LEXICAL PHONOLOGY

by

KARUVANNUR PUTHANVEETTIL MOHANAN

M.S., Massachusetts Institute of Technology
1981

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 1982

Signature of Author _____
Department of Linguistics and Philosophy
January 4, 1982

Certified by_____
Morris Halle
Thesis Supervisor

Accepted by_____
Samuel Jay Keyser
Chairman, Departmental Committee

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

APR 27 1982

Hum.

LIBRARIES

LEXICAL PHONOLOGY

by

Karuvannur Puthanveettil Mohanan

Submitted to the Department of Linguistics and Philosophy
on 4 January 1982 in partial fulfillment of the requirements for
the Degree of Doctor of Philosophy

ABSTRACT

This thesis develops the model of Lexical Phonology, in which
a subset of phonological rules applies in the lexicon as part of the
word formation component. Phonological rules apply after every morpho-
logical operation, to the output of which morphological operations may
apply once again. The lexicon consists of ordered lexical strata which
function as the domains of application of these phonological and mor-
phological rules. This model eliminates the need for the use of distinct
boundary symbols by allowing phonological rules to have direct access
to morphological information, and imposes severe restrictions on the
class of possible grammars.

The model of Lexical Phonology yields three levels of phonolo-
gical representation, namely, the underlying, the lexical, and the
phonetic. The lexical level of representation is the output of the
lexical rule applications, which is also the input to lexical inser-
tion. It is shown that several interesting formal and psychological
properties converge on this level. Pauses are assigned to lexical
representations after lexical insertion, and therefore, lexical rule
applications are unaffected by pauses, while post lexical rule app-
lications are blocked by intervening pauses. Speakers' judgments on
the 'sameness' and 'distinctness' of speech sounds are based on lexi-
cal representations. Secret code languages, such as Pig Latin and the
Aib language, take the lexical representation as the input, and so do
speech errors which permute phonological segments. It is suggested that
the lexical level may also have interesting consequences for theories
of speech acquisition, speech recognition, and speech production.

Thesis Supervisor: Morris Halle

Title: Institute Professor

I dedicate this thesis to my four gurus: my
father, from whom I imbibed intellectual
restlessness; N.S. Prabhu, who introduced
me to the excitement of linguistics ;
M.V. Nadkarni, who led me step by step to
the doors of current linguistic research;
and Morris Halle, who made me a linguist.

CONTENTS

PROLOGUE

The nagging feeling that there was some truth in the taxonomic phonemic representation was what set me working on this thesis. My first attempt was to revise the conditions that Taxonomic Phonemics imposed on the phonemic-phonetic mapping, so that the level of representation between the underlying and the phonetic could be brought back into phonological theory. After several frustrating attempts, I gave up.

Later, when I started investigating the relation between morphology and phonology, and developed a model in which phonological rules applied in the word formation component, I discovered that this model yielded just the intermediate level of representation I was looking for. It was then that I realised why reviving the taxonomic phonemic level of representation by repairing the conditions imposed on the phonemic-phonetic mapping was impossible: the conditions should be imposed not from the bottom end of phonetics, but from the top, from the interaction between morphology and phonology.

Several people, and several incidents, helped me move towards the discovery. An initial impetus came from the course on computational linguistics given by Joan Bresnan and Ron Kaplan in 1978, where I was exposed to the idea that syntactic rules which have lexical exceptions are lexical rules. What would happen if the same principle applied in phonology as well, I asked myself: the traditional 'morpho-

phonemic' rules would become lexical rules, and 'allophonic' rules would become post lexical rules. The idea that phonological rules could apply in the lexicon took seed in my mind.

A conversation with David Pesetsky about his ideas of cyclic rule application in the lexicon, another with Mark Liberman about lexical access and taxonomic phonemics, and several other bits and pieces slowly formed a pattern around the core of my search for an intermediate level of representation. When I started writing exploratory papers on the subject, several people gave me the help and reassurance I badly needed: Paul Kiparsky, Jay Keyser, Mark Liberman, Bruce Hayes, Dave Shipman, Joan Bresnan, Alec Marantz, Phil Lesourd, Mohamed Guerssel. I remember a long discussion I had with Morris Halle about one of my early immature papers. He tore the paper to bits, and reduced me to the point of tears, then, as I was leaving, called me back and said, "Listen, there is some truth in what you say, so don't give up. Continue working." If Morris hadn't said that then, I may have written a thesis in syntax.

Several people patiently read the drafts of the thesis and helped me with suggestions and criticism: Jim Harris, Ken Hale, Joan Bresnan, Douglas Pulleyblank, Donca Steriade, Lisa Selkirk, Noam Chomsky, Haj Ross, Wayne O'neil. I have also had useful discussions with Vijayakrishnan, Moira Yip, Janet Pierrehumbert, Jane Simpson, Tim Stowell, Joe Perkell, Ken Stevens, Victor Zue, Dennis Klatt, Ed Walker, Will Leben, Ellen Broselow, Wilson Gray, Raj Singh. An initial version

of the stratum theory was presented at the Trilateral Conference

on Nonlinear Phonology at the University of Texas at Austin, 1981.

I benefited a great deal from the discussion that followed the

presentation.

