TUPI STUDIES I

Edited by:

David Bendor-Samuel

THE MORPHOPHONOLOGY OF ASURINI WORDS

Carl H. Harrison

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

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

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

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

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

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

this paper. Here we are basically concerned with both syn-
tactically and phonologically conditioned alternation be-
tween, in the main, contrastive segment types corresponding
in most cases to those types which would result from tradi-
tional segmental phonemic analysis.

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

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

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

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

The actual phonological variants (allomorphs) of

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

The symbols of the output alphabet are taken from
articulatory phonetics, one symbol per contrastive segment
type, each representing a bundle of contrasting articula-
tory features. The input alphabet is somewhat larger, con-
taining some ad hoc cover symbols for smaller sets of con-
trastive features as well as for possible null realization
(∅).

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

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

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

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

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

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

Consonants[6]

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

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

Vowels

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

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

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

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

   Note that for Ŋ there is no corresponding resonant.
This represents a lacuna in the phonological pattern. An
ŋ fills in for the missing resonant in the pattern of alter-
nation.

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

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

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

	alp	resonant	sibilant	C
Y	1	( )	( )	1

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

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

Along with the alphabets, certain other upper case
symbols are used.[7]

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

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

S stands for a syllable in the string. It is used in
the rules that indicate accent placement. In order to de-
termine accent, one needs to know how many syllables there
are, which ones have inherent accents, and the kinds and
locations of junctures in the combinations. We will intro-
duce the following convention for determining the number of
syllables in a string. For any string, every vowel is a
syllable center and can be assigned an S; consonants may be
disregarded. Accent placement is then determined according
to the inherent accent marks, syllable symbols, and junc-
ture marks in the string by applying Accent Rules 24-29.
Syllables in Asurini have the structure V, CV, VC_1, and
CVC_1, where C_1 is a variable for m,n,ŋ, and y. There are
no other restrictions or combinations, except that y never
occurs word medially before V. Syllable boundary shifts
occur to bring the structure of a string as closely as
possible into a series of CV syllables. If the syntax
gives a CVC_1 followed by a V, the structure imposed is
$CV.C_1V$. Some of the rules for null realization seem to re-
flect this pressure.

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

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

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

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

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

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

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

     (1) Single parenthesized letters of the phonological
alphabets are used. These constitute phonological condi-
tioning for the most part.

     (2) Two or more concatenated symbols of the alphabets
are used. Such a substring is often part of one and only
one form and hence is not a strictly phonological environ-
ment, though historically it may have been.

     (3) Syllable structure dominates a number of the pro-
cesses. It must be remembered that such symbols as C and
V in a rule represent an abbreviation for a number of rules.

     (4) Junctures appear in all of the rules and can be
thought of as orienting the rule so that it applies in the
proper place in the string. In the sense that junctures
mark syntactical boundaries, none of the rules represents
a purely phonological process. A possible qualification
would be to say that once a grammatical boundary is estab-
lished, such and such a process is phonological.

     (5) The symbol ∞ is used to indicate that other
environments have been mentioned in the same rule, and that
if the string being scanned contains the symbol to be re-
placed as specified in the rule but does not match any of
the other environment specifications in the rule, then the
part of the rule containing ∞ is to be applied.

     (6) When the name for a class of lexical forms
appears, the conditioning is clearly syntactic. The

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

  1   first person singular

  2   second person singular

  12   first person plural, inclusive of addressee

  13   first person plural, exclusive of addressee

  22   second person plural

 3(3) third person, singular or plural

    Aff 1   (affix class 1)

      ce- '1', ne- '2', cene- '12', ore- '13',

      pe- '22', I- '3(3)'.

    Aff 2

      we- '1', e- '2', cere- '12', ore- '13',

      pece- '22', O- '3(3)'.

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

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

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

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

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

$$AX \Rightarrow BX$$

$$AY \Rightarrow CY$$

we can conflate them in the following fashion

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

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

Given the rules

$$AX \Rightarrow BX \qquad AY \Rightarrow EY$$

$$CX \Rightarrow DX \qquad CY \Rightarrow FY$$

we can use subscripted brackets to reduce them to

$$\begin{bmatrix} A \\ C \end{bmatrix}_{1\ 1} \begin{bmatrix} X \\ Y \end{bmatrix}_{2\ 2} \Rightarrow \begin{bmatrix} \begin{bmatrix} B \\ D \end{bmatrix}_{1\ 1} & X \\ \begin{bmatrix} E \\ F \end{bmatrix}_{1\ 1} & Y \end{bmatrix}_{2\ \ \ \ \ 2}$$

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

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

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

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

has often been given as

$$A \longrightarrow \begin{Bmatrix} B/\_Y \\ C/\_Z \end{Bmatrix}$$

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

$$A \longrightarrow \begin{Bmatrix} B \\ C \end{Bmatrix}$$

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

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

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

(1) epenthesis: 1

(2) vowel harmony: 13

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

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

(5) accent rules: 24-29

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

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

Rule 1

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

This is a rule for epenthesis.

