text{cnt}\end{bmatrix} \rightarrow \text{null} \;/\; \left\{\begin{matrix}[+\text{cns}]\\\#\end{matrix}\right\} \left(\left(\begin{bmatrix}-\text{cns}\\-\text{voc}\\-\text{cnt}\end{bmatrix}\right)\text{CV}\left(\begin{bmatrix}-\text{cns}\\-\text{voc}\\-\text{cnt}\end{bmatrix}\right)\text{CV}\right)_1 \_\_\; \text{C}$$

i.e., the cluster ?C is reduced to C in a word if the cluster closes an even numbered syllable, counting from the preceding closed syllable or from the beginning of the word. Glottal stop in this case does not count as a syllable closure.

Clitics and suffixes will not be separated from the preceding noun or verb by word boundaries, though in the examples below clitics are separated by a space to facilitate reading.

NP:

pĭ?i ta? ki 'it is a leaf'

pĭ?i ra? ta ki 'it is probably a leaf'

pĭ?i ya?pa ta? ki 'he has wings (leaf-with)'

pĭ?i ya?pa ra? ta ki 'he probably has wings'

pĭ?i ya?pa ma ra ta? ki 'he probably has no wings'

?očiti ta ki 'it is a dog'

?očiti ra ta? ki 'it is probably a dog'

?očiti ma ra? ta ki 'it is probably not a dog'

Some morphemes with ?C clusters do not reduce and are marked as exceptions, [-23].

?očiti ya?pa ta ki 'he has a dog'
[-23]

ʔočiti ya?pa ra ta? ki 'he probably has a dog'
[-23]

ʔočiti ya?pa ma ra? ta ki 'he probably does not have a dog'
[-23]

Verb:

rakati? ki 'he lies down'

pĭraka?ti ki 'he lies down on his back'

ča?čikin 'poked him'

pičačikin 'poked him in the ribs'

rakawĭ 'lie down'

pĭraka?wĭ 'lie down on your back'

MP 24 Flatting (optional)

$$\begin{bmatrix}\text{p, m}\\ +\text{obs}\\ -\text{cmp}\\ +\text{grv}\end{bmatrix} \rightarrow \begin{matrix}\text{p}^{\text{w}}\text{, m}^{\text{w}}\\ [+\text{flt}]\end{matrix} \;/\; \underline{\quad} \begin{bmatrix}\text{ĭ}\\ +\text{syl}\\ +\text{dif}\\ +\text{grv}\\ -\text{flt}\end{bmatrix}$$

i.e., p and m are labialized preceding the non-flat back vowel ĭ. It is a regular feature with some speakers, but absent with others or used sporadically.

kap$^{\text{w}}$ĭ 'alligator'

p$^{\text{w}}$ĭʔi 'leaf'

m$^{\text{w}}$ĭʔi 'bark'

ʔam$^{\text{w}}$ĭn 'capybara'

MP 25 h deletion

a. (obligatory) $\begin{bmatrix}\text{h}\\ -\text{cns}\\ -\text{voc}\\ +\text{cnt}\\ +\text{cmp}\end{bmatrix}$ → null / [ ] ___

b. (optional) [+syl] # $\begin{bmatrix}-\text{cns}\\-\text{voc}\\+\text{cnt}\\+\text{cmp}\end{bmatrix}$ [+syl]

1 2 3 4

SC 1: Cond: 4 = [+flt] 3,4 → 0

SC 2: Cond: 1 = 4 3 → 0

i.e., h is deleted in any word medial position. If the preceding word is vowel final and the vowel following h is [+flt], h and the vowel are both lost. Or if the vowel following h is identical to that of the vowel preceding h, h is lost.

a.	hoʂ̆ohakĭ?ti	hoʂ̆oakĭ?ti 'to turn white'
	?aibohoma	?aibooma 'without a wife'
	pĭ?ihoma	pĭ?ioma 'without leaves'
b.	hawa hanĭ ?in	hawa anĭ ?in (what name is) 'What is his name?'
	mapo hoʂ̆oni	mapo oʂ̆oni (head white) 'white-headed'
	hiwi hisĭ	hiwi isĭ (tree see) 'see the tree'
	hisi howĭ	hisi wĭ 'come and see'
	pi?i howĭ	pi?i wĭ 'come and eat'
	hawa honi ?in	hawa ni ?in 'what man is it?'

MP 29 s becomes h (optional)

a. Many speakers aspirate a sibilant when it precedes r, especially in fast speech.

s,š,ṣ		h		r
[+std]	→	-cns -syl +cnt +cmp	/ ___	+cns -syl -grv

ba?kihri?bi 'tomorrow, too'

hisihri?bi 'ant also'

marahri?bi 'poison also'

b. Other speakers drop the sibilants before r, optionally.

s,š,ṣ				r
[+std]	→	null	/ ___	[+cns -syl -grv]

ba?kiri?bi 'tomorrow also'

hisiri?bi 'ant also'

marari?bi 'poison also'

VII. STRESS AND PITCH RULES

Two cultural facts support the analysis of tone as phonemically distinctive in Capanahua on a taxonomic phonemic level. First, limited conversations are carried on in whistle-speech, each participant of the conversation whistling the tone contours of the related segmental strings. This is much more limited and infrequent than that reported for Mazatec by Cowan (1948), but I have observed it in use.

Second, there is a strong tradition for the use of tonal signal drums by past generations, though none currently exist. Legends and personal narratives abound with references to the drums and their use, and some of the older present-day speakers claim to remember participating in the fabrication of the drums. Their descriptions of them indicate that the drums very closely resembled the hollow-log drums still used by the Bora tribe in northeastern Peru. Each drum was constructed to produce two tones, and they were used singly or in pairs to communicate with neighboring villages, sending challenges, invitations to work parties, etc.

These facts do not prove tone to be phonemically distinctive, nor do they necessarily indicate that tone, not

stress, is the significant feature distinguishing morphemes. But it is a fact that tonal signal drums are generally regarded as a characteristic only of tone languages.

In a taxonomic phonemic approach such as that of Pike's Tone Languages frames are set up in which items can be tested, first for identification of contrastive pitch, and second for classification of the pitches so identified to assign them phonemic status.

Using this technique, Capanahua is found to have three phonetic levels of pitch: high, mid, and low, of which mid is assigned as a variant of the high toneme in final position. Other perturbations between high and low are then considered to be substitutions of one toneme for another in given contexts. But such an approach gives no means of determining decisively for a language like Capanahua whether it is tone or stress that is distinctive.

Stress and tone coincide in citation forms so that one is largely predictable from the other. If stress is regarded as basic, a statement can be made for nouns that predicts the occurrence of high tone on stressed syllables and on non-final post-stressed syllables:

'nía 'trumpeter bird'
nĩ'á 'this'
'sóntáko 'young girl'
yo'sánbo 'old lady'

In verbs the rule does not apply, for final unstressed syllables do not always have low tone:

'nánĭkin 'he put it in the canoe'

'nánĭwĭ 'put it in the canoe'

As seen in the above example, tone in final position does not coincide with a stress.

By taking the opposite approach, it is possible to conclude that stress is predictable on the basis of tone: the first high tone of a word is stressed. This would explain most of the stress-tone patterns of the following, for which examples abound.

'nánĭwĭ 'put it in the canoe'

na'nĭwĭ 'get in the canoe'

'bánĭti ta 'aiki 'it is necessary to leave it'

ba'nĭti ta 'aiki 'it is necessary to stop here'

'?ónámati ta 'pĭšíni?kĭ 'one should cause her to sew'

?o'námáti ta 'pĭsíni?ki 'one should cause her to learn'

There are other frames, however, where stress and tone are not predictable on the basis of one another. Sometimes only stress is distinctive, as in frames where the tones have been perturbed by the context:

ha'wĭn 'mápo 'his head'

ha'wĭn má'po 'his clay'

and in other frames only tone is distinctive, with stress not predictable from the tone. This is especially true in frames involving monosyllabic forms.

