TOPICS IN SANSKRIT PHONOLOGY

by

ARNOLD M. ZWICKY, JR.

B.A., Princeton University (1962)

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August, 1965

Signature of Author.....Arnold M. Zwicky, Jr.........
Department of Modern Languages
August 23, 1965

Certified by..
Thesis Supervisor

Accepted by...
Chairman, Departmental Committee
on Graduate Students

TOPICS IN SANSKRIT PHONOLOGY

by

ARNOLD M. ZWICKY, JR.

Submitted to the Department of Modern Languages on Aug. 23, 1965, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

A set of ordered rules accounting for a wide variety of Sanskrit sandhi phenomena are presented and defended. It is demonstrated that these phenomena, which have traditionally been considered as quite separate and unrelated processes, are in fact intimately connected.

Chapter 1 contains a brief formal account of the linguistic principles underlying the analyses in the following chapters.

In Chapter 2 the rules of vowel sandhi are examined. It is shown that these rules must apply cyclically.

In Chapter 3 the traditional rules for the retroflexion of *s* and *n* and for assimilation of retroflexion are restated in terms of distinctive features.

Chapter 4 deals with the relationship of the retroflexion rules to other morphophonemic processes--to internal sandhi rules for roots in *ś* *ṣ*, and *kṣ* and to general spirant sandhi rules. Another set of cyclical rules is discovered, and the traditional treatment of these rules is examined and criticized.

Chapter 5 is concerned with the features of aspiration and voicing. Grassman's Law and Bartholomae's Law are formulated as synchronic rules of Sanskrit, and evidence for considering voiceless aspirates as clusters is adduced.

In Chapter 6 the relationship of sandhi rules for *r* and *s* to the rules in the preceding chapters is examined.

Thesis Supervisor: Morris Halle
Title: Professor of Modern Languages

Acknowledgments

The author is indebted to Morris Halle, whose perceptive criticisms of the material in earlier drafts of this dissertation cannot be valued too highly; to students in the Department of Modern Languages, M.I.T., especially to R. P. V. Kiparsky and Theodore Lightner, who commented upon many specific points; to the U.S. Office of Education, which provided the author with fellowship funds; to Stephen Isard, Phoebe Acheson, T. M. Scanlon, and Lucy Scanlon, whose proof-reading abilities were much appreciated; and to Ann Zwicky, who spent long hours typing this work and encouraging its author.

Table of Contents

CHAPTER 1

Introduction

1.1. Purpose and Scope

In this work we present a set of rules for the description of certain aspects of Sanskrit phonology. It is our aim to demonstrate that if rules are formulated in terms of phonological features, rather than in terms of phonemes or similar units, and if the set of rules is considered as having an order imposed upon it, then significant economies of description can be achieved and apparently unrelated phenomena can often be accounted for by a single very general rule.

We shall show that a number of Sanskrit sandhi rules--particularly those dealing with vowel-vowel sandhi, retroflexion, voicing, and aspiration--are, in fact, intimately connected. In the course of this demonstration we shall have occasion to comment on various formal features of phonological representations and rules.

1.2 Sources

There is no lack of Sanskrit data. On the contrary, there is an immense amount of venerable material, both linguistic and metalinguistic--the Vedas, the epics, fables, drama, legal and philosophical treatises, treatises on phonetics and grammar, commentaries on these, and commentaries on the commentaries. This body of literature was composed over a period of not less than 3500 years, during which time the language changed considerably and many regional dialects developed. One form of Sanskrit, fixed (more or less) by the great grammarians as a learned language, continues in use today. However, true native speakers of Classical Sanskrit do not exist, and have not existed for over two millenia. The language described here is the Classical Sanskrit of the great grammarians.

The purpose of this work is not the elucidation, interpretation, or analysis of the texts. Nor is it a commentary on Pāṇini, or even a reinterpretation of Pāṇini in terms of generative phonology. Sanskrit textual analysis has a long history, as does the discussion of grammar (chiefly morphology and phonology) in the Pāṇinean framework. However, the principal results of such investigations are available in numerous handbooks, and we shall have little to add to them. Rather we shall take the handbook rules as our starting point. Our examples are either drawn from the handbooks (hence, ultimately from the texts) or else created according to well-known rules. The sources of these examples will not be indicated, although doubtful forms and particularly rare phenomena will be marked as such.

Our principal sources are Whitney 1960 and Emeneau 1958 (for statements of the traditional rules) and Whitney 1885 and Monier-Williams 1899 (for lists of forms). Whitney's rules, like those of Renou 1961, are notable for their exhaustiveness; nearly every type of exception is noted, and most are exemplified.

Emeneau's rules have the opposite virtue: they are concise and quotable. In matters of phonetic detail we are indebted to Allen 1961.

1.3. Linguistic Framework

The general linguistic framework for our approach to Sanskrit phonology is the Jakobsonian system of distinctive features as incorporated into Chomsky's system of generative grammar by Halle and Chomsky.[1] In this framework a _grammar_ consists of three components--a _syntactic_ component that generates a set of _deep_ structures for the sentences of a language and converts these deep structures into an (infinite) set of _surface_ structures, a _semantic_ component that operates upon the deep structures to yield semantic interpretations, and a _phonological_ component that operates upon the surface structures to yield phonological interpretations (transcriptions in a universal phonetic alphabet).

1.3.1. The Input to the Phonological Component

We assume that the surface structure of a sentence can be represented as a labeled bracketing of a string of elements called _formatives_. Some formatives have associated with them sets of function-values, or _markings_.

Some markings are at least partially syntactic

---

[1] The discussion of generative phonology in Chomsky 1964 includes a bibliography of this development.

or morphemic in character; the animateness of nouns and the conjugation class of verbs are examples. Let us call these the unit markings of a formative. The set of unit markings of a formative has its source in sets of markings associated with lexical items inserted into deep structures. We leave open the possibility that in the mapping of deep structures into surface structures unit markings may be added, changed, or referred to.

The remaining markings of a formative--the segmental markings--are the specifically phonological ones. Some formatives are divided into one or more segments, each segment being a (possibly null[2]) set of segmental markings. Segmental markings, like unit markings, derive from markings associated with lexical items; segmental markings, however, are not affected by the operation of transformational rules, or referred to by such rules.

So that we can conveniently refer to unit markings in phonological rules, we make the convention that if y is a unit marking of the formative x, then every segment in x has y as a member. Thus, if a root has the marking 3declension (i.e., "third declension") every segment in the root also has this marking.

---

[2] It is certainly possible that there are languages in which all the features of a segment are predictable in certain positions, although (to our knowledge) no such example has been pointed out in the literature.

A function of which a marking is a value is a <u>feature</u>; a <u>(classificatory) distinctive feature</u> is a function having as its value a segmental marking of a formative. Classificatory distinctive features are binary; that is, they are functions onto a two-element set. The other features are not necessarily binary.

We have left open the possibilities that some formatives have no markings and that some formatives have only unit markings. Such grammatical elements as tense or case morphemes may be of one or the other of these types in some languages. The <u>word boundary</u> # is often assumed to be a formative with no markings; however, we shall present some arguments for considering word boundaries to be sets of segmental markings. # sometimes separates formatives that are not words.

A string given as input to the phonological component will include labeled brackets specifying the higher constituent structure of the formatives. Some phonological rules will refer to these labeled brackets. However, many rules will not, so that it is convenient to be able to refer to parts of a string without having to indicate which particular brackets might occur within the string. To this end the <u>morpheme boundary</u> + is used as a general separator of formatives; it stands for any sequence (including the null sequence) of brackets between two formatives. Thus, for many purposes we can consider each input to be a string consisting of +'s, formatives without segmental markings, and segments. + sometimes separates formatives that are not constituents.

1.3.2. The Classificatory Distinctive Features

Following Jakobson, Fant, and Halle 1961, the set of classificatory distinctive features in any particular language is assumed to be a subset of a small universal set of features. The features to which we shall refer in this work are the following:

1. Consonantality, abbreviated cons. Liquids and true consonants are +cons; glides and vowels are -cons.

2. Vocalicity, abbreviated voc. Liquids and vowels are +voc; glides and true consonants are -voc.

3. Obstruence, abbreviated obst. Non-nasal true consonants are +obst; all other segments are -obst.

4. Continuance, abbreviated cont. Segments with oral closure (stops, nasals, the Sanskrit liquids) are -cont; all others (spirants, vowels, glides) are +cont.

5. Gravity, abbreviated grv. Velar and labial consonants and back vowels are +grv; dental and palatal consonants and front vowels are -grv.

6. Compactness, abbreviated comp. Back (velar and palatal) consonants and low vowels are +comp; front (dental and labial) consonants and mid and high vowels are -comp.

7. Diffuseness, abbreviated diff. High vowels and noncompact consonants are +diff; mid and low vowels and compact consonants are -diff.

8. Nasality, abbreviated nas.

9. Voicing, abbreviated vcd.

10. Tenseness, abbreviated tns. Segments with notable muscular tenseness (aspirated stops, some so-called "long" vowels, and some spirants) are +tns; all other segments are -tns.

11. Flatness, abbreviated fl. Segments with secondary constrictions at the periphery of the oral cavity (labialized, pharyngealized, retroflected segments) are +fl; all other segments are -fl.

12. Stridency, abbreviated str. Segments with high turbulence (affricates, s as opposed to θ, etc.) are +str; all other segments are -str.

1.3.3. Phonemes and Segments

In general a segment is not marked for every distinctive feature; phonological rules operate to fill in (and change) markings. From the set of all segments in lexical entries (or, equivalently, from the set of all segments in strings given as input to the phonological component) we can abstract a set of segment-types called systematic phonemes (or, for short, phonemes):

> If x occurs as a segment in a lexical entry, and if no other segment in a

> lexical entry includes x (in the set-theoretic sense), then x is a *(systematic) phoneme*.

Every segment can then be viewed as an instance of at least one phoneme.

As a general requirement on the phonological system of any language we require that each phoneme be *distinct* from every other phoneme, in the following sense:

> The segment x is *distinct* from the segment y if there is a distinctive feature f such that $\underline{+f}$ is in x and $\underline{-f}$ is in y or such that $\underline{-f}$ is in x and $\underline{+f}$ is in y.

The notation for a segment with the markings $m_1, m_2, \ldots, m_n$ is

$$\begin{bmatrix} m_1 \\ m_2 \\ \cdot \\ \cdot \\ \cdot \\ m_n \end{bmatrix}$$

It is often convenient, however, to refer to segments by single symbols. For this purpose we use underlined Latin letters with various diacritics. It is to be understood that these symbols are abbreviations for sets of markings.

1.3.4. The Operation of the Phonological Component

The phonological component consists of an ordered set of rules. These rules map the input to the phonological component into a narrow phonetic transcription. In the course of this mapping the binary classificatory features are replaced by many-valued phonetic features. The nature of the particular Sanskrit problems dealt with in this work, however, is such that the distinction between the original binary features and the resultant many-valued features is not important. Moreover, there is considerable (and not very well understood) dialectal and historical variation in the phonetic details of Sanskrit, so that it would be pointless to attempt to carry the mapping beyond a certain stage.

The rules may be grouped into three sets: (a) redundancy and morpheme structure rules, which fill in a marking for every feature in every segment, either on the basis of markings already in a segment or on the basis of markings in adjoining segments; (b) morphophonemic rules, which change markings in segments, insert segments, delete segments, and combine segments, sometimes using the information in unit markings of labeled bracketing; (c) late phonetic rules, which produce narrow phonetic transcriptions. Our interest in this work is almost entirely with the rules of set (b), although we shall have a few comments to make about the rules of set (a).

1.3.4.1. The Form of the Rules

1.3.4.1.1. General Conventions

A typical rule is of the form

$$s_1 \longrightarrow s_2 \,/\, e_1 \_\_\_ e_2 ,$$

read "s_1 is rewritten as s_2 in the environment after e_1 and before e_2." A rule may lack a left environment or a right environment, or both; a rule that lacks both is written as

$$s_1 \longrightarrow s_2 .$$

The s_2 in a rule is a segment. The s_1 is either (a) a formative without segmental markings or (b) a segment. In case (a) the rule prescribes that the formative s_1 be given the phonological shape s_2; in case (b) the rule prescribes an addition to or modification of the markings in s_1. We shall be concerned primarily with rules of type (b). The e_1 and e_2 in a rule are strings consisting of segments, boundaries, labeled brackets, and formatives without segmental markings.

A type (b) rule

$$s_1 \longrightarrow s_2 \,/\, e_1 \_\_\_ e_2$$

is said to be applicable to a segment s_3 occurring in a string S if the following conditions are satisfied:

1. s_3 is not distinct from s_1, in the sense of 1.3.3 above.

2. Immediately to the left of s_3 there is a substring e_3 of S which is <u>subsumed</u> by e_1, in the following sense:
   a. If x is a boundary, a labeled bracket, or a formative without segmental markings, then it is subsumed by x, by +x, and by x+.
   b. If x is a segment and y is a segment not distinct from x, then x is subsumed by y, +y and y+.
   c. If x_1 is subsumed by y_1 and x_2 is subsumed by y_2, then x_1x_2 is subsumed by y_1y_2.
3. Immediately to the right of s_3 there is a substring e_4 of S which is subsumed by e_2.

When the rule is applied to s_3, the markings in s_2 replace the corresponding markings in s_3; that is, if s_2 contains a marking m for some feature f, then m replaces whatever marking f has in s_3. When a rule is applied to a string S, it is applied (simultaneously) to every segment in S to which the rule is applicable and yields a resultant string S'. We indicate that Rule X when applied to S yields S' by writing

$$S \xrightarrow[X]{} S' \ .$$

1.3.4.1.2. Abbreviatory Notations

In addition to the general conventions on the form and interpretation of rules, there are a number of conventions about abbreviations that we shall use throughout this work.

First, curly brackets are used to indicate a logical disjunction. For example, $\left[\begin{Bmatrix}\text{-cons}\\ \text{+voc}\end{Bmatrix}\right]$ is used to specify that a segment be marked either -cons or +voc-- that is, that the segment be a non-vowel. And the environment

$$/ \begin{Bmatrix} \begin{Bmatrix} + \\ \# \end{Bmatrix} \ \_\_ \ [\text{-voc}] \\ [_{V} \ \_\_ \ \begin{Bmatrix} \text{PST} \\ \text{PF} \end{Bmatrix} \end{Bmatrix}$$

is an abbreviation of the four environments

$$/ \ + \ \_\_ \ [\text{-voc}]$$
$$/ \ \# \ \_\_ \ [\text{-voc}]$$
$$/ \ [_{V} \ \_\_ \text{PST}$$
$$/ \ [_{V} \ \_\_ \ \text{PF}$$

There is no particular difficulty of interpretation when curly brackets occur within segments.[3]

[3] It is possible to imagine situations in which different sets of curly brackets within different segments in a rule might be required to be "expanded" in

When curly brackets are used in other situations, however, they must be assumed to be a device for abbreviating a set of ordered rules into a single "rule." In order to apply such a "rule," we must first expand the "rule" into its component rules and then apply each of these in turn. In the general case the principles of expansion are rather complex. Our use of curly brackets is sufficiently restricted that a correct expansion of any given "rule" will usually follow from the application of the principles (a) expand outer brackets first, and (b) $x\begin{Bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{Bmatrix}z$ expands into

$$\begin{matrix} xy_1z \\ xy_2z \\ \cdot \\ \cdot \\ \cdot \\ xy_nz \end{matrix} .$$

In the ambiguous cases--when one pair of brackets

---

a particular order. A hypothetical example is cited in McCawley 1965, p. 39. For us, however, curly brackets within segments are merely a device for referring to the complement of a class of segments; $\begin{bmatrix} \begin{Bmatrix} -\text{cons} \\ +\text{voc} \end{Bmatrix} \end{bmatrix}$ is used to refer to the complement of the class of all segments not distinct from $\begin{bmatrix} +\text{cons} \\ -\text{voc} \end{bmatrix}$.

neither includes nor is included in another pair--it is usually immaterial, for our purposes, which of the possible expansions is chosen. When we find it necessary to specify a particular order in these cases, we will number the brackets; brackets with the same number are expanded together (so that $\left\{\begin{matrix}x\\y\end{matrix}\right\}_n \left\{\begin{matrix}u\\v\end{matrix}\right\}_n$ expands into $\begin{matrix}xu\\yv\end{matrix}$), and a pair of brackets numbered n is expanded before a pair of brackets numbered n+1 (so that $\left\{\begin{matrix}x\\y\end{matrix}\right\}_n \left\{\begin{matrix}u\\v\end{matrix}\right\}_{n+1}$ expands into $\begin{matrix}xu\\xv\\yu\\yv\end{matrix}$, and $\left\{\begin{matrix}x\\y\end{matrix}\right\}_{n+1} \left\{\begin{matrix}u\\v\end{matrix}\right\}_n$ expands into $\begin{matrix}xu\\yu\\xv\\yv\end{matrix}$).

Second, parentheses are used to indicate optional parts of environments. A "rule" with parentheses in its environments must be expanded into a set of ordered rules. It is sufficient for our purposes to require that the rules be ordered by decreasing length of environment, so that the environment

$$/\ (x(y))\ \_\_\_\ z$$

expands into

$$/\ xy\ \_\_\_\ z$$
$$/\ x\ \_\_\_\ z$$
$$/\ \ \_\_\_\ z\ .$$

Parentheses may be numbered in the same way as curly brackets.

Third, numerical subscripts and superscripts are used to indicate a sequence of segments all of the same type. A subscript specifies a lower bound on the number of segments, a superscript an upper bound. If no subscript is used, the lower bound is 1. If a subscript, but no superscript, is used, the upper bound is (in principle) infinite. Thus, $[+\text{cons}]_0$ indicates a sequence of zero or more segments marked +cons, $\begin{bmatrix}-\text{cons}\\+\text{voc}\end{bmatrix}^2$ indicates either one or two vowels, $[\quad]_2^4$ indicates a sequence of from two to four segments of any kind. The use of subscripts and superscripts is equivalent to the use of parentheses, and a "rule" containing subscripts or superscripts in environments must be expanded into an ordered set of rules.[4] For example, the environment

$$/ \;\_\_\; [-\text{cons}]_0^3\ [+\text{obst}]$$

expands into

$$/ \;\_\_\; [-\text{cons}]\,[-\text{cons}]\,[-\text{cons}]\,[+\text{obst}]$$
$$/ \;\_\_\; [-\text{cons}]\,[-\text{cons}]\,[+\text{obst}]$$
$$/ \;\_\_\; [-\text{cons}]\ [+\text{obst}]$$
$$/ \;\_\_\; [+\text{obst}]$$

---

[4] If the upper bound on the number of occurrences of some segment in the environment of a "rule" is infinite, the "rule" then abbreviates an infinite number of rules. In this case the use of a subscript is no longer a simple notation for abbreviating a set of rules. However, the applicability of the "rule" to any given string is still decidable by the principles discussed above.

From this point on we shall cease distinguishing between *rules* (in the original sense of 1.3.4.1.1) and "*rules*" that are, presumably, abbreviations for ordered sets of rules. The reference of 'rule' in the following sections should be clear from the context.

Fourth, Greek letters are used as variables over + and - to indicate the agreement (or nonagreement) of markings within a segment or in different segments and also to indicate assimilation and dissimilation of segments with respect to certain features. Examples of the use of Greek variables: (a) $\begin{bmatrix}\alpha\text{cons}\\ -\alpha\text{voc}\end{bmatrix}$ indicates a segment that is either $\begin{bmatrix}+\text{cons}\\ -\text{voc}\end{bmatrix}$ or $\begin{bmatrix}-\text{cons}\\ +\text{voc}\end{bmatrix}$--i.e., a vowel or a true consonant; (b) $\begin{bmatrix}\alpha\text{voc}\\ +\text{cons}\end{bmatrix} \rightarrow [+\text{vcd}]\ /\ \_\_\_\ [\alpha\text{cons}]$ is a rule specifying that liquids are voiced before liquids and true consonants and that true consonants are voiced before vowels and glides; (c) $[+\text{obst}] \rightarrow [\alpha\text{vcd}]\ /\ \_\_\_[\alpha\text{vcd}]$ is a rule specifying that obstruents are voiced before voiced segments, and voiceless before voiceless segments; (d) $\begin{bmatrix}+\text{voc}\\ \alpha\text{grv}\end{bmatrix} \rightarrow [-\alpha\text{grv}]\ /\ \_\_\_\ [+\text{cons}]_0 \begin{bmatrix}+\text{voc}\\ -\text{comp}\end{bmatrix}$ is a rule specifying that a back vowel is fronted, and that a front vowel is backed, when a mid or high vowel follows it, either immediately or with a consonant cluster intervening. In cases like (a), (b), and (c) the use of Greek letters is eliminable in favor of an expanded set of rules. In cases like (d), however, there is no general procedure for expanding the rule into an ordered set of rules that will have the same effect as the original rule, so that the use of Greek letters is not entirely a notational convention.

Fifth, some rules will insert, delete, combine, or rearrange segments. The interpretation of such rules does not follow immediately from the principles in 1.3.4.1.1. The insertion of the segment s in the environment after e_1 and before e_2 is indicated by

$$\emptyset \rightarrow s \,/\, e_1 \_\_ e_2 ,$$

and the deletion of s in the same environment by

$$s \rightarrow \emptyset \,/\, e_1 \_\_ e_2 .$$

The symbol Ø then stands for a special <u>null</u> segment. The conventions and definitions in the previous section could be adjusted to take account of the null segment. However, the use of Ø can be combined with the use of Greek letters--in, for example, the rule

$$\emptyset \rightarrow \begin{bmatrix} +\text{obst} \\ -\text{cont} \\ \alpha\,\text{comp} \\ \beta\,\text{grv} \\ \gamma\,\text{fl} \end{bmatrix} \Big/ \begin{bmatrix} +\text{cons} \\ +\text{nas} \\ \alpha\,\text{comp} \\ \beta\,\text{grv} \\ \gamma\,\text{fl} \end{bmatrix} \_\_ \begin{bmatrix} +\text{obst} \\ +\text{cont} \end{bmatrix} ,$$

which inserts, between a nasal consonant and a following spirant, a stop homorganic with the nasal--in such a way that the replacement of a rule with Greek letters by a set of rules without them can be achieved only at the cost of great complexity.