Examples:

#I-ɨwɨrápaN-A# ⇒ #I-c-ɨwɨrápaN-A#

3(3)-bow-noun marker

'his/her/their bow(s)'

#I-(ð)pe# ⇒ #I-c-ðpe#

3(3)-to/at

Rule 2

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

This describes a certain lexically conditioned alternation.

Examples:

    #ce-(ð)pe# ⟹ #ce-ðpe#

     1-to/for

    #ʔɨ-A-(ð)pe# ⟹ #ʔɨ-Aᴸpe#

    water-n-to/for

Rule 3

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

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

Example:

    #ʔɨ-Aᴸpe# ⟹ #ʔɨᴸpe#

    water-n-to

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

Rule 4

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

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

Examples:

    #máʔe-A# ⟹ #máʔe-e#

    something-n

#coówi-A# ⟹ #coówi-a#

grass-n

Rule 5

I - (e) ⟹ h - (e)

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

Example:

#I-(e)ro-ʔaN-D# ⟹ #h-(e)ro-ʔaN-D#

3(3)-com-fall-dependent

Rule 6

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

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

Examples:

#n(a)-O-KeN-ihi# ⟹ #n-O-KeN-ihi#

negative-3(3)-sleep-negative

#t(e)-pe-ʔo# ⟹ #te-pe-ʔo#

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

'in order for you all to eat it'

Rule 7

e - (e) ⟹ e - ∅

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

Example:

$\quad$ #cene-(e)ro$^\angle$(to)N-$\underline{D}$# $\Rightarrow$ #cene-ro$^\angle$(to)N-$\underline{D}$#

$\quad$ x/12-commitative-come-dependent verb marker

Rule 8

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

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

Example:

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

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

Example:

$\quad$ #n-O$^\angle$?aN-ihi# $\Rightarrow$ #n-o$^\angle$?aN-ihi#

$\quad$ neg-3(3)-fall-neg

Rule 9

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

Examples:

    #pe-(e)ro-áta# ⟹ #pe-ro-áta#
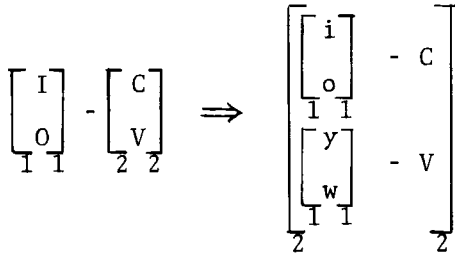
    22-commitative-walk

    #oro-(e)ro-áta# ⟹ #oro-ero-áta#

    13-com-walk

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

Rule 10

$$
\begin{bmatrix} I \\ O \end{bmatrix}_1 - \begin{bmatrix} C \\ V \end{bmatrix}_2 \Rightarrow \left[\begin{array}{cc} \begin{bmatrix} i \\ o \end{bmatrix}_1 & - C \\[2ex] \begin{bmatrix} y \\ w \end{bmatrix}_1 & - V \end{array} \right]_2
$$

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

Examples:

    #O-PaM# ⟹ #o-PaM#

    3(3)-finished

    #O-áta# ⟹ #w-áta#

    3(3)-walked

    #I-PaM-D# ⟹ #i-PaM-D#

    3(3)-finish-dep

    #I-áta-I# ⟹ #y-áta-I#

    3(3)-walk-post relational marker

Rule 11

$$\begin{bmatrix} \underline{\text{Aff 3}} \\ \infty \end{bmatrix} - \text{(to)} \implies \begin{bmatrix} \underline{\text{Aff 3}} & - \text{ to} \\ \infty & - \emptyset \end{bmatrix} \quad \begin{matrix} (\text{Aff 2 - Aff 3} \\ = \text{o-)} \end{matrix}$$

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

Examples:

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

1-come-dependent marker

#o-(to)N# $\implies$ #o-N# (eventually ón)

3-come

Rule 12

$$\begin{bmatrix} \# \\ \underline{\text{Aff 2}} - \\ \text{i-} \\ \text{pe-} \\ \text{o-} \\ \text{mo-} \\ \infty \\ // \end{bmatrix} \quad N \implies \begin{bmatrix} \# \, t \\ \underline{\text{Aff 2}} - t \\ \emptyset - h \\ \text{pe-n} \\ \text{o-}\emptyset \\ \text{w-}\emptyset \\ \infty - r \\ // r \end{bmatrix}$$

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

Examples:

#konómi-a//Néha-a# ⇒ #konómi-a//réha-a#

boy-n//eye-n

#we-Néha-a# ⇒ #we-téha-a# (Aff 2)