Jane Simpson, Susan Rothstein, Diana Archangeli, Donca

Steriade, and Malka Rappaport proofread the thesis with patience

and care, and saved me from the mistakes that a nonnative speaker

of English makes.

Through three difficult years, the Kiparsky family gave us

moral support, and many dinners. Amrit, Jay, and Meena lightened

homesickness, and kept us cheerful. Little Peter brought sunshine

into our home.

Unitrode gave me sustenance and inspiration during my life

at MIT. I cannot forget the kindness of my boss, Sgt. Plouffé, and

the friendship of my colleague, Wilson Gray, who cheerfully granted

all my unreasonable requests.

I cannot adequately thank my committee : Morris Halle, Paul

Kiparsky, and Jay Keyser. I learnt how to do phonology from Morris,

who praised me, criticised me, scolded me, and sermonised me into

shape. This thesis is as much his work as mine. Paul taught me the

need to examine facts meticulously from all possible angles before

rushing to a conclusion. From Jay, I learnt to translate thinking

into lucid argumentation.

In addition to the academic help they provided, I remain indebted to my committee for being quick to recognise a personal crisis in my life, and executing a solution for it. Without their sensitiveness and understanding, I wouldn't have been able to complete this thesis.

I have freely drawn upon the work on Malayalam that Tara did for her M.Litt. thesis. She went through my thesis with a red pen and rewrote it, retaining the parts she found not objectionable. She disclaims all responsibility for the errors, and since the errors were not intentional on my part, so do I. Only Malushka doesn't!

Chapter I : INTRODUCTION

1.1. The Overview

This thesis seeks to answer two important questions in
phonological theory: What is the nature of the relation between
phonology and morphology, and what are the levels of phonological
representation? The answer to both these questions emerges from
the close integration of phonology and morphology in the model of
Lexical Phonology I shall develop, in which a subset of phonological
rules applies in the lexicon as part of the word formation compo-
nent,[1] yielding three levels of phonological representation, the
underlying, the lexical, and the phonetic.

Phonological rules apply in the lexicon after every morpho-
logical operation such as affixation or compounding, to the output
of which morphological operations may apply once again, in a cyclic
fashion. The lexicon consists of ordered lexical strata which func-
tion as the domains of application for these phonological and morpho-
logical rules. By allowing phonological rules to have direct access
to morphological information, the model of Lexical Phonology elimi-
nates the need for boundary symbols proposed in Chomsky & Halle (1968,
henceforth SPE), and imposes severe restrictions on the class of
possible grammars.

The output of the phonological and morphological operations in the lexicon, which is also the input to lexical insertion, is the level of lexical representation. It is shown that several interesting formal and psychological properties converge on this level. Pauses are assigned to lexical representations after lexical insertion, and therefore, lexical rule applications are unaffected by pauses, while post lexical rule applications are blocked by intervening pauses. Speakers' judgments on the 'sameness' and 'distinctness' of speech sounds are based on lexical representations. Secret code languages, such as Pig Latin and the Aib language, take the lexical representation as the input, and so do speech errors which permute phonological segments. It is suggested that the lexical level may also have interesting consequences for theories of speech acquisition, speech recognition, and speech production.

In the sections that follow, I shall sketch a brief outline of the model of Lexical Phonology, with illustrations from English, reserving the details and the burden of argumentation for subsequent chapters.

1.2. Lexical and Post Lexical Domains

It is well known that rules such as trisyllabic laxing and t → s/ — i require morphological information for their correct application, while rules such as the aspiration of voiceless stops and the flapping of t do not. Laxing, for example, applies across + in

<u>divinity</u> (divīn+iti → <u>diviniti</u>), but not across # in <u>maidenhood</u>

(mǣden # hud → *<u>mǣdenhud</u>). t → s/ — i applies only when t is

followed by + : <u>rezident + i</u> → <u>rezidensi</u> (residency), but

<u>kritik</u> → *<u>krisik</u> (critic). In contrast, the rules of aspiration

and flapping are governed by factors which are strictly phonological,

such as the foot and the syllable (Kahn (1976)).

In SPE, the two kinds of phonological processes intermingle

freely, and apply as part of the derivation from the underlying to

the phonetic representation: no theoretical status is assigned to

the distinction between the two types of processes. The model of

Lexical Phonology assumes that rules which require word internal

morphological information, such as trisyllabic laxing and velar soft-

ening, apply in the lexicon, as part of the word formation component;

aspiration and flapping apply post lexically. The principle govern-

ing the distinction between lexical and post lexical applications of

rules is:

> (1) Post lexical operations are blind to the internal
> structure of words.

1.3. The Organisation of the Lexicon

Having stated that morphologically sensitive phonological

rules apply in the lexicon, I shall now illustrate the precise mode

of application of these rules in the lexicon. I assume that phono-

logical rules apply in word formation immediately after each

morphological operation, and that the output of the phonological

processes then becomes the input to further morphological opera-

tions. This can be schematised as follows:

(2)
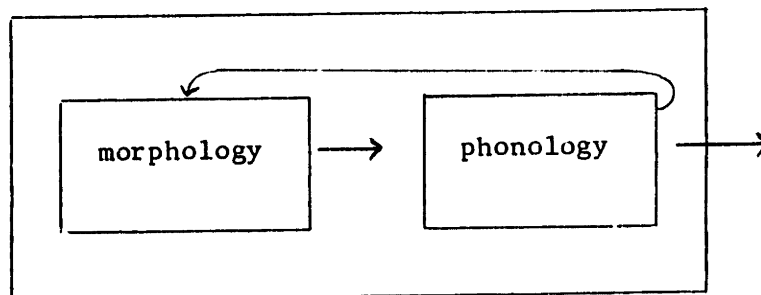
The derivation of <u>presidentiality</u> illustrates the way (2)

works. (For ease of exposition, stress is ignored.)