Combination and rearrangement are two processes that cannot, in general, be described by rules of the ordinary sort. Simple rearrangement (metathesis) is described by rules like

$$\underset{1}{\begin{bmatrix}+\text{voc}\\-\text{cons}\end{bmatrix}}\underset{2}{\begin{bmatrix}-\text{voc}\\-\text{cons}\end{bmatrix}} \rightarrow 2\ 1 / \ \underline{\quad}\begin{bmatrix}+\text{cons}\\+\text{nas}\end{bmatrix},$$

which specifies the metathesis of a vowel and a glide before a nasal consonant. This technique of numbering segments can be carried over to the description of combinatory processes, in which several segments each contribute markings to the resultant segment; for example,

$$\underset{1}{\begin{bmatrix}+\text{voc}\\-\text{cons}\end{bmatrix}}\underset{2}{\begin{bmatrix}+\text{cons}\\+\text{nas}\end{bmatrix}} \rightarrow \begin{bmatrix}1\\+\text{nas}\end{bmatrix}$$

specifies the combination of a vowel and following nasal consonant into the nasalized vowel corresponding to the original vowel, and

$$\underset{1}{\begin{bmatrix}+\text{voc}\\-\text{cons}\\\alpha\text{grv}\\\beta\text{comp}\\\gamma\text{fl}\end{bmatrix}}\ 1\ (1) \rightarrow \begin{bmatrix}1\\+\text{tns}\end{bmatrix}$$

specifies the coalescence of two or three vowels with the same gravity, compactness, and flatness into a single tense vowel.

1.3.4.2. The Ordering of the Rules

Let us assume as a first approximation that the set of rules is linearly ordered--in other words, that consecutive positive integers are assigned to the rules and that the operation of the phonological component proceeds as follows:

1. Begin with an input string as the <u>current string</u> and with Rule 1 as the <u>current rule</u>.
2. Test the current string S for the applicability of the current rule, Rule n;
   a. If the rule is not applicable, Rule n+1 (if there is one) the current rule.
   b. If the rule is applicable, and if

   $$S \xrightarrow[n]{} S' ,$$

   then make S' the current string and Rule n+1 (if there is one) the current rule. Note that Rule n may have applied <u>vacuously</u>, i.e., that S and S' may be identical.

3. If there is no Rule n+1, the current string is an output string; otherwise, go back to step 2.

We shall argue in later sections for a somewhat more complicated ordering, in which a subset of the full set of rules is distinguished as a cycle. Suppose that Rule i is the first rule of the cycle, that Rule j is the last rule, and that S is a string to which Rule i is to apply. Then the operation of the cyclical rules proceeds as follows:

1. S is the current string.
2. Underline the largest subparts of the current string which contain no labeled brackets. Rule i is the current rule.
3. Apply the rules in the cycle, as above, but only to the underlined portions of the current string. The resultant of an underlined portion is also underlined.
4. If the current rule is j+1, then
   a. If the current string contains no labeled brackets, the operation of the cycle is completed.
   b. Otherwise, erase each pair of brackets surrounding an under-

lined portion of the current string, and go back to Step 2.

In the operation of the cycle, then, the same set of rules is applied again and again to larger and larger constituents.

It is possible to make rules applicable only during certain passes through the cycle. A cyclical rule can be restricted to application to <u>words</u> by the addition of the environment

$$/ \_\_ [\ ]_0 \#\#$$

to the rule, or to application on the last pass through the cycle by the addition of the environment

$$/ \_\_ [\ ]_0 \$ \quad ,$$

where $ is a sentence boundary.

In addition to the cycle, there is a set of <u>pre-cyclic</u> rules and (possibly) a set of <u>post-cyclic</u> rules. However, a presumably post-cyclic rule can be included within the cycle by adding to it the "last pass" environment above. Or, better, the entire set of post-cyclic rules can be included within the cycle by surrounding the set by curly brackets and adding the "last pass" environment only once for the entire set. There seems to be no substantive issue involved in the choice between post-cyclic rules and added environments. We have chosen the first of the alternatives.

More general questions about the nature of the ordering of rules and about the justification of particular orderings will be discussed in later sections.

1.4. Organization

In Section 1.3. we presented formal characterizations of some important phonological concepts and discussed, in less detail, the notations we will use in later sections. It is not our purpose to justify the linguistic theory underlying this work, or even to justify our particular version of that theory. Rather we are concerned with the application of the theory to a set of problems.

Our procedure will be to discuss one problem, provide a tentative solution to that problem, adduce evidence (or counter-evidence) for the solution, and consider related problems. In the course of this procedure any rule previously set up is liable to be changed, as are the set of Sanskrit phonemes (in the sense of Section 1.3.3.) and the content of lexical entries.

Five problem areas will be discussed--vowel sandhi (Chapter 2), retroflexion (Chapter 3), roots in *ṣ*, *s*, and *kṣ* (Chapter 4), aspiration and voicing (Chapter 5), and the sandhi of s and r (Chapter 6). The rules are grouped into six sets, each set having a mnemonic tag:

VIS:	visarga
VOW:	vowels
RET:	retroflexion
PAL:	palatals
SP:	spirants
AV:	aspiration and voicing
RS:	r and s

Within a set the rules are numbered in the order of their presentation. The order in which rules are to be applied is indicated in the text. A statement of the ordering relations of the final set of rules is given in an appendix.

1.5. Transcription

The segments to which we shall have reference are listed below with their markings.

+cons -voc									
+comp -diff			-comp +diff						
+grv -fl	-grv -fl	-grv +fl	-grv -fl	+grv -fl					
k	c	ṭ	t	p	-tns	-vcd	-nas	+obst	-cont
k^h	c^h	$ṭ^h$	t^h	p^h	+tns	-vcd	-nas	+obst	-cont
g	j	ḍ	d	b	-tns	+vcd	-nas	+obst	-cont
g^h	j^h	$ḍ^h$	d^h	b^h	+tns	+vcd	-nas	+obst	-cont
n̄	ń	ṇ	n	m	-tns	+vcd	+nas	-obst	-cont
x	ś	ṣ	s	f	+tns	-vcd	-nas	+obst	+cont

+cons
+voc
-obst
-nas
+vcd
-tns

+comp -diff -grv +fl +cont	-comp +diff +grv -fl -cont
<u>r</u>	<u>l</u>

-cons
-voc
-obst
-nas
+cont

+tns +comp -diff +grv -fl		-tns -comp +diff	
+vcd	-vcd	-grv -fl +vcd	+grv +fl +vcd
<u>h</u>	<u>ḥ</u>	<u>y</u>	<u>w</u>

-cons
+voc
-obst
+cont
+vcd

-grv -fl -comp			+grv				
			-fl +comp		+fl -comp		
+diff		-diff	-diff		+diff		-diff
-tns	+tns	+tns	-tns	+tns	-tns	+tns	+tns
<u>i</u>	<u>ī</u>	<u>e</u>	<u>a</u>	<u>ā</u>	<u>u</u>	<u>ū</u>	<u>o</u>

Also: 1. Nasalized (+nas) forms of all vowels; e.g., *ã*

2. Diphthongs *āw* and *āy*.

3. The boundary #, having at least the segmental markings +obst and -cont, and distinguished from all other segments by having the unit marking +BND where other segments have -BND. The boundary + has no markings. The boundary $ may be considered to have segmental markings, although such markings would play no part in our rules.

4. A distinction between segments occurring in a verbal root, indicated by the unit marking +ROOT, and other segments. Our use of ROOT in rules could be replaced by reference to additional segments and brackets or boundaries.

The markings in the charts above are the markings in segments when the first morphophonemic rule applies. However, not all the segments in the charts above will actually occur at this point in the operation of the phonological component. It is well known that all occurrences of *x*, *f*, *ḥ*, and the nasalized vowels are predictable. In later chapters we shall argue that most occurrences of retroflex consonants and all occurrences of e, o, and the tense vowels are also predictable, and that the voiceless aspirated stops should be considered as being derived from clusters. Finally, ñ and the palatal stops are undoubtedly predictable, although we shall not treat these problems.

A number of the markings in the charts merit some discussion. The following sections treat these markings individually.

1.5.1. The Retroflex Consonants and r

The retroflex consonants and r are flat palatals, not flat dentals (or alveolars). With respect to the retroflex consonants, it is difficult to determine from the statements of the Indian phoneticians and grammarians exactly where the closure-point was located. The usual term for the retroflex consonants is mūrdhanya[5] 'on or in the head, highest'; although the turning back of the tip of the tongue is often remarked upon, the point at which the tongue makes contact with the roof of the mouth is not specified.

The traditional use of mūrdhanya suggests that the closure point may have been the highest part of the palate. Allen, commenting upon this interpretation as

---

[5] Native grammatical terminology and the names of grammarians and grammatical treatises are cited in a traditional system of transcription that differs to some extent from the transcription we use in citing forms.

found in the Tribhāṣyaratna, declares that "there is no evidence that the word was ever used in this special sense."[6] On the other hand, neither is there any evidence that the word was not used in this special sense, which is the interpretation placed upon mūrdhanya by most Western commentators: Thus Whitney, in describing the Sanskrit palatals, says that "they [the palatals] seem to have been... brought forward in the mouth from the guttural point, and made against the hard palate at a point not far from the lingual [retroflex] one..., but with the upper flat surface of the tongue instead of its point."[7]

On general phonetic grounds it is to be expected that the closure-point would be closer to the top of the hard palate than to the teeth, for as the closure-point is moved back from the teeth to the top of the hard palate, the distinctive quality of retroflexion becomes increasingly apparent (as a result of an increase in the degree of the pharyngeal constriction that accompanies retroflexion). A retroflex stop made with closure

---

6 Allen 1961, p. 53.

7 Whitney 1960, p. 16.

at the top of the hard palate is, in a sense, maximally retroflected, whereas the acoustic difference between a retroflex alveolar stop and an ordinary alveolar stop is not very great. In fact, a palatal closure-point for retroflex stops seems to be universal.[8] We shall therefore consider the Sanskrit retroflex stops to be phonetically palatal.

Determining the phonetic nature of r is a more difficult problem, for the texts prescribe a wide variety of articulations for r. According to Varma, the Prātiśākhyas agree in prescribing a dental, gingival, or alveolar articulation, while some of the Śikṣās (in particular, the Pāṇinīya Śikṣā) describe r as mūrdhanya.[9] Undoubtedly, each type of articulation was represented in some dialect at some time.

When we consider the status of the retroflex consonants and r with respect to morphophonemic alternations in Sanskrit, we find good evidence for viewing these segments as retroflected palatals. In Chapter 3 we show that both the retroflex consonants and the (ordinary) palatals should be marked +comp, -grv. Then the interchanges of r and ṣ described in Chapter 6 indicate that r should also have these markings; for some dialects a

---

[8] Heffner 1964, p. 128.

[9] Varma 1929, pp. 6-9.

late phonetic rule must be added to change the position markings.

We should add that since the closure-point for the retroflex consonants probably was somewhat to the front of the closure-point for the (ordinary) palatals, the two sets of segments could be distinguished by their markings for an additional feature of position, rather than by their markings for flatness, a feature of secondary articulation. Such an approach would entail only minor changes in our rules.

1.5.2. Stridency

The only segments for which a stridency marking might be relevant are the spirants and the palatal stops.

The nongrave spirants are undoubtedly strident. About the grave spirants, however, there is some question. Since they are sandhi-variants of *s*, they will be marked as strident unless their stridency marking is specifically adjusted. Allen suggests that at least the labial spirant is nonstrident, because the native term for this segment is *upadhmānīya*, "literally 'blowing upon'--the consecrated description of the voiceless bilabial fricative,"[10] nonstrident φ as opposed

---

10 Allen 1961, p. 50.

to strident f; the native term for the velar spirant--*jihvāmūlīya*, 'formed at the root of the tongue' (i.e., velar)--is less informative. At any rate, the question is not of great moment, since a decision one way or the other involves only a very minor change in our rules.

The palatal stops, although pronounced as affricates by most Western students of Sanskrit, seem to have been nonstrident.[11] Whitney concurs in this view, although he adds that

> Such sounds, in all languages, pass easily into the (English) *ch*- and *j*-sounds. The value of *ch* as making the preceding vowel "long by position"... and its frequent origination from *t* + *ç*... lead to the suspicion that it, at least, may have had this character from the beginning.[12]

If any one of these segments is phonetically strident it acquires this marking at a rather late stage in the grammar. For morphophonemic purposes, we consider the palatal stops to be nonstrident.

1.5.3. Tenseness

Four sets of segments are marked as tense:

---

11 Ibid., p. 52.

12 Whitney 1960, p. 16. Whitney's *ç* is the same as our *ś*.

(a) the aspirated stops, (b) the spirants, (c) the glides *h* and *ḥ*, and (d) the "long vowels" and the "diphthongs" *e* and *o*.

Tenseness is the natural feature for distinguishing aspirate stops from nonaspirate stops,[13] and this use of the feature requires no special comment.

Ceteris paribus, strident segments are tenser than nonstrident segments, and voiceless strident segments tenser than voiced strident segments. On this basis we have chosen to mark the (voiceless strident) spirants of Sanskrit as tense. It then follows that the spirants and the aspirate stops form a "natural class" specifiable as $\begin{bmatrix}+\text{obst}\\+\text{tns}\end{bmatrix}$; in Chapter 5 we shall have occasion to refer to this class.

In the Jakobsonian theory of distinctive features, tense *ḥ* is opposed to the even onset or decay of a vowel (our boundary #) or to the glottal catch.[14] Following this general principle, and noting that *ḥ* occurs as a predictable sandhi-variant of (tense) *s* (Chapter 6), we mark *ḥ* as tense. In Chapter 5 we shall identify the aspiration of voiceless aspirated stops with this *ḥ*. As for (voiced) *h*, it may or may not

---

13 Jakobson, Fant, and Halle 1961, pp. 38-9.

14 Ibid., p. 39.

have been phonetically tense. However, its role in morphophonemic processes (see the next section) indicates that throughout the part of the grammar with which we are concerned it should be marked as tense.

In the case of the "long" vowels <u>ā</u>, <u>ī</u>, <u>ū</u>, as opposed to <u>a</u>, <u>i</u>, <u>u</u>, we could choose either tenseness or length as the distinguishing feature. The decision is of no great importance for our purposes, because we shall derive <u>ā</u>, <u>ī</u>, <u>ū</u> from sequences of "short" vowels at a fairly late stage in the grammar (Chapter 2). Evidence that the "long" vowels differed in quality from the corresponding "short" vowels could be construed as evidence that tenseness rather than length is the relevant distinguishing feature, although there is no <u>a priori</u> reason why secondary differences in quality should not result from primary differences in quantity. Sanskrit <u>a</u> and <u>ā</u>, at least, exhibited a quality difference.[15]

We shall derive the "diphthongs" <u>e</u> and <u>o</u> from the clusters <u>ai</u> and <u>au</u>, so that the remarks in the previous paragraph apply to these segments as well as to the "long" vowels. We need only add that "from

---

15 Allen 1961, pp. 57-8; Whitney 1960, p. 10. Allen finds no evidence for a quality difference in the case of the diffuse vowels.

the phonetic standpoint e/o are represented at a still comparatively early period by simple long vowels intermediate in quality between aa and ii/uu."[16]

1.5.4. The Glides

The glides h/ḥ, y, and w have the gravity, flatness, compactness, and diffuseness markings of a, i, and u, respectively.

It seems clear that y was a phonetic glide. The nature of w (usual transcription: v) is less clear: "whilst its earlier pronunciation was doubtless as a bilabial [w], it had by the time of our treatises acquired, at least in some dialects, the labio-dental articulation [v] which is typical of many modern Indo-Aryan languages."[17] However, the parallel between the alternations i/y and u/w (Chapter 2) indicates that until a very late stage in the grammar w is a glide.

That ḥ was voiceless (IPA h) and h voiced (IPA ɦ) is well established by the remarks of the Indian phoneticians[18]--in particular by their identification of ḥ with the aspiration of voiceless aspirates

---

16 Allen 1961, p. 63. Allen cites some evidence on p. 64.

17 Ibid., p. 57.

18 Ibid., pp. 33-9.

and of *h* with the aspiration of voiced aspirates. The *h*, in a grammar of Sanskrit more complete than the set of rules offered in this work, would in fact be derived from g^h, with which it alternates. Although *h* is marked as a sonorant in our charts, this marking is purely phonetic; from a morphophonemic point of view, *h* is an obstruent.

1.5.5. The Syllabic Liquids

Our charts do not include markings for *ṛ*, *ṝ*, and *ḷ*--the traditional "syllabic liquids." We treat these segments as occurrences of *r*, *rr*, and *l*, respectively, which are syllabic by position (that is, by virtue of their occurrence in contexts in which they are more sonorous than the adjacent segments).

Chapter 2

Vowel Sandhi

2.1. Internal and External Sandhi

It is customary in Sanskrit grammar to make a distinction between external sandhi--processes operating between successive words in a sentence and between successive elements of a compound word--and internal sandhi--processes operating between parts of an (uncompounded) word and possibly also[1] within a morpheme. We carry over this terminology, with the following formal correlates of the internal/external distinction: (a) morphemes within an (uncompounded) word are separated by +; (b) successive elements of a compound are separated by + # +; (c) successive words are separated by + # + # + (one # contributed by each word). We shall ignore instances of + flanking a #. Then an internal sandhi rule applies across +, if it is specifically intermorphemic (see footnote 1 below), or across null, if it is both intermorphemic and intra-

---

1 See our definition of subsumed under in Section 1.3.4.1.1., which includes a general convention that a rule that applies within a morpheme also applies across morpheme buundaries. The converse of this convention does not generally obtain. However, many of our rules will be stated as intramorphemic rules, which are then by extension also intermorphemic rules; this approach contrasts with that of the Indian grammarians, who were primarily concerned with junctural phenomena and did not normally provide for the application of rules within morphemes.

morphemic, while an external sandhi rule applies across #, if it applies only in compound words, or across #(#) if it is general.

2.2. Internal Vowel Sandhi

We begin by discussing the rules of internal vowel sandhi. These rules are somewhat simpler than those for external vowel sandhi. Moreover, the principal rules are the same in both situations. Emeneau's formulation of the rules[2]:

E35 =E1. "Like simple vowels coalesce to the corresponding long vowel."

E36 =E2. "a or ā plus an unlike simple vowel, yields guṇa of the second vowel."

E37 =E3. "a or ā plus a diphthong, yields the corresponding long diphthong."

E40. "...short or long i and u... before a vowel are replaced by the corresponding semivowels."

E41. "A diphthong before a vowel is replaced by short or long a (depending on whether it is a short or long diphthong) plus the corresponding semivowel."

(E38 deals with a set of exceptions to E40. E39 deals with the so-called "roots in r̥̄". We will not discuss either of these cases.)

---

2 E35 refers to Rule 35 in Emeneau 1958. Similarly W492b refers to Section 492b in Whitney 1960.

First we note that it is not necessary to make any special provision for E37/E3 if all diphthongs are represented as vowel sequences; the application of E35/E1 automatically yields $\underline{\bar{a}}$ followed by a noncompact vowel.

The remaining processes are then, in essence, three: coalescence of like vowels, glide formation, and mid vowel formation. Moreover, glide formation takes place in two different environments: prevocalically (E40 and E41), and postvocalically (the formation of the long diphthongs $\underline{\bar{a}y}$ and $\underline{\bar{a}w}$ from $\underline{\bar{a}i}$ and $\underline{\bar{a}u}$). A semi-formal statement of the rules, using V as a general cover symbol for vowels, would then be:

$$1.\quad \begin{bmatrix} V \\ x \end{bmatrix}_2^4 \rightarrow \begin{bmatrix} V \\ +tns \\ x \end{bmatrix}$$

$$2.\quad \begin{bmatrix} V \\ -comp \end{bmatrix} \rightarrow [-voc] / \left\{ \begin{array}{l} \_\_\_\ V \\ \underline{\bar{a}}\ \_\_\_ \end{array} \right\}$$

$$3.\quad \underset{1}{\breve{\underline{a}}}\ \underset{2}{\begin{bmatrix} V \\ -comp \end{bmatrix}} \rightarrow \begin{bmatrix} 2 \\ -diff \\ +tns \end{bmatrix}$$

There are several possible variants of this set of rules. Rule 2 could be split into two rules, 2a (the prevocalic case) applying <u>before</u> Rule 3, 2b (the postvocalic case) applying *after* Rule 3; then the environment for 2b would be simply "after a vowel" and not "after *ā*." Or, the second environment in Rule 2 could be changed to "after a vowel," and Rule 3 could create *e* and *o* from *ay* and *aw* before a consonant. Somewhat arbitrarily (but see the following sections) we choose to split Rule 2 into two cases.

Before completely formalizing the rules, however, we must consider the form *dewāw* 'two gods', derived from *daiu* + *a* + *aau*: root *daiu* 'god' (itself from *diu* 'heaven', which appears in forms as both *diw* and *dyu*), stem-vowel *a*, and nom. du. ending *aau*. Rule 2a does not specify whether, in the cluster *iu*, it is the *i* or the *u* that is to become a glide. In such cases it is always the *second* vowel that becomes a glide. To achieve this result, we make the environment of 2a

$$\left(\begin{bmatrix} \mathrm{V} \\ \mathrm{-comp} \end{bmatrix}\right) \underline{\quad\quad} \mathrm{V}$$

Then by a general convention on application of rules (Section 1.3.4.1.2.), 2a will apply first to the *u* in the cluster *iu*, and, as a result, will not apply at all to the *i* (because it is no longer followed by a vowel).

Rules 1-3 apply in external, as well as internal, sandhi. We take this fact into account by inserting (#(#)) into the rules in the appropriate places. The formalized versions of these rules are:

$$\text{VOW1.}\quad \underset{1}{\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ \alpha\,\text{grv} \\ \beta\,\text{comp} \\ \gamma\,\text{fl} \end{bmatrix}}\ (1)\ (\#(\#))\ 1\ (1) \longrightarrow \begin{bmatrix} 1 \\ +\text{tns} \end{bmatrix}$$

$$\text{VOW2.}\quad \begin{bmatrix} -\text{cons} \\ -\text{comp} \end{bmatrix} \longrightarrow \begin{bmatrix} -\text{voc} \\ -\text{tns} \end{bmatrix} \Big/ \left(\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ -\text{comp} \end{bmatrix} \right) \_\_\_\ (\#(\#)) \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix}$$

$$\text{VOW3.}\quad \underset{1}{\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{comp} \\ -\text{tns} \end{bmatrix}}\ (\#(\#))\ \underset{2}{\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ -\text{comp} \end{bmatrix}} \longrightarrow \begin{bmatrix} 2 \\ -\text{diff} \\ +\text{tns} \end{bmatrix}$$

$$\text{VOW4.}\quad \begin{bmatrix} -\text{cons} \\ -\text{comp} \end{bmatrix} \longrightarrow \begin{bmatrix} -\text{voc} \\ -\text{tns} \end{bmatrix} \Big/ \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix}\ (\#(\#))\ \_\_\_$$

Examples:

āsīt, 3 sg. act. impf. of *as* 'be' (augment *a*, root *as*, union *ī*, ending *t*):

a + *as* + *ii* + *t* $\xrightarrow[\text{VOW1}]{}$ *ās* + *ī* + *t*

b^hawet, 3 sg. act. opt. of *b^hū* 'be' (guṇa[3] of root *b^hū*, theme *a*, mode-sign *ī*, ending *t*):

b^hauu + *a* + *ii* + *t* $\xrightarrow[\text{VOW1}]{}$ *b^haū* + *a* + *ī* + *t*

$\xrightarrow[\text{VOW2}]{}$ *b^haw* + *a* + *ī* + *t* $\xrightarrow[\text{VOW3}]{}$

b^haw + *e* + *t*

dewāw:

daiu + *a* + *aau* $\xrightarrow[\text{VOW1}]{}$ *daiu* + *āu* $\xrightarrow[\text{VOW2}]{}$

daiw + *āu* $\xrightarrow[\text{VOW3}]{}$ *dew* + *āu* $\xrightarrow[\text{VOW4}]{}$

dew + *āw*

3 In the traditional grammars, certain vowel alternations (the source of which will not be discussed here) are described by the terms *guṇa* and *vṛddhi*. The basic vowels are *a*, *i/ī*, and *u/ū*; the *guṇa* vowels are *a*, *e*, *o*, respectively; the *vṛddhi* vowels are *ā*, *āy*, *āw*, respectively.