1 (own)-eye-n

#Néha-a# ⇒ #téha-a#

eye-n

#i-Néha-a# ⇒ #héha-a#

3(3)-eye-n

#pe-Néha-a# ⇒ #pe-néha-a#

22-eye-n

#ce-Néha-a# ⇒ #ce-réha-a#

1-eye-n

#o-NópiN# ⇒ #O-ópiN#

3(3)-lift

#o-mo-Náhɨ# ⇒ #o-w-áhɨ#

3(3)-causative-hurt

Rule 13

$$
\begin{bmatrix} ero \\ \infty \\ \underline{1} \quad \underline{1} \end{bmatrix} - \begin{bmatrix} ha \\ ka \\ \underline{2} \quad \underline{2} \end{bmatrix} \Rightarrow \begin{bmatrix} \begin{bmatrix} era \\ \infty \\ \underline{1} \quad \underline{1} \end{bmatrix} - ha \\ \begin{bmatrix} ere \\ \infty \\ \underline{1} \quad \underline{1} \end{bmatrix} - ka \\ \underline{2} \qquad\qquad \underline{2} \end{bmatrix}
$$

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

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

Examples:

#o-ha# ⟹ #a-ha#

3(3)-go

#w-ero-ha# ⟹ #w-era-ha#

3(3)-commitative-go

#h-ero-ka# ⟹ #h-ere-ka#

3(3)-com-be present

Rule 14

$$
\begin{bmatrix} mo \\ \infty \\ 1 \ 1 \end{bmatrix} - \begin{bmatrix} P \\ T \\ K \\ 2 \ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} mo - \begin{bmatrix} m \\ n \\ \eta \\ 2 \ 2 \end{bmatrix} \\ \infty - \begin{bmatrix} p \\ t \\ k \\ 2 \ 2 \end{bmatrix} \\ 1 \qquad\qquad 1 \end{bmatrix}
$$
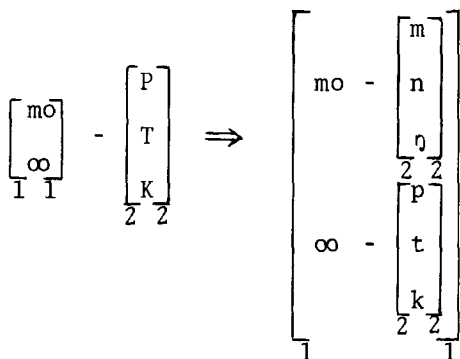
Certain verb stems which under most conditions begin with
stops, begin with corresponding nasals when preceded by mo-
'causative'.

Examples:

#o-mo-KeN#   #o-mo-ŋeN#

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

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

#o-́KeN# $\Rightarrow$ #o-́keN#

3(3)-sleep

Rule 15

$$C_3 = \Rightarrow \emptyset = \qquad (C — C_3 = y)$$

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

Examples:

#o-manáhaŊ=paM# $\Rightarrow$ #o-manáha=paM#

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

#o-́kay=paM# $\Rightarrow$ #o-́kay=paM#

3(3)-burn-finish

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

Example:

#i-́qaN=qaN# $\Rightarrow$ #I-́qa=qaN#

3(3)-has several holes

Rule 16

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

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

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

Examples:

$$\#a\text{-}ha\text{-}\underline{D}\# \;\Rightarrow\; \#a\text{-}ha\text{-}w\#$$

3(3)-go-dependent

$$\#o\text{-}keN\text{-}\underline{D}\# \;\Rightarrow\; \#o\text{-}keN\text{-}(t)a\#$$

3(3)-sleep-dep

Rule 17

$$
\begin{bmatrix} (N) \\ (\text{\textipa{N}}) \end{bmatrix}_{1\ 1}
-
\begin{bmatrix} \# \\ \infty \end{bmatrix}_{2\ 2}
\;\Rightarrow\;
\begin{bmatrix}
\begin{bmatrix} \emptyset \\ \emptyset \end{bmatrix}_{1\ 1} - \# \\[2ex]
\begin{bmatrix} N \\ \text{\textipa{N}} \end{bmatrix}_{1\ 1} - \infty
\end{bmatrix}_{2 \qquad 2}
$$

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

Examples:

$$\#o\text{-}p\text{í}h\text{i}(\text{\textipa{N}})\# \;\Rightarrow\; \#o\text{-}p\text{í}h\text{i}\#$$

3(3)/3(3)-catch

$$\#o\text{-}p\text{í}h\text{i}(\text{\textipa{N}})\text{-}(r)\text{á}me\# \;\Rightarrow\; \#o\text{-}p\text{í}h\text{i}\text{\textipa{N}}\text{-}(r)\text{á}me\#$$

-continuative

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

Rule 18

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

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

Examples:

#a⌐ʔaN-(r)áme# ⟹ #a⌐ʔaN-áme#

1-fall-now   (eventually aʔaráme)

#o⌐ʔo-(r)áme# ⟹ #o⌐ʔo-ráme#

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

Rule 19

$$\begin{bmatrix} C \\ V \end{bmatrix}_{1\ 1} - \begin{bmatrix} (po) \\ (c) \end{bmatrix}_{2\quad 2} \implies \begin{bmatrix} C - \begin{bmatrix} \emptyset \\ \emptyset \end{bmatrix}_{2\ 2} \\ V - \begin{bmatrix} po \\ c \end{bmatrix}_{2\quad 2} \end{bmatrix}_{1\qquad\qquad 1}$$

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

potan   potat   potar   pota

tan   tat   tar   ta

(c) occurs in a small number of suffixes.