?ĭn 'hóno 'rĭ?tĭkin 'he killed my wild pig'

?ĭn 'hóno 'rĭ?tĭkin 'I killed a wild pig'

In a taxonomic phonemic analysis, stress and pitch are not usually both regarded as phonemic in any one language. Stress and pitch as suprasegmentals are abstracted from the segmental strings, and if one is predictable, the other is made phonemic. In the Capanahua case, neither stress nor tone is fully predictable, but an analysis is permitted in which tone is given phonemic status and stress is relegated to the morphology. In this treatment class I words have stress on the ultima, class II words have stress on the penult, and class III words on the antepenult. This is in fact how some Panoan languages have been treated, including my own earlier version of Capanahua phonemics.

On a systematic phonemic level, however, both pitch and stress are predictable, providing a considerable simplification of the grammar and at the same time a solution to the enigma of how to decide on some principled basis whether tone or stress is the significant feature.

First, basic stress and pitch patterns are assigned to the underlying forms on the basis of their segmental, and in some cases grammatical, features. Then, after the segment deletion rules have operated on the underlying forms, the cyclical pitch rules assign pitch contours to the bracketed strings. In this way the close relation between pitch and stress in citation forms is explained, as well as the seeming independence of pitch from stress by the later pitch rules.

For convenience, the basic stress and pitch assigning

rules will be referred to as stress rules, and the later group of pitch affecting rules as pitch rules. The stress rules assign only an initial "+" or "-" pitch feature and later rules will assign a phonetic value of high to "+" and low to "-". In the examples, [+stress] will be marked 'CV, [+pitch] will be marked V́, and the [-] values of both stress and pitch will be left unmarked.

Stress rule 1

$$[+syl] \rightarrow \begin{bmatrix} \alpha\,strs \\ \alpha\,pit \end{bmatrix} / \alpha \left\langle \# \; [-syl] \; ([+syl] \; [-syl])_{o}^{1} \; \_\_ \right\rangle$$

i.e., all nouns, verbs, and adjectives receive a stress and pitch on the first two syllables. All other syllables are at this point [-stress] and [-pitch].

Stress Rule 1 provides the proper tone contour for words that do not have a closed second syllable.

'ṣ́ón'táko 'young girl'

'čí'číka 'knife'

'má'pó 'head'

However, the stress on the second syllable must be removed in these case, and words with a closed second syllable must be modified so that the initial syllable has no pitch or stress.

Stress Rule 2

$$[+syl] \rightarrow \left\{ \begin{bmatrix} -strs \\ -pit \end{bmatrix} / \#[-syl] \; \_\_ \left\{ \begin{matrix} \# \\ [-syl] \end{matrix} \begin{bmatrix} +syl \\ +strs \\ +pit \end{bmatrix} [+obs] \left\{ \begin{matrix} \# \\ [-syl] \end{matrix} \right\} \right\} \right.$$

$$\left. [-strs] / \begin{bmatrix} +strs \\ +syl \end{bmatrix} \; [-syl]_{o} \_\_ \right\}$$

i.e., the initial syllabic of words which are monosyllabic or which have a closed second syllable becomes unstressed and "minus" for pitch, and any syllable which is preceded by a stressed syllabic becomes unstressed.

All nouns, verbs, and adjectives will now have the proper base form, but this base form is not always the same as that shown in citation forms. The latter are formed by later rules.

Most nouns, verbs, and adjectives are bisyllabic. Those which have a closed second syllable will have a contour in which stress and tone coincide on the second syllable.

ma'póp 'clay'
hi'sís 'ant'
wa'rán 'squash'
ra'bíč 'two'
ba'kóš 'foam'

All bisyllabic words which are stressed on the first syllable will carry a level pitch contour:

'mápó 'head'
'písí 'rotten'
'bárá 'bullet'
'hónó 'wild pig'
'híwí 'tree'

Trisyllabic words receive the same contour on the first two syllables as the corresponding group above, but the third syllable will have no stress or high pitch.

The stress-assigning rules cannot be incorporated into the MS component because they must follow the morpheme lengthener rules (MP 6, MP 7, MP 9). Words that have been lengthened will carry a different contour than the corresponding short form, for the lengthener rules change the segmental shape of the nouns. For comparison with the bisyllabic forms above, the same entries will be shown with their contours on their lengthened forms, following the examples of underlying forms of tri-syllabic morphemes.

Basic trisyllabic forms:

'ṣ́óntáko	'girl'	'bĭ́mánan	'face'
'?óčíti	'dog'	'ɸá?óri	'a small variety of bird'
'číčíka	'knife'	'mááro	'to buy'

Lengthened bisyllabic forms:

'mápópan	'clay'	ma'pón	'head'
'hísísin	'ant'	pi'sín	'rotten'
'wáráman	'squash'	ba'rán	'bullet'
'rábĭ́čin	'two'	ho'nón	'wild pig'
'bakosĭ́n	'foam'	hi'wín	'tree'

The stress rules must precede the consonant deletion rules, for the deletion rules will remove the segments needed by the stress rules to assign the proper stress-pitch contours, e.g., the underlying form mapop 'clay' would become by the deletion rules mapo to which the stress rules would assign the contour 'mápó erasing the distinction between 'clay' ma'pó and 'head' 'mápó.

Thus far then it is seen that the pitch-stress contrast between many minimal pairs of nouns in the surface structure is explained as a difference in the underlying forms upon which the stress rules operate and not a difference of inherent pitch or stress in the (taxonomic) phonemic forms.

The syllable-final consonants of underlying forms are needed independently of the stress rules, for they are the basis for the proper functioning of the lengthener rules. The stress rules therefore reduce the complexity of the grammar considerably, since no stress or pitch features need be marked in the lexicon, and the pitch rules will be needed in any case to shift (according to syntactic factors) the pitches initially assigned.

The stress rules are not morpheme level rules, hence, cannot be part of the MS component, for morpheme junctures do not affect them.

'mápĭt+i (mapĭt 'to climb up') 'he climbed up'

pi'+kín (pi 'to eat') 'he ate it'

wĭ'rán+kin (wĭran 'to push') 'he pushed it'

'nĭṣá+tan+kin (nĭsa 'to tie', tan 'over a distance') 'went and tied it'

nĭ'ṣĭ́+t+tan+i (nĭṣa 'to tie'; t 'reflexive'; tan 'over a distance') 'went and entangled himself'

Monosyllabic words remain without stress or high pitch. These are pronouns and clitics:

?ĭn	'I'	man	'you (pl)'
min	'you'	ta?	'decl Md'

non 'we' ra? 'perhaps'

han 'he' rĭs 'just'

Words with the second syllable closed by a glottal stop do not undergo rule 2a, either because the glottal deletion rule precedes 2a, or because the glottal does not function as syllable closure:

raka?ti 'he lies down'

'ráká?ti (stress rule 1 and 2)

'rákáti (glottal deletion rule)

ča?čikin 'to spear' (note glottal)

pičа?čikin 'to poke him in the ribs'

'pícá?čikin (stress rule)

'píčáčikin (glottal deletion rule)

The glottal deletion rule (MP 23) deletes a glottal which has been produced by a syllable-final č (MP 15), but the č is needed by the stress rules to put the proper contour on words where it has been deleted:

ka+rič+i (go-soon-pres) 'he goes soon'

'káríči (stress rule 1 and 2)

ka+rič+wĭ (go-soon-impv) 'go soon'

ka'ríčwĭ (stress rules 1, 2, and 3)

ka'rí?wĭ (č → ? rule)

ka'ríwĭ (glottal deletion rule)

The stress rules must therefore precede the rule that changes č to ?, but it is not yet clear that the glottal deletion rule must precede the pitch rules. Pitch rule 1

requires a closed syllable, but crucial data to show that the glottal deletion rule must follow it is lacking. In the case of the stress rules, the facts that the ɤ → ? rule must follow the stress rules, and the glottal deletion rule must follow the ɤ → ? rule, show conclusively that the glottal deletion rule must follow the stress rules.