VOW1-4 appear to account for all the facts described in Emeneau's rules above, with one exception. E36/E2 prescribes a <u>guṇa</u> vowel as the result of <u>a</u> or <u>aa</u> before <u>i</u>, <u>u</u>, or <u>ṛ</u>. Our rules predict a <u>guṇa</u> vowel from <u>a</u> and a <u>vṛddhi</u> vowel from <u>aa</u> in this context. To account for the falling together of the two cases we add a laxing rule (applying only before junctures), ordered after VOW1:

$$\text{VOW5.} \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{comp} \end{bmatrix} \longrightarrow [-\text{tns}] \quad / \quad \underline{\quad\quad} \begin{Bmatrix} + \\ \#(\#) \end{Bmatrix} [+\text{voc}]$$

It is not necessary to specify <u>i</u>, <u>u</u>, and <u>r</u> as a class in any direct way, because a following <u>a</u> would have coalesced (by VOW1) with the <u>ā</u> in question.

Examples of external combination with <u>tat^{h}ā</u> 'thus':

$$\underline{\text{tat}^h\text{aa}} \ \#\# \ \underline{\text{icc}^h\text{ati}} \xrightarrow[\text{VOW1}]{} \underline{\text{tat}^h\bar{\text{a}}} \ \#\# \ \underline{\text{icc}^h\text{ati}} \xrightarrow[\text{VOW5}]{}$$

$$\underline{\text{tat}^h\breve{\text{a}}} \ \#\# \ \underline{\text{icc}^h\text{ati}} \xrightarrow[\text{VOW3}]{} \underline{\text{tat}^h\text{ecc}^h\text{ati}}$$ 'thus he desires'

$$tat^{h}aa \text{ \#\# } ukta \xrightarrow[\text{VOW1}]{} tat^{h}\bar{a} \text{ \#\# } ukta \xrightarrow[\text{VOW5}]{}$$

$$tat^{h}\breve{a} \text{ \#\# } ukta \xrightarrow[\text{VOW3}]{} tat^{h}okta \quad \text{'thus spoken'}$$

VOW5 does <u>not</u> apply to the final cluster in *daiu* + *āu* (see the example above), because there is no boundary between the *ā* and the *u*.

The system of vowel rules that we have set up thus far is the following:

VOW1. Coalescence of Like Vowels.

VOW5. Laxing of *ā*.

VOW2. Prevocalic Glides.

VOW3. Mid Vowels.

VOW4. Postvocalic Glides.

2.3. Cyclical Rules

Let us examine VOW1 and VOW2 a bit more closely. First we consider the example *iti* ## *i* + *anti* (*iti* 'thus'; *yanti*, 3 pl. act. pres. of *i* 'go'). Because we have ordered VOW1 before VOW2, we would expect the derivation

$$iti \text{ \#\# } i + anti \xrightarrow[\text{VOW1}]{} it\bar{i} + anti \xrightarrow[\text{VOW2}]{} ityanti$$

But, as a matter of fact, the correct output is <u>itiyanti</u>.

Even if VOW1 were ordered after VOW2, there are difficulties. VOW2 is applicable to this example in two places--to the final <u>i</u> in <u>iti</u> (external sandhi) and to the initial <u>i</u> in <u>i</u> + <u>anti</u> (internal sandhi). The rule applies first to the external sandhi case, because that case satisfies the larger environment of the rules (Section 1.3.4.1.2.). We then obtain the incorrect output <u>ityyanti</u>.

It appears that the example provides evidence (a) for the separation of VOW2 into two rules, one applying internally and the other externally, and (b) for the ordering of VOW1 after VOW2. But if we must split VOW2 into two rules, then we cannot capitalize (by saving symbols) on the similarity of the two processes. Moreover, serious difficulties arise if VOW1 is ordered after VOW2; from <u>iti ## i + tas</u> we obtain <u>ityitas</u> instead of the correct output <u>itītas</u> 'thus they two go'. From the forms <u>itiyanti</u> and <u>itītas</u> we conclude that the required ordering is

VOW2I (Internal)

VOW1

VOW2E (External)

Yet even this ordering is not satisfactory. Consider the examples <u>strii ## iti</u> and <u>strii ## ati</u> (<u>strī</u> 'woman', <u>ati</u> 'over').[4] We would expect the derivations

$$\underline{strii}\ \#\#\ \underline{iti} \xrightarrow[\text{VOW2I}]{} \underline{stryi}\ \#\#\ \underline{iti} \xrightarrow[\text{VOW1}]{} \underline{stry\bar{i}ti}$$

$$\underline{strii}\ \#\#\ \underline{ati} \xrightarrow[\text{VOW2I}]{} \underline{stryi}\ \#\#\ \underline{ati} \xrightarrow[\text{VOW2E}]{} \underline{stryy}\ \#\#\ \underline{ati}$$

But, in fact, the correct outputs are <u>strīti</u> and <u>stryati</u>. We are thus forced either to constrain VOW2I to application before <u>unlike</u> vowels only, or to split VOW1 into two rules, one internal (ordered before VOW2I) and the other external (ordered between VOW2I and VOW2E). The second of these alternatives is particularly interest-

---

4 In our examples of external sandhi we try to use short words, so that the operation of the rules under discussion is not obscured by other details. As a result the examples do not always appear to make sense, although the cited collocations of words could occur within sentences.

ing, for the system of rules would be

VOW1I.

VOW2I.

VOW1E.

VOW2E.

That is, VOW1 and VOW2 would apply, in that order, first in cases of internal sandhi and then again in cases of external sandhi. But we already have at our disposal a device--the cyclical application of rules--that achieves this effect and that does not require us to state what is essentially the same rule in two different places.

If we assume that VOW1 and VOW2 apply in a cycle, then we need not make any changes at all in the system of rules set up in the previous section.

Sample derivations:[5]

	iti ## i + anti	ati ## iti
FIRST PASS		
VOW1	*	*
VOW5	*	*
VOW2	iti ## y + anti	*

[5] In sample derivations the asterisk * indicates that the rule in question does not apply, 'vac' that the application is vacuous.

SECOND PASS		
VOW1	*	atīti
VOW5	*	*
VOW2	vac	*

	strii ## iti	strii ## ati
FIRST PASS		
VOW1	strī ## iti	strī ## ati
VOW5	*	*
VOW2	*	*
SECOND PASS		
VOW1	strīti	*
VOW5	*	*
VOW2	*	stry ## ati

2.4. External Vowel Sandhi

In this section we seek to account for the facts of external vowel sandhi by adding to or amending the rules we have already set up. Emeneau's rules are as follows:

E1 = E35.

E2 = E36.

E3 = E37.

E4. "Other simple vowels before unlike vowels are replaced by the corresponding semivowels." (Same as E40.)

E5. "After a short diphthong initial a is elided."

E6. "e before any other vowel than a is replaced by a and hiatus remains."

E7. "āi before a vowel is replaced by ā and hiatus remains."

E8. "o before any other vowel than a and ā and āu before all vowels, are replaced by av and āv respectively."

(E9 deals with the special treatment of c^h in combination with certain vowels. E10 lists forms in which there is no sandhi at all.)

We have already discussed E1-E5.

E6-E8 are concerned with the final clusters a(a)i and a(a)u. The rules we have already would predict glide formation in most cases. Thus, in the external sandhi of tāw 'the two of them' and ewa 'indeed', the derivation proceeds as follows:

ta + *aau* ## *aiua* $\xrightarrow[\text{VOW1}]{}$ *tāu* ## *aiua* $\xrightarrow[\text{VOW2}]{}$

tāu ## *aiwa* (end of first pass)

$\xrightarrow[\text{VOW2}]{}$ *tāw* ## *aiwa* $\xrightarrow[\text{VOW3}]{}$ *tāwewa*

Glide formation occurs regularly in the case of *aau* and frequently in the case of *au* (E8). *Au* is, however, an infrequent final--except as the result of *auz* (see Chapter 6)--and is variously treated in the texts (W134a). We will assume with Emeneau that a glide is the normal output in all cases except *au(z)* ## *ă*, where the result is *o*.

This glide formation should occur even when the following vowel is *u*--thus *tāwupa* from *ta* + *aau* ## *upa* (*upa* 'towards'). As our rules now stand, we would expect the derivation

taau ## *upa* $\xrightarrow[\text{VOW1}]{}$ *tāu* ## *upa* (end of

first pass) $\xrightarrow[\text{VOW1}]{}$ *tāūpa* $\xrightarrow[\text{VOW4}]{}$

tāwpa

Hence, VOW1 must be blocked from applying during the second pass when the sequence of like vowels is preceded by a vowel. This can be accomplished by adding to the environment of VOW1:

$$\text{VOW1}'.\quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ \alpha\,\text{grv} \\ \beta\,\text{comp} \\ \gamma\,\text{fl} \end{bmatrix}_{1} (1)(_{1}\#(\#))1(1) \rightarrow \begin{bmatrix} 1 \\ +\text{tns} \end{bmatrix} \Big/ \quad (_{1}\left[\begin{Bmatrix} +\text{cons} \\ -\text{voc} \end{Bmatrix}\right]) \underline{\qquad} \qquad {}^{6}$$

The cases of *au* and *ai* before *ă* are special; instead of glide formation (in the case of *au*) or hiatus (in the case of *ai*), the result is a mid vowel with loss of the following *ă*. Now the rules that we have already set up would predict a mid vowel as the resultant of *au* or *ai* if the *ă* were dropped *before* the formation of prevocalic glides (VOW2), and hence before the formation of the mid vowels (VOW3). That is to say, the appearance of *e* or *o* is not exceptional if there is an early rule prescribing the deletion of initial *ă* after final *au* or *ai* (instead of a late rule prescribing the deletion of *ă* after *o* or *e*, as it is usually conceived--cf. W135):

---

[6] That is, a sequence of like vowels coalesces in any position in internal sandhi, and when preceded by a non-vowel in external sandhi. The conventions concerning numbered parentheses (Section 1.3.4.1.2.) insure that on the second pass the sequence of like vowels must be preceded by a non-vowel.

$$\text{VOW6.}\quad \begin{bmatrix}+\text{voc}\\-\text{cons}\\+\text{comp}\\-\text{tns}\end{bmatrix} \longrightarrow \emptyset \;/$$

$$\begin{bmatrix}+\text{voc}\\-\text{cons}\\+\text{comp}\\-\text{tns}\end{bmatrix}\begin{bmatrix}+\text{voc}\\-\text{cons}\\-\text{comp}\end{bmatrix} \quad \#(\#) \;\_\_\_\_$$

Thus, from *taí ## atí* (*té* 'they' masc.), we obtain by VOW6 (on the second pass) *taí ## tí*; VOW2 is then inapplicable, and VOW3 applies to yield the correct output *tetí*.

The remaining cases of vowel-vowel sandhi are the hiatus cases--*aaí* before any vowel, and *ai* before any vowel except *ǎ* (E6 and E7). The rule creating vowels in hiatus (VOW7) must follow both VOW3 and VOW4, for (a) If VOW7 preceded VOW3, the resultant of *taí ## ití* would be *teti* instead of the correct form *taiti*, and (b) If VOW7 preceded VOW4, the resultant of *tasmaai ## iti* (*tasmāy* 'to him') would be *tasmāyti* instead of the correct form *tasmāiti*. Hence, VOW7 is a *y*-deletion rule:

$$\text{VOW7.}\quad \begin{bmatrix}-\text{voc}\\-\text{cons}\\-\text{comp}\\-\text{grv}\end{bmatrix} \longrightarrow \emptyset \;/$$

$$\begin{bmatrix}+\text{voc}\\-\text{cons}\\+\text{comp}\end{bmatrix} \;\_\_\_\_\; \#(\#) \begin{bmatrix}+\text{voc}\\-\text{cons}\end{bmatrix}$$

It is not necessary to specify "before any vowel except *ă*" in the case of final *ai*, because this situation will already have been handled by an earlier application of VOW6.

In the case of the example *tai* ## *iti* we have ordered VOW3 and VOW7 so as to avoid the derivation

$$\underline{tai}\ \#\#\ \underline{iti} \longrightarrow \underline{ta}\ \#\#\ \underline{iti} \longrightarrow \underline{teti}\ .$$

However, if VOW3 is permitted to apply on the first pass, we will have the derivation

$$\underline{tai}\ \#\#\ \underline{iti} \longrightarrow \underline{te}\ \#\#\ \underline{iti}\ ,$$

with no application of VOW7 at all. Similar problems arise in the case of VOW4. The solution to this difficulty is to make VOW3 and VOW4 applicable only during the second pass; VOW7, referring as it does to word boundaries, is already applicable only during this pass. The revised forms of VOW3 and VOW4 are:

$$\text{VOW3}'.\ \underset{1}{\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{comp} \\ -\text{tns} \end{bmatrix}}\ (\#(\#))\ \underset{2}{\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ -\text{comp} \end{bmatrix}} \longrightarrow \begin{bmatrix} 2 \\ -\text{diff} \\ +\text{tns} \end{bmatrix}\ /$$

$$\underline{\quad\quad}\,[\ ]_{0}\ \#\#$$

$$\text{VOW4'.}\quad \begin{bmatrix}-\text{cons}\\-\text{comp}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{voc}\\-\text{tns}\end{bmatrix} \;/\; \begin{bmatrix}+\text{voc}\\-\text{cons}\end{bmatrix} (\#(\#)) \;\_\_\_\; [\;]_0 \;\#\#$$

The resultant system of vowel sandhi rules is then

VOW1'. Coalescence of Like Vowels.

VOW6. <u>ă</u>-Deletion.

VOW5. Laxing of <u>ā</u>.

VOW2. Prevocalic Glides.

VOW3'. Mid Vowels.

VOW4'. Postvocalic Glides.

VOW7. <u>y</u>-Deletion.

Of these rules, only VOW1', VOW5, and VOW2 will be applicable during the first pass. All seven rules are applicable during the second pass.

Sample derivations (involving <u>sā</u> 'she'; the voc. sg. <u>indra</u> 'O Indra'; <u>ehi</u>, 2 sg. act. impv. of <u>aa</u> # <u>i</u> 'go to, approach')

	tasmaai ## ati	ta + aau ## ati
FIRST PASS		
VOW1'	tasmāi ## ati	tāu ## ati
VOW5	*	*
VOW2	*	*
SECOND PASS		
VOW1'	*	*
VOW6	*	*
VOW5	*	*
VOW2	tasmāy ## ati	tāw ## ati
VOW3'	*	*
VOW4'	*	*
VOW7	tasmā ## ati	*

	saa ## iua	saa ## aiua
FIRST PASS		
VOW1'	sā ## iua	sā ## aiua
VOW5	*	*
VOW2	sā ## iwa	sā ## aiwa

SECOND PASS

VOW1'	*	sāiwa
VOW6	*	*
VOW5	sa ## iwa	*
VOW2	vac	vac
VOW3'	sewa	*
VOW4'	*	sāywa
VOW7	*	*

---

indra ## aa # ihi

FIRST PASS

VOW1'	indra ## ā # ihi
VOW5	*
VOW2	*

SECOND PASS

VOW1'	indrā # ihi
VOW6	*
VOW5	indra # ihi

VOW2		*
VOW3'	indrehi	
VOW4'		*
VOW7		*

At this point we abandon the topic of vowel sandhi and turn to some aspects of consonant sandhi. In Chapter 6 we shall examine the vowel sandhi rules once again, this time in relation to several rules of consonant sandhi. In the following chapters we shall transcribe all tense vowels as clusters.

Chapter 3

Retroflexion

Let us consider the occurrence of retroflex consonants within words. We shall see that retroflexion is predictable, with only a few exceptions.

3.1. The Retroflex Spirant ṣ

It is well known that nearly all occurrences of ṣ are predictable. This segment occurs neither in word-final position nor (with only one important exception, the numeral ṣaṣ 'six') in word-initial position. Nearly every occurrence of ṣ immediately follows i, u, r, or k; and s hardly ever occurs in this environment. Thus, in combination with the loc. pl. ending su:

With ṣ:	With s:
agniṣu (agni 'Agni')	jaasu (jaa 'progeny')
śatruṣu (śatru 'enemy')	marutsu (marut 'wind')
waakṣu (waac 'voice')	apsu (ap 'water')
swasrṣu (swasr 'sister')	
kaamaiṣu (kaama 'love')	

and in the citation-forms of verbal roots:

With ṣ:		With s:	
kṣud^h	'crush'	tsar	'creep up on'
kṣam	'endure'	psaa	'devour'
akṣ	'attain'	b^{h}arts	'revile'
dwiṣ	'hate'	was	'clothe,' 'dwell,' 'shine'
d^{h}rṣ	'dare'	tras	'be terrified'
kuṣ	'tear'	kas	'open'
hraiṣ	'neigh'	hras	'shorten'
b^{h}uuṣ	'adorn'	b^{h}aas	'shine'

The class of segments that cause the modification of s to ṣ can be specified as $\begin{bmatrix}\alpha\text{cons}\\ \alpha\text{comp}\end{bmatrix}$. This class includes all velar and palatal consonants, the vowels i and u, the glides y and w, and the liquid r. We can then write:

$$\text{RET1.}\quad \begin{bmatrix}+\text{obst}\\ +\text{cont}\\ -\text{comp}\end{bmatrix} \longrightarrow \begin{bmatrix}+\text{comp}\\ +\text{fl}\end{bmatrix} \Big/ \begin{bmatrix}\alpha\,\text{cons}\\ \alpha\,\text{comp}\end{bmatrix}\ (\#)\ \underline{\quad\quad}$$

This rule applies across # as well as +: that is, it applies after the verbal prefixes ab^{h}i 'unto, against', prati 'back (to)', ni 'down, into', wi 'apart, away', anu 'after, along', and su 'well, good' (e.g., anuṣṭup 'following in praise; a kind of meter', from the verbal root stub^h 'praise') and in other compounds as well (e.g., the proper name yud^{h}iṣṭ^{h}ira 'firm in battle', from yud^{h}i 'battle' and st^{h}ira 'firm'). The conversion usually takes place with the verbal prefixes. In other compounds the best that can be said is that it often takes place (see W186); the factors that determine whether or not an s is converted in compounding are not well understood.

There is a regular class of exceptions in which the effect of RET1 is canceled: whenever an r immediately follows the s, and frequently whenever there is any later r within the same word, there is no retroflexion. We could account for this "protective" influence of an r either by adding

$$\left[\begin{Bmatrix}\text{-fl}\\\text{-voc}\\\text{-cons}\end{Bmatrix}\right]_{o} \#\#$$

as a right-hand environment in RET1, or by adding the deretroflexion rule

$$\text{RET2.}\quad \begin{bmatrix}\text{+obst}\\\text{+cont}\\\text{-comp}\end{bmatrix} \rightarrow [\text{-fl}] \;/\; \underline{\quad\quad} \;[\;]_{o} \begin{bmatrix}\text{+voc}\\\text{+cons}\\\text{+fl}\end{bmatrix}$$

Since RET1 is a cyclical rule (see Chapter 4), we must use the deretroflexion rule instead of the added environment, because the protecting environment may not appear until a pass after the one in which the affecting environment appears.

RET1 predicts most occurrences of ṣ in Sanskrit. Other occurrences result from s or kṣ before a dental stop (see Chapter 4) or from an assimilation of retroflexion (see Section 3.4.). There remain only a few unpredictable occurrences, principally the numeral ṣaṣ and four verbal roots. There is also a handful of forms--e.g., kusuma 'flower'--in which s occurs where ṣ is predicted by RET1.

These facts suggest that there should be no ṣ distinct from s in lexical entries, and that exceptional forms should be handled by special rules that list the forms (if they show an unpredictable ṣ) or by marking the forms as not being subject to certain rules (if the forms show an s where the rules predict an ṣ). The four verbal roots that show an unpredictable ṣ present more of a problem, however, because to each one there corresponds a different root in s; we have:

<u>kaṣ</u> 'scratch'	<u>kas</u> 'open'
<u>b^haṣ</u> 'bark'	<u>b^has</u> 'devour'
<u>b^haaṣ</u> 'speak'	<u>b^haas</u> 'shine'
<u>laṣ</u> 'desire'	<u>las</u> 'be lively'

The roots in the two columns can be distinguished by listing those in the first column with a special mark--either a lexical mark or a phonological one. We shall suggest a phonological mark in Section 3.3.

3.2. The Retroflex Nasal ṇ

The predictability of most occurrences of ṇ, like the predictability of most occurrences of ṣ, is well known. Whitney's very prolix statement of the rule (W189) is:

> The dental nasal... n, when immediately followed by a vowel or by... n or... m or... y or... v, is turned into the lingual... ṇ if preceded in the same word by the lingual sibilant or semi-vowel or vowels--that is to say, by... ṣ,... r, or... ṛ or... ṝ--:
>
> and this, not only if the altering letter stands immediately before the nasal, but at whatever distance from the letter it may be found: unless, indeed, there intervene (a consonant moving

the front of the tongue: namely) a palatal (except... y), a lingual, or a dental.