(3)

	/ prezīd /	underlying
	[[prezīd] ent]	affixation
	[[prezid] ent]	laxing
	[[[prezid] ent] y]	affixation
	[[[prezid] ens] y]	t → s
	[[[[prezid] ens] y] æl]	affixation
	[[[prezid] enš] y] æl]	s → š
	[[[[prezid] enš] y] æl] iti]	affixation
	(no phonological rules applicable)	

lexical rule applications

output of the lexicon →

	[prezidenši æliti]	y → i
	[phrezidenšiæliti]	aspiration
	[phrezidenšiæliDi]	flapping
	[phrezidēnšiæliDi]	nasalisation of V
	.	
	.	
	.	
	[. . . . .]	phonetic represen- tation

post lexical rule applications

Given that phonological rules apply after each morphological operation, it follows that phonological rules apply cyclically in the lexicon, each morphological operation creating a new cycle (Pesetsky (1979)). In other words, we derive (4) from (2):

(4) All phonological rule applications in the lexicon
     are cyclic.

The conjunct of (1) and (4) makes a powerful empirical prediction, stated in (5):

(5) All rules sensitive to word internal morphological
     information are cyclic rules.

(5) constitutes a significant restriction that Lexical Phonology imposes on the class of possible grammars.

1.4. The Strata

I assume that the lexicon consists of an ordered set of domains that I shall call strata. The notion of the stratum was first suggested by Siegel (1974), and developed further by Allen (1978).[2] Siegel observes that affixes like in-, -ion, -ity, and -al behave differently from affixes like un-, -ness, -less, and -ful. The former affixes, which she calls Class I affixes, can be attached to roots which are not independent words (e.g. ert in inert), but the latter affixes do not allow themselves to be attached to such roots:

(6) *ert   inert   *unert

     *trepid   intrepid   *untrepid

     *sipid   insipid   *unsipid

     *maculate   immaculate   *unmaculate

(Examples from Allen (1978))

Class I affixes cannot be attached to stems containing class
II affixes. Class II affixes, on the other hand, have no such restriction:

(7) a. inorganic   b. unorganic
       incoherent   uncoherent

    but

    c. *inclassifiable   d. unclassifiable
       *incomfortable   uncomfortable

Both un- and in- can be attached to stems containing class
I affixes like -ic and -ent. Only un-, and not in-, can be attached
to stems containing class II affixes like -able. Siegel proposes to
account for this asymmetry by ordering all class I affixations prior
to all class II affixations. Class I affixation takes place at stratum
1 (one may think of a stratum as an abstract domain or a submodule of
the lexicon), class II affixation takes place at stratum 2, and
stratum 1 is ordered prior to stratum 2. The following derivations
illustrate the effect of stratum ordering:

(8)

	inorganic	unorganic	uncomfortable	
Stratum 1	[organ]	[organ]	[comfort]	
	[[organ]ic]	[[organ]ic]	____	-ic suf- fixation
	[in[[organ]ic]]	____	____	in- pre- fixation
Stratum 2	____	____	[[comfort]able]	-able suf- fixation
	-____	[un[[organ]ic]]	[un[[comfort]able]]	un- pre- fixation

in- affixation cannot apply to comfortable , because comfort

and -able are put together at stratum 2, at which stage in- prefixation

can no longer apply. This accounts for the ungrammaticality of (7c).

We may state the Stratum Ordering Hypothesis as follows:

(9)   The lexicon consists of an ordered set of strata.

1.5.   Phonological Rules and Ordered Strata

Siegel and Allen point out that the morphologically motivated

stratum differences correlate with interesting phonological differen-

ces. Class II affixes are what SPE calls stress neutral affixes in that

they do not change the stress of the stem, while class I affixes may

change the stress of the stem. Thus, when -al is attached to díalect

at stratum 1, we derive dialéctal, but when -hood is attached at

stratum 2, the stress does not shift: díalecthòod.

Allen observes that in-, and not un-, undergoes nasal

deletion: [in[legal]] → illegal, but [un[lawful]] → *ullawful. In

order to account for such contrasts, Allen assumes that each stratum

assigns its boundary to the morphological structure created at that

stratum (Allen (1978:17). Thus, stratum 1 assigns the boundary + to

illegal, and stratum 2 the boundary # to unlawful. This assignment of

boundary symbols within phonological strings is necessary because

of Allen's assumption that phonological rules apply to the final out-

put of morphological and syntactic operations, following the SPE

position that the input to the phonological component is the surface

structure. Given this position, it becomes necessary to record diffe-

rent kinds of morphological operations in terms of different boundary

symbols for subsequent refernce in phonology (see Pesetsky (1979) for

a discussion on this issue).