Stress rule 3

Among the verb suffixes are a few that have been developed in the transformational component from syntactic features. These now also receive a plus pitch feature.

$$\begin{bmatrix} +\text{syl} \\ \begin{Bmatrix} +\text{B} \\ +\text{Imp} \\ +\text{Decl} \end{Bmatrix} \end{bmatrix} \rightarrow [+\text{pit}] \; / \; \_\_ \; \left\{ [\neq -\text{syl}]_0 \; [+\text{Decl}] \right\}$$

i.e., a vowel is given high pitch if it precedes the terminal declarative mood marker or the sentence juncture and carries one of the "B", imperative, or declarative features.

"B" is an arbitrary designation for a feature carried by some of the tense suffixes and is used by the late T rules (T 41, T 47) to put the tense suffixes in proper sequence with respect to the plural marker. It therefore has independent justification and is now used to assign a high pitch to just those suffixes for which the segmental shape does not provide a basis for assigning pitch.

+B suffixes: ipi past

ɤa? future

ɤi? future

'ʔóšáʔipíkí 'I slept'
[+B]

'ʔóšášíʔkí 'I will sleep'
[+B]

'ʔóšášáʔnwĭ 'go to sleep (sometime soon)'
[+B]

The imperative morpheme wĭ always carries high tone, as does the declarative marker.

'ʔóšáwĭ 'sleep'
[+Imp]

'ʔóšáʔikí 'he sleeps'
[+Decl]

'ʔóšáʔí 'I sleep'
[+Decl]

Pitch rule 1 (word level)

$$[+syl] \rightarrow [+pit] \;/\; [\begin{bmatrix} +syl \\ +strs \\ +pit \end{bmatrix} [-syl]\ [-syl]\ \underline{\quad}\ ]_{N,A,Vb}$$

i.e., a vowel is given high pitch if the preceding syllabic is stressed, has high pitch, and is in a closed syllable.

Before rule 1	After rule 1	
[ko?'kóštobi]$_N$	[ko?'kóštóbi]$_N$	'butterfly'
[bo'tánnawĭ]$_{Vb}$	[bo'tánnáwĭ]$_{Vb}$	'take it away'
[ti'nintinin]$_N$	[ti'nintínin]$_N$	'governonr'

Pitch rule 1 must follow the consonant deletion rule (MP 18). The deletion rule will resolve many consonant clusters that would otherwise cause this rule to operate incorrectly by putting a high tone on the syllabic following

a closed syllable that the deletion rule will make into an open syllable.

mapɨ̌t+ni+ʂ̌+ki (climb up-past-3rd pers-Decl)
'he climbed up'

ma'pɨ́tniʂ̌kí stress rules

*ma'pɨ́tníʂ̌kí by pitch rule 1 (ill formed)

*ma'pɨ́níǰ̌kí deletion rules

correct form:

ma'pɨ́niǰ̌kí by deletion rule operating before pitch rule 1.

The deletion rule, needed independently of the stress and pitch rules, must therefore follow the stress rules and precede the pitch rules. The ǰ̌ → ? rule must also precede the pitch rules, so the pitch rules cannot form a continuous sequence with the stress rules (i.e., they can not immediately follow the stress rules).

Pitch rule 2 (word level)

$$\begin{bmatrix}+\text{syl}\\+\text{pit}\end{bmatrix} \rightarrow [-\text{pit}] \;/\; \_\_\_ \begin{bmatrix}-\text{cns}\\-\text{voc}\\-\text{cnt}\end{bmatrix} \begin{bmatrix}+\text{syl}\\-\text{pit}\end{bmatrix}$$

i.e., a syllabic with minus stress and high pitch becomes low in pitch if it is followed by a glottal stop and a syllabic with low pitch. This rule must be a word level rule because it will not operate across word boundaries.

'hónó # ?i'nón # 'kámákin 'the tiger chased the wild pig'

'čáʂ̌ó # '?á?wɨ́ 'shoot the deer'

'ʂ̌óbó # ?o'nánkin 'he knew the house'

ʂóbó # ʔokĭ 'beyond the house'

Within word boundaries the rule operates across morpheme junctures.

'ʂóbo?o 'in the house'

'ʂóbo?on 'about the house'

'bákĭ?on 'about the child'

'bána?a? 'which he planted'

'rĭra?a? 'which he chopped'

'wína?a? 'which he paddled'

'píba?inwĭ 'eat and leave'

'káka?inwĭ go quickly'

If the syllable is stressed, the tone does not drop.

ma'pó?o 'in the mud'

ma'pĭ?o?ki 'ascended'

bo'tó?o?ki 'descended'

rĭ'rá?o?ki 'chopped ycurself'

Pitch rule 3a

$$\begin{bmatrix}+\text{syl}\\-\text{strs}\end{bmatrix} \rightarrow [-\text{pit}] \;/\; \_\_ \left\{ \begin{matrix} [+\text{sk}] \\ \left(\begin{bmatrix}-\text{cns}\\-\text{syl}\\-\text{cnt}\end{bmatrix}\right) \end{matrix} \begin{bmatrix}+\text{syl}\\-\text{past}\\-\text{fut}\\ \left\{\begin{matrix}+\text{Decl}\\+\text{Q}\end{matrix}\right\}\end{bmatrix} \right\}$$

i.e., an unstressed syllabic preceding the morpheme ska 'already, yet, next' or preceding the interrogative marker or the first and second person present tense declarative morpheme (with optional ? intervening) is assigned low pitch.

The morpheme ska is irregular. It is the only morpheme in the language that has an initial consonant

cluster, and it alone requires a rule that it always carries a pitch identical to that of the following syllable. ska is marked with the arbitrary designation [+sk].

'bána+ska+wĭ 'bánaskawĭ 'plant, then'

'báná+rĭs+ska+wĭ 'bánárĭsskawĭ 'plant, now then'

The present tense morpheme carrying the features of first and second person will also always carry the feature [+Decl], assigned by the T rules that spell out the shape of the person, plural, tense, and mood features. This morpheme is always preceded by a low-pitch vowel if that vowel has no stress.

'bána?í 'I plant'
[+Decl]

'mápĭtaí 'I climb up'
[+Decl]

'rĭra?í 'I chop'
[+Decl]

'ǧĭ?a?í 'I drink'
[+Decl]

'káísaí 'I choose'
[+Decl]

Pitch rule 3b

$$\begin{bmatrix}+\text{syl}\\+\text{sk}\end{bmatrix} \rightarrow [\alpha\text{pit}] \;/\; \_\_ \; +C_0 \; \begin{bmatrix}+\text{syl}\\ \alpha\text{pit}\end{bmatrix}$$

i.e., ska always takes the same pitch as the following syllable.

bánaskáwĭ 'well, plant'

bánaskatanwĭ 'go and plant'

karĭsskáwĭ 'go on and leave'

kámárĭsĭ́ 'just chase him'

kámárĭsskáwĭ́ 'just chase him, then'

pirĭ́sskáwĭ́ 'go ahead and eat'

Pitch Rule 4 (word level)

$$[+\text{syl}] \rightarrow [-\text{pit}] \;/\; [\; X \_\_ (C\acute{V})_{o} C\acute{V}C\acute{V}C\acute{V}_{Vb}\;] \qquad C = [-\text{syl}]$$

i.e., all syllabics are [-pit] if they are followed in the verb by three consecutive high pitch syllabics.

The verb consisting of verb stem, aspectual and adverbial suffixes, person-marking and tense suffixes, and closure suffixes has a limit of three consecutive high pitches. If the verb is expanded by a series of adverbial or aspectual suffixes, it will not usually have more than three consecutive pitches because the adverbs and aspects (apart from ska) do not receive a high pitch. But if a monosyllabic stem is chosen, the stress rules will include one syllable of any affix that follows, and the maximum would then be exceeded if no low-pitch syllables intervene between the stem and the closure affixes.

	ka+no+šą?n+wĭ (go-soon-fut-imp) 'go soon'
Stress 1	'ká'nóšą?nwĭ
Stress 2	'kánóšą?nwĭ
Stress 3	'kánóšą́?nwĭ́
Pitch 4	'kanóšą́?nwĭ́

pi+ya+šaʔn+wĭ (eat-fut-fut-imp) 'eat sometime later'

Stress 1 'pí'yáša?nwĭ

Stress 2 'píyáša?nwĭ

Stress 3 'píyášá?nwĭ

Pitch 4 'piyášá?nwĭ

Pitch rule 5

$$[+\text{syl}] \rightarrow [+\text{pit}] \;/\; [\,[\,(CV)_o\; C_o \begin{cases} VC_o \,\#\, [\; C\_\_X\;] & \langle NP,V,A\rangle \\ \_\_C_o\;]\,\#\,C_o\,V & \langle NP,N,A\rangle \end{cases} X\,]_{Vb\langle NP\rangle}$$

i.e., (a) high pitch is assigned to a syllabic if it is the first syllable of the second word of a verb, or if it is the first syllable of a noun, embedded noun phrase, or adjective, which is the second word of an NP.
(b) high pitch is assigned to a syllabic if it is the final syllabic of a verb stem followed by a word juncture in a Vb, or if it is the final syllabic of a NP, N, or A followed by another word in a NP.