Examples:

#o⌐keN-(pó)ta# ⟹ #o⌐keN⌐ta#

3(3)-sleep-future

#a⁻ha-(pó)ta# ⟹ #a⁻ha-póta#

1-go-future

#o⁻keN-(c)awa# ⟹   #o⁻keN-awa#

3-sleep-perhaps

#a⁻ha-(c)awa# ⟹ #a⁻ha-cawa#

Rule 20

$$
\begin{bmatrix} \begin{bmatrix} M \\ N \\ \text{Ņ} \end{bmatrix}_{1\ 1} - \begin{bmatrix} (t) \\ \\ \begin{bmatrix} \# \\ C_4 \end{bmatrix}_{2\ 2} \\ \\ V \end{bmatrix}_3 \end{bmatrix}_3 \Rightarrow \begin{bmatrix} \begin{bmatrix} p \\ t \\ k \end{bmatrix}_{1\ 1} - (t) \\ \\ \begin{bmatrix} m \\ n \\ \text{ŋ} \end{bmatrix}_{1\ 1} - \begin{bmatrix} \# \\ C_4 \end{bmatrix}_{2\ 2} \\ \\ \begin{bmatrix} w \\ r \\ \text{ŋ} \end{bmatrix}_{1\ 1} - V \end{bmatrix}_3 \quad (C - C_4 = (t)\ )
$$
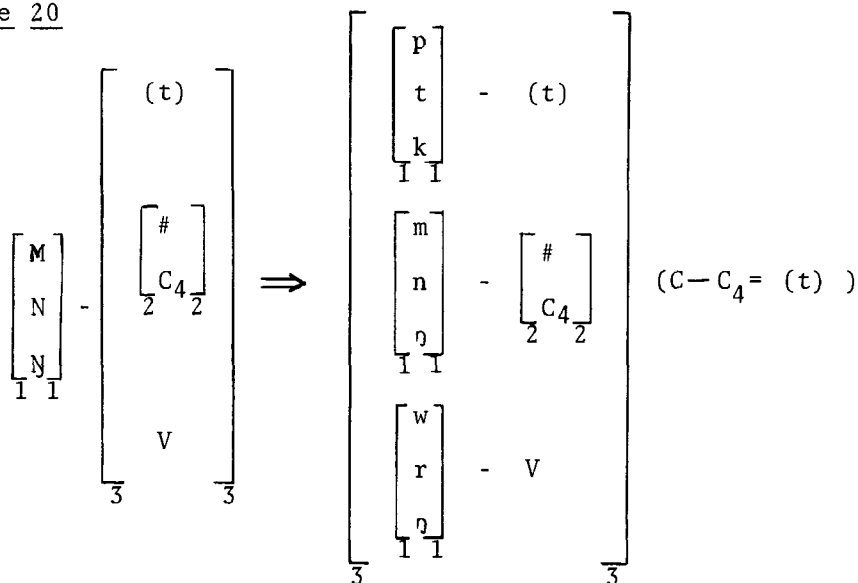
The parenthetic notation indicates that C_4 has the same
elements as C except for (t) which belongs only to C. A
large number of stems end in one of the three upper-case
nasals, and hence the rule will appear often in actual
derivations. A real problem in ordering occurs here. If
we allow the rule for deletion of (t) to precede this one,
we are left with the vowel that follows it as an environ-
ment which would then invalidate the generalization we wish
to make about the nature of change before -V. If we put
ͺthe (t) rule after this one we must indicate that it is not
to be considered one of the consonants of the language for
this rule. Clearly the conditioning is grammatical for p,
t,k and phonological for the rest (with the exception of
the environment #). Hence ordering problems are a result

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

Examples:

 #oˊpaM-(t)a# ⟹   #oˊpap-(t)a#

 3(3)-finish-dep

 #oˊpaM# ⟹   #oˊpam#

 #oˊpaMˊta# ⟹   #oˊpamˊta#

  -future, desiderative

 #oˊpaM-áme# ⟹   #oˊpaw-áme#

  -now

 #oˊkeN-(t)a# ⟹   #oˊket-(t)a#

  -sleep-

 #oˊkeN# ⟹   #oˊken#

 #oˊkeNˊta# ⟹   #oˊkenˊta#

 #oˊkeN-áme# ⟹   #oˊker-áme#

 #iˊʔaŊ-(t)a# ⟹   #iˊʔak-(t)a

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

 #oˊʔaŊ# ⟹   #oˊʔaŋ#

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

 #oˊʔaŊ-áme# ⟹   #oˊʔaŋ-áme#

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

 #o-pótaN# ⟹   #o-pótan#

#i-pótaN-(t)a# ⟹ #i-pótat-(t)a#

#o-pótaN-áme# ⟹ #o-pótar-áme#

Rule 21

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

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

Examples:

#o-kaY-(t)a# ⟹ #o-kay-ta#

3(3)-burn-dep

#i-ʔak-(t)a# ⟹ #i-ʔák-a#

x/3(3)-dig up-dep

Rule 22

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

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

Examples:

    #o∸kaY# ⟹ #o∸kay#

    3(3)-burn

    #o∸kaY∸ta# ⟹ #o∸kay∸ta#   .

        -future

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

    #o∸kaY-óho# ⟹ #o∸kac-óho#

        -fiercely

Rule 23

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

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

Examples:

    #i∸ha-I# ⟹ #i∸ha-y#

    3(3)-go-agreement

    #i∸ker-I# ⟹ #i∸ker-i#

    3-sleep-agreement

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

each other and junctures.

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

Rule 24

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

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

Examples:

    $\acute{\#}$ʔaŋ-a# ⟹ #ʔáŋ-a#

    house-n

    #cáwar-a//ˊci-a# ⟹ #cáwar-a//cí-a#

    dog-n//nose-n

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

Rule 25

     S $\overset{\prime}{=}$ S ⟹ Ś = S

= juncture, though essentially affecting the bordering

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

Example:

    #o-ápo⁼pam# ⟹   #o-ápó=pam#

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

Rule 26

        S ⁻ S ⟹ Ś - S

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

Example:

    #o⁻ken# ⟹   #ó-ken#

    3(3)-sleep

Rule 27

        ...Ś...S̀S# ⟹   #...Ś...SS#

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

Example:

    #ʔáŋ-a-pɨpe# ⟹   #ʔáŋ-a-pɨpe#

    house-n-in

Rule 28

        #...S̀S# ⟹   #...ŚS#

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

Example:

    #i-pɨpe# ⟹ #i-pɨpe#

    3(3)-in

Rule 29

$$...\acute{s}...\acute{s}...\begin{bmatrix} // \\ \# \end{bmatrix} \implies ...s...\acute{s}...\begin{bmatrix} // \\ \# \end{bmatrix}$$

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

Examples:

        #o-ápo-póta# ⟹ #o-apo-póta#

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

        #o-ápo⁻pam≠tar-áme#

Rule 25   #o-ápo⁻pám=tar-áme#

Rule 26   #o-ápó-pám=tar-áme#

Rule 29   #o-ápó-pam=tar-áme#

Rule 29   #o-ápo-pam=tar-áme#

Rule 29   #o-apo-pam=tar-áme#

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

        #ce-rá?ɨr-a//ráŋ-a#

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

Rule 24   #ce-rá?ɨr-a//ráŋ-a# (no other change)

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

Example:

  #ɨwoʔoetɨmakɨŋdŋawa#   (secondary accent underlined)

  soap-n

Rule 30[14]

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

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

   4.   The Rules Applied to Selected Texts.

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

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

   With each sentence in the text the following infor-
mation is presented:

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

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

Examples:

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

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

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

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

Symbols:

  I: independent   part: particle

  D: dependent   (r)ire phrase: phrase ending in

  i: intransitive   (r)ire

  t: transitive   (r)amo phrase: phrase ending in

  N: noun phrase   (r)amo

  V: verb phrase   Dsc: descriptor

  Aux: auxiliary verb dep: dependence marker

  com: commitative   PP: post positional

  n: noun marker   Aff: affix

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

#3-sleep-future-continuative#

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

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

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

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

    #IiV#Aux iV#

    #3-walk-future#3-go#

    #0-áta-(pó)ta(N)#0-ha#

10   #w-áta-(pó)ta(N)#0-ha#

17   #w-áta-(pó)ta#0-ha#

19   #w-áta-póta#0-ha#

29   #w-ata-póta#0-ha#

30   #wata-póta#0-ha#

30   #watapóta#0-ha#

10   #watapóta#o-ha#

13   #watapóta#a-ha

26   #watapóta#á-ha#

30   #watapóta#áha#

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

Text 1

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

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

        #coówi-A#a-manáhaŊ#we-ka#I-manáhaN-D̲#I-péhiŊ-D̲#

        4,30 #coówia#

        20,30   #amanáhaŋ#

        26,30   #wéka#

        10   #i-manáhaŊ-D̲#

        16   #i-manáhaŊ-(t)a#

        20   #i-manáhak-(t)a#

        21   #i-manáhak-a#

        30,30   #imanáhaka#

        10   #i-péhiN-D̲#

        16   #i-péhiN-(t)a#

        20   #i-péhit-(t)a#

        21,30,30   #ipéhita#

        #coówia#amanáhaŋ#wéka#imanáhaka#ipéhita#

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

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

        #a-moíriŊ#a-péhiN#we-ka#I-manáhaŊ-D̲#

        20   #A-moíriŋ#

        20   #a-péhin#

        26   #we-ka#

10	#i-manáhaŊ-<u>D</u>#
16	#i-manáhaŊ-(t)a#
20	#i-manáhak-(t)a#
21	#i-manáhak-a#

30,...,30. #amoɪriŋ#apéhin#wéka#imanáhaka#

1.3.   #ItV#Part#N#

#1/3-cut-future#again#grass-n#

#a-manahaŊ-(po)ta(N)#no#coowi-A#

17	#a-manáhaŊ-(pó)ta#
19	#a-manáhaŊ⸍ta#
20	#a-manáhaŋ⸍ta#
26	#a-manáháŋ-ta#
29	#a-manaháŋ-ta#
4	#no#coówi-a#

30...   #amanaháŋta#no#coówia#

1.4.   #Adv#(r)amo-phrase#IiV#Aff-PP#

#there#sun-n-go down-when#I-1eave-future#3-from#

#pe#qaN-A-ha-(r)amo#a-pohin-(po)ta(N)#I-(o)hi#

24	#pé#
4	#qaN-a⸍ha-(r)amo#
18	#qaN-a⸍ha-ramo#
20	#qar-a⸍ha-ramo#
24,etc.	#qár-a⸍ha-ramo#
29	#qar-á-ha-ramo#
30,-	#qaráharamo#

17   #a-póhin-(pó)ta#

19   #a-póhin⁻ta#

26   #a-póhín-ta#

29   #a-pohín-ta#

30,-   #apohínta#

1   #I-c-(ð)hi#

2   #I-c-ðhi#

10   #i-c-ðhi#

28,30,-   #icóhi#

#DiV#DtV#N#

#1-eat-dep#X/3-eat-dep#game meat#

#we-káro-D#I⁻ʔo-D#maʔeáʔa-A#

16   #we-káro-w#

30,- #wekárow#

10   #i⁻ʔo-D#

16   #i⁻ʔo-w#

26   #í-ʔo-w#

30   #íʔow#

4   #maʔeáʔe-a#

30   #maʔeáʔaa#

#wekárow#íʔow#maʔeáʔaa#

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

Line 1: word types

Line 2: dictionary glosses

Line 3:   input form

Line 4:   output form (terminal string)

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

1.5.   #DiV#Part#

        #I-bathe-fut-dep#also#

        #we-cáhoN-(pó)ta(N)-D#we#

        #wecahóŋtata#we#

        #16,17,19,20,21,28,29,30...# —— #

1.6   #DSC#Part#Part#

        #I-hot-fut#emphatic#again#

        #ce-pirí'ay-(pó)ta(N)#ca#no#

        #cepiri'áyta#ca#no#

        #17,19,28,29,30# —— # —— #

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

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

        #I-manáhaN≠paM-(r)íre#iwo'oetimakiŋáŋaM-A-po#

        #imanahapawíre#iwo'oetimakiŋáŋawa-po#

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

        #IiV#

        #1-bathe-future#

        #a-cáhoN-(pó)ta(N)#

        #acahóŋta#

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

Free translation of Text 1.

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

The following text will be given in this form:

Line 1: word/phrase names

Line 2: dictionary glosses

Line 3: input form

Line 4: output form

2.1. # I iv # Aff=PP # DiV # N #

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

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

#aporoŋéta#neópe#wétoyna#ícee#.

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

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

#t(e)-ere-mo-cíwaŊ#pápe-A#e⁻ʔɨyn-D#ce-(e)ro-wáke#

#teremocíwaŋ#pápee#éʔɨyna#cerowáke#

pápee ⟨ papel (Ptg.)

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

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

#t(e)-a-máʔe-ne#I-Nèhe#we⁻toyn-D#ne-mo-cíwaŊ⁻(t)aM-A-
                                                   -Nèhe#