The rule may be inferred from the following forms:

Root		3 Sg. Pres Act. Ind.	Derived Noun/Adjective	
naṭ	'dance'	naṭati	naṭana	'dance'
jwal	'burn'	jwalati	jwalana	'fire'
tap	'heat'	tapati	tapana	'heat'
g^has	'eat'	g^hasati	g^hasana	'devouring'
sru	'flow'	srawati	srawaṇa	'flowing; sweat; urine'
hrṣ	'be excited'	hrṣyati	harṣaṇa	'erection'
jrmb^h	'gape'	jrmb^{h}ati	jrmb^{h}aṇa	'yawn'
pluṣ	'burn'	plauṣati	plauṣaṇa	'burning'
kṣap	'be abstinent'	kṣapati	kṣapaṇa	'abstinence'
caiṣṭ	'stir'	caiṣṭati	caiṣṭana	'motion, effort'
muurc^h	'thicken, become violent'	muurc^{h}ati	muurc^{h}ana	'violence, growth'
nard	'bellow'	nardati	nardana	'roaring'
kṣwail	'play'	kṣwailati	kṣwailana	'play'
ras	'roar'	rasati	rasana	'roaring, noise'

Conjugation of *rud^h* 'oppress' in the sg. pres.:

	Active	Middle
1	ruṇad^{h}mi	rund^{h}ai
2	ruṇatsi	runtsai
3	ruṇadd^{h}i	rundd^{h}ai

That is, an n is retroflected if there is an r or ṣ preceding it anywhere within a word, so long as no palatal or dental intervenes and so long as the n is followed by a sonorant. In distinctive feature terms:

$$\text{RET3.}\quad \begin{bmatrix}\text{+nas}\\ \text{-grv}\end{bmatrix} \longrightarrow \begin{bmatrix}\text{+comp}\\ \text{+fl}\end{bmatrix} \Big/$$

$$\begin{bmatrix}\text{+cons}\\ \text{+cont}\\ \text{+fl}\end{bmatrix} \left[\begin{Bmatrix}\text{-cons}\\ \text{+grv}\end{Bmatrix}\right]_0 \underline{\quad\quad} \begin{bmatrix}\text{-obst}\end{bmatrix}$$

Like RET1, this rule applies generally to n following prefixes (in this case, paraa 'forth', pari 'around', pra 'forward', niṣ 'not, without', antar 'within', and duṣ 'bad'), and frequently in other compounds. A few roots--e.g., nab^h 'burst' and nand 'rejoice'--invariably occur without retroflexion, however, and must be marked as [-RET3] in the lexicon.

Again like RET1, RET3 may have its effect canceled by a following retroflex sound. In this case an *r* immediately following the *n* always prevents retroflexion, and any *r* or *ṣ* later in the same word usually prevents it. This protective effect could be accounted for by changing the right-hand environment of RET3 to

$$\begin{bmatrix}\text{-obst}\\ \text{-voc}\end{bmatrix}\left[\begin{Bmatrix}\text{-cons}\\ \text{-cont}\\ \text{-fl}\end{Bmatrix}\right]_0 \#\# \qquad ;$$

but we choose, as before, to add a dissimilation rule:

$$\text{RET4.}\quad \begin{bmatrix}\text{+nas}\\ \text{-grv}\end{bmatrix} \longrightarrow [\text{-fl}] \;/\; \_\_\_\; [\quad]_0 \begin{bmatrix}\text{+cons}\\ \text{+cont}\\ \text{+fl}\end{bmatrix}$$

The root *nrt* 'dance', for example, never occurs with retroflexion. And the past part. and 3 sg. act. fut. of *pra* # *naś* 'reach, attain' are *pranaṣṭa* and *pranañkṣyati*, respectively, and not *praṇaṣṭa* and *praṇañkṣyati*, respectively.

Most occurrences of *ṇ* are predicted by RET3, or by assimilation rules. However, there are a fair number of occurrences that cannot be predicted on those grounds--e.g., the substantives *guṇa* 'quality' and *puṇya* 'prosperous, happy', and the following (complete) list of verbal roots, many of which are of rare occurrence:

j^{h}aṇ	'sound'	*maṇḍ*	'deck'
paṇ	'bargain'	*hiṇḍ*	'be empty'
p^{h}aṇ	'go, move'	*waṇṭ*	'divide'
b^{h}aṇ	'speak, call'	*aṇṭ^h/aṭ^h*	'go, move'
maṇ	'sound'	*kuṇṭ^h*	'dull'
kwaṇ	'sound'	*guṇṭ^h*	'cover up'
kuuṇ	'shrink'	*luṇṭ^h/luṭ^h*	'rob'

Unpredictable *ṇ* also occurs rather often in proper names--e.g., *Paaṇini*, *Saayaṇa*, *Paaṇḍu*, and *C^{h}aaṇakya*. Since our analysis of these remaining cases of *ṇ* is closely connected to our analysis of the retroflex oral stops, we turn next to these segments.

3.3. The Retroflex Oral Stops *ṭ*, *ṭ^h*, *ḍ*, *ḍ^h*

Although some occurrences of the retroflex oral stops can be explained as the result of assimilation rules (Section 3.4) or terminal sandhi rules for *ṣ*, *s*, *kṣ*, *j* and *h* (Chapter 4), there remain many occurrences not accounted for by these rules. The retroflex oral stops occur in both verbal roots and substantives, and (like *ṇ*) are rather common in proper names--e.g., *Wiraaṭa*, *B^{h}aṭṭi*. In the following discussion we will confine ourselves to verbal roots as representatives of the entire Sanskrit lexicon.

In initial position retroflex stops (oral or nasal) are very rare. Whitney 1885 lists four roots beginning with *ḍ* (of which two are doubtful), one in *ḍ*^{h}, four in *ṭ* (of which three are doubtful), and none in *ṭ*^{h} or *ṇ*. In final position (and when not explainable as assimilations) retroflex stops are more common. Whitney's list contains seventeen roots ending in *ṭ* or *ṭṭ*, five in *ṭ*^{h}, nineteen in *ḍ*, thirteen in *ṇ*, *ṇḍ*, *ṇṭ*, or *ṇṭ*^{h}, and none in *ḍ*^{h}.

Now we notice that the occurrence of retroflex stops in roots correlates with a gap in the distribution of *l*. In root-initial clusters *l* can be preceded by velars or labials or by *ś*--e.g., *klam* 'be weary', *glah* 'gamble', *hlaad* 'refresh', *plu* 'float', *mlaa* 'relax', *ślaag*^{h} 'extol'--but not by dentals. In the four quotable root-final clusters *l* is followed by labials or velars: *galb*^{h} 'dare', *gulp*^{h} 'bunch', *jalp* 'murmur', and *walg* 'spring'. Finally, the only root containing the "vowel *l*" is *klp* 'be adapted', in which *l* is preceded by a velar and followed by a labial.

These facts suggest that root-final retroflex stops might be considered as the product of an original dental stop preceded by <u>l</u>. A similar explanation could be given for the root-initial retroflex stops; however, the fact that they are so very infrequent suggests that they are truly exceptional.

The remarks in the preceding paragraphs are only suggestive. They do not provide a very strong argument for the representation of retroflex stops (and "unpredictable" <u>ṣ</u>) as clusters. There is no available direct evidence in the morphophonemic processes of Sanskrit, because:

1. There is no situation in which a retroflecting <u>l</u> is <u>not</u> deleted; no output form contains a retroflecting <u>l</u> fortuitously preserved for our inspection.
2. The combination of an <u>l</u> and a following dental in the ordinary juxtaposition of morphemes is systematically avoided in Sanskrit. No root ending in <u>l</u> belongs to either the root class, the reduplicating class, or the nasal class in the present system. Hence, an

> l can never come to precede a dental in conjugation. Moreover, whenever an l would ordinarily come to precede a dental in derivation (in the formation of the past participle, for example), a union vowel is inserted.

Some indirect evidence can be obtained by considering the formation of past participles, gerunds, and infinitives from roots ending in retroflex consonants. This evidence is discussed in Chapter 5.

It should be emphasized that one of the major supports for our suggestion is the relative rarity of "unpredictable" retroflex consonants. That is, the attractiveness of the suggested representation depends crucially on the number of (lexical) forms containing the clusters. We are trading the use of two segments against the postulation of additional phonemes (with the attendant necessity of distinguishing these new phonemes from all other phonemes by at least one feature). If the number of (lexical) forms containing the two-segment clusters increases during a period of time, eventually a point will be reached at which it will be more economical--in terms of the number of features it is necessary to specify in the lexicon--

to postulate new phonemes.

A final criticism of our treatment of the "unpredictable" retroflex consonants is that we have not sufficiently motivated the choice of *l* as the cause of the retroflexion (given the assumption that the retroflexion does indeed *have* a cause). It is possible to devise other treatments of the problem. One method--which is attractive because it accords with the facts of linguistic history, but which cannot be used in every instance of "unpredictable" retroflection--is to postulate an *s* as the affecting environment. Thus, the root *liḍ* 'implore' would be represented as *lisd* (or *isd*); by RET1 the *s* would become *ṣ*, the *d* would assimilate to the *ṣ* in retroflexion (see the next section), the *ṣ* would assimilate to the *ḍ* in voicing (see Chapter 5), and the resulting *ẓ* would be deleted by a new rule (possibly with compensatory lengthening of the preceding vowel). Our observations about the distribution of *l* would find expression in a morpheme-structure rule. However, this treatment suffers from many of the same deficiencies as the treatment we have suggested, without the latter's virtue of generality. On the other hand, there is no reason to assume only *one* source of "unpredictable" retroflexion.

To account for root-final retroflex stops, then, we add a new retroflexion rule (RET5) and an l-dropping rule (RET6):

$$\text{RET5.}\quad \begin{bmatrix}+\text{cons}\\-\text{comp}\\-\text{grv}\end{bmatrix} \rightarrow \begin{bmatrix}+\text{comp}\\+\text{fl}\end{bmatrix} \;/\; \begin{bmatrix}+\text{cons}\\+\text{voc}\\-\text{fl}\end{bmatrix} \;\underline{\quad\quad}$$

$$\text{RET6.}\quad \begin{bmatrix}+\text{cons}\\+\text{voc}\\-\text{fl}\end{bmatrix} \longrightarrow \emptyset \;/\; \underline{\quad\quad} \begin{bmatrix}+\text{cons}\\+\text{fl}\end{bmatrix}$$

Then kuuṇ 'shrink' will be represented in the lexicon as kuuln and paṭ^h 'read' as palt^h. Exceptional occurrences of ṣ can be accounted for by RET5 and RET6; b^{h}aṣ 'bark' will be represented as b^{h}als.

3.4. Assimilation of Retroflexion

We have already alluded to "assimilation of retroflexion" rules and quoted several relevant forms--anuṣṭup from anu # stub^h, and yud^{h}iṣṭ^{h}ira from yud^{h}i # st^{h}ira. Further examples:

> dwiḍḍ^{h}wai from dwiḍ + d^{h}wai (ultimately from dwiṣ 'hate' and 2 pl. midd. ending d^{h}wai: see Chapter 4)
>
> dwiṣṭas from dwiṣ 'hate' and 3 du. act. ending tas

liṭṭai from *iiḍ* 'implore' and 3 sg. midd. ending *tai*

From these examples, we see that in internal combination a dental consonant is retroflected after a retroflex consonant:

$$\text{RET7.}\quad \begin{bmatrix}\text{+cons}\\ \text{-voc}\\ \text{-grv}\end{bmatrix} \rightarrow \begin{bmatrix}\text{+comp}\\ \text{+fl}\end{bmatrix} \Big/ \begin{bmatrix}\text{+cons}\\ \text{-voc}\\ \text{+fl}\end{bmatrix} \underline{\qquad}$$

(but see the further discussion in the next chapter).

In external combination there is the same assimilation of retroflexion, but in the opposite direction:

taḍ ## ḍayatai from *tat* 'it' and 3 sg. midd. pres. of *ḍii* 'fly'

triiṇ ## ṭañkati from acc. sg. masc. of *tri* 'three' and 3 sg. act. pres. of *ṭañk* 'cover'

taṭ # ṭiikaaḥ from *tat* and nom. sg. of *ṭiikaa* 'commentary'

paadaṣ ## ṭalati from nom. sg. of *paada* 'foot' and 3 sg. act. pres. of *ṭal* 'be confused' *taaṣ* ## *ṣaṭ* from *taas* 'they' (fem.) and *ṣaṣ* 'six'

The rule for this assimilation is then

$$\text{RET8.}\quad \begin{bmatrix}+\text{cons}\\ -\text{voc}\\ -\text{grv}\end{bmatrix} \longrightarrow \begin{bmatrix}+\text{comp}\\ +\text{fl}\end{bmatrix} \;/\; \underline{\quad\quad}\ \#(\#) \begin{bmatrix}+\text{cons}\\ -\text{voc}\\ +\text{fl}\end{bmatrix}$$

(but see the further discussion in the next chapter).

A final retroflex consonant remains before an initial non-retroflex consonant: *jiiwanaḍ* ## *dewaaya*, from *jiiwanaś* 'living sacrifice' and dat. sg. of *daiwa* 'god' (for *ḍ* in place of *ś*, see Chapters 4 and 5).

There are several problems remaining in connection with these assimilation rules and their relationship to the other retroflexion rules. The first arises when we consider the past participle *kṣuṇṇa*, from the root *ksud*[1] 'crush'. This root is one of approximately seventy that have past participles in *na* instead of *ta*. Moreover, *na* is peculiar in that a stop assimilates in nasality to it; ordinarily in internal sandhi there is no such assimilation (cf. the noun *ratna* 'jewel'), while in external sandhi there is (cf. *tan* ## *namas*

[1] From this point on we will not indicate predictable retroflex consonants in citation forms of roots; thus *ksud* instead of the usual *kṣud*.

from tat and namas 'bow, obeisance'). Hence we assume that the participial endings are + ta and # na, so that the relevant rule of external sandhi will apply to roots before # na. However, if the ending is given as # na, the boundary will prevent RET7 from applying to the n of the ending.

A similar problem arises when we consider the locative plurals dwiṭsu and hawiṣṣu (or hawiḥṣu), from dwis 'hate' (a verbal root used directly as a substantive) and haw + is (sacrifice', actually from the verbal root hu 'sacrifice' with guṇa vowel[2]), respectively. From dwiṭsu and similar forms we conclude that the locative plural ending is # su: ṣ is converted to ṭ before this ending, just as it is in word-final position (see Chapter 4), and there is no assimilation of retroflexion. But in hawiṣṣu, cakṣuṣṣu (from caks + us 'eye', ultimately from caks 'notice'), and similar forms, the application of RET7 seems to be necessary for the retroflexion of the s of the ending.

---

2 Although the glide w in this form and the w in dewa can be predicted by our vowel sandhi rules, we continue to cite forms with w (and y) rather than u (and i). We have not extended our rules to predict all occurrences of glides, and the transcription with y and w is closer to the standard transcriptions (although sometimes less perspicuous for our purposes).

One way to account for these forms would be to assume that RET7 applies across # as well as +. It would then be necessary to add a rule to account for the non-occurrence of ṭ # ṣ. A simpler solution, however, is to assume that RET1 applies across an intervening spirant and that RET3 applies across an intervening nasal:

$$\text{RET1}'.\quad \begin{bmatrix}\text{+obst}\\ \text{+cont}\\ \text{-comp}\end{bmatrix} \longrightarrow \begin{bmatrix}\text{+comp}\\ \text{+fl}\end{bmatrix} \;/$$

$$\begin{bmatrix}\alpha\,\text{cons}\\ \alpha\,\text{comp}\end{bmatrix}\left(\begin{bmatrix}\text{+obst}\\ \text{+cont}\end{bmatrix}\right)\ (\#)\ \underline{\quad\quad}$$

$$\text{RET3}'.\quad \begin{bmatrix}\text{+nas}\\ \text{-grv}\end{bmatrix} \longrightarrow [\text{+fl}] \;/$$

$$\begin{bmatrix}\text{+cons}\\ \text{+cont}\\ \text{+fl}\end{bmatrix}\left[\begin{Bmatrix}\text{-cons}\\ \text{+grv}\end{Bmatrix}\right]_0\ \left([\text{+nas}]\right)\underline{\quad\quad}[\text{-obst}]$$

Then

$$\underline{\text{ksud}}\ \#\ \underline{\text{na}} \longrightarrow \underline{\text{ksun}}\ \#\ \underline{\text{na}} \xrightarrow[\text{RET1}']{} \underline{\text{kṣun}}\ \#\ \underline{\text{na}}$$

$$\xrightarrow[\text{RET3}']{} \underline{\text{kṣuṇ}}\ \#\ \underline{\text{ṇa}}$$

$$\underline{\text{haw}} + \underline{\text{is}}\ \#\ \underline{\text{su}} \xrightarrow[\text{RET1}']{} \underline{\text{haw}} + \underline{\text{iṣ}}\ \#\ \underline{\text{ṣu}}$$

Chapter 4

Roots in ṣ, ś, and kṣ

4.1. Internal Sandhi

4.1.1. The Problem

In the following sections we will examine the behavior of roots ending in ṣ (that is, s or ls), ś, and kṣ (that is, ks). All such roots (with the exception of four roots in ś--diś 'point', dṛś 'see', spṛś 'touch', and sometimes naś 'attain'--and of one root in kṣ--jakṣ 'eat', which is discussed in Chapter 5) have the same treatment in internal sandhi before obstruents. That is (W218, 221, 226),

Finally and before # su: ṭ

Before b^h and d^h: ḍ

Before t and t^h: ṣ

Before s: k

Examples are provided in the following chart:

		dwis	daaś	caks	endings
		'hate'	'make offering'	'see'	
	1 sg. act.	dwaiṣmi	daaśmi	cakṣmi	+ mi
	2 sg. act.	dwaikṣi	daakṣi	cakṣi	+ si
Verb	3 sg. act.	dwaiṣṭi	daaṣṭi	caṣṭi	+ ti
in Present	2 du. act.	dwiṣṭ^{h}as	daaṣṭ^{h}as	caṣṭ^{h}as	+ t^{h}as
	2 pl. midd.	dwiḍḍ^{h}wai	daaḍḍ^{h}wai	caḍḍ^{h}wai	+ d^{h}wai
	nom. sg.	dwiṭ	daaṭ	caṭ	+ s[1]
Derived	inst. pl.	dwiḍb^{h}is	daaḍb^{h}is	caḍb^{h}is	+ b^{h}is[2]
Root Noun	loc. pl.	dwiṭsu	daaṭsu	caṭsu	# su

1 This s is, however, deleted by an early rule for the simplification of final clusters; see Sections 4.1.4., 4.2. and 4.3. Hence for our purposes the nom. sg. ending is null for roots ending in a consonant.

2 In Chapter 6 we suggest that this ending is preceded by # rather than +.

4.1.2. Spirant-Shift Rules

We begin by considering the roots ending in *ṣ* and *ś*. The spirant in these roots becomes dental before stops and before # (including before the # *su* of the loc. pl.), and velar before *s*. The dental is a stop in some cases, a spirant in others; the velar is a stop.

The change can be split into two steps--first the shift of *ṣ* and *ś* to *k* before *s* and to *ṣ* before stops and #, and then the shift of *ṣ* to *ṭ* before voiced stops and #. Since we have supposed that # has some of the markings of ordinary phonological segments--in particular, that it has the markings +obst, -cont--the first rule can be stated as

$$\text{SP1.} \quad \begin{bmatrix} +\text{obst} \\ +\text{cont} \\ +\text{comp} \end{bmatrix} \dashrightarrow \begin{bmatrix} \alpha\ \text{grv} \\ -\alpha\ \text{cont} \\ -\alpha\ \text{fl} \end{bmatrix} \;/\; \underline{\qquad} \begin{bmatrix} +\text{obst} \\ \alpha\ \text{cont} \end{bmatrix}$$

That is, *ṣ* and *ś* become *k* (which is +grv, -cont, -fl) before continuants and become *ṣ* (which is -grv, +cont, +fl) before stops.

The second change (the shift in continuancy) is simpler:

$$\text{SP2.}\quad \begin{bmatrix} +\text{obst} \\ +\text{fl} \\ +\text{ROOT} \end{bmatrix} \rightarrow [-\text{cont}] \; / \; \underline{\quad\quad} \left\{ \begin{matrix} \begin{bmatrix} +\text{obst} \\ -\text{cont} \\ +\text{vcd} \end{bmatrix} \\ \# \end{matrix} \right\}$$

The specification that the spirant be a segment of a root is necessary to account for the difference between <u>hawiṣṣu</u> (or <u>hawiḥṣu</u>) and <u>dwiṭsu</u> (see the next section).

It is important that # have the markings +obst, -cont, for if we could not refer to those markings in SP1, then SP1 would have to be split into two rules:[3]

$$\begin{bmatrix} +\text{obst} \\ +\text{cont} \\ +\text{comp} \end{bmatrix} \rightarrow \left\{ \begin{matrix} \begin{bmatrix} \alpha\,\text{grv} \\ -\alpha\,\text{cont} \\ -\alpha\,\text{fl} \end{bmatrix} / \; \underline{\quad\quad} \begin{bmatrix} +\text{obst} \\ \alpha\,\text{cont} \end{bmatrix} \\ \\ \begin{bmatrix} -\text{grv} \\ +\text{cont} \\ +\text{fl} \end{bmatrix} / \; \underline{\quad\quad} \; \# \end{matrix} \right\}$$

This difficulty does not arise in the case of SP2, even though the voiced stop environment and the # environment are stated separately, because SP2 involves no assimilation or dissimilation.

4.1.3. <u>Cyclical Rules</u>

Let us now consider the ordering of SP1 and

---

3 We cannot assume that the <u>s</u> of the nom. sg. remains when SP1 applies, for if the ending were <u>s</u>, we would expect <u>daaš</u> + <u>s</u> to become <u>daak</u> + <u>s</u> and then <u>daak</u>. But the correct output is <u>daaṭ</u>.

SP2 with respect to some of the other rules we have established--in particular, with respect to RET1'.

On the one hand, RET1' must precede SP1 and SP2, for it creates instances of ṣ to which SP1 and SP2 will apply. But on the other hand, SP2 (and hence SP1) must precede RET1', for if the ordering were reversed we would have the derivation

$$\underline{\text{dwis}}\ \#\ \underline{\text{su}} \xrightarrow[\text{RET1}']{} \underline{\text{dwiṣ}}\ \#\ \underline{\text{ṣu}} \xrightarrow[\text{SP1}]{}$$

$$\underline{\text{dwiṣ}}\ \#\ \underline{\text{ṣu}} \xrightarrow[\text{SP2}]{} \underline{\text{dwiṭ}}\ \#\ \underline{\text{ṣu}}$$

That is, the ṭ would be created too late to block the retroflexion of the second s in this form.