In a noun phrase, verbs do not take the same contours as nouns, although the compound verb stem formed by stem reduplication has the same contour as a noun-modifier combination in a NP. Nominalized verbs, which by T 11 and T 12 lose their Vb bracketing but retain NP brackets, receive the contours of nouns. Thus rule 5 will assign a high contour over the juncture of adjective-noun, adjective-adjective, noun-noun, noun-nominalized verb, adjective-nominalized verb, and verb stem-verb stem, but not to

noun-verb or vice versa. The need to distinguish verb stems from nouns and adjectives requires that the labeled bracketing shall not yet have been erased.

Reduplicated verb stems will be separated by a word juncture and will be in Vb brackets. No other verb stems are affected by the rule.

[ɸa?ot] # [ɸa?ot] i sit-sitting 'to sit a long while'
[ɸa'?ó] # ['ɸá?ót] i

[yono] # [yono] kin command-commanding
['yóno] # ['yóno] kin 'to command repeatedly'

[wina] # [wina] kin 'addle-paddling
['wíná] # ['wíná] kin 'to paddle a long while'

[?onan] # [?onan] kin know-kowing
[?o'nán] # [?ó'nán] kin 'to know thoroughly'

[rakat] # [rakat] wĭ lie-lies
[ra'ká] # [rá'káwĭ] 'lie down all the time'

Noun-adjective sequences in a NP:

[ʂontako] # [piškap] girl-small 'young girl'
N A

['ʂóntáko] # [piš'káp] stress rules
['ʂóntákó] # [píšká] rule 5

[čičika] # [siɾip] knife-good 'a good knife'
['číčíka] # [si'ɾíp] stress rules
['číčíká] # [sí'ɾí] rule 5

[siriɨp] # [ʔaibo] good-woman 'a good woman'
A N

[si'rɨ́p] # ['ʔáíbo] stress rules

[si'rí] # ['ʔáíbo] rule 5 (vacuous)

Noun-noun sequences:

[hawɨ̄n] # [ʔawap] he-tapir 'his tapir'
N N

[ha'wɨ́n] # [ʔa'wáp] stress rules

[ha'wɨ́n] # [ʔá'wá] rule 5

[ʔino] # [mɨ̄nǿis] tiger-claw 'a tiger claw'

['ʔínó] # [mɨ̄n'ǿís] stress rules

['ʔínó] # [mɨ́n'ǿís] rule 5

Noun-nominalized Vb in NP:

[ha'wɨ́n] # ['nónti] his-floater 'his canoe'
N NP

Verb stems not nominalized in NP do not undergo the rule:

[[kiši] # [tobitka ʔinai] [honi]] leg break suddenly man
N Vb N NP
'the man who suddenly broke his leg'

[['kíší] # [to'bítka ʔinai] ['hóní]] stress rules

[['kíší] # [to'bíka ʔinai] ['hóní]] rule 5 (vacuous)

[[mapɨ̄tai] # [honi]] he goes up-man 'the man who ascends'
Vb N NP

[['mápɨ́tai] # ['hóní]] stress rule

[['mápɨ́tai] # ['hóní]] rule 5 (vacuous)

The effect of rule 5 is to put a high contour over the word boundary between the first and second words of a phrase, except when one of the words is a verb, and over the

internal word boundary of a verb, produced when the verb stem has been reduplicated.

Words with a closed second syllable will be affected by the first part of rule 5, since they will otherwise have a low pitch on the first syllable and words of more than two syllables will be affected by the second part, for the final syllables would in such cases be low in pitch. Bisyllabic words with an open final syllable will not be changed by 5, since they already have high pitch on both syllables.

Pitch rule 6

$$\begin{bmatrix} -\text{syl} \\ \langle -\text{Vb af}\rangle_1 \\ \langle -\text{Vb}\rangle_2 \end{bmatrix} \rightarrow [-\text{pit}] \; /$$

$$\left[\left\{ \begin{array}{c} X\,[X]_{Vb}\,(X)_o \\ C_o\ V\ C_o\ V\ C_o\ V\ (C_o\ V)_o \\ C_o\ V_o^2\ C_o\ V\ C_o\ \#\ CV\ (C_oV)_o \end{array} \right\} \_\_ (C_o\ V\ C_o)_o \begin{bmatrix} \\ \langle -\text{pit}\rangle_2 \end{bmatrix} \langle Vb\rangle_1 \langle NP\rangle_2 \right.$$

V represents a syllabic unmarked for pitch.

i.e., if the bracketing of the domain of the rule is labeled Vb, the pitch of a syllabic is lowered in a verb stem if:

1) it is preceded by a verb within the brackets (vacuous)

2) it is preceded by three or more successive high pitches within the brackets

3) it is preceded by the first syllable of the second word within the brackets.

If the bracketing is a NP, the same expansion is followed by a high pitch. Since nominalized verbs no longer

carry Vb bracketings, they will undergo the rule in the same way as nouns.

As in rule 5, only reduplicated verb stems will be affected by this rule: non-reduplicated verb stems will not have an internal # juncture and will not exceed three syllables.

ɸa'ʔó # 'ɸáʔoti 'to sit a long while'

'yónó # 'yónokin 'to command repeatedly'

'wínáˊ # 'wínakin 'to paddle a long while'

ʔo'nán # ʔó'nankin 'to know thoroughly'

ra'ká # rá'katai 'to lie down all the time'

'ṣóntákó # pis'ka 'young girl'

'číčíká # si'ri 'good knife'

si'rí # 'ʔáibo 'good woman'

ha'wĩn # ʔá'wa 'his tapir'

'ʔínó # mĩn'ɸis 'tiger claw'

Stress and pitch no longer coincide in many cases, for pitch has been shifted but stress has not. The second line of the environment of the rule amounts to the mirror image of rule 4 which applies only to verbs. Rule 6 applies to verb stems only if there has been stem reduplication but does not apply to post-stem suffixes. In verbs, therefore, the post-stem suffixes will remain with high pitch.

ʔoá kátánipí that-went 'the one who went'

But in noun phrases formed by a string of nouns or adjectives, all the syllables following the first syllable of the second word will be lowered.

hawĭn # ʔočiti # sirip # piškap 'his nice little dog'

ha'wĭn # 'ʔóčiti # si'rip # piš'kap stress rules

ha'wĭn # 'ʔociti # si'ri # pis'ka stress rule 6

Following a verb, all highs are also reduced to low.

'mápĭtai 'honi ascend-man 'the man who ascends'

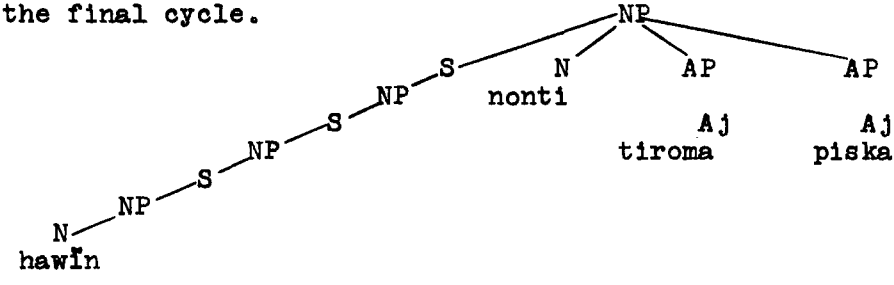
'mápĭtipí 'honi ascended-man 'the man who ascended'

mín 'bákĭ 'wĭsaʔipí 'ʔino your-child-clawed-tiger 'the tiger that clawed your child'

ʔi'nón 'píʔipí 'ʔočiti si'ri piš'ka tiger-bit-dog-good-small 'the nice little dog that the tiger bit'

It is the restrictive relative clause that requires the pitch rules to be cyclical. Modifiers that have no verb constituent produce the same result by cycling or by a single application of the rules to the largest expansion defined by rules 5 and 6.