As stated earlier, the model proposed in this thesis deviates

from the SPE position in allowing phonological rules to apply in the

lexicon. Given that phonological rules can apply as part of word for-

mation, it is no longer necessary to encode morphological information

in terms of boundary symbols. I propose an alternative to the boundary

theory by allowing the domains of phonological rules to be defined

in terms of morphological strata. In the boundary theory, the fact

that nasal deletion applies in illegal, but not in unlawful, is account-

ed for by representing them as in+legal and un#lawful, and stipulating

that all boundaries except + block the application of phonological

rules unless stated in the structural description of the rule (SPE;67).
Instead, I define the domain of nasal deletion as stratum 1, and
derive the same result:

(10) Nasal Deletion (Domain; Stratum 1)

[+nasal] → $\emptyset$ / —— [+sonorant]

(11) illustrates the application of the rule:

(11)

	illegal	unlawful	
Stratum 1	[legal]	[law]	
	[in[legal]]	——	in- prefixation
	[i [legal]]	——	nasal deletion
Stratum 2	——	[[law]ful]	-ful suffixation
	——	[un[[law]ful]]	un- prefixation

Nasal deletion applies in <u>illegal</u> because the relevant
environment <u>nl</u> is available at stratum 1, where the rule can apply.
<u>Un</u> and <u>lawful</u> are not put together at stratum 1, and therefore, the
sequence <u>nl</u> is not available at stratum 1. At stratum 2, where <u>un-</u>
is prefixed to <u>lawful</u>, the rule is no longer applicable as its domain
is restricted to stratum 1.

We noted earlier that trisyllabic laxing takes place across +,
but not across #. I account for this property of laxing by specifying
the domain of the rule as stratum 1:

(12)   divinity   maidenhood

	[divīn]	[m æd]	underlying
	[[divīn]iti]	____	-ity suffixation
	[[divin]iti]	____	laxing
	____	[[m æd]en]	-en suffixation
	____	____	laxing
	____	[[m æden]hud]	-hood suffixation

Stratum 1 / Stratum 2

We may state the assumption that allows us to eliminate

boundary symbols as the Stratum Domain Hypothesis:

(13)   The domain of a rule is specified as a set of

       continuous strata.

1.6.   The Opacity Principle

Consider the rule of sonorant syllabification in English,

formulated in SPE as follows:

(14)   sonorants become syllabic/ C —— # (p.85)

We may formally state (14) as (15):[3]

(15) $[\text{+son}] \rightarrow \overset{V}{\underset{|}{}} / \overset{C}{\underset{|}{}} \text{---} \#$

(15) accounts for words like hinder, cylinder, and burgle, in

which the final sonorant is syllabic, in contrast to the r and l in

hindrance, cylindrical, and burglar. SPE notes that the noun twinkling

(in the sense of 'instant') is bisyllabic, while the participle

twinkling is trisyllabic: $[\text{twiŋkliŋ}]_N$ vs $[\text{twiŋkliŋ}]_V$. In order to account for the contrast, SPE assumes the following rep esentations:

(16) a. twinkl + ing$_N$ b. twinkl # ing$_V$

(15) applies in (16b), but not in (16a), as the l in the latter is not followed by #.

Let us assume that English morphology requires four strata: stratum 1 (class I derivation); st atum 2 (class II de rivation); stratum 3 (compounding); stratum 4 (productive inflections); and that the participial -ing is attached at stratum 4, while the nominalising -ing is a class II de rivational affix attached at stratum 2.[4]

We may now state the sonorant syllabification rule as applying at st ratum 4:

(17) Sonorant Syllabification (Domain : St atum 4)

$$[\text{+son}] \rightarrow \overset{V}{|} \; / \; \overset{C}{|} \underline{\quad}]$$

Instead of referring to #, (17) simply refe rs to the bracketing, i.e., the constituent structure of the word. The derivations of twinkling $_N$ and twinkling $_V$ are given below:

(18)

	twinkling $_N$	twinkling $_V$	
Stratum 1	[twinkl]	[twinkl]	underlying
	(no rule)	(no rule)	
Stratum 2	[[twinkl]ing]	_____	-ing affixa tion
Stratum 3	(no rule)	(no rule)	
Stratum 4	_____	[[twinkl]ing] $_V$	-ing affixa tion
	_____	[[twinkl]ing]	rule (17)

It is important that (17) should not apply to twinkling $_N$. We can guarantee this result by assuming that this form has no internal brackets at stratum 4: $[[twinkl]_V ing]_V$, but $[twinkling]_N$. (17) demands] after the sonorant, and in $[twinkling]_N$, this structural condition is not met.

The general principle assumed in (18) can be stated as the Opacity Principle:

(19) The internal structure at one stratum is invisible to the processes at another.

We may look at the Opacity Principle as a device that erases all the internal brackets at the end of all the morphological opera-

tions at each stratum. Principle (1), that post lexical rule appli-
cations are blind to the internal structure of words, can now be
seen as deriving from the Opacity Principle. All word internal bra-
ckets are erased at the end of the last lexical stratum: post
lexical operations, as a result, have no access to word internal
structure, which is what (1) says.