The correct output can be obtained only if we assume that the three rules in question apply in a cycle--first to the root or stem, then again to the root or stem in combination with conjugational or declensional endings. The relative ordering of the rules within this cycle is

SP1

SP2

RET1'

RET7

Then RET1' will apply twice in some forms, once before SP1 and SP2 and again after these rules.

Both passes through the retroflexion and spirant-shift rules are included within the "first pass" through the cycle discussed in Chapter 2; that is, we have now argued (in Chapter 2 and in the current section) for at least three passes through a cycle--once for roots and stems, once for (uncompounded) words, and once for complete sentences.[4] These three passes are crucial. There will, of course, be more than three passes when the cyclical rules are applied to a sentence of any complexity.

In the tables below seven cyclical derivations are illustrated. We use (unlabeled) parentheses, instead of the usual labeled brackets, to indicate the relevant constituent structure of the forms.

FIRST PASS	((dwis) # su)	((dwais) + si)
SP1	*	*
SP2	*	*
RET1'	((dwiṣ) # su)	((dwaiṣ) + si)
RET7	*	*

4 The "first pass" of Chapter 2 now includes two passes, and the "second pass" is a third pass.

SECOND PASS	(dwiṣ # su)	(dwaiṣ + si)
SP1	vac	(dwaik + si)
SP2	(dwiṭ # su)	*
RET1'	*	(dwaik + ṣi)
RET7	*	*

FIRST PASS	((dwis) + t^has)	((haw + is) # su)
SP1	*	*
SP2	*	*
RET1'	((dwiṣ) + t^has)	((haw + iṣ) # su)
RET7	*	*
SECOND PASS	(dwiṣ + t^has)	(haw + iṣ # su)
SP1	vac	vac
SP2	*	*
RET1'	vac	(haw + iṣ # ṣu)
RET7	(dwiṣ + $ṭ^h$as)	*

FIRST PASS	((daaś) # su)	((daaś) + si)
SP1	*	*
SP2	*	*

RET1'	*	*
RET7	*	*
SECOND PASS	(daaś # su)	(daaś + si)
SP1	(daaṣ # su)	(daak + si)
SP2	(daaṭ # su)	*
RET1'	*	(daak + ṣi)
RET7	*	*

FIRST PASS	((daaś) + $b^h$is)
SP1	*
SP2	*
RET1'	*
RET7	*
SECOND PASS	(daaś + $b^h$is)
SP1	(daaṣ + $b^h$is)
SP2	(daaṭ + $b^h$is)
RET1'	*
RET7	*

For the voicing assimilation in daaḍ$b^h$is, see Chapter 5.

4.1.4. Roots in *kṣ*

When we consider root-final *ks*, we see that SP1 and SP2 predict forms like *cakṣṭi*, *cakṭsu*, and *cakṭ*, when in fact the actual forms are *cakṣi*, *caṭṣu*, and *caṭ*.

It is difficult to see how such forms can be accounted for if we continue to assume that *ks* occurs in root-final position. In particular, the form *caṭ* is troublesome, because by a general rule of Sanskrit--the reduction of a cluster of consonants before # to the first consonant of the cluster--*caks* before ## reduces to *cak*.

$$\text{ABBR.} \quad [+\text{obst}]_0 \longrightarrow \emptyset \;/\; [+\text{cons}] \;\_\_\_\; \#(\#)$$

Hence, any special rule for roots ending in *ks* must precede ABBR. The obvious solution would be to suppose that RET1' and a special *k*-dropping rule precede ABBR; then we would have

caks $\xrightarrow[\text{RET1'}]{}$ *cakṣ* $\longrightarrow$ *caṣ* $\xrightarrow[\text{SP2}]{}$ *caṭ*

Rather than assume the existence of a special *k*-dropping rule, let us instead search for an alternative representation of root-final *ks*.

An alternative representation must satisfy at least the following conditions:

a. It must reduce to *ṭ* in word-final position; this *ṭ*, however, could result from an *ṣ*, *ś*, *ṭ*, or *ḍ* in that position.

b. It must become *kṣ* before sonorants.

If we assume just the rules discussed so far, then the only clusters that suit both these conditions are *ṣṣ*, *ṣs*, *śṣ*, and *śs*. Now each of the clusters *ṣṣ*, *ṣs*, and *śṣ* contains an *ṣ* in a position where retroflexion is unpredictable. Rather than assume new instances of unpredictable *ṣ*, we take *śs* to be the cluster from which *ks* is derived. Then

	(# (*caśs*) + *mi* #)	(# (*caśs*) + *s* #)
ABBR	*	(# (*caś*) #)
FIRST PASS		
SP1	(# (*caks*) + *mi* #)	*
SP2	*	*
RET1'	(# (*cakṣ*) + *mi* #)	*
RET7	*	*
SECOND PASS	(# *cakṣ* + *mi* #)	(# *caś* #)
SP1	*	(# *caṣ* #)

SP2	*	(# caṭ #)
RET1'	vac	*
RET7	*	*

Next we consider the remaining environments--that is, non-finally before an obstruent. The output forms in these cases will be predicted correctly if there is a rule (ordered before SP1 and SP2) deleting s between ś and an obstruent. In fact, we could generalize the rule to delete s between any two obstruents:

$$\text{SP3.}\quad \begin{bmatrix}\text{+obst}\\ \text{+cont}\end{bmatrix} \longrightarrow \emptyset \;/\; [\text{+obst}]\ (\#)\ \_\_\_\_\ (\#)\ [\text{+obst}]$$

SP3 is in fact a well-known rule of Sanskrit (W233b-e). Some of the evidence for this rule is found in the forms of the s-aorist, active voice. When we compare the following examples--

	nii 'lead'	c^{h}id 'cut off'	tap 'heat'
1 du.	anaaiṣwa	acc^{h}aaitswa	ataapswa
2 du.	anaaiṣṭam	acc^{h}aaittam	ataaptam
3 du.	anaaiṣṭaam	acc^{h}aaittaam	ataaptaam
1 pl.	anaaiṣma	acc^{h}aaitsma	ataapsma
2 pl.	anaaiṣṭa	acc^{h}aaitta	ataapta

we see that although the forms of *nii* and the 1 du. and 1 pl. of *c^{h}id* and *tap* have the structure

$$((\underline{a} + \text{ROOT} + \underline{s}) + \text{ENDING})$$

the 2 du., 3 du., and 2 pl. of *c^{h}id* and *tap* are missing the *s*. But these are exactly the forms in which the *s* would appear between two obstruents. That is to say, SP3 would be necessary even if we did not postulate roots ending in *śs*. And although there are no quotable forms, we can suppose that the *s*-aorist of roots in *śs* will also show the deletion of *two* occurrences of *s* in the 2 du., 3 du., and 2 pl; from *raśs* 'protect' (which does have an *s*-aorist, although only a few forms occur in texts) we would expect

$$\underline{a} + \underline{\text{raa\'{s}s}} + \underline{s} + \underline{\text{tam}} \xrightarrow[\text{SP3}]{} \underline{a} + \underline{\text{raa\'{s}}} + \underline{\text{tam}}$$

$$\xrightarrow[\text{SP1}]{} \underline{a} + \underline{\text{raa\d{s}}} + \underline{\text{tam}} \xrightarrow[\text{RET7}]{}$$

$$\underline{a} + \underline{\text{raa\d{s}}} + \underline{\text{\d{t}am}}$$

Other derivations using SP3:

caśs + *si* $\xrightarrow{\text{SP3}}$ *caś* + *si* $\xrightarrow{\text{SP1}}$ *cak* + *si*

$\xrightarrow{\text{RET1}'}$ *cak* + *ṣi*

caśs + *d^hwam* $\xrightarrow{\text{SP3}}$ *caś* + *d^hwam* $\xrightarrow{\text{SP1}}$

caṣ + *d^hwam* $\xrightarrow{\text{SP2}}$ *caṭ* + *d^hwam*

$\xrightarrow{\text{RET7}}$ *caṭ* + *$ḍ^h$wam*

There is further support for our postulation of the final cluster *śs* instead of *ks*. This evidence is to be found in a few pairs of roots:

b^haa 'shine'	*b^haas* 'shine'
śru 'hear'	*śrus* 'hear'
haa 'leave, go forth'	*haas* 'go'

and possibly also

b^huu 'be'	*b^huus* 'attend upon, adorn'

The roots in the second column are clearly related to those in the first column and should not be considered as morphemically unanalyzable. Let us say that the roots in the second column contain a final morpheme *s*, of obscure semantic import.[5]

To the first list we now add

aś/āś 'attain'	*aks* 'attain'
naś/nāś 'attain'	*naks* 'attain'

and possibly also

naś 'be lost'	*niks* 'pierce'

These pairs are analogous to the pairs in the first list. It is then natural to say that *aks* should be represented as *aś* + *s*.

Note that what we are now suggesting is not simply the final cluster *śs*, but rather a final *ś* followed by a monosegmental morpheme *s*. Such a suggestion is, in fact, more reasonable than the postulation of an underlying cluster *śs*, because (a) final clusters of unlike spirants without a separating morpheme-boundary are unusual in any language, and (b) the cluster *śs* would seriously complicate the statement of morpheme-structure rules for Sanskrit.

There are other phonological phenomena and relationships that suggest that not *every* root-final

---

[5] Whitney 1885 refers to the roots with *s* simply as "secondary forms" of the roots without *s*.

ks be represented as *ś* + *s*. The two roots *jaks* 'eat' and *jaks* 'laugh' are discussed in the next chapter. Other instances can be adduced from the following examples:

mrj 'wipe'	*mrks* 'stroke'
b^haj 'divide, share'	*b^haks* 'eat, partake of' possibly *b^hiks* 'beg'

Here again we have a final morpheme *s*, and *mrks* should be represented as *mrj* + *s*. There is no evidence or counter-evidence to be found for this representation, because in every attested form of *mrks*, *b^haks*, and *b^hiks*, the root is followed by a sonorant.

Before turning our attention to assimilation rules, we note that SP3, as it stands, is not correct; because # is an obstruent, SP3 will delete the *s* in the initial clusters *st*, *st^h*, *sk*, *sp*, etc. We could prevent the deletion of initial *s* by changing the environment in SP3 to

$$\begin{bmatrix}\text{+obst}\\ \text{-BND}\end{bmatrix} (\#) \ \underline{\qquad}\ (\#) \begin{bmatrix}\text{-obst}\end{bmatrix}$$

A related problem arises in the operation of the abbreviation rule. Since # is an obstruent, ABBR deletes # when it occurs after a final obstruent and before #; that is, ABBR reduces every instance of ## after a final obstruent to #. We could change ABBR to require that the deleted segments be +obst, -BND. However, this type of problem is liable to occur again and again--perhaps so often that the number of times we are forced to refer to the marking -BND will offset the economies gained by having # contain segmental markings. It seems to be the case, though, that the problem ordinarily arises when rules refer to #, as do SP3 and ABBR. The necessity of mentioning -BND can be eliminated in such cases if we make the following general convention: If a rule mentions # or the marking +BND, then every segment (except a # or a segment already marked +BND) is assumed to be marked -BND. The effect of this convention is to restrict the usefulness of the markings on # to the expression of similarities between processes internal to words and processes in word-initial or word-final position.

4.2. Some Assimilations of Position

4.2.1. Internal Palatality Assimilation

We have been assuming that the retroflex consonants are retroflex palatals, even though they are derived in most cases from dentals. The alternation of ś with ṣ, and the identical treatment of ś and ṣ in internal sandhi provides some support for this assumption. Certainly, if ṣ is a flat dental, then the statement of SP1 is less simple than our rule in Section 4.1.2.:

$$\begin{bmatrix} +\text{obst} \\ +\text{cont} \\ -\text{grv} \\ \beta\ \text{comp} \\ -\beta\ \text{fl} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\ \text{comp} \\ \alpha\ \text{grv} \\ -\alpha\ \text{fl} \\ -\alpha\ \text{cont} \end{bmatrix} / \ \underline{\qquad} \ \begin{bmatrix} +\text{obst} \\ \alpha\ \text{cont} \end{bmatrix}$$

In addition, if the retroflex consonants are marked as dentals during the operation of the morphophonemic rules, then we must add a late phonetic rule to make them palatals (see Section 1.5.1.).

There is further evidence to be found for the assumption. We have already discussed the assimilation of retroflexion in internal (RET7) and external (RET8) sandhi. As a matter of fact, palatality assimilates in exactly the same fashion as retroflexion.

The only case that happens to occur in internal sandhi is the change of n to ń after a palatal stop

(E51)[6]--thus, *yaacńaa* 'an entreaty' (from *yaac* 'ask, beg') *yajńa* 'sacrificial' (from *yaj* 'sacrifice'), *raajńii* 'queen' (from *raajan* 'king'). However, *n* does not assimilate to a preceding *ś*--thus *aśna* 'voracious' (from *aś* 'eat'), and not *aśńa*. This apparent irregularity of *ṇ* would seem to prevent us from generalizing RET7 to

$$\begin{bmatrix}+\text{cons}\\-\text{voc}\\-\text{grv}\end{bmatrix} \rightarrow \begin{bmatrix}+\text{comp}\\\alpha\,\text{fl}\end{bmatrix} / \begin{bmatrix}+\text{cons}\\-\text{voc}\\+\text{comp}\\\alpha\,\text{fl}\end{bmatrix} \underline{\quad\quad}$$

But now we notice that it is not necessary for RET7 to apply to instances of *n*, for RET3' is applicable in every case. Any application of RET7 would be vacuous. For example, the occurrence of *ṇ* in *wrṣṇi* 'manly, strong' (from *wrs* '(to) rain'), *dʰrṣṇu* 'bold, fierce' (from *dʰrs* 'be bold'), and *jiṣṇu* 'victorious' (from *ji* 'win') can be explained on the basis of RET3' alone. Hence we can restrict the generalized rule to assimilation of non-nasal consonants:

$$\text{PAL1. (= RET7')} \quad \begin{bmatrix}+\text{cons}\\-\text{voc}\\-\text{nas}\\-\text{grv}\end{bmatrix} \rightarrow \begin{bmatrix}+\text{comp}\\\alpha\,\text{fl}\end{bmatrix} / \begin{bmatrix}+\text{cons}\\-\text{voc}\\-\text{grv}\\+\text{comp}\\\alpha\,\text{fl}\end{bmatrix} \underline{\quad\quad}$$

---

6 There is also a general assimilation of *n* to a following stop (E48).

4.2.2. External Palatality Assimilation

In external sandhi the situation is very similar, although the process is obscured by the operation of a number of other rules. The facts are as follows:

1. Final *t* or *d* becomes palatal before palatal stops and retroflex before retroflex stops (E28).
2. Final *n* becomes palatal before voiced palatal stops and retroflex before voiced retroflex stops (E24).
3. When final *n* precedes a voiceless nongrave stop a spirant homorganic with the stop is inserted after the *n*, which then becomes *anusvāra*[7] (E23, E49).
4. The combinations *t* ## *ś* and *n* ## *ś* become *c* ## *c*^{h} and *n* ## *c*^{h}, respectively (E31, E25).
5. Final *n* before *ṣ* or *s* remains, with the optional insertion of a *t* after the *n* (W207).
6. Final *t* before *s* or *ṣ* remains.

---

[7] Nasalization of the preceding vowel.

7. Final (non-radical) *s* or *ṣ* assimilates in position to any following voiceless stop or spirant. *Visarga* (*ḥ*) is often prescribed in place of *x* and *f*, whenever these would occur; similarly *visarga* is often prescribed in place of the first spirant in the combinations *s* ## *s*, *ṣ* ## *ṣ*, and *ś* ## *ś* (E11, E12, W170, W172).
8. Final *s* or *ṣ* before a voiced segment has a special treatment, which will be discussed in Chapter 6.

In tabular form the sandhi-products of the combinations in question (final dental before initial palatal or retroflex) are

Initial:

		c	*j*	*ṭ*	*ḍ*
Final:	*t*	*cc*	*jj*	*ṭṭ*	*ḍḍ*
	n	*nśc*	*ńj*	*nṣṭ*	*ṇḍ*
	s	*śc*	See Chap. 6	*ṣṭ*	See Chap. 6

		Initial:		
		ś	ṣ	s
Final:	t	cc^h	tṣ	ts
	n	$\acute{n}c^h$	n(t)ṣ	n(t)s
	s	śś	ṣṣ	ss
		or	or	or
		ḥś	ḥṣ	ḥs

If we postulate an external assimilation rule PAL2 that is a near analogue of PAL1 (applicable to nasals but not caused by spirants, and regressive rather than progressive) and if we assume that PAL2 _fol-lows_ a number of other rules, then the phenomena *listed* above can be described quite simply.

$$\text{PAL2. (= RET8')} \quad \begin{bmatrix} +\text{cons} \\ -\text{voc} \\ -\text{grv} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{comp} \\ \alpha\,\text{fl} \end{bmatrix} \Big/ \; \_\_\_ \; \# \; (\#) \begin{bmatrix} +\text{cons} \\ -\text{voc} \\ -\text{cont} \\ +\text{comp} \\ -\text{grv} \\ \alpha\,\text{fl} \end{bmatrix}$$

In the next three sections we examine the rules that obscure the operation of PAL2.

4.2.3. External Spirant Assimilation

According to PAL2, spirants do not cause assimilation. This is certainly true for the combinations *t* ## *ṣ* and *n* ## *ṣ*. There remain the combinations of *s* with a spirant and the combinations *t* ## *ś* and *n* ## *ś*--cases in which there apparently is assimilation.

The sandhi of *s* before an initial spirant is only a subcase of the sandhi of *s* before any initial voiceless segment (case 7 in the previous section). The appropriate rule is

$$\text{SP4.}\quad \begin{bmatrix}+\text{obst}\\+\text{cont}\end{bmatrix} \rightarrow \begin{bmatrix}\alpha\,\text{comp}\\\beta\,\text{grv}\\\gamma\,\text{fl}\end{bmatrix} \;/\; \underline{\quad\quad}\ \#\ (\#) \begin{bmatrix}+\text{obst}\\-\text{vcd}\\\alpha\,\text{comp}\\\beta\,\text{grv}\\\gamma\,\text{fl}\end{bmatrix}$$

that is, the assimilation of position in this case is covered not by PAL2 but by a general rule for spirants.

4.2.4. The ś ⟶ c^h Rule

For the combinations *t* ## *ś* and *n* ## *ś*, we add a rule that changes initial *ś* to *c^h* after *t* and *n*:

$$\text{SP5.}\quad \begin{bmatrix}+\text{obst}\\+\text{cont}\\+\text{comp}\\-\text{fl}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{cont}\\+\text{tns}\end{bmatrix} \;/\; \begin{bmatrix}+\text{voc}\\-\text{cons}\\-\text{cont}\\-\text{comp}\\-\text{grv}\end{bmatrix} \#\ (\#)\ \underline{\quad\quad}$$

If SP5 is ordered before PAL2, then *t* and *n* will assimilate (by PAL2) to the *c^h* created by SP5.

Examples (involving *daiwaan*, acc. pl. of *daiwa* 'god', and *śrṇauti*, 3 sg. act. pres. of *śru* 'hear'):

daiwaan ## *śrṇauti* $\xrightarrow[\text{SP5}]{}$ *daiwaan* ## *c^hrṇauti*

$\xrightarrow[\text{PAL2}]{}$ *daiwaań* ## *c^hrṇauti*

tat ## *śrṇauti* $\xrightarrow[\text{SP5}]{}$ *tat* ## *c^hrṇauti*

$\xrightarrow[\text{PAL2}]{}$ *tac* ## *c^hrṇauti*

4.2.5. Spirant Insertion

SP4 and SP5 handle the cases in which there is assimilation not directly predicted by PAL2. There are also the cases in which there is no assimilation although PAL2 predicts that there should be--the combination of *n* with a voiceless nongrave stop (case 3 in Section 4.2.2.). For these situations we add a spirant insertion rule:

$$\text{SP6.}\quad \emptyset \rightarrow \begin{bmatrix}+\text{cons}\\-\text{voc}\\+\text{obst}\\+\text{cont}\end{bmatrix} / \begin{bmatrix}+\text{cons}\\-\text{grv}\\+\text{nas}\end{bmatrix} \_\_\_ \;\#\; (\#) \begin{bmatrix}+\text{obst}\\-\text{cont}\\-\text{vcd}\end{bmatrix}$$

It is not necessary to specify the inserted spirant any further than this, for if SP6 is ordered before SP4, then SP4 will correctly mark the segment for compactness, gravity, and flatness. In our examples we indicate this partially unspecified spirant by $\underline{S}$.

SP6 must also precede SP5. For if SP5 were orderedbefore SP6, then the original combinations $\underline{n}$ ## $\underline{c^h}$ and $\underline{n}$ ## $\underline{\acute{s}}$ would fall together (as $\underline{n}$ ## $\underline{c^h}$) and both yield $\underline{n\acute{s}}$ ## $\underline{c^h}$, whereas the actual sandhi-forms are $\underline{n\acute{s}}$ ## $\underline{c^h}$ and $\underline{\acute{n}}$ ## $\underline{c^h}$, respectively. SP6 must follow the abbreviation rule, of course. We have so far established the relative ordering

ABBR.	Abbreviation.
SP6.	Spirant Insertion.
SP5.	$\underline{\acute{s}} \longrightarrow \underline{c^h}$
PAL2.	External Palatality Assimilation

with the further restriction that SP4 (External Spirant Assimilation) follow SP6.

Examples (involving *b^hawants* and *b^hawat*, masc. and neut. nom. sg., respectively, of pres. act. part. of *b^huu* 'be', and *ca* 'and'):

b^hawants ## *ca* $\xrightarrow{\text{ABBR}}$ *b^hawan* ## *ca*

$\xrightarrow{\text{SP6}}$ *b^hawanS* ## *ca* $\xrightarrow{\text{SP4}}$

b^hawanś ## *ca*

b^hawants ## *śṛnauti* $\xrightarrow{\text{ABBR}}$

b^hawan ## *śṛnauti* $\xrightarrow{\text{SP5}}$

b^hawan ## *c^hṛnauti* $\xrightarrow{\text{PAL2}}$

b^hawañ ## *c^hṛnauti*

b^hawat ## *ca* $\xrightarrow{\text{PAL2}}$ *b^hawac* ## *ca*

b^hawat ## *śṛnauti* $\xrightarrow{\text{SP5}}$ *b^hawat* ## *c^hṛnauti*

$\xrightarrow{\text{PAL2}}$ *b^hawac* ## *c^hṛnauti*

4.2.6. Visarga Rules

We remarked in Section 4.2.2. above that *ḥ* is an optional variant of *f* and *x* and of a final spirant before an initial spirant. The only other position in which *ḥ* occurs is absolute final; there it is an obligatory alternant of *s*, *r*, *ṣ*--thus, we have *manuḥ*, *punaḥ*, and *daiwaḥ* (from *manuṣ*--that is, *manus*--the nom. sg. of *manu* 'Manu', *punar* 'again', and *daiwas*, nom. sg. of *daiwa* 'god') before absolute pause (the sentence boundary \$).