The tree structure for hawĭn nonti tiroma piška (his canoe worthless small) 'his small worthless canoe' is cycled up through the <u>hawĭn</u> branch, then the adjectives, and finally the total string. The contour produced will be the same as when the total string is cycled only once since there are no pairs of words to apply the rules to until the final cycle.



But NP's with embedded relative clauses cannot be accounted for without a cyclical application of the rules. If the rules apply to the following example simply as a linear string, an incorrect contour will result:
'the little snake that the man who climbed up stepped on with his foot'

	climbed	man	his	foot	stepped
	mapĭtipi #	honin #	hawĭn #	ta?ĭn #	hamakipi
strs 1	ˈ/ˌˈ/	ˈ/ˌˈ/	ˈ/ˈ/	ˈ/ˌˈ/	ˈ/ˈ/
strs 2	ˈ/ /	ˈ/	ˈ/	ˈ/	ˈ/ /
strs 3	ˈ/ / /	ˈ/	ˈ/	ˈ/	ˈ/ / /
pit 5	ˈ/ / /	/ ˈ/	ˈ/	ˈ/	ˈ/ / /
pit 6	ˈ/ / /	/ ˈ/	ˈ/	ˈ/	ˈ/ /

	snake	small
	rono #	piškap
strs 1	ˈ/ˈ/	ˈ/ ˈ/
strs 2	ˈ/ /	ˈ/
strs 3	ˈ/ /	ˈ/
pit 5	ˈ/ /	ˈ/
pit 6	ˈ	ˈ

Applied cyclically, the rules will apply to mapĭtipi honin 'the man who climbed up' independently of the cycle on hawĭn ta?ĭn 'his foot'; and to the verb hamakipi 'stepped on' independently of the two mentioned NP's.

[[[mapĭt] ipi # [honin]] # [[hawĭn] #[ta?ĭn]] #
S NP Vb N NP N N NP

strs 1 ˈ/ ˈ/ ˈ/ ˈ/ ˈ/ ˈ/ ˈ/ ˈ/

strs 2 ˈ/ / ˈ/ ˈ/ ˈ/

strs 3 ˈ/ / / ˈ/ ˈ/ ˈ/

pit 5 ˈ/ / / / ˈ/ ˈ/ / ˈ/

pit 6 ˈ/ / / ˈ ˈ/ / ˈ

[hamak] ipi] #
Vb S

strs 1 ˈ/ ˈ/

strs 2 ˈ/ /

strs 3 ˈ/ / /

pit 5 ˈ/ / /

pit 6 ˈ/ / /

Then the AP and head noun are cycled (only strs 1 and 2 apply).

[rono]#[piškap]
A

strs 1 ˈ/ ˈ/ ˈ/ ˈ/

strs 2 ˈ/ / ˈ/

Finally, with the whole NP as domain, the pitch rules cycle again. (From this it is evident that only the pitch rules require cyclical application: the stress rules do not require labeled brackets). In this case only pitch rule 6 applies:

```
[  [  [ [mapɨt][ipi] # [honin]  ] # [ [hawɨn] # [ta?ɨn]  ] #
NP S        Vb   tns       N NP           N          N  NP

        1/ /    /       1              1/       /1

pit 6   1/ /    /       1              1/       /1
```

```
[hamak] [ipi]  ]#[rono]#[pišKap]  ]
     Vb   tns S      N       A NP
 1/ /     /      1/ /     1/

 1/ /     /      1        1
```

Final cycle pitch rules

a. $\begin{bmatrix}+\text{syl}\\+\text{pit}\end{bmatrix} \rightarrow [\ 3\ \text{pit}]$

b. $\begin{bmatrix}+\text{syl}\\-\text{pit}\end{bmatrix} \rightarrow [\ 1\ \text{pit}]$

i.e., the abstract features of "+" or "-" pitch are now assigned a relative phonetic value, #1 being low pitch and #3 high. The strings are now ready for the rules specifying the mid tone.

Pitch rule 8

$$\begin{bmatrix}+\text{syl}\\+\ 3\ \text{pit}\end{bmatrix} \rightarrow [\ 2\ \text{pit}] \ /\ \begin{bmatrix}+\text{syl}\\1\text{pit}\end{bmatrix} C_0 \_\_\_ \#$$

i.e., a sentence-final high tone becomes mid if it is preceded by a low tone

```
'haa ta 'mapo ki  'that is a head'
  33  1   3 1  2

'mapo ?in  'is it a head?'
  3 1  2

'haa ta ma'po ki  'that is clay'
  33  1  1  3  3

ma'po ?in  'is it clay?'
 1  3  3
```

'nĭa ?in 'is it a trumpeter bird?'
 31 2

nĭ'a ?in 'is it this?'
 1 3 3

ʂ̌o'o ?in 'is it wind?'
 1 3 3

'ʂ̌oo ?in 'is it green?'
 31 2

ka'tanwĭ 'go on'
 1 3 3

'banawĭ 'plant it'
 3 3 3

'banatanwĭ 'go plant it'
 3 3 1 2

Pitch rule 9

$$\begin{bmatrix} +\text{syl} \\ -\text{strs} \\ +\text{N} \end{bmatrix} \rightarrow [\ 1\ \text{pit}] \ / \ \underline{\quad} \neq$$

i.e., a final unstressed vowel of a noun is low in pitch if it occurs sentence final. This will also cover citation forms.

[?o'a ta? ki, 'honi] 'there he is, the man'
 1 3 1 3 3 1 S

ma'po 'clay'
 1 3

'mapo 'head'
 3 1

nĭ'a 'this'
 1 3

'nĭa 'trumpeter bird'
 31

'ʂ̌ontako 'girl'
 3 3 1

Rule 10

$$[-\text{cns}] \rightarrow [+\text{nas}] \;/\; [\; [+\text{Irr}]\; X \;\underline{\quad}\;]_{S}$$

i.e., all syllables in a sentence are nasalized if a [+irritation] feature has been assigned to the first part of the sentence.

As the stress and pitch rules indicate, stress is fixed by the segmental shape of the underlying forms, and pitch is determined by the segmental shape of the underlying forms, later modification of those forms as they are shaped by the morphophonemic rules, and by the syntactic bracketing and cyclical rules, but not by "intonational" meanings.

Interrogative, exclamatory, imperative, emphatic, and doubt meanings in Capanahua are represented in the base component by features that are spelled out as segments (morphemes) by the T rules.

Anger and irritation are not given morphemic shape in the string, but are expressed as rule 10 indicates by nasalization of the whole sentence.

The point at which the feature for a rule such as 10 should be introduced is an open question, and there is no compelling reason why 10 must follow the pitch rules, but it should follow any rules that introduce segments since the nasalization would not otherwise affect a vowel inserted between non-nasal obstruents and would therefore leave an oral vowel in the string.

The cyclical pitch rules must precede the nasal deletion rule MP 19, for the nasal-closed syllable is needed by pitch rule 1. They must follow the consonant deletion rules, as has already been observed, and the glottal gain rule MP 17. They could come before or after MP 18 without any difference in output. Since MP 18 and MP 19 both deal with the nasals, the cyclical rules are tentatively located just following MP 17.

VIII. SUMMARY

The problem of evaluating alternate analyses involves the assumption that if each analysis is formulated as rigorously as possible for minimization of feature specification, one will capture generalizations that the other will not. The one which is most effective in capturing linguistically significant generalizations will be the correct theory. Within any one theory of a language the simplicity metric should lead to the selection of the most valued formulation.