We may note that the principle of Lexical Integrity, first
proposed in Chomsky (1970), follows directly from the Opacity Princi-
ple. The Lexical Integrity Principle says that syntactic rules can-
not move material into or out of lexical items, or modify their struc-
ture. What this means is that syntactic rules are blind to the inter-
nal structure of words. If word internal brackets are not present in
the output of the lexicon, it follows that syntactic rules cannot
refer to word internal structure.

1.7. Boundary to Stratum Translation

In SPE, the use of boundary symbols serves two functions:
(i) If a rule contains a boundary symbol such as + or # in its
structural description, it applies to a string only if the string
contains the required boundary. (17) shows how this function of
the boundary symbol is taken care of in Lexical Phonology by de-
fining the domain of the rule as the stratum associated with the
boundary, and replacing the boundary with brackets. (ii) If the
structural description of a rule does not contain a boundary, but

the phonological string does, the presence of the boundary in the string blocks the application of the rule. Nasal deletion (as in (11)), and trisyllabic laxing (as in (12)), show how this effect is taken care of in Lexical Phonology by defining the domain of the rule as a stratum that precedes the one which is associated with the boundary that blocks the rule. Thus, the use of boundary symbols is eliminated in Lexical Phonology by allowing phonological rules to have direct access to morphological information, that is, by defining the domains of phonological rules in terms of morphological strata, and allowing them to apply at the relevant stratum.

1.8. Rules with Lexical and Post Lexical Domains

The Stratum Domain Hypothesis (1.5.) allows phonological rules to have more than one stratum as their domain. Rules with multiple strata do exist in natural languages, as will be demonstrated as their domains in 2.1. There are rules, in fact, which apply at both the lexical and the post lexical domains. An example is the rule of palatalisation which changes s to š, the relevant part of which may be stated as: s → š/ —— yV̊ (where V̊ is an unstressed vowel). The rule applies lexically, in facial (f æs] y] æl → feyšəl), and post lexically (in some dialects), in miss you (mis yu → mišyə). Other examples from English are the rhythm rule, and r deletion in British English (3.3.1.2.). In Malayalam, the rule that inserts a glide between two

vowels applies at all the lexical strata, and at the post lexical stratum as well (see 4.2.).

As has been pointed out already, Lexical Phonology differs from SPE in setting up the lexicon as a separate module for the application of phonological rules. It is logically possible to distinguish between the lexical and phrasal modules in terms of the partitioning of the set of phonological rules, and say that rules R_1, R_2,...R_i belong to module A, and rules R_j, ... R_n belong to module B. It is important to note that this not the option adopted by Lexical Phonology. Rules do not belong to different modules; rather, they have one or more modules assigned to them as their domain. This conception of the stratum may be schematised as follows:

(20)

Lexicon   Phonology

Rule 1

Stratum 1   Rule 2

Stratum 2

.
.
.

Stratum n

Output of
the Lexicon

Post lexical
stratum   Rule j

Phonetic representation
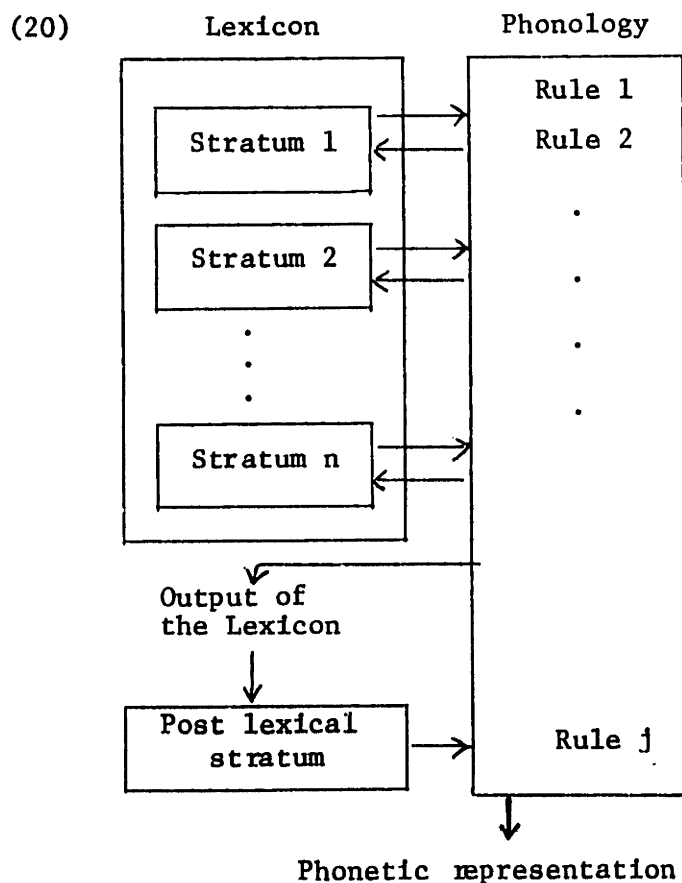
As can be seen from this diagram, the phonological compo-
nent is a module independent of the lexicon, containing rules which
may apply lexically, post lexically, or both.

1.9. Morphology, Syntax, and Ordered Strata

I assume that the morphological component contains rules
such as the following:

> (21)a. N → Adj -ity   (e.g. civility)
>
>   b. N → Adj -ness   (e.g. goodness)
>
>   c. N → N N   (e.g. powerhouse)
>
>   d. N → N plur.   (e.g. boys)

How are morphological rules of this kind associated with
the stratum structure? I suggest that, like phonological rules, mor-
phological rules can have their domains defined in terms of lexical
strata. Thus, (21a) will have stratum 1 as its domain; (21b), stratum
2; (21c), stratum 3; and (21d), stratum 4.

The syntactic component generates syntactic structures, the
lexicon generates words, and the operation of lexical insertion
puts the two together. If so, one may expand (20) to incorporate
syntax and morphology into the model as follows:

(22) Syntax   Lexicon   Phonology

```
         Syntax              Lexicon                    Phonology
       ┌────────┐   ┌──────────────────────┐   ┌──────────────┐
       │        │   │        Morphology     │◄─ │   rule 1     │
       │        │   │  ↓    ┌───────────┐   │   │   rule 2     │
       │        │   │stratum 1→│ rule 1 │   │   │      .       │
       │        │   │  ↓    │   rule 2  │---┼──►│      .       │
       │        │   │stratum 2→│   .    │   │   │      .       │
       │        │   │  .    │   .       │   │   │              │
       │        │   │  .    │   .       │   │   │              │
       │        │   │stratum m→│  rule n│   │   │              │
       └────────┘   └──────────────────────┘   │              │
            │              words                │              │
     syntactic                                  │              │
     structures                                 │              │
            │      lexical insertion            │              │
            ▼              ▼                     │              │
       ┌──────────────────────────┐             │              │
       │   post lexical stratum    │────────────►│   rule j    │
       └──────────────────────────┘             └──────────────┘
                    │
          phonetic representation
```
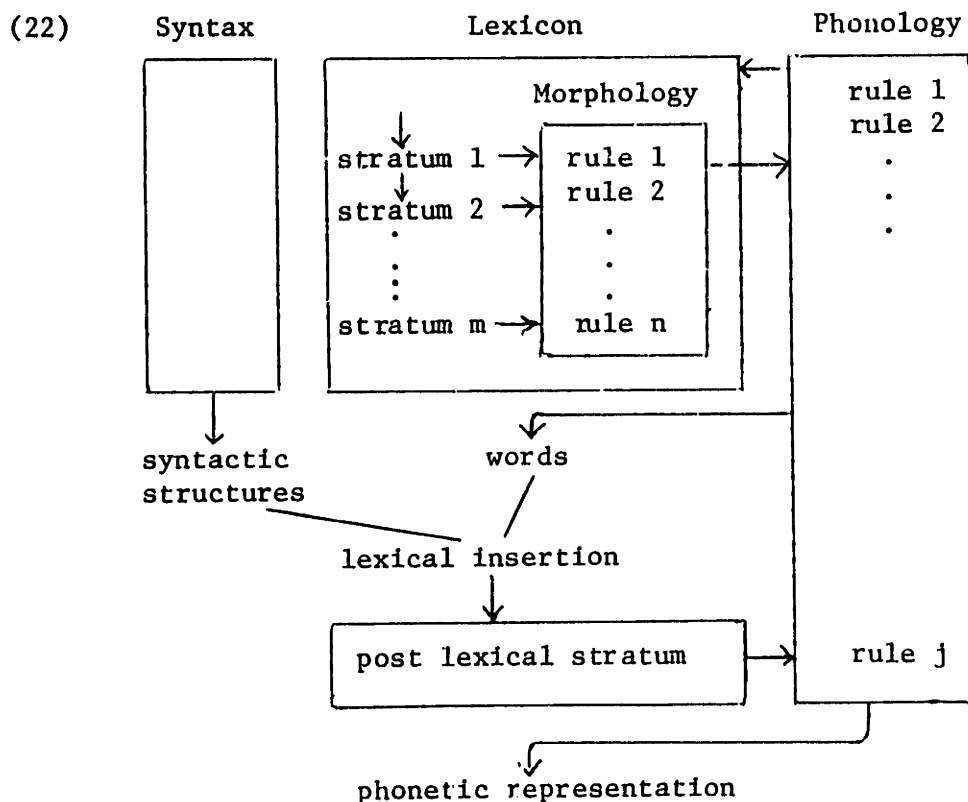
1.10. The Lexical Level of Representation

An essential feature of (22) is that phonological rules can apply in the lexicon, before lexical insertion, as well as to the output of the lexicon, after lexical insertion. I assume that what is entered in the lexicon as a lexical entry is a word, and refer to the phonological representation of words as the Lexical Representation. (22) derives a model of phonology that has three levels of representation, namely, the underlying, the lexical, and the phonetic:

(23)   Lexicon

underlying
representation
                                                    phonological
                              .←──────────────   rules

lexical
lexical   representation
insertion ──────→.

                              .←──────────────   phonological
                              .   rules

                 phonetic representation
                 rep esentation
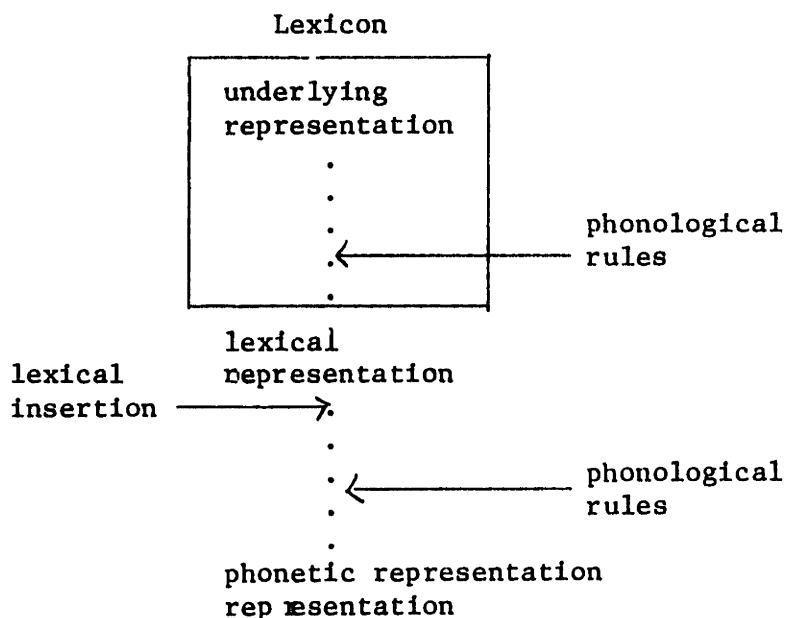
The derivation of <u>presidentiality</u> in (3) illustrates the three

levels of representation: /prezīd/ (underlying), prèzidenši ǽlity

(lexical), and [phrèzidẽnšiǽliDi] (phonetic).

There are several interesting properties that converge on

the level of lexical representation. The judgments of the native spea-

kers of a language on the sameness and distinctness of sounds appear

to be based on this level, not on the underlying or the phonetic.

Thus, speakers of English judge the p in <u>pin</u> and <u>spin</u> to be the same

sound, in spite of the phonetic difference in aspiration ([phin] vs

[spin])[5], and the first vowel in <u>metric</u> and <u>met</u> to be the same, in

spite of the underlying distinction (/ē/ vs /e/). They also judge

the second vowels in <u>divine</u> and <u>divinity</u> to be different, in spite of

the underlying identity: /ī/. The lexical representations of these

words reflect the judgments: <u>pin</u>, <u>spin</u>, <u>metrik</u>, <u>met</u>, <u>divayn</u>, <u>diviniti</u>.

Another phenomenon that the model in (23) accounts for is the behaviour of phonological rules with respect to pauses. Compare, for example, the rule of trisyllabic laxing with the rule of flapping. When asked to pronounce words slowly, syllable by syllable, for example,
English speakers say, /[di...vi...ni...ti], not *[di..vay..ni..ti] or *[di..vi..ni..Di]. This shows that the rule of flapping that changes t to D is blocked by an intervening pause, but not the rule of laxing that changes ī to i.

We can account for this difference in behaviour between pause sensitive (e.g. flapping) and pause blind (e.g. laxing) rules by assuming that pauses are assigned to phonological strings after lexical insertion. It follows that rules applying before lexical insertion will be insensitive to pauses, since pauses do not exist at that stage, and that rules applying after lexical insertion will be blocked by the presence of pauses.

There are many other phenomena which converge on the level of lexical representation. Secret code languages, such as Pig Latin, the Aib language of English, and the pa language of Malayalam, take the lexical representation as the input, perform the operation of the code on it, and apply the post lexical rules to the output of the code. Speech errors which permute segments take place at the level of lexical representation. Phenomena of this kind illustrate that the output of the lexical component of phonology is not like other stages

in a de rivation, but constitutes a level of linguistic representation
which has important psychological consequences.

The lexical level of representation is the level at which
lexical items (= words) are represented in the mental lexicon. Given
this assumption, the behaviour of code languages, speech errors, etc.
can be viewed as operations on the mental representations which are
the input to on line processing. It is also speculated that what
learners do when they come ac ross a new wo rd is to enter it in the
lexicon in terms of the lexical level, as this level is available to
the learner without ~the~ knowledge of the morphological structure of
the word. The acquisition of the underlying representation of the mor-
phemes involved takes place at a more relaxed pace, as and when mor-
phological information becomes available to the learner from related
wo rds.

Given this approach to the acquisition of phonology, under-
lying representations, and the lexical rules which derive lexical
representations from them, can be looked upon as serving to reduce
the burden of storing lexical ent ries. What lexical rules do, in other
words, is to encode the regula rities of the lexicon.

Lexical Phonology has consequences for speech production
and recognition as well. Given that words are stored in the lexicon
in terms of lexical representations, it is possible to conceive of
a model of speech production which takes lexical representations as

the input, and computes the phonetic representations from them, thereby eliminating the complex computations involving lexical processes. Speech recognition can be seen as involving (among other things) a bottom up computation of lexical representations from phonetic representations, thereby eliminating the need to determine the morphological structure of a word in order to identify its underlying representation.

In sum, by assuming a closer integration of phonology and morphology than has been assumed in currently available models of grammar, we arrive at a model of phonology that is intuitively more satisfying, and explanatorily more powerful, and which, at the same time, has the right consequences for a wider range of language related phenomena.

Footnotes for Chapter I

1. See Bresnan (1971), and Pesetsky (1979).

2. Siegel and Allen use the term 'level' to refer to what I call
   stratum. I replace 'level' with 'stratum' in order to avoid
   confusion with the more common use of the term 'level' to refer
   to a level of linguistic representation. (Chomsky (1957, 65).

3. See Halle & Vergnaud (1980) for the notation. See also 4.1.

4. Whether nominalising -_ing_ is a class I or class II suffix is
   irrelevant for my analysis. Note, however, that -_ing_ is a stress
   neutral suffix, indicating that it belongs to class II rather
   than class I.

5. This is not to say that human beings do not perceive sublexical
   phonetic distinctions. To the extent that all human beings have
   some phonetic skills to a greater or lesser degree, they do per-
   ceive gross phonetic distinctions. It is not, however, difficult
   to separate language based judgments from language independent
   phonetic judgments, as the latter tend to exhibit as much indi-
   vidual variation, say, as the aptitude for music, while the for-
   mer are shared by all speakers of the language community.

Chapter II : THE STRATUM THEORY

This chapter develops an important component of Lexical

Phonology, namely, the theory of ordered strata. The basic ques-

tion in our enquiry will be : how best is the relation between

morphology and phonology characterised? All theories agree that

phonological rules require information about the different junc-

tures between morphological forms. SPE makes morphological

information available to phonology by encoding it in distinct

boundary symbols. The stratum theory developed in this chapter

investigates the alternative of more direct access to morpho-

logical information. I shall show that the stratum theory is

both explanatorily and descriptively preferable.

The essential elements of the Stratum Theory are the Stratum

Ordering Hypothesis, the Stratum Domain Hypothesis, the Opacity

Principle, the cyclic application of rules, and the device of the

Loop. The Stratum Ordering Hypothesis says that the lexicon con-

sists of an ordered set of strata. The Stratum Domain Hypothesis

states that the domains of rules are specified in terms of conti-

nuous strata. The Opacity Principle makes the internal structure

at one stratum invisible at another.

The Loop is a device that allows two adjacent strata to be

inputs to each other. In English, for example, class II affixes

can be attached to compound stems ([ex[frogman]]); compou: , in turn, can be formed out of stems that contain class II affixes ([[township][committee]]). Given that the stratum of class II affixation p recedes that of compounding, we can account for forms like exfrogman with a loop from the compounding stratum to the class II stratum.

The basic structu re of the stratum theory is developed and illustrated using facts of Malayalam phonology. Malayalam is of particular inte rest to the theory because of the presence of two kinds of productive compounds in the language, subcompounds and cocompounds. These compounds raise extremely interesting issues about the relation between phonology and morphology, and therefore , I shall examine the phonology of compounding in Malayalam in detail.

The analysis of Malayalam compounding is followed by the application of the model to English and Dakota. In each case, the stratum theory is found to be explanatorily more adequate than the bounda ry theory.

2.1. Subcompounds and Cocompounds

Malayalam has two kinds of productive compounds, which, adopting the terminology of Sweet (1871), I shall call sub(ordinate) compounds and co(ordinate) compounds.

(1) Subcompounds

    a.   taaraakaantanmaarə   'Tara's husbands'

        (taara 'a name'; kaantan 'husband'; maar 'plural')

    b.   taamarakkannan   'lotus-eyed person'

        (taamara 'lotus'; kan 'eye'; -an 'masc. singular'

(2) Cocompounds

    a.   acchanammamaarə   'parents'

        (acchan 'father'; amma 'mother'; maar 'pl.')

    b.   yakṣakinnaragandharwaadikaḷə 'Yaksa's, Kinnara's,
                                         Gandharwa's, etc.'

        (yakṣan; kinnaran; gandhawan; aadi 'etc.'; kaḷ 'pl.')

Subcompounds have the structure 'modifier + head'; cocompounds have the coordinate structure of 'stem + stem + stem...', corresponding to the 'dwanda' compounds in Sanskrit. There are various reasons why the two kinds of compounds should be distinguished from each other, and also from inflectional and derivational affixation, and phrasal concatenation. I shall briefly discuss below some of these reasons.

2.1.1. Preliminary Exploration

2.1.1.1. Gemination

One reason why subcompounds should be distinguished from cocompounds is that the former, and not the latter, exhibit the process of stem final and stem initial gemination of obstruents in Dravidian stems.[1]

(3) Subcompounds:

    a.   [[kaat]$_2$[maram]] → kaattəmaram ^{2}

        forest   tree   forest tree

    b.   [[kayar]$_2$[kattil]] → kayattəkattilə   (rr → tt)

        rope   cot   cot made with rope

(4) Cocompounds:

    a.   [[[aat]$_3$[maat]]kal] → aatəmaatəkalə

        goat   cow pl.   cattle

    b.   [[[kayar]$_3$[kampi]$_3$[caŋŋala]]$_2$[wargam]] →

        rope   wire   chain   category

                kayarəkampicaŋŋalawargam

                the category of ropes, wires, and chains

The stem final t and r geminate in (3), but not in (4).

Like stem final gemination, stem initial gemination also takes place in subcompounds, but not in cocompounds:

(5) Subcompounds

    a.   [[[petti]$_2$[pattaayam]]kal] → pettippattaayaŋŋalə

        box   grain bin   pl.   grainbins used as boxes

    b.   [[aana]$_2$[kutira]] → aanakkutira

        elephant horse