All possible occurrences of *ḥ* can therefore be predicted by three late rules:

$$\text{VIS1. (optional)} \quad \begin{bmatrix} +\text{obst} \\ +\text{cont} \\ +\text{grv} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{obst} \\ -\text{cont} \\ +\text{comp} \end{bmatrix}$$

$$\text{VIS2. (optional)} \quad \begin{bmatrix} +\text{obst} \\ +\text{cont} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{obst} \\ -\text{cons} \\ -\text{fl} \\ +\text{comp} \\ +\text{grv} \end{bmatrix} / \ \_\_\_\_ \ \#(\#) \begin{bmatrix} +\text{obst} \\ +\text{cont} \end{bmatrix}$$

$$\text{VIS3.} \quad \begin{bmatrix} +\text{cons} \\ +\text{cont} \\ -\text{grv} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{obst} \\ -\text{cons} \\ -\text{fl} \\ +\text{comp} \\ +\text{grv} \end{bmatrix} / \ \_\_\_\_ \ \# \ \$$$

Examples using these rules will be given in Chapter 6.

4.3. Excursus on Relative Ordering

4.3.1. Types of Ordering

In earlier sections we determined the relative ordering of several rules, on the assumption that the set of rules is linearly ordered--with one kind of exception: a subset of the rules may be linearly ordered within a cycle. At any rate the rules are ordered in some fixed way.

Contrasting with such a <u>fixed ordering</u> is a <u>free ordering</u>. When a set of rules is freely ordered (or, as it is sometimes expressed, <u>unordered</u>), any rule that is applicable to a given form may be applied to it; when none of the rules applies, the derivation of an output form has been completed.

Often when a set of rules is freely ordered, certain ordering relations will hold between sets of rules in every derivation. For example, every application of Rule X might entail a later application of Rule Y, simply because Rule X introduces a symbol or configuration of symbols to which Rule Y applies. Or more elaborate restrictions might obtain: e.g., if both Rule X and Rule Y apply in any derivation, then Rule X always applies before Rule Y.

The effect of all such restrictions on a freely-ordered set of rules is to define a (partial) ordering of the set. Let us call this partial ordering the natural ordering of the set.

In an attempt to translate a freely ordered set of rules into a set with fixed ordering the most obvious strategy is to take advantage of the natural ordering. For example, our placement of RET1' (Retroflexion of s) before RET3' (Retroflexion of n) follows the natural ordering, as does our placement of both RET1' and RET3' before PAL1.

Of course, our object is not merely the translation of a freely ordered set of traditional rules (as expressed in terms of distinctive features) into a set with a fixed order. However, many of the traditional rules seem to be essentially correct (although often clumsily stated), so that our principal interests in them are (a) their elegant restatement in terms of distinctive features and (b) their ordering.

It is sometimes the case that two or more rules concerned with similar or related phenomena are stated so as to be mutually exclusive--and hence so as to bear no particular ordering relation to each other. In such a situation it often happens that the imposition of a fixed order on the rules allows for a considerable simplification in their statement[8]--often to the point of making one or more of the rules superfluous.

An example of such a simplification can be seen in our treatment of SP6 (Spirant Insertion) and SP4 (External Spirant Assimilation). If we order SP6 before SP4 it is not necessary to specify the inserted spirant for compactness, gravity, or flatness. Whitney treats the two cases quite separately--on the one hand there is spirant assimilation divided into three rules, one before initial spirants (W172):

> Before an initial sibilant... s is either assimilated, becoming the same sibilant, or it is changed into visarga.

one before initial voiceless dentals and palatals (W170c):

> Before the palatal lingual surd mutes... it is assimilated, becoming the sibilant of either class respectively ...

and one before initial voiceless velars or labials (W170d):

> Before the guttural and labial surd mutes... it is also theoretically assimilated, becoming respectively the jihvāmūlīya and upadhmānīya spirants ...; but in practice these breathings are unknown, and the conversion is to visarga.

---

[8] Some examples are discussed in Halle 1958, Halle 1962, and Chomsky 1964.

On the other hand, there is spirant insertion, combined with the anusvāra rule in W208:

> Before the surd palatal lingual, and dental mutes, there is inserted after final n a sibilant of each of those classes respectively, before which the n becomes anusvāra.

Emeneau follows Whitney--first two rules for spirant assimilation (E11 and E12), then the spirant insertion rule combined with the anusvāra rule (E23). Allen, like Renou,[9] considers the spirant assimilation as a single rule, but does not connect it (except historically) with spirant insertion.[10] MacDonell, although he separates the two rules, inserts a footnote leading from the first to the second.[11] But no one seems to have seen clearly the relation between the two rules.

4.3.2. Whitney and Ordering

In Section 4.1.3. we argued that SP1, SP2, and RET1' should apply cyclically.Now although the ordering of rules within a cycle is fixed, the operation of

---

9. Renou 1961, pp. 35, 37.

10 Allen 1962, pp. 74-6, 86-8.

11 MacDonell 1927, pp. 18, 21.

the cycle is such that a rule can apply to different stages in the derivation of a single form. This possibility also arises when a set of rules is freely ordered.

We should expect Whitney, whose rules are for the most part freely ordered, to make some reference to the double application of RET1' in forms like <u>dwaikṣi</u>. But Whitney's statements on this matter are not clear. In one section he states that "the occurrence of <u>ṣ</u> in Sanskrit words is nearly limited to cases falling under this rule" (W182) and cites, among other forms, the four exceptional roots <u>kaṣ</u>, <u>laṣ</u>, <u>b^{h}aṣ</u>, and <u>b^{h}aaṣ</u> (see Section 3.1.). In a later section, however, he says that "roots having a final sibilant (except <u>ç</u>) after an alternant vowel are ... regarded as ending in <u>ṣ</u>, not <u>s</u>." (W184c). That is, on the one hand he suggests that RET1' could apply twice--once in the conversion of stem-final <u>s</u> to <u>ṣ</u> and once in the conversion of suffixal <u>s</u> to <u>ṣ</u>--while on the other hand he discounts the first of the two possible applications.

For another example of Whitney's skirting of ordering problems we turn to his treatment of RET1' with respect to SP1 and SP2. To account for the double ṣ in hawiṣṣu Whitney says that "the alter[ation [RET1'] takes place in the initial s of an ending after the final s of a stem, whether the latter is to be regarded as also changed to ṣ or as converted into visarga" (W183). Whitney seems to realize that RET1' and PAL1 should follow SP1 and SP2; in W218 he says:

> Final... ç reverts to its original ... k, in internal combination, only before the... s of a verbal stem or ending (whence by 180 [RET1'] ,... kṣ); before... t and... th, it everywhere becomes... ṣ (whence by 197 [PAL2],... ṣṭ and... ṣṭh)

However, the "whence"s can reasonably be taken to refer to the natural ordering of the rules rather than to a crucial fixed ordering. The crucial cases are forms like dwiṭsu, which show that RET1' must follow SP1 and SP2, and that these rules cannot be freely ordered. Whitney seems to miss this point entirely, for he cites dwiṭsu not in connection with RET1' but rather in connection with PAL1; to his statement of the internal palatality assimilation rule he appends the remark that "those cases in which final ṣ becomes ṭ before su ... do not, of course, fall under this rule" (W197c). The force of the "of course" in this section is quite inexplicable, although two sections later he provides a perfectly good explanation for the absence of assimilation: "In external combination ... an initial dental after a final liquid usually remains unchanged; and <u>su</u> of the loc. pl. follows the same rule" (W199).

We have no way of determining whether the "whence"s in W218 refer to a fixed ordering or to a natural ordering, although the evidence points to the second interpretation: in addition to Whitney's failure to cite or discuss the crucial forms in their correct context, there is his general failure to notice simplifications that could be achieved by a fixed ordering.

There is one instance in which Whitney can be said to have placed a fixed ordering on his rules. The rules in question are the "rules of permitted finals"--the abbreviation rule, the rule of final devoicing and deaspiration (see the next chapter) and a "reversion rule" for palatals and <u>h</u>. These rules, according to Whitney, are to be applied before any rules of external sandhi (W152):

> For all the processes of external combination--that is to say, in composition and sentence-collocation--a stem-final or word-final is in general to be regarded as having, not its etymological form, but that given it by the rules as to permitted finals. From this, however, are to be excepted the s and the r ...

This use of "basic alternants," which Whitney has adopted from the Indian grammarians,[12] is the only clear instance of his intention to place a fixed ordering on his rules--and this only to a very restricted extent.

---

12 And which Allen 1962 attempts to expand into a theory of phonology. Allen applies a sort of simplicity principle: the simplicity of a set of rules is measured by the number of basic alternants that must be assumed for the correct operation of the rules. The basic alternants may, however, be created from underlying forms by the application of an initial set of rules. On this basis, it is desirable to have an initial rule that devoices final obstruents (thus reducing the number of basic alternants) plus a later external voicing assimilation rule. Cf. our discussion in the next chapter.

Chapter 5.

Aspiration and Voicing

There are a number of sandhi rules that account for alternations in aspiration and a number of rules that account for voicing assimilations. Voicing and aspiration are, however, related in several ways, so that we shall discuss both kinds of rules in this chapter.

5.1. Grassmann's Law

An inspection of Whitney 1885 shows that no Sanskrit root contains two aspirates. There are roots like c^{h}id 'cut', mat^{h} 'shake', and bud^{h} 'know', but none like c^{h}id^{h} or b^{h}ud^{h}. This absence of "double aspirates" is also evidenced in reduplication, for an aspirate is always reduplicated as a non-aspirate (W590a). For example,

		Present	Perfect
p^{h}al	'burst'	p^{h}alati	pap^{h}aala
b^{h}r	'bear'	bib^{h}arti	bab^{h}aara
d^{h}rs	'dare'	d^{h}ṛṣṇauti	dad^{h}arṣa
ḍ^{h}aauk	'approach'	ḍ^{h}aaukatai	ḍuḍ^{h}aaukai

c^{h}id	'cut'	c^{h}inatti	cic^{h}aida
k^{h}aad	'chew'	k^{h}aadati	cak^{h}aada
g^{h}us	'sound'	g^{h}auṣati	jug^{h}auṣa

Instead of accounting for these facts in two separate places--once in a morpheme structure rule and elsewhere in the rule(s) creating the syllable of reduplication--we can employ a single rule, ordered after the reduplication rule(s). From the cited forms it is clear that this rule, the familiar Grassmann's Law now operative in synchronic phonology, should be stated so as to deaspirate the first of two aspirates:

$$\text{GL.}\quad [\quad] \rightarrow [-\text{tns}] \;/\; \_\_\_\; [+\text{cont}]_1\; [+\text{tns}]$$

The operation of GL must be restricted to permit the generation of such forms as bib^{h}rt^{h}a and bib^{h}rd^{h}wai (2 pl. act. pres. and 2 pl. midd. pres., respectively, of b^{h}r), in which an aspirate in an ending has no effect upon a preceding aspirate in a root. The restriction can be accomplished by requiring that the second of the two aspirates be in a root (cf. SP2 in Section 4.1.2.):

$$\text{GL}'.\ [\quad] \rightarrow [-\text{tns}] \ / \ \underline{\quad\quad} \ [+\text{cont}]_1 \begin{bmatrix} +\text{tns} \\ +\text{ROOT} \end{bmatrix}$$

Now that we have added Grassmann's Law as a synchronic rule for Sanskrit, it is possible for the lexical representations of roots to contain more than one aspirate. And in fact, several roots (W155) have a "mobile aspirate" that indicates that they should be represented as having two aspirates. Compare:

		Present	Perfect	Future
bud^h		baud^{h}ati	bubaud^{h}a	b^{h}autsyati
band^h	'bind'	bad^{h}naati	baband^{h}a	b^{h}antsyati / band^{h}iṣyati

with

prac^h	'ask'	precc^{h}ati	papracc^{h}a	prakṣyati
krud^h	'be angry'	krud^{h}yati	cukraud^{h}a	krutsyati

Similarly, for the corresponding root nouns we have:

	nom. sg.	acc. sg.	inst. pl.	loc. pl.
bud^h	b^{h}ut	bud^{h}am	b^{h}udb^{h}is	b^{h}ut u
band^h	b^{h}ant	band^{h}am	b^{h}andb^{h}is	b^{h}antsu

but

	nom sg.	acc. sg.	inst. pl.	loc. pl.
pat^h 'go'	pat	pat^ham	$padb^his$	patsu
$krud^h$	krut	$krud^ham$	$krudb^his$	krutsu

In the case of such forms as b^hut and $b^hautsyati$ (as opposed to bud^ham and $baud^hati$) we seem to have the following situation: the root contains two aspirates--and thus is represented as b^hud^h--and when the second aspirate loses its aspiration, GL' is inapplicable. The rule for this loss of aspiration is quite general; any stop is deaspirated before a following obstruent (including the obstruent #):

$$\text{AV1.} \quad [+\text{obst}] \rightarrow [-\text{tns}] \;/\; \underline{\quad\quad} \; [+\text{obst}]$$

If Grassmann's Law is stated as GL', it must follow AV1: in the case of b^hud^h + am, AV1 would not apply, but GL' would, so that we would obtain bud^ham; but in the case of b^hud^h #, AV1 would apply to yield b^hud #, so that the operation of GL' would be blocked.

5.2. Bartholomae's Law

The situation is by no means as simple as it appears to be, however. Problems arise when we consider suffixes that begin with a voiceless dental stop--e.g., ta in the past participle, tum in the infinitive, twaa in the gerund, t^has in the 2 du. act., t^ha in the 2 pl. act., and tas in the 3 du. act. In many cases no unexpected forms occur; thus, from mad 'be exhilarated' + ta we obtain, by the regressive assimilation of voicing to be discussed in Section 5.5.1., matta.

But when a root-final voiced aspirate combines with a following t or t^h, the result is unexpected. voiced non-aspirate followed by d^h (Bartholomae's Law, W160). Thus,

rud^h 'oppress' + $\left\{\begin{matrix} t^h as \\ tas \end{matrix}\right\}$ yields rundd^{h}as ;

lab^h 'take' + ta yields labd^{h}a

The most obvious way of treating these cases is to postulate a rule that aspirates and voices t and t^h when they follow a voiced aspirate:

$$\text{BL.} \quad \begin{bmatrix} +\text{obst} \\ -\text{cont} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{tns} \\ +\text{vcd} \end{bmatrix} \Big/ \begin{bmatrix} +\text{obst} \\ -\text{cont} \\ +\text{vcd} \\ +\text{tns} \end{bmatrix} + \underline{\quad\quad}$$

BL would have to precede both AV1 and the ordinary voicing assimilation rule, and we would have such derivations as

$$\underline{rund^h} + \left\{\begin{matrix}\underline{t^has}\\ \underline{tas}\end{matrix}\right\} \xrightarrow[BL]{} \underline{rund^h} + \underline{d^has} \xrightarrow[AV1]{}$$

$$\underline{rund} + \underline{d^has}$$

$$\underline{lab^h} + \underline{ta} \xrightarrow[BL]{} \underline{lab^h} + \underline{d^ha} \xrightarrow[AV1]{} \underline{lab} + \underline{d^ha}$$

When we consider some further facts, however, the solution above (which is essentially a formalization of the traditional descriptions) becomes less attractive. The complicating factor is the behavior of double-aspirate roots like $\underline{b^hud^h}$ and $\underline{b^hand^h}$ in combination with suffixes beginning with $\underline{t}$ and $\underline{\ ^h}$. From $\underline{b^hud^h}$ + $\underline{ta}$ we obtain $\underline{budd^ha}$, and from $\underline{b^hand^h}$ + $\underline{t^has}$, $\underline{band d^has}$. It then follows that GL' must precede BL. But we have already indicated that BL must precede AV1, and that AV1 must precede GL'; hence, BL must precede GL'.

In other words, either these three rules apply in a cycle, or else they must be reformulated. They cannot, in their present form, be ordered with respect to each other. It can be shown that no arrangement of of the rules (as they now stand), even in a cycle, will result in the generation of the correct forms $\underline{buddh_i}$ ($\underline{b^hud^h}$ + $\underline{ti}$) and $\underline{b^hutsi}$ ($\underline{b^hud^h}$ + $\underline{si}$); see Section 5.6. for the demonstration. Some reformulation is necessary.

5.3. Reformulations

Instead of maintaining a general form of Grassmann's Law and attempting to block unwanted applications of it by means of the previous application of the deaspiration rule, we suggest that the operation of Grassmann's Law (as it applies in Sanskrit) is restricted to those situations in which the second aspirate is followed by a sonorant or by $\underline{t}$ or $\underline{t^h}$. If this restricted form of Grassmann's Law, which we can refer to as GL", is ordered before BL and AV1 (and BL before AV1, as before), then we have the following derivations:

$$\underline{b^hud^h} + \underline{ta} \xrightarrow[\text{GL}'']{} \underline{bud^h} + \underline{ta} \xrightarrow[\text{BL}]{} \underline{bud^h} + \underline{d^ha}$$

$$\xrightarrow[\text{AV1}]{} \underline{bud} + \underline{d^ha}$$

$$\underline{b^hud^h}\ \# \xrightarrow[\text{AV1}]{} \underline{b^hud}\ \#$$

$$\underline{b^hud^h} + \underline{am} \xrightarrow[\text{GL}'']{} \underline{bud^h} + \underline{am}$$

$$\underline{b^hud^h} + \underline{d^hwai} \xrightarrow[AV1]{} \underline{b^hud} + \underline{d^hwai}$$

$$\underline{b^hud^h} + \underline{si} \xrightarrow[AV1]{} \underline{b^hud} + \underline{si}$$

We note here that the form of BL, as given in Section 5.2. above, is considerably simplified by virtue of a special feature of Sanskrit endings. Because the only obstruents that begin endings are $\underline{s}$, $\underline{d^h}$, $\underline{b^h}$, $\underline{t}$,, $\underline{t^h}$, and #, the class consisting of $\underline{t}$ and $\underline{t^h}$ alone can be specified simply as the voiceless stops,

$$\begin{bmatrix} +\text{obst} \\ -\text{cont} \\ -\text{vcd} \end{bmatrix} .$$

The new version of Grassman's Law can be formalized as

$$\text{GL}''.\ [\ \ ] \rightarrow [-\text{tns}]\ /\ \_\_\_ [+\text{cont}]_1 \begin{bmatrix} +\text{tns} \\ +\text{ROOT} \end{bmatrix} \left\{ \begin{matrix} [-\text{obst}] \\ \begin{bmatrix} -\text{vcd} \\ -\text{cont} \\ -\text{BND} \end{bmatrix} \end{matrix} \right\}$$

The marking -BND prevents GL" from deaspiratin, the initial stop in $\underline{b^hud^h}$ #. This formulation is clumsy, at best, for the sonorants, _t_, and _t^h_ do not constitute a natural class. The clumsiness of this formulation does not depend upon our decision to assign markings to #; the two issues are quite separate. Let us search for a more natural formulation.

We begin by noting the relative rarity of the voiceless aspirates in lexical entries. _K^h_, _c^h_, and _p^h_ occur very infrequently. _T^h_ occurs somewhat more often, although it is nevertheless uncommon. In running text (as indicated by the figures given in W75) the same generalization holds, except in the case of _t^h_, which occurs fairly often in verbal endings. Now when a particular segment occurs only infrequently in the lexical entries of a language, the phonemic status of that segment should be reconsidered; _ceteris paribus_, it is likely either that the features of the segment are predictable from the features of other segments, or that the segment is better represented as a cluster of segments with a more certain phonemic status. When an entire natural class of segments is of infrequent occurrence, the likelihood is even stronger. We have already suggested that some retroflex consonants should be represented as clusters. We now suggest that the voiceless aspirates should also be so represented.

If voiceless aspirates are represented as clusters consisting of voiceless stop plus $\underset{\bullet}{h}$--that is, if their aspiration is viewed not as a marking, but rather as a separate phonological segment--then a natural formulation of GL" becomes possible, and moreover our formulation of BL can be simplified. What we shall do is to oppose the sonorants and $\underline{t(\underset{\bullet}{h})}$ to $\underline{s}$, $\underline{d^h}$, and $\underline{b^h}$ on the basis of the laxness of the former segments as against the tenseness of the latter segments. The final form of Grassmann's Law is then

$$\text{AV2 (= GL"'). } [-\text{cont}] \rightarrow [-\text{tns}] \ / \ \_\_\_ \ [+\text{cont}]_1 \begin{bmatrix}-\text{cont}\\+\text{tns}\\+\text{ROOT}\end{bmatrix} \begin{Bmatrix}-\text{tns}\\ \\ -\text{BND}\end{Bmatrix}$$

Furthermore, since in this reformulation all aspirate stops are voiced (at least until the operation of a later rule creating voiceless aspirate stops from clusters), BL can be simplified to

$$\text{AV3 (= BL'). } [+\text{obst}] \rightarrow [+\text{tns}] \ / \ [+\text{tns}] + \_\_\_$$

Two new rules must be added: one specifying that all aspirates are voiced, and the other creating aspirated stops from clusters. That is,

$$\text{AV4.}\quad \begin{bmatrix}+\text{tns}\\ -\text{cont}\end{bmatrix} \longrightarrow \begin{bmatrix}+\text{vcd}\end{bmatrix}$$

$$\text{AV5.}\quad \underset{1}{\begin{bmatrix}-\text{cont}\end{bmatrix}} \underset{2}{\begin{bmatrix}-\text{vcd}\\ -\text{cons}\\ -\text{vcd}\end{bmatrix}} \longrightarrow \begin{bmatrix}1\\ +\text{tns}\end{bmatrix}$$

The order of the rules is AV2-5 followed by AV1. Sample derivations:

	b^hud^h + t(ḥ)a	b^hud^h #	b^hud^h + am
AV2.	bud^h + t(ḥ)a	*	bud^h + am
AV3.	bud^h + t^h(ḥ)a	vac	*
AV4.	bud^h + d^h(ḥ)a	vac	vac
AV5.	bud^h + d^ha	*	*
AV1.	bud + d^ha	b^hud #	*

---

	b^hud^h + d^hwai	b^hud^h + si
AV2.	*	*
AV3.	vac	vac
AV4.	vac	vac
AV5.	*	*
AV1.	b^hud + d^hwai	b^hud + si

5.4. Evidence

5.4.1. The Union Vowel

Evidence for the correctness of our analysis of the voiceless aspirates (and also the retroflex stops) as clusters can be found by examining the occurrence of the union vowel i before three suffixes: ta of the past participle, tum of the infinitive, and twaa of the gerund. In general, a root takes the union vowel either with all three suffixes or with none of them. The correspondence is not perfect, however; some roots have alternative forms--e.g., from śak 'be able', both śakta and śakita--and a few have the union vowel with one suffix but not with another--e.g., from ja(m)b^h 'chew up, crush', jab^{h}itum but jabd^{h}a.