One difficulty encountered is establishing on some principled grounds just what are the linguistically significant generalizations that are desired.

There are generalizations to be captured in both the MP rules and the lexicon. In the MP rules, the generalizations are captured by correctly specifying the phonetic output from the input. While the output is fixed so that in most cases there is little doubt about what constitutes the goal of a rule (i.e., what is a valued generalization) there are problems introduced because of irregularities of certain forms, making it sometimes desirable to use morpheme features or to change the form of lexical entries to increase the regularity of the MP rules.

Adjusting the form of the lexical entries can affect the generalizations that are captured for the lexicon by the MS rules, increasing or decreasing the claims made by them. This is one point where the problem of giving preference to one or the other subcomponent arises.

If it is granted that a generalization captured by a rule using only phonological features is predictive while a rule using arbitrary morpheme features is not, we have a basis on which to give a higher value to a formulation that uses only phonological features over one which has recourse to arbitrary morpheme features. This criterion would not apply to cases where the disparity of phonological forms would mean merely listing irregular forms, or where a feature motivated by T rules independent of the irregular forms accounts for their behavior.

Morpheme features which simply mark an entry as an exception to a rule fit the simplicity metric better than features that call for rule operation. In the case of exception-marking features, the rule will presumably have been independently motivated on phonological grounds. The rule itself is therefore unaffected in its evaluation by the simplicity metric when there is an exception, for the exception will be listed in the lexicon and will not be part of the rule. Morpheme features that call for rule operation, however, do not contribute to predictability, and they prevent any meaningful comparison of competing formulations. The cost of such morpheme features should therefore be high,

but it is difficult to give a relative value to the two types of morpheme features such that one is more or less highly valued than the other by the simplicity metric. Historical changes that must be described as the spread of a feature (calling for rule application) to include more lexical entries may be a fact, but in terms of the simplicity metric such a change means increased generality by increased feature specification rather than by decreased feature specification. Without giving a specific feature value to one or the other type, then, a general preference for exception-marking features can be established.

Such a preference for exception-marking morpheme features automatically gives priority to the MP component over the lexicon by shifting the unevaluable features to the lexicon. It also helps define what is a linguistically significant generalization by putting a high value on predictiveness and a high cost on ad hoc morpheme features.

The T rules spell out many morphemes which should undergo the MS rules along with the lexical items. It therefore seems that the convention which assigns the phonological features to the string produced as output should be assumed to apply to all formatives at once, both to lexical items contained in the output of the T component and to morphemes spelled out by T rules. The "spelling" of the latter morphemes can be thought of as abstract notations which have phonological features only after the convention has applied. Likewise, there are segments introduced in the

MP rules that, if they do not undergo the MS rules, will need to be spelled out in great detail. The MS rules apply to the output of the T component, immediately following the feature-assigning convention, but should also be assumed to apply once to every new segment inserted in the MP rules.

Thus, the MS rules (as well as MP rules) should be conceived of as constraints of well-formedness on segments and not as a mechanical production model. They need not be required to be confined to a point in the series of rules as they would be in an analog of the speaker or hearer's speech mechanism. Used as a general theory of a native speaker's linguistic competence, they apply once to all segments wherever the segments arise. Then epenthetic vowels or consonants for which the features are not predictable from the context need not be fully specified for all features by the rules that insert them: i.e., an epenthetic a vowel need only be marked for syllabicity and compactness in Capanahua. The convention that MS rules apply to each new segment automatically supplies all redundant features.

In order to cover cases where a new segment not previously defined is inserted by MP rules, say glottal stop in English, it will be necessary to provide for its specification by previous segment structure rules. In some cases, however, it may not be desirable to place these later SS rules with the other MS rules if the features they add to other of the segments would interfere with generalizations

to be captured. Presumably such late SS rules could still be counted with other SS rules in the simplicity metric, even though they might actually be listed among the MP rules. Clear evidence to argue one way or the other is lacking. Sequential constraint rules, however, are clearer. For example, the Capanahua rule for assimilation of nasals to following stops covers redundancies of the lexicon but is also needed for concatenations arising from the syntax, and must therefore be in the MP component.

If each different arrangement of distinctive features is regarded as a different theory of a language, the competing theories can be compared with respect to their ability to capture generalizations, first in the MP rules, then in the MS rules. If they are apparently equivalent in this respect, they can be compared in economy. In the phonological component of Capanahua, Consonant Solution I differed from Consonant Solution II only in the feature of obstruence assigned to the nasals. Solution II captured generalizations in the MS subcomponent that I did not and was therefore more highly valued than I. If they had been equivalent in capturing generalizations, II would still have been preferred by economy in the MP subcomponent, though by only a slight margin (two features). Vowel Solution C required five fewer features in the MP rules than D to capture the same generalizations and was also more highly valued in the MS rules than D.

In both the consonant and vowel solutions, then, the simplicity metric correlated to the way generalizations were captured if rule economy was regarded as more highly valued than economy of lexicon specification. On the other hand, Solution I which was incapable of capturing some generalizations would have been regarded as the most highly valued if absolute specification of the lexicon were more highly valued.

The effect that the initial feature arrangement had in determining what generalizations could be captured seems to confirm the position that any difference of feature arrangement in the initial inventory (i.e. the feature matrix) constitutes a different theory. If so, the simplicity metric does not serve to compare the results except in the event that generalizations captured are equivalent in some sense, and then economy of lexicon specification would be a consideration only after economy of rule specification had been compared and found to be indecisive.

The tone rules show how important syntax is to the phonology. The tone contour on reflexive verbs (nanĭ́wĭ́ 'get in the canoe' vs. nánĭ́wĭ́ 'put it in the canoe') can be traced back to the base where an object NP was identical to the subject NP and was deleted by the T rules. This caused a consonant to be added to the verb stem. Then the pitch was assigned because of the consonant. The consonant was deleted by later MP rules so that a pitch contour remains as the surface marker of the underlying NP identity

of a reflexive verb.

On universals

Chomsky has stated that the search for language universals seems to be the most promising way of constraining the form of possible grammars in order to construct the strongest possible theory of language. It would seem that one way to establish universals is to posit them on the basis of known facts, then attempt to refute them. The present study has shown that we are not yet to the point where it is of no interest to test our present theory by applying it to new languages, because our concepts of what may be language universals are not yet based on a sufficiently wide variety of languages to give assurance that future study will not likely refute them. Evidence has been set forth showing that morpheme level rules not in the MS component are in fact needed, that current notions of noun phrase and sentence conjunction with "and" and "or" are not necessarily universal, and that the notion of the cyclical operation of transformational rules (recently called in question) seems to be necessary to explain the operation of the lengthener rules on embedded causative sentences.

Paragraph structure

The previous treatment of Capanahua paragraph structure (Loos, 1963) in general stated that the topic of

the paragraph was the noun or noun phrase around which the narration centered. It was identified as the subject NP of the first sentence of the paragraph. Thereafter the NP was rarely mentioned, and when it was, it occurred in the form of a pronoun rather than a repetition of the NP. During a narration, a change of topic (i.e., a new paragraph) could be signalled by a paragraph marker and identification of a new topic.

Referential identification of the topic when it was the unmentioned subject of a sentence was effected through the terminal suffixes of subordinate verbs. The suffixes marked each verb as having the same or different subject referent as the following clause. Time sequence between clauses was also marked by the same terminal suffixes, signalling that the action of the subordinate clause preceded, followed, or was simultaneous to that of the independent clause.

The terminal suffixes of the independent clauses of each paragraph were different from independent sentences outside the paragraph in that they showed no distinction of tense or subject-person, but were instead marked for transitive concord with the verb stem. Sentences were viewed as linked together in a loose way by object, locative, and time phrases placed at the beginning of each sentence. Thus, if a sentence had the same subject referent as the preceding sentence, the subject NP was omitted and the locative, if different from any locative in the preceding

sentence, would be placed at the beginning of the sentence to signal "same subject but different place."