#tamáʔene#héhe#wétoyna#nemocíwákawarehe#

2.4. #ItV#DiV#N#DtV#

    #2-write-write-fut#2-sit-dep#paper-n#X/3-write-dep#

    #ere-kánaM=kánaM-(pó)ta(N)#e-toyn-D#pápe-A#I-mo-cíwaŋ-D#

    #erekanakanámta#étoyna#pápee#imocíwaka#

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

2.5.   #IiV#DiV#DiV#Aff=PP#

    #I-sit-fut#I-sit-dep#I-look-dep#2-at#

    #a-ápiŋ-(pó)ta(N)#we-toyn-D#we-má?e-D#ne-Nèhe#

    #aapíŋta#wétoyna#wemá?ew#neréhe#

2.6.   #IiV#Aff=PP#DiV#

    #I-see-fut#2-at#I-sit-dep#

    #a-má?e-(pó)ta(N)#ne-Nèhe#we-toyn-D#

    #ama?epóta#neréhe#wétoyna#

Text 2   Free Translation

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

Text 3

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

Text 4

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

3.1.   #N#ITV#Aux iV#

    #grass-n#i-cut-fut#I-go#

    #coówi-A#a-manáhaN-(pó)ta(N)#we⸴ha#

    #coówia#amanahánta#wéha#

3.2.   #Adv#ItV#Aux iV#

    #a little#I-cut-fut#I-go#

    #pípi#a-manáhaN-(pó)ta(N)#we⸴ha#

    #pípi#amanahánta#wéha#

3.3.   #Dsc#IiV#Aff=PP#DiV#

    #sun-small#I-leave-fut#3-from#I-bathe-dep#

    #qaráhɨ-pípi#a-póhin-(pó)ta(N)#I-(ð)hi#we-cáhoN-D#

    #qarahɨpípi#apohínta#icóhi#wecáhoka#

3.4.   #IiV#

    #3-finish#

    #O⸴paM#

    #ópam#

4.1.   #IiV#Aux iV#DtV#

    #I-walk-fut#I-go#X/3-kill-dep#

    #a-áta-(pó)ta(N)#we⸴ha#I-cóka-D#

    #aatapóta#wéha#icókaw#

4.2.   #N#part#DtV#ItV#

    #deer-n#indef.#X/3-com.-come-dep.#12/3-eat#

    #mícaN-A#mo#I-(e)ro⸴(to)N-D#ca⸴ʔo#

    #mícara#mo#hérota#cáʔo#   ((e)ro is not Aff 3)

4.3.   #anohi#N#N#

        #there is no#food#game meat#

        #anóhi#kóme-A#ma?eá?a-A#

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

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

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

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

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

        #á?eramo#ahapóta#weátaw#hérota#

                        #I-cóka-D#mícaN-A#ciwá?a-A#

                        #icókaw#mícara#ciwá?aa#

4.5.   #Adv#ItV#DtV#

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

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

        #eómi#acokapóta#hérota#

Text 5

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

        #there#Trocara-n-to#rec-part#12-go#again#

        #qeY#tokára-A-Nòpi#raka#ca‑ha#no#

        #qéy#tokáraaropi#raka#cáha#no#

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

        #it=to#hortative#12-go#X/3-kill-dep#thing-n#indef#

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

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

5.3.   #Adv#part#N#IiV#

#there#rec-past#motor-n#12-com-go#

#pe#raka#móto-A#ca-(e)ro⸗ha#

#pé#raka#mótoa#caráha#   mótoa < Ptg. motor

5.4.   #N=PP#part#ItV#DtV#N#

#that-n-at#hortative#12-com-arrive#X/3-com-go#motor-n#

#á?e-A-(ð)pe#ke#ce-(e)ro⸗hem#I-(e)ro⸗ha-D#móto-A#

#á?epe#ke#ceróhem#heráhaw#mótoa#

5.5.   #DtV#DtV#N#part#DtV#

#X/3-kill-dep#X/3-com-come-dep#game meat#indef#X/3-eat-

                                                    dep#

#I-cóka-D#I-(e)ro⸗(to)N-D#ma?eá?a-A#mo#I⸗?o-D#

#icókaw#hérota#ma?eá?aa#mo#í?ow#

5.6.   #N#part#part#DtV#

#fish-n#indef#also#X/3-catch-D#

#ipíra-A#mo#we#I-píhɨ(Ŋ)-D#

#ipíraa#mo#we#ipíhɨka#

Text 5.   Translation

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

Text 6

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

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

                         manioc-finisher-n#

#karótaM-A-(r)imo#O-ápo#I-(e)ro-ka#mani?aŋópaM-A#

#karótawarimo#oápo#heréka#mani?aŋópawa#

                         (karótawa < Ptg. Garotão)

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

       #stream-n#3/3-divide#3/3-commitative-here/now#

               manioc-n//holder-to become#

       #ɨharápe-A#O-mánaŊ#I-ero⸱ka#maní?aŋ-A//Nɨro-(r)amo#

       #ɨharápee#ománaŋ#heréka#maní?aŋa//rɨroramo#

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

       #stream-n#3/3-divide-indef#X/3-commitative-here/now#

               Teteŋoa#manioc-n//holder-become#

       #ɨɨqápi-A#O-mánaŊ-(r)imo#I-(e)ro⸱ka#tetéŋo-A#maní?aŋ-
               A//Nɨro-(r)amo#