For many roots it is difficult to find any explanation for the occurrence or non-occurrence of the union vowel. Some very interesting generalizations can be made, nevertheless. Whitney characterizes many of the roots that take the union vowel as "presenting difficulties of combination"(W956a); of these roots, three types are of special interest to us:

1. Roots ending in two consonants, with the exception of those that drop a penultimate nasal. Thus, *ubjita* from *ubj* 'force' but *badd^{h}a* from *band^{h}* (actually, *b^{h}and^{h}*, as we have seen). The generalization is quite reliable for this class of roots, and the "difficulties of combination" are obvious. However, roots ending in *ks* (that is to say, *ś* + *s*) do not invariably take the union vowel: of the thirteen roots in *ks* for which Whitney 1885 cites at least one of the three forms in question, one (*kaañks*) clearly ends in a consonant cluster and takes the union vowel, one (*jaks*) is a special case that we will discuss in a following section, eight take the union vowel, and three (*aks*, *caks*, and *taks*) do not. It is not unexpected that there should be some variation here, for *ś* + *s* + *t* is not, strictly speaking, a "difficult" combination. There is always the possibility of transforming the combination into *ṣṭ* (as in *aṣṭa*, *caṣṭum*, *taṣṭa*) by means of the rules in the previous chapter.

2. "All [roots] that end in linguals (including ṣ after a or ā)" (W956a2). Whitney's judgment that these roots present difficulties of combination is puzzling. If these roots end in a single retroflex consonant (as Whitney assumes they do), we should expect an assimilation of retroflexion (Section 4.2.1.); aṭṭa would be the ordinary resultant of aṭ + ta (from aṭ 'wander'). However, we have argued that root-final retroflex stops should be represented as clusters with l as first member. The fact that all roots ending in these postulated clusters invariably take the union vowel provides further support for our argument. We can provide an explanation for the occurrence of the union vowel with these roots: Such roots are merely additional examples of roots ending in two consonants.

Of the four roots ending in "unpredictable" ṣ (i.e., in the postulated cluster ls), three (kaṣ, bʰaṣ, and laṣ) occur only with the union vowel. For the fourth, bʰaaṣ, the forms bʰaaṣita and bʰaaṣitwaa are the only attested past participle and gerund, respectively, although both bʰaaṣitum and bʰaaṣṭum are attested as infinitives. We would expect the union vowel in every case, so that bʰaaṣṭum is an exceptional form.

In the case of roots ending in "predictable" ṣ, we would expect no union vowel. There is no particular difficulty of combination here, even though Whitney appears to have included these roots under the heading "all that end in linguals." Of the thirty roots with at least one of the relevant forms attested, sixteen never take the union vowel, ten always take it, and four are variable (dʰrs 'dare': dʰr̥ṣṭa or dʰr̥ṣita; mus 'steal': muṣṭa or <u>muṣita</u>, but <u>muṣitwaa</u>; <u>rus</u> 'be vexed': <u>ruṣṭa</u> or <u>ruṣita</u>; <u>wrs</u> 'rain': <u>wrṣṭa</u> and <u>wrṣṭwaa</u>, but <u>warṣitum</u>).

The treatment of "predictable" <u>ṣ</u> is thus quite variable. However, the treatment of roots ending in a single consonant is in general variable--e.g., of the twenty-two roots ending in <u>p</u> (preceded by a vowel) for which the relevant forms are attested, ten never take the union vowel, five always take it, and seven are variable. There seems to be no generalization possible for roots ending in a single consonant.

3. All roots ending in voiceless aspirates. Once again, Whitney's judgment that we are confronted with a difficult combination is incomprehensible. If these roots end in a single stop (as Whitney assumes they do), then we should expect the stop to be deaspirated (by AV1); <u>gratta</u> would be the ordinary resultant of

<u>grat^h</u> + <u>ta</u> (from <u>grat^h</u> 'tie'). However, we have argued that voiceless aspirate stops should be represented as clusters with <u>ḥ</u> as a second member. The fact that all roots ending in these postulated clusters invariably take the union vowel provides further support for this argument.

5.4.2. <u>Consonants of Reduplication</u>

We have presented positive evidence that voiceless aspirates should be represented as clusters. It remains for us to present non-negative evidence--that is, we must show that such a representation, while simplifying some descriptions and permitting new explanations for certain phenomena, does not at the same time necessitate complication in the description of other phenomena. In the case at hand, the method of formation of reduplicating syllables provides relevant non-negative evidence.

Consonants are reduplicated in the same way in each of the formations involving reduplication (the present system of the reduplicating class, the perfect system, the reduplicative aorist, the intensive, and the desiderative). The only complication occurs in the intensive, in which both the final and initial consonants of the root are reduplicated. We are here concerned only with the reduplication of initial consonants. The general principle (W589) is as follows:

1. If the root begins with a single consonant followed by a vowel, that consonant is reduplicated.
2. If the root begins with a consonant cluster, *one* of the consonants in the cluster is reduplicated. In the case of a spirant-plus-stop cluster, the stop is reduplicated. Otherwise, the initial consonant of the cluster is reduplicated.

Thus,

<table>
<tr><th>Cluster</th><th>Reduplicating Consonant</th></tr>
<tr><td>st</td><td>t</td></tr>
<tr><td>sk</td><td>k</td></tr>
<tr><td>śc</td><td>c</td></tr>
<tr><td>sn</td><td rowspan="4">s</td></tr>
<tr><td>sm</td></tr>
<tr><td>sr</td></tr>
<tr><td>sl</td></tr>
<tr><td>ks</td><td>k</td></tr>
<tr><td>gr</td><td>g</td></tr>
</table>

Two further conditions are traditionally added to this principle:

3. An aspirate is reduplicated as a non-aspirate.
4. A velar or <u>h</u> is reduplicated as a palatal.

Let us consider condition (3). When the consonant to be reduplicated is a voiced aspirate, deaspiration is accomplished by Grassmann's Law (in any of its versions). In the early versions of Grassmann's Law, this rule also accounted for the deaspiration of voiceless aspirates. However, in its final form (as AV2) Grassmann's Law is no longer related to the deaspiration of voiceless aspirates, which are now clusters. Instead, condition (2) above explains the lack of aspiration: Because <u>th</u> is a cluster, only the initial consonant, <u>t</u>, is reduplicated. No complication is introduced by the new representation of <u>t^h</u>.

5.4.3. <u>The Root jaks</u>

Having established rules AV1-5, we are now in a position to discuss the root <u>jaks</u> 'eat', an exception to our treatment of roots in <u>ks</u> in Chapter 4.

In most of the attested forms of <u>jaks</u> the ending or suffix begins with a sonorant, so that the root shows *ks*. However, there are a number of attested forms showing the root in combination with a *t*--3 sg. act. pres. *jagd^{h}i*, past participle *jagd^{h}a*, infinitive *jagd^{h}um*, gerund *jagd^{h}waa*. In every example we observe that the *t* has become *d^h*--a state of affairs that could only have resulted from an application of AV3 (Bartholomae's Law). But AV3 is applicable only when the root ends in a voiced aspirate. We conclude that the root is correctly represented as *jag^{h}s* (and not *jaks*, *jags*, or *jaś* + *s*), and that we have another example of the application of SP3 (Spirant Deletion):

$$\underline{jag^h s} + \underline{ta} \xrightarrow[\text{SP3}]{} \underline{jag^h} + \underline{ta} \xrightarrow[\text{AV3}]{} \underline{jag^h} + \underline{t^h a}$$

$$\xrightarrow[\text{AV4}]{} \underline{jag^h} + \underline{d^h a} \xrightarrow[\text{AV1}]{} \underline{jag} + \underline{d^h a}$$

In a few forms of the *s*-aorist we can observe SP3, AV3, AV4, and AV1 again operating together (W881d). Thus, from the root *ru(n)d^h* we have 1 sg. midd. *arutsi* and 3 sg. midd. *arudd^{h}a*:

$$\underline{a} + \underline{rud^h} + \underline{s} + \underline{i} \xrightarrow[\text{AV1}]{} \underline{a} + \underline{rud} + \underline{s} + \underline{i}$$

$$\underline{a} + \underline{rud^h} + \underline{s} + \underline{ta} \xrightarrow[SP3]{} \underline{a} + \underline{rud^h} + \underline{ta}$$

$$\xrightarrow[AV3]{} \underline{a} + \underline{rud^h} + \underline{t^ha} \xrightarrow[AV4]{}$$

$$\underline{a} + \underline{rud^h} + \underline{d^ha} \xrightarrow[AV1]{} \underline{a} + \underline{rud} + \underline{d^ha}$$

Our decision about the representation of jaks 'eat' is supported by the relationship between this root and the root g^{h}as, also 'eat'; jaks is then a reduplicated form of g^{h}as: ja + g^{h}s. There is a similar relationship between another root jaks 'laugh' and the root has, also 'laugh'.

Another form with "zero-grade" of a root, as in ja + g^{h}s is of some interest in connection with the aspiration rules: bapsati, 3 pl. act. pres. of b^{h}as 'devour', in contrast to the 3 sg.act. pres. bab^{h}asti These are forms of the reduplicating present; the underlying forms are presumably b^{h}a + b^{h}s + a + ti and b^{h}a + b^{h}as + ti, respectively. However, from b^{h}a + b^{h}s + a + ti we would expect to derive b^{h}apsati, because the s following the second b^h blocks the application of AV2. We must assume that the reduction of the root occurs after AV2 and before AV1, so that the derivation proceeds as follows:

$$\underline{b^ha} + \underline{b^has} + \underline{a} + \underline{ti} \xrightarrow[\text{AV2}]{}$$

$$\underline{ba} + \underline{b^has} + \underline{a} + \underline{ti} \longrightarrow \underline{ba} + \underline{b^hs} + \underline{a} + \underline{ti}$$

$$\xrightarrow[\text{AV1}]{} \underline{ba} + \underline{bs} + \underline{a} + \underline{ti}$$

For the 2 pl. act. impv. of $\underline{d^haa}$ 'place'-- $\underline{d^hatta}$, presumably from the reduplicated form $\underline{d^ha} + \underline{d^haa} + \underline{ta}$ with reduction of the root to $\underline{d^h}$--our rules predict $\underline{dadd^ha}$. Lehmann[1] suggests that a root-final laryngeal (our $\underline{ḥ}$) would block AV3. Indeed it would; it would also block AV2 (because $\underline{ḥ}$ is tense) so that from $\underline{d^ha} + \underline{d^hḥ} + \underline{ta}$ we would have the derivation

$$\underline{d^ha} + \underline{d^hḥ} + \underline{ta} \xrightarrow[\text{AV5}]{} \underline{d^ha} + \underline{d^h} + \underline{ta}$$

$$\xrightarrow[\text{AV1}]{} \underline{d^ha} + \underline{d} + \underline{ta}$$

(whence $\underline{d^hatta}$)

But in this example the root must be reduced <u>before</u>

---

[1] Lehmann 1955, p. 83.

AV2 (otherwise AV2 would apply), whereas in the previous example we required that the root be reduced _after_ AV2. We have at present no way of resolving this apparent contradiction. Certainly Lehmann's postulation of _ḥ_ must be examined more carefully.

5.5. Voicing Assimilation

5.5.1. Internal Voicing Assimilation

We have already cited many forms illustrating the operation of the rule of internal voicing assimilation: an obstruent assimilates in voicing to a following obstruent. Formally:

$$\text{AV6.} \quad [+\text{obst}] \rightarrow [\alpha\ \text{vcd}] \ / \ \_\_\_ \begin{bmatrix} +\text{obst} \\ \alpha\,\text{vcd} \end{bmatrix}$$

There are very few forms that provide any sort of evidence about the ordering relations between AV6 and other rules. The examples _jag^h^s_ + _ta_ and _a_ + _rud^h^_ + _s_ + _ta_ in the previous section indicate that AV6 must follow SP3; if the ordering were reversed, the outputs would be _jakta_ and _arutta_, respectively. AV3 and AV4 must also precede AV6, since together AV3 and AV4 have the effect of a progressive voicing assimilation. That is the extent of the evidence on ordering.

5.5.2. External Voicing Assimilation

The assimilation in external, as in internal, sandhi is regressive. It is more general, however: a final obstruent assimilates in voicing to any following initial segment, obstruent or sonorant:

AV7. $[+\text{ obst}] \rightarrow [\alpha\text{ vcd}]$ / ___ $\#(\#)$ $[\alpha\text{ vcd}]$

The traditional discussion of this rule is somewhat complicated by the postulation of a special "permitted finals" rule that devoices all word-final obstruents. There is no necessity for any sort of special rule, for AV7 will adjust the voicing correctly in every context except before absolute pause. And this case can also be handled if we (quite reasonably) assign the feature -vcd to the boundary $.

We have as yet very little evidence about the position of AV7 within the set of rules. This matter will be discussed in the next chapter.

Adjustments in the form of AV7 will be made in Section 6.1.1.

5.6. Appendix

In this appendix we show that no arrangement of Grassmann's Law (GL), Bartholomae's Law (BL) and the deaspiration rule (AV1) yield only correct outputs.if these three rules are formulated as in Sections 5.1. and 5.2. This result holds even if it is assumed that the rules apply in a cycle.

There are six possible arrangements of the three rules. In the tables below each possible arrangement is examined to determine the effect of the rules on the two underlying forms b^hud^h + *ti* and b^hud^h + *si*. Incorrect output forms are marked with the sign ¢ at the bottom of the column. Assimilation of voicing is disregarded.

1. Order: GL, BL, AV1.

GL.	bud^h + *ti*	bud^h + *si*
BL.	bud^h + d^hi	*
AV1.	*bud* + d^hi	*bud* + *si*
		¢

2. Order: GL, AV1, BL.

GL.	bud^h + *ti*	bud^h + *si*
AV1.	*bud* + *ti*	*bud* + *si*
BL.	vac	vac
	¢	¢

3. Order: BL, GL, AV1.

BL.	$\underline{b^hud} + \underline{d^hi}$	*
GL.	$\underline{bud^h} + \underline{d^hi}$	$\underline{bud^h} + \underline{si}$
AV1.	$\underline{bud} + \underline{d^hi}$	$\underline{bud} + \underline{si}$
		¢

4. Order: BL, AV1, GL.

BL.	$\underline{b^hud^h} + \underline{d^hi}$	*
AV1.	$\underline{b^hud} + \underline{d^hi}$	$\underline{b^hud} + \underline{si}$
GL.	*	*
	¢	

5. Order: AV1, GL, BL.

AV1.	$\underline{b^hud} + \underline{ti}$	$\underline{b^hud} + \underline{si}$
GL.	*	*
BL.	*	*
	¢	

6. Order: AV1, BL, GL.

AV1.	$b^hud + ti$	$b^hud + si$
BL.	*	*
GL.	*	*
	¢	

It is of course possible that the rules might apply cyclically. In each example GL would apply only on the first pass; we indicate below the effects of the second pass on the forms bud^h + ti and bud^h + si.

7. Order: GL, BL, AV1

GL.	vac	vac
BL.	bud^h + d^{h}i	*
AV1.	bud + d^{h}i	bud + si
		¢

8. Order: GL, AV1, BL.

GL.	vac	vac
AV1.	bud + ti	bud + si
BL.	*	*
	¢	¢

9. Order: BL, GL, AV1.

BL.	bud^h + d^{h}i	*
GL.	vac	vac
AV1.	bud + d^{h}i	bud + si
		¢

10. Order: BL, AV1, GL

BL.	<u>bud</u>^{h} + <u>d^{h}i</u>	*
AV1.	<u>bud</u> + <u>d^{h}i</u>	<u>bud</u> + <u>si</u>
GL.	*	*
		¢

11. Order: AV1, GL, BL.

AV1.	<u>bud</u> + <u>ti</u>	<u>bud</u> + <u>si</u>
GL.	*	*
BL.	*	*
	¢	¢

12. Order: AV1, BL, GL.

AV1.	<u>bud</u> + <u>ti</u>	<u>bud</u> + <u>si</u>
BL.	*	*
GL.	*	*
	¢	¢

In each of the twelve methods of ordering, at least one of the output forms is incorrect.

Chapter 6

The Sandhi of r and s

6.1. External Sandhi

In Chapter 4 we treated the external sandhi of s and ṣ before voiceless initials. In the following sections we shall discuss the remaining cases and their relation to the rules of vowel sandhi (Chapter 2).

Emeneau's rules for the sandhi of s and r:

E11. "-s and -r, final in phrase or before a sibilant or a voiceless labial or velar stop, are replaced by ḥ."

E12. "-s and -r, before a voiceless palatal, retroflex, or dental stop, are replaced by the sibilant corresponding to the stop."

E13. "-as plus initial a yield o."

E14. "-as before any other vowel than a loses s and hiatus remains."

E15. "-as before a voiced consonant is replaced by o."

E16. "-ās before a voiced sound loses s, and if the voiced sound is a vowel, hiatus remains."

E17. "Any other case of -s and any case of -r, before r is dropped with lengthening of the preceding vowel if it is short."

E18. "Any other case of -s before any voiced sound other than -r, is replaced by r."

E21. "Any other case of -r remains, i.e. before other voiced sounds than itself."

(E19 and E20 deal with three exceptional forms.)

6.1.1. Sandhi After Noncompact Vowels

Rules E11 and E12, as they apply to s/ṣ, have already been discussed. It remains to adjust or add to the rules we already have so that r before voiceless initials will be treated in exactly the same way as s and ṣ. That is, we must either adjust PAL2 (External Palatality Assimilation) and SP4 (Spirant Assimilation) so that r will be affected by these rules, or we must make r indistinguishable from ṣ by the time these two rules apply. The second of these two courses is the simpler, if we assume that AV7 (External Voicing Assimilation) applies to *r* and that AV7 is ordered before PAL2 and SP4. Then AV7 creates a voiceless *r* (which we shall denote by *R*) before voiceless initials. A rule must then be added to shift *R* to *ṣ*--a change in the features of obstruency and vocalicity only.

Next we notice that a rule to shift *ẓ* to *r* is also needed. This rule will have the same effect as E18, since "any other case of -s before a voiced sound" will be retroflex (E13-16 handle those cases of final *s* to which RET1' is not applicable) and voiced (by AV7). The shift of *R* to *ṣ* and the shift of *ẓ* to *r* can be combined into a single rule, RS1 (an "*r* ⟷ *s*" rule). The revised form of AV7 and the new rule RS1 are as follows:

$$\text{AV7}'.\quad \begin{bmatrix}+\text{cons}\\ -\text{nas}\end{bmatrix} \rightarrow [\alpha\,\text{vcd}] \;/\; \underline{\qquad}\ \#(\#)\ [\alpha\,\text{vcd}]$$

$$\text{RS1}.\quad \begin{bmatrix}+\text{cons}\\ +\text{cont}\\ \alpha\,\text{vcd}\end{bmatrix} \rightarrow \begin{bmatrix}-\alpha\,\text{obst}\\ \alpha\,\text{voc}\end{bmatrix}$$

AV7' cannot be generalized to affect all +*cons* segments, since the nasals are not devoiced before voiceless initials.

Examples (involving *punar* 'again'; *ca* ''and'; *śṛṇauti*, 3 sg. act. pres. of *śru* 'hear'; *agnis*, nom. sg. of *agni* 'Agni'; and *ati* 'beyond')

punar ## *ca* $\xrightarrow[\text{AV7}']{}$ *punaR* ## *ca* $\xrightarrow[\text{RS1}]{}$

punaṣ ## *ca* $\xrightarrow[\text{PAL2}]{}$ *punaś* ## *ca*

punar ## *śṛṇauti* $\xrightarrow[\text{AV7}']{}$ *punaR* ## *śṛṇauti*

$\xrightarrow[\text{RS1}]{}$ *punaṣ* ## *śṛṇauti* $\xrightarrow[\text{SP4}]{}$

punaś ## *śṛṇauti* $\xrightarrow[\text{VIS2 (opt)}]{}$

punaḥ ## *śṛṇauti*

agnis ## *ati* $\xrightarrow[\text{RET1}']{}$ *agniṣ* ## *ati*

$\xrightarrow[\text{AV7}']{}$ *agniẓ* ## *ati* $\xrightarrow[\text{RS1}]{}$

agnir ## *ati*

We turn now to E17. If this rule is ordered after RS1, then there is no need to specify "any other case of -s"; the rule will apply simply to any r followed by initial r. Since the natural source of a tense vowel is a cluster of two like vowels, we order this rule before the rules for vowel sandhi.

$$\text{RS2.}\quad \underset{1}{\begin{bmatrix}+\text{voc}\\-\text{cons}\end{bmatrix}}\ \underset{2}{\begin{bmatrix}+\text{voc}\\+\text{cons}\\+\text{comp}\end{bmatrix}}\ \underbrace{\#(\#)\ \begin{bmatrix}+\text{voc}\\+\text{cons}\\+\text{comp}\end{bmatrix}}_{3} \longrightarrow 1\ 1\ 3$$

Examples (involving the voc. sg. of ratna 'jewel'):

$$\text{punar } \#\# \text{ ratna} \xrightarrow[\text{RS2}]{} \text{punaa } \#\# \text{ ratna}$$

$$\text{agnis } \#\# \text{ ratna} \xrightarrow[\text{RET1}']{} \text{agni}\d{s} \ \#\# \text{ ratna}$$

$$\xrightarrow[\text{AV7}']{} \text{agni}\d{z} \ \#\# \text{ ratna} \xrightarrow[\text{RS1}]{}$$

$$\text{agnir } \#\# \text{ ratna} \xrightarrow[\text{RS2}]{} \text{agnii } \#\# \text{ ratna}$$

We have so far established the following ordering:

RET1'. Retroflexion of s.

AV7'. External Voicing Assimilation.

RS1. r $\longleftrightarrow$ s.

RS2. Simplification of rr.

Rules for the sandhi of az and aaz (as and aas before voiced initials will al-

ready have been affected by AV7')
and for vowel sandhi; PAL2. External
Palatality Assimilation; SP4. Spirant Assimilation.

VIS1-3. *Visarga* rules.