Among the limitations of such an approach was the fact that it was concerned primarily with surface structure so that significant relationships between sentence types were missed and some features were attributed to paragraph structure that were not properly related to paragraphs, i.e., the sentence "links" of locative, time, and object tagmemes.

The phrase structure and transformational rules of this study present evidence showing that most of the characteristics which I previously associated with paragraph structure are accounted for as characteristics of sentences embedded within sentences. The narrative sentence, formerly considered to be a distinguishing feature of narrative speech and a contrastive feature marking a paragraph, is an embedded conjoined sentence. Supporting evidence for this claim is found in the fact that the following characteristics mark both the narrative sentence and the subordinate clauses (non-restrictive relative clauses):

1. Subject NP deletion, with delay of the deletion necessary.

2. Object NP deletion, with delay of the deletion necessary.

3. Time NP deletion.

4. Md marker deletion.

5. Conditions on tense features required for embedding.

6. Assignment of suffixes to the verbs.

7. Allomorphs of verb suffixes conditioned by transitivity.

8. Independent but similar pronominalization rules.

9. Topic marker affects non-restrictive relative clauses as well as conjoined S's.

10. No subject person-markers are given forms in non-restrictive relative clauses or in conjoined S's.

These facts argue that since the embedded conjoined sentences are defined as well-formed by essentially the same types of rules as are used for non-restrictive relative clauses, the former are simply instances where the base rules generate sentences of unlimited length by successive embeddings. The narrative sentences are not in fact "independent," for they are constrained in form and dominated by another S in the same way as non-restrictive relative clauses. The only feature then which can be equated with paragraph structure is focus, so that I am compelled to relinquish my former claim for the contrastive features of paragraph structure in Capanahua. However, it may still be necessary to posit a paragraph or some such structure to explain certain characteristics (e.g., subject NP deletion) of fully independent sentences.

FOOTNOTES

[1]John Ross and George Lakoff, in a presentation to the March, 1967 Linguistics Symposium, La Jolla, University of San Diego.

[2]Robert T. Harms, "The measurement of phonological economy," Language, Vol. 42, No. 3, July-September, 1966, pp. 602-611. In this article Harms discusses the interesting cases of Russian and Bengali where economy in the lexicon conflicts with economy in the MP rules.

[3]Robert T. Harms (1966). Final paragraph. "Factor" cannot be taken in any precise mathematical sense, but rather as meaning some principled basis for weighting components or otherwise complementing the simplicity metric.

[4]Note Noam Chomsky, Aspects of the Theory of Syntax. Cambridge, Massachusetts: The MIT Press, 1965, p. 46.

"In brief, it is clear that no present-day theory of language can hope to attain explanatory adequacy beyond very restricted domains. In other words, we are very far from being able to present a system of formal and substantive linguistic universals that will be sufficiently rich and detailed to account for the facts of language learning. To advance linguistic theory in the direction of explanatory adequacy, we can attempt to refine the evaluation measure for grammars or to tighten the formal constraints on grammars so that it becomes more difficult to find a highly valued hypothesis compatible with primary linguistic data. There can be no doubt that present theories of grammar require modification in both of these ways, the latter in general, being more promising."

[5]The following recent conventions have been used in addition to those customarily employed in a generative phonology:

a) Omission of the environment bar of a context-sensitive rule is to be taken as first, a bar following the environment

condition, then as a bar preceding the mirror image of the environment condition (following Bach, 1966 and Harms, 1967).

$$\begin{bmatrix}+\text{cns}\\ -\text{voc}\\ +\text{cnt}\\ +\text{std}\end{bmatrix} \rightarrow [\ \text{cmp}] \ / \ \begin{bmatrix}+\text{cns}\\ -\text{voc}\\ -\text{cnt}\\ +\text{std}\\ \propto\text{cmp}\end{bmatrix} \quad \begin{bmatrix}-\text{cns}\\ +\text{voc}\end{bmatrix}$$

i.e., a sibilant either following or preceding an affricate, with an intervening vowel, agrees in compactness with the affricate.

b) A labeled bracketing marks the boundary of a syntactic category: ___) means "preceding the word boundary of a N noun" (following McCawley, 1965).

c) Angled brackets ⟨ ⟩ mark the first pairs to be chosen in a disjunctive series. e.g.,

$$\begin{bmatrix}+\text{cns}\\ -\text{voc}\\ +\text{std}\end{bmatrix} \rightarrow \begin{bmatrix}+\text{cnt}\\ \propto\text{cmp}\\ \langle+\text{flt}\rangle\end{bmatrix} \ / \ \begin{bmatrix}+\text{cns}\\ -\text{voc}\\ +\text{std}\\ \propto\text{cmp}\\ \langle+\text{flt}\rangle\end{bmatrix}$$

i.e., if the second consonant is [+flt] (in addition to the other features noted), the first will be [+flt] as well as [∝ cmp] and [+cnt]. Otherwise, only the features not included in the diagonals are relevant to the rule.

[6]So reported by McCawley, a position with which he apparently agrees (McCawley, 1965, 54). So also Harms (1967, 79).

[7]Gilbert Prost, "Fonemas de la lengua chacobo," Notas Linguisticas de Bolivia, II, Cochabamba, Bolivia: SIL, 1960. Same author, "Signalling of transitive and intransitive in Chacobo (Pan)," IJAL, 28, (April, 1962), pp. 108-118.

[8]Data from Olive Shell, Reconstructed Pano, Ph.D. dissertation, Univ. of Penn., 1965.

APPENDIX: A SAMPLE LEXICON

+N, +AP

+count, -An, +Num

wǐstič one

rabič two

kimiša three

čosko four [+foreign]

+N, +NP

-count

mai earth

maši sand

naip sky

niwǐ wind

ǐniǐ water

ni?i jungle

+count, -An, -Num

hiwi tree

makan stone

bǐro eye

ta?ǐ track

čičika knife

maya beach

yami axe

mačito machete

+count, +An, -Pro

?ino tiger

?atapa chicken

šontako girl

bǐronan boy

yošin spirit

?awap tapir

?i?sap bird

ro?o monkey

bakĭ child
bĭ+bo man, male
?ai+bo woman, female
papa father
mama mother
?amĭn capybara

+count, +An, +Pro, -S__

?ĭ I [+spkr]
a he [-spkr, -addr]
mi you [-spkr, +addr]

+N, +NP, +Aux

nĭtĭ day
simana week

+A, +AP

oṣo+ni white
?ani big
piškap small
toron+ni round
nĭnkĭp long
ṣana hot
maɸi cold
mĭ?čat wet

+V (middle verbs)

ɸa?o to set [+n][+[+NP,+N,+AN]...__],[+[+NP,+N+count]__]
raka to lay [+n] " "
kaya to jump [+tan] " "
po?kĭ to cross a river[+ha] " "
rakĭ to fear [+ha] " [+[+NP,+N,+An]__]
mapo to burn [+ha]

+V, +[NP $\underset{VP}{\_\_}$] -[$\underset{VP}{\_\_X}$] (transitive verbs)

is to see [+[+NP,+N,+AN] ... __]

ninkak to hear "

ʔonan to know "

kĭĭn to want " , [+K], [+S __]

bana to plant " [+[+NP,+N,-An,+count] __]

amak to step on "

bo to carry away "

biǩ to grasp, take "

nĭʂa to tie " [+[+NP,+N,+count,-An] __]

pakĭ to fell

ʂoka to skin [+[+NP,+N,+count] __]

taran to roll "

yoʔi to tell [+[+NP,+N,+An]...__][+[+NP,+N,+An] __]
[+S __] [+J]

+V,+[$\underset{VP}{\_\_X}$] -[NP $\underset{VP}{\_\_}$] (intransitive verbs)

ʔoʂa to sleep [+[+NP,+N,+An] ... __]

yowan to talk "

wini to mourn "

kĭsa to tell a lie "

biǩi to run "

niǩ to walk "

ʔoʔko to cough "

sĭso to whistle "

mai to fall over [+[+NP,+N,+count] ... __]

BIBLIOGRAPHY

Bach, E. (1965). "On some recurrent types of transformation," MSLL 18. pp. 3-18.