       #ɨɨqápia#ománaŋimo#heréka#tetéŋoa#maní?aŋa//rɨroramo#

6.4.   #DtV#Aff-PP#N#

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

       #I-mána=mána-(pó)ta(N)-D#I-pɨpe#maní?aŋ-A#

       #imanamanapótata#ipɨpe#maní?aŋa#

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

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

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

       #qéy#ipɨkóytata#océ?iwe#itáwaramo#

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

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

#O⁻(to)N-(pó)ta(N)-(r)imo#I-(ð)hi#I-ápo⁼paM-(r)íre#

#óntarimo#icóhi#iapopawíre#

6.7.   #Itv#N#IiV#N#

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

indef#Garatao-n#

#O-ápo⁼paM-(pó)ta(N)-(r)imo#maní?aŋ-A//Níro-A#O⁻(to)N-

(pó)ta(N)-(r)imo#karótaM-A#

#oapopámtarimo#maní?aŋa//ríroa#óntarimo#karótawa#

Text 6.   Translation

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

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

#DtV#

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

#I-kótoŊ-(pó)ta(N)-D̲#

#ikotóŋtata#

#N-PP-part.#

#what-for-indefinite#

#má?e-A-Nèhe-(r)imo#

#má?eereherimo#

#N-PP#

#3-hair-with#

#I⌐?aM-A-Nèhe#   4,10,12,20

#í?awarehe#

(Note that í?awa 'his hair';. í?am 'he has much hair' from
#I⌐?aM-A# and #I⌐?aM# respectively.)

#night-during#

#ipíton-(r)imo#

#ipítonimo#

#13-dance=finish-when#

#oro-poráhay=paM-(r)íre#

#oroporahaypawíre#   (See Rule 22)

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

#qeY#I⌐ha-I#

#qéy#íhay#   (See Rule 23)

#I-come-dep#

#we⌐(to)N-D#

#wétota#   11,16,20

#12-sit=sit=dep#

```
#cere-ápɨŋ=ápɨŋ-D#      (reduplication)

#cereapɨápɨka#          (See Rule 15)


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

#I-Nópiⁿ=I-Nópiⁿ-D#     (reduplication)

#hopihópita#            (See Rule 15)
```

FOOTNOTES

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

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

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

[3]In a more complete grammar for Asurini, these would be the terminal strings of the syntactic component of a generative

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

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

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

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

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

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

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

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

[10]Such a rule may also be written
              A →   B/ X__Y
We will adopt the other form. We will also adopt the con-
vention that when we find that environments for one side of
the symbol to be rewritten do not need to be specified, no
symbol need appear on that side and 'all possible environ-
ments' is understood for the blank (lack of symbol).

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

[12]The fact that a choice is necessary is a consequence of
the form of the descriptive apparatus. See Lamb (1966).

[13]Historical linguistics generally recognizes this as sup-
pletion or replacement by a different morph in certain en-
vironments in which "...the phoneme discrepancies between
suppletive co-allomorphs (e.g., E go;wend) are not signifi-
cantly repeated in allomorphs making up the other morphemes
of the language." (Hoenigswald 1960.48). This comes about
as a result of syncretism. "...morphs which originally
contrast recur, at a later stage, in phonemically corres-
ponding shape but in complementary distribution and there-
fore as co-allomorphs..." (Hoenigswald 1965.36).

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

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

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

A 1,2,3,4   ordered within the group
                    1,2 IMP; 2,3 IMP; 3,4 INT
B 5,6,7,8,9,10   ordered within the group
                    5,6 IMP; 6,7 IMP; 7,8 INT;
                    8,9 NOT; 9,10 INT
C 11   highly idiosyncratic
D 12
E 13
F 14
G 15
H 16,17,18,19,20,21 (ordering important for the whole set)
                    16,17 NOT; 17,18 INT; 18,19 NOT; 16,17,18,
                    19,20,21 INT
I 22
J 23
K 24-29   24,25 NOT; 25,26 NOT; 26,27 IMP; 27,28 IMP;
                    28,29 INT; 24-27,29 INT
A,B IMP; B,C IMP; A,C IMP; C,D NOT; D,E NOT; E,F NOT;
C,E NOT; C,F NOT; F,G NOT; G,H INT; H,I IMP;
I,J INT; J,K NOT

Bibliography

Bach, Emmon
    1964 An Introduction to Transformational Grammars.
          Holt, Rinehart and Winston, Inc.
Bloch, Bernard
    1941 Phonemic Overlapping. American Speech 16.278-84.
Chomsky, Noam
    1957 Syntactic Structures. Mouton & Co.
    1964 Current Issues in Linguistic Theory. Mouton & Co.
    1965 Aspects of the Theory of Syntax. The MIT Press.
Chomsky, Noam and M. Halle
    1965 Some Controversial Questions in Phonological
          Theory. Journal of Linguistics, Vol. 1, No. 2.

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