The effects of E11 and E12 together are produced by AV7', RS1, PAL2, SP4, and the *visarga* rules together. The effect of E18 is produced by RET1', AV7', and RS1. The effect of E17 is produced by RET1', AV7', RS1, and RS2. E21 does not need to be explicitly stated, of course. It thus appears that we have replaced five rules of Emeneau's by nine rules of our own (RET1', AV7', RS1, PAL2, SP4, RS2, VIS1-3). However, RET1', AV7', and PAL2 are motivated on quite independent grounds. Moreover, Emeneau's statement of the conditions on the appearance of *visarga* is oversimplified; our rules take account of the variation between *visarga* and spirants. Finally, our rules are simpler than Emeneau's would be if his were directly formalized in terms of distinctive features.

6.1.2. Sandhi after a and ā

We are left with four of Emeneau's rules (E13-E16) to account for; in slightly more formal versions these rules are

$$\text{E13.}\quad \underline{\breve{a}z}\ \#\#\ \underline{\breve{a}} \longrightarrow \underline{o}$$

$$\text{E14.}\quad \underset{1\,2}{\underline{\breve{a}z}}\ \ \underset{3}{\#\#}\ \underset{4}{\begin{bmatrix}+\text{voc}\\ -\text{cons}\\ \begin{Bmatrix}+\text{tns}\\ -\text{comp}\end{Bmatrix}\end{bmatrix}} \rightarrow 1\ 3\ 4$$

$$\text{E15.}\quad \underset{1\,2}{\underline{\breve{a}z}}\ \ \underset{3}{\#\#}\ \underset{4}{\begin{bmatrix}\begin{Bmatrix}+\text{voc}\\ -\text{cons}\end{Bmatrix}\end{bmatrix}} \rightarrow \underline{o}\ 3\ 4$$

$$\text{E16.}\quad \underline{z} \longrightarrow \emptyset\ /\ \underline{\bar{a}}\ \_\_\_\ \#\#$$

All four rules involve the deletion (or at least the disappearance) of a z, a segment that happens not to occur anywhere in Sanskrit anyway. E13 and E15 are similar in that they involve the appearance of an o. We have already discussed the deletion of initial $\underline{\breve{a}}$ called for by E13 (VOW6 in Section 2.3.)

It then appears that E13-E16 could be translated into two rules--the creation of o in certain environments followed by the general deletion of z. Let us first consider the deletion of z. On one hand (a) there is the case of the combination $\underline{\breve{a}z}$ ## $\underline{\breve{a}}$.

Here the deletion of *z* must precede VOW6, for otherwise the initial *ǎ* would not be deleted; the resultant of *nalaz* ## *ati* (*nalas*, nom. sg. of *nala* 'Nala') is *nalati*, not *nalawati*. But on the other hand (b) there is the case of *ǎz* before any vowel except *ǎ*, and also the case of *āz* before any vowel. Here the deletion of *z* must follow VOW3' (Mid Vowels) and VOW4' (Postvocalic Glides)--and hence must follow VOW6--because the hiatus resulting from the deletion remains; the resultant of *nalaz* ## *iti* is *nalaiti*, not *naleti*, and the resultant of *kamaaz* ## *iti* (*kāmās*, nom. pl. of *kaama* 'love') is *kāmāiti*, not *kāmeti*.

We are forced to postulate two *z*-deletion rules; for case (b) a late rule prescribing the general deletion of *z*, for case (a) an earlier rule prescribing this deletion only in certain circumstances. The late rule is simply

$$\text{RS3.}\quad \begin{bmatrix} +\text{obst} \\ +\text{cont} \\ +\text{vcd} \end{bmatrix} \longrightarrow \emptyset$$

For case (a)--and the related case of *ǎz* before an initial consonant--we note that the natural source of an *o* is the cluster *au*, so that the most obvious method of creating an *o* would be to insert a *u* after the *ă* in *ăz*; the *z* would be deleted, and VOW6 would apply. However, several considerations weigh against this approach.

First, the *u*-insertion rule is by no means simple. The rule would be "Insert *u* between *ă* and *z* when the *z* is final and when an initial non-vowel or an initial vowel other than *ă* follows." A completely formalized version of the rule would take one of two forms, depending upon its ordering with respect to VOW1'. If *u*-insertion precedes the coalescence of like vowels, the simplest version of the rule is

$$\emptyset \rightarrow \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{grv} \\ -\text{comp} \end{bmatrix} /$$

$$\left[\begin{Bmatrix} +\text{cons} \\ -\text{voc} \end{Bmatrix}\right] \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{comp} \end{bmatrix} \_\_\_ \begin{bmatrix} +\text{cons} \\ -\text{voc} \\ +\text{cont} \end{bmatrix}$$

$$\#\#\left(\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{comp} \end{bmatrix}\right) \left[\begin{Bmatrix} +\text{cons} \\ -\text{voc} \end{Bmatrix}\right]$$

If *u*-insertion follows the coalescence of like vowels (so that we are dealing with lax vs. tense vowels, instead of single segments vs. clusters), the simplest version of the rule is

$$\emptyset \rightarrow \begin{bmatrix}\text{+voc}\\ \text{-cons}\\ \text{+grv}\\ \text{-comp}\end{bmatrix} \Big/ \begin{bmatrix}\text{+voc}\\ \text{-cons}\\ \text{+comp}\\ \text{-tns}\end{bmatrix} \underline{\qquad} \begin{bmatrix}\text{+cons}\\ \text{-voc}\\ \text{+cont}\end{bmatrix} \#\# \left(\begin{bmatrix}\text{+voc}\\ \text{-cons}\\ \text{+comp}\\ \text{-tns}\end{bmatrix}\right) \left[\begin{Bmatrix}\text{+cons}\\ \\ \text{-voc}\end{Bmatrix}\right]$$

Second, the early *z*-deletion rule is not a general rule, for the instances of *z* in case (b) must *not* be deleted. In fact, *z*-deletion would take place if and only if the *z* followed a *u*, so that the applicability of the *z*-deletion rule would be logically equivalent to the applicability of the *u*-insertion rule that it immediately follows. In such a situation it is pointless (and indefensible, on the grounds of simplicity considerations) to postulate *two* rules.

In the absence of evidence for postulating two rules or for an alternative analysis, the most direct method of accounting for case (a) is to add the rule "Replace final *z* by *u* when it is preceded by *ă* and followed by a non-vowel or by a vowel other than *ǎ*." The change of *z* to *u* is certainly an unnatural one, since *z* and *u* share hardly any markings (they are both continuant, noncompact, and voiced). Nevertheless, this single *z* $\longrightarrow$ *u* rule is simpler than a *u*-insertion rule followed by the deletion of *z* after *u*. As was the case with the *u*-insertion rule, the *z* $\longrightarrow$ *u* rule could be ordered either before or after VOW1'. We provide here a formalized version of the rule as it applies *after* VOW1':

$$\text{RS4.}\quad \begin{bmatrix} +\text{cons} \\ -\text{voc} \\ +\text{cont} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{cons} \\ +\text{voc} \\ -\text{obst} \\ +\text{grv} \\ +\text{fl} \\ -\text{tns} \end{bmatrix} \Big/$$

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{comp} \\ -\text{tns} \end{bmatrix} \underline{\qquad} \ \#\# \left(\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{comp} \\ -\text{tns} \end{bmatrix} \right) \left[\begin{Bmatrix} +\text{cons} \\ \\ -\text{voc} \end{Bmatrix} \right]$$

Examples:

$$\underline{\text{nalaz}}\ \#\#\ \underline{\text{ati}} \xrightarrow[\text{RS4}]{} \underline{\text{nalau}}\ \#\#\ \underline{\text{ati}} \xrightarrow[\text{VOW6}]{}$$

$$\underline{\text{nalau}}\ \#\#\ \underline{\text{ti}} \xrightarrow[\text{VOW3'}]{} \underline{\text{nalo}}\ \#\#\ \underline{\text{ti}}$$

nalaz ## naama $\xrightarrow[\text{VOW1'}]{}$ nalaz ## nāma

$\xrightarrow[\text{RS4}]{}$ nalau ## nāma $\xrightarrow[\text{VOW3'}]{}$

nalo ## nāma

Compare:

śatrauẓ ## ati $\xrightarrow[\text{VOW3'}]{}$ śatroẓ ## ati

$\xrightarrow[\text{RS1}]{}$ śatror ## ati

śatruẓ ## ati $\xrightarrow[\text{RS1}]{}$ śatrur ## ati

(nāma, nom. sg. of naaman 'name'; śatros and śatrus, gen. sg. and nom. sg., respectively, of śatru 'enemy')

6.2. Internal Sandhi

In internal sandhi the treatment of s before obstruents is as follows (W166, 414):

(a) Before (#)s, t(ḥ): s

(b) Before d^h: deleted[1]

---

1 Or possibly, in some cases, converted to d (W166, 232).

(c) Before *b^h*: *as* becomes *o*; after any other vowel, *s* becomes *r*.[2]

The rules in previous sections do not affect *s* in case (a). Case (b) is already handled by AV6 (Internal Voicing Assimilation) and RS3 (*z*-Deletion): *śād^{h}i*, 2 sg. act. impv. of *śaas* 'order' is obtained by the derivation

$$\text{śaas} + d^hi \xrightarrow[\text{VOW1'}]{} \text{śās} + d^hi \xrightarrow[\text{AV6}]{}$$

$$\text{śāz} + d^hi \xrightarrow[\text{RS3}]{} \text{śā} + d^hi$$

In case (c) we have in fact the same phenomena as in external sandhi before an initial voiced consonant. The instr. pl. forms *manob^{h}is* (*man* + *as* 'mind') *hawirb^{h}is* (*haw* + *is* 'offering'), *cakṣurb^{h}is* (*caś* + *s* + *us* 'eye')--compare the loc. pl. forms *manassu*, *hawiṣṣu*, *cakṣuṣṣu*--are analogous to *mano* ## *b^{h}awati*, *hawir* ## *b^{h}awati*, *cakṣur* ## *b^{h}awati*. Hence *b^{h}is*, like *su*, should be represented with an initial #. We then have the deriva-

---

2 In the case of noun stems ending in *as*, *is*, or *us*. Whitney lists no "normal" treatment of root-final *s*, only two sets of presumably exceptional phenomena (W167, 168).

vations (disregarding any changes in the <u>s</u> of <u>b^h^is</u>)

$$\underline{\text{man}} + \underline{\text{as}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}} \xrightarrow[\text{AV7}']{} \underline{\text{man}} + \underline{\text{az}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}}$$

$$\xrightarrow[\text{RS4}]{} \underline{\text{man}} + \underline{\text{au}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}} \xrightarrow[\text{VOW3}']{}$$

$$\underline{\text{man}} + \underline{\text{o}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}}$$

$$\underline{\text{haw}} + \underline{\text{is}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}} \xrightarrow[\text{RET1}']{} \underline{\text{haw}} + \underline{\text{iṣ}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}}$$

$$\xrightarrow[\text{AV7}']{} \underline{\text{haw}} + \underline{\text{iẓ}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}} \xrightarrow[\text{RS1}]{} \underline{\text{hawir}} \; \# \; \underline{\text{b}^{\text{h}}\text{is}}$$

Appendix:

Summary of Rules and Ordering

The rules presented in this work are listed below in the order of their appearance. With each rule is given the number of the section in which it is first discussed. Final versions of rules are starred.

VOW1. Coalescence of Like Vowels (2.2)

*VOW2. Prevocalic Glides (2.2)

VOW3. Mid Vowels (2.2)

VOW4. Postvocalic Glides (2.2)

*VOW5. Laxing of *ā* (2.2)

VOW1'. Revision of VOW1 (2.4)

*VOW6. *ă*-Deletion (2.4)

*VOW7. *y*-Deletion (2.4)

*VOW3'. Revision of VOW3 (2.4)

*VOW4'. Revision of VOW4 (2.4)

RET1. Retroflexion of *s* (3.1)

*RET2. Deretroflexion of *ṣ* (3.1)

RET3. Retroflexion of *n* (3.2)

*RET4. Deretroflexion of *ṇ* (3.2)

*RET5. Retroflexion by *l* (3.3)

*RET6. *l*-Deletion (3.3)

RET7. Internal Retroflexion Assimilation (3.4)

RET8. External Retroflexion Assimilation (3.4)

*RET1'. Revision of RET1 (3.4)

*RET3'. Revision of RET3 (3.4)

*SP1. First Spirant-Shift (4.1.2)

*SP2. Second Spirant-Shift (4.1.2)

*ABBR. Abbreviation (4.1.4)

*SP3. *s*-Deletion (4.1.4)

*PAL1 (RET7'). Internal Palatality Assimilation (4.2.1)

*PAL2 (RET8'). External Palatality Assimilation (4.2.2)

*SP4. External Spirant Assimilation (4.2.3)

*SP5. *ś* → *ch* (4.2.4)

*SP6. Spirant Insertion (4.2.5)

*VIS1. *f*/*x* → *ḥ* (4.2.6)

*VIS2. *Visarga* before Spirant (4.2.6)

*VIS3. *Visarga in pausa* (4.2.6)

GL. Grassmann's Law (5.1)

GL'. Revision of GL (5.1)

*AV1. Deaspiration (5.1)

BL. Bartholomae's Law (5.2)

GL''. Revision of GL' (5.3)

*AV2. Revision of GL'' (5.3)

*AV3. Revision of BL (5.3)

*AV4. Tense Stops (5.3)

*AV5. Aspirates from Clusters (5.3)

*AV6. Internal Voicing Assimilation (5.5.1)

AV7. External Voicing Assimilation (5.5.2)

*AV7'. Revision of AV7 (6.1.1)

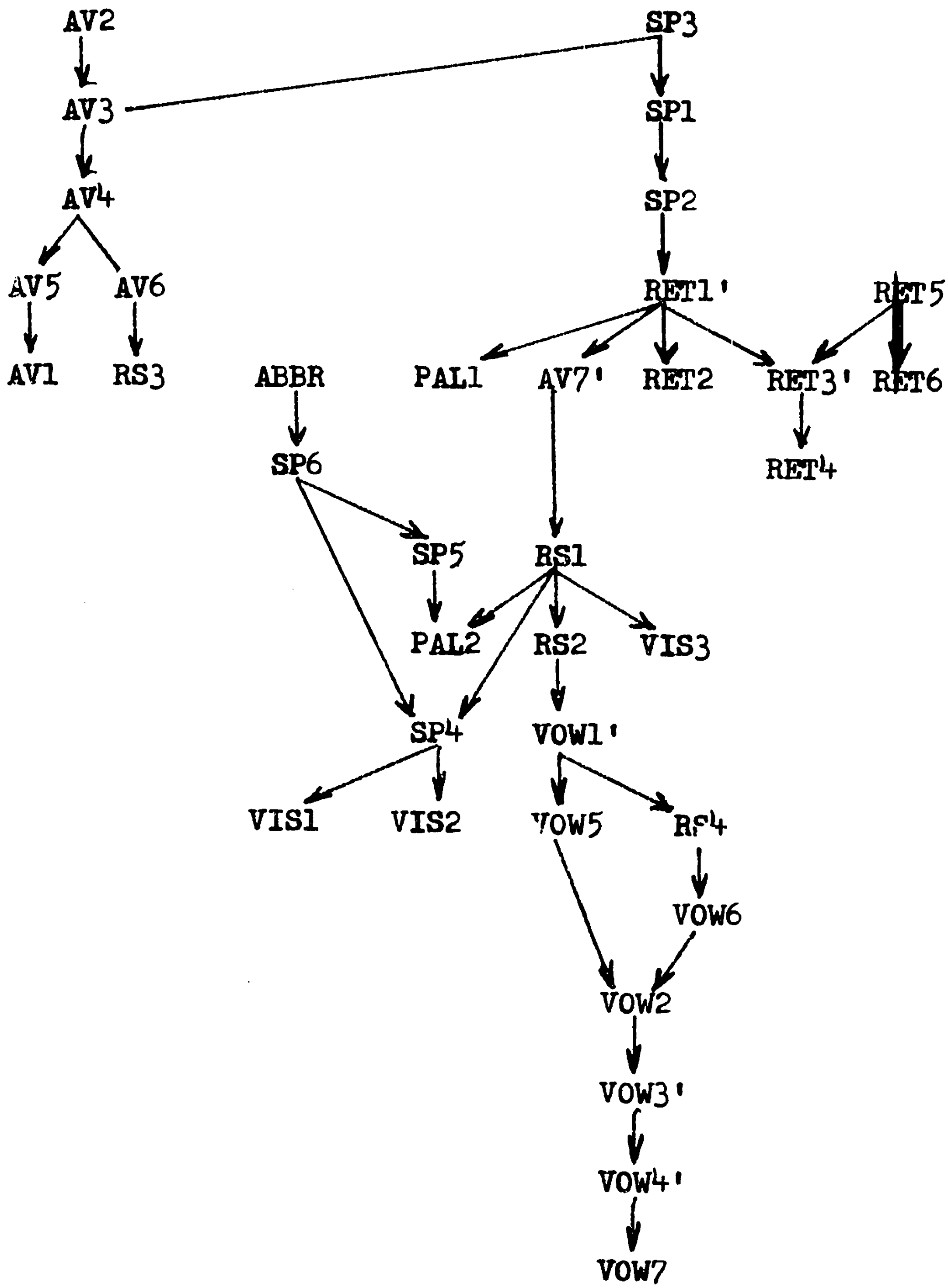
*RS1. <u>r</u> ⟷ <u>s</u> (6.1.1)

*RS2. Simplification of <u>rr</u> (6.1.1)

*RS3. <u>r</u>-Deletion (6.1.2)

*RS4. <u>z</u> → <u>u</u> (6.1.2)

The ordering restrictions on the final versions of these rules are summarized in the chart on the next page.


AV2
AV3
AV4
AV5
AV6
AV1
RS3
SP3
SP1
SP2
RET1'
RET5
ABBR
PAL1
AV7'
RET2
RET3'
RET6
SP6
RET4
SP5
RS1
PAL2
RS2
VIS3
SP4
VOW1'
VIS1
VIS2
VOW5
RS4
VOW6
VOW2
VOW3'
VOW4'
VOW7

BIBLIOGRAPHY

Allen, W. Sidney. 1961. Phonetics in Ancient India. London: Oxford Univ. Press.

_____. 1962. Sandhi. 'S-Gravenhage: Mouton.

Burrow, T. 1959. The Sanskrit Language. London: Faber and Faber.

Chomsky, Noam. 1964. "Current Issues in Linguistic Theory," in J. A. Fodor and J. J. Katz (eds.) The Structure of Language (Englewood Cliffs, N.J.: Prentice-Hall), pp. 50-118.

Emeneau, Murray B. 1958. Sanskrit Sandhi and Exercises. Berkeley and Los Angeles: Univ. of Calif. Press.

Halle, Morris. 1958. "On the Bases of Phonology," Il Nuovo Cimento 13:494-517.

_____. 1962. "Phonology in Generative Grammar," Word 18:54-72.

Heffner, R-M. S. 1964. General Phonetics. Madison, Wisc.: Univ. of Wisc. Press.

Jakobson, R., Fant, C.G.M., and Halle, Morris. 1961. Preliminaries to Speech Analysis. Cambridge, Mass.: M.I.T. Press.

Lehmann, Winfred. 1955. Proto-Indo-European Phonology. Austin, Tex.: Univ. of Texas Press and LSA.

MacDonell, Arthur A. 1927. A Sanskrit Grammar for Students. Oxford: Oxford Univ. Press.

McCawley, James. 1965. The Accentua[illegible]tem of Standard Japanese. M.I.T. [illegible] Dissertation.

Monier-Williams, Monier. 1899. [illegible]krit-English Dictionary. Oxford: Clarendon Press.

Renou, Louis. 1961. Grammaire Sanscrite. Paris: Adrien-Maisonneuve.

Uhlenbeck, C. C. 1960. A Manual of Sanskrit Phonetics. Delhi: Jayyad Press.

Varma, Siddheshwar. 1929. Critical Studies in the Phonetic Observations of Indian Grammarians. London: Royal Asiatic Society.

Whitney, W. D. 1885. The Roots, Verb-Forms, and Primary Derivatives of the Sanskrit Language. Leipzig: Breitkopf and Härtel.

_____. 1960. Sanskrit Grammar. Cambridge, Mass.: Harvard Univ. Press.

BIOGRAPHICAL NOTE

The author was born in Allentown, Pennsylvania, on September 6, 1940. He received his primary and secondary education in West Lawn, Pennsylvania. In 1958 he entered Princeton University, where he held a General Motors National Scholarship and majored in mathematics. In June, 1962, he received a B.A., magna cum laude, and was elected to Phi Beta Kappa and Sigma Xi.

In September, 1962, the author entered the Ph.D. in Linguistics program at M.I.T., under an NDEA Title IV Fellowship. He is an Honorary Woodrow Wilson Fellow and a member of the Research Laboratory of Electronics, M.I.T.

The author was employed in the Language Processing Techniques Subdepartment of the MITRE Corporation, Bedford, Mass., during the summers of 1963, 1964, and 1965. In 1964 and 1965 he assisted in teaching the Field Methods course at M.I.T. During the summer of 1965 he was a member of a National Science Foundation project on mathematical linguistics. He has accepted an assistant professorship in linguistics at the University of Illinois.

The author is a member of the Linguistic Society of America and the Linguistic Circle of New York. His publications include:

In the Quarterly Progress Report of the R.L.E., M.I.T.:

"Some Languages that are not Context-Free," QPR No. 70 (July 15, 1963), pp 290-3.

"Three Traditional Rules of Sanskrit," QPR No. 74 (July 15, 1964), pp 203-4.

MITRE Corporation documents:

'Grammars of Number Theory: Some Examples," MITRE W-6671, Nov. 20, 1963.

(with S. Isard) "Some Aspects of Tree Theory," MITRE W-6674, Dec. 3, 1963.

"Grammars Representing Sequences," MITRE W-07295, Dec. 1, 1964.

(with J. B. Fraser, M. L. Geis, B. C. Hall, S. Isard, J. W. Mintz, P. S. Peters) "English Preprocessor Manual," MITRE SR-132 and W-07287, Dec. 1, 1964. Revisions and additions (with J. Friedman and B. C. Hall), March 8, April 27, and May 18, 1965.

"A LISP Program for the General ANL Restriction," to appear.

"A LISP-to-FORTRAN Tree Conversion Program," to appear.

"On the Writing of Transformational Grammars," to appear.

"A LISP Program for the Kay Parsing Procedure," to appear.

(with B. C. Hall, J. Friedman, and D. E. Walker) "The MITRE Syntactic Analysis Procedure for Transformational Grammars," paper to be presented before the Fall Joint Computer Conference of the AFIPS, November, 1965.

To appear in <u>Studia Grammatica</u>: "Umlaut and Noun Plurals in German."