Bever, T. G. (1963). "Theoretical implications of Bloomfield's 'Menomini morphophonemics,'" Quarterly Progress Report No. 68, Cambridge, Mass.: Research Lab. of Electronics, MIT pp. 197-203.

Bloomfield, L. (1939). "Menomini morphophonemics," Travaux du cercle linguistique de Prague, VIII, 1939, pp. 105-115.

_____ (1962). The Menomini Language. New Haven: Yale University Press.

Chomsky, Noam (1957). Syntactic Structures. The Hague: Mouton & Co.

_____ (1961). "Some methodological remarks on generative grammar," Word, 17, pp. 219-239. Reprinted in part in Fodor and Katz (1964).

_____ (1962). "A transformational approach to syntax." In A. A. Hill (ed.), Proceedings of the 1958 Conference on Problems of Linguistic Analysis in English, pp. 124-148. Austin, Texas. Reprinted in Fodor and Katz (1964).

_____ (1964). Current Issues in Linguistic Theory. The Hague: Mouton & Co. A slightly earlier version appears in Fodor and Katz (1964).

_____ (1965). Aspects of the Theory of Syntax. Cambridge, Massachusetts: The MIT Press, Massachusetts Institute of Technology.

_____, and G. A. Miller (1963). "Introduction to the formal analysis of natural languages," In R. D. Luce, R. Bush, and E. Galanter (eds.), Handbook of Mathematical Psychology, Vol. II, pp. 269-322. New York: Wiley.

De la Grasserie, Raul (1890). "De la famille linguistique Pano," 7 Congres International de Americanistes, Berlin: 1888, 438-449.

Fillmore, C. J. (1963). "The position of embedding transformations in a grammar," Word, 19, pp. 208-231.

_____ (1967). "The case for case," mimeographed paper presented to the Symposium on Language Universals, University of Texas, April, 1967.

Fodor, J. A., and J. J. Katz (eds.) (1964). The Structure of Language: Readings in the Philosophy of Language. Englewood Cliffs, N.J.: Prentice-Hall.

Gleitman, L. (1961). "Conjunction with and," Transformations and Discourse Analysis Projects, No. 40, mimeographed. Philadelphia: University of Pennsylvania.

Halle, M. (1959). The Sound Pattern of Russian. The Hague: Mouton & Co.

_____ (1961). "On the role of the simplicity in linguistic description." In R. Jakobson (ed.), Structure of Language and Its Mathematical Aspects, Proceedings of the Twelfth Symposium in Applied Mathematics, pp. 89-94. Providence, R. I.: American Mathematical Society.

_____ (1962). "Phonology in generative grammar," Word, 18. pp. 54-72. Reprinted in Fodor and Katz (1964).

_____ (1964). "On the bases of phonology," In Fodor and Katz (1964).

_____, and N. Chomsky (1960). "The morphophonemics of English," Quarterly Progress Report, No. 58, Research Laboratory of Electronics, MIT, pp. 275-281.

Harms, Robert T. (1966). "The Measurement of phonological economy," Language, Vol. 42, No. 3, July-September, 1966, pp. 602-611.

_____ (1967). The ABC's of Generative Phonology. University of Texas, mimeographed, January, 1967. Revised edition, Univ. of Texas (mimeographed) June, 1967.

Humboldt, Alexander von (1852). Personal Narrative of Travels to the Equinoctial Regions of America, during the years 1799-1804. 3 volumes. Translated by Thomasina Ross. London: Henry G. Bohn, (orig. in French, Paris: Casa de Rosa, 1826), Vol. I, 505 pp.

Jakobson, Fant, and Halle (1961). Preliminaries to Speech Analysis. The Distinctive Features and their Correlates. Cambridge, Massachusetts: The MIT Press, Massachusetts Institute of Technology, reprinted 1963.

Katz, J. J., and J. A. Fodor. "The structure of a semantic theory," Language, 39, pp. 170-210. Reprinted in Fodor and Katz (1964).

_____, and P. Postal (1964). An Integrated Theory of Linguistic Descriptions. Cambridge, Massachusetts: MIT Press.

Klima, E. S. (1964). "Negation in English," In Fodor and Katz (1964).

_____ (1965). "Current developments in generative grammar," Kybernetika 1, 1965, pp. 184-197.

Lakoff, George, and Stanley Peters (1966). "Phrasal conjunction and symmetric predicates," Report #17 to the NSF of the Harvard Computation Laboratories.

Lees, R. B. (1957). Review of Chomsky (1957), Language, 33, pp. 375-407.

_____ (1960). The Grammar of English Nominalizations. The Hague: Mouton & Co.

Loos, Eugene E. (1963). "Capanahua Narration Structure," Texas Studies in Literature and Language, Vol. IV Supplement, 1963, pp. 697-742.

Matthews, P. H. (1961). "Transformational grammar," Archivum Linguisticum, 13, pp. 196-209.

_____ (1965). "Problems of selection in transformational grammar," J. L. 1, 1965, pp. 35-47

McCawley, James D. (1965). The Accentual System of Standard Japanese. Ph.D. dissertation, Cambridge, Mass.: Massachusetts Institute of Technology (mimeographed).

McQuown, Norman A. (1955). "Indigenous languages of Latin America," American Anthropologist 57, (1955) 501-570.

Postal, P. M. (1964a). Constituent Structure. Pub. #30 of the Research Center in Anthropology, Folklore, and Linguistics. IJAL 30, I: III.

_____ (1964b). "Underlying and superficial linguistic structure," Harvard Educational Review, 34, pp. 246-266.

_____ (1964c). "Limitations of phrase structure grammars." In Fodor and Katz (1964).

Prideaux, Gary Dean (1966). The Syntax of Japanese Honorifics. Unpublished Ph.D. dissertation, Austin: University of Texas.

Williamson, Kay (1965). A Grammar of the Kolokuma Dialect of Ijo. Cambridge University Press, West African Language Monograph Series II.

A SELECT BIBLIOGRAPHY OF

THE PANOAN LANGUAGES

Kensinger, Kenneth M. "The phonological hierarchy of Cashinahua (Pano)," Studies in Peruvian Indian Languages I, (1963) 207-217, Mexico, Summer Institute of Linguistics.

Loos, Eugene E. "Capanahua narration structure," Texas Studies in Literature and Language, Vol. IV Supplement (1963), University of Texas Press, Austin, Texas, 697-742.

Mason, John Alden. "The Languages of South American Indians," in Handbook of South American Indians, Bureau of American Ethnology, Bulletin 143, Vol. 6, 1950 (1946) (1963) pp. 157-317.

McQuown, Norman A. "Indigenous Languages of Latin America," American Anthropologist 57 (1955), 501-570.

Prost, Gilbert. "Fonemas de la lengua chacobo," Notas Lingüísticas de Bolivia, II, Cochabamba, Bolivia, Summer Institute of Linguistics, 1960.

_____. "Signalling of transitive and intransitive in Chacobo (Pano)," IJAL 28 (April, 1962), 103-118.

_____. "Chacobo," in Gramáticas Estructurales de Lenguas Bolivianas II, Riberalta, Bolivia, Summer Institute of Linguistics, pp. 1-130.

Russell, Robert. "Algunos morfemas de Amahuaca (Pano) que equivalen a la entonación del Castellano," Peru Indigena 8 (1958), 29-33.

_____. A Transformational Grammar of Amahuaca. unpublished M.A. thesis, Ohio State University, 1965.

_____ and Delores. "Syntactotonemics in Amahuaca (Pano) and Amahuaca vocabulary," Serie Linguistica Especial, No. 1, Publicacoes do Museo Nacional, Rio de Janeiro, (1959), 128-170.

Scott, Eugene and Eunice Pike. "The phonological hierarchy of Marinahua," Phonetica, 8 (1962), 1-8.

Shell, Olive A. "Cashibo I: Phonemes," IJAL 16 (1950), 198-202.

_____. "Cashibo II: Grammemic analysis of transitive and intransitive verb patterns," IJAL 23 (1957), 178-218.

_____. Pano Reconstruction. Unpublished Ph.D. dissertation, University of Pennsylvania, 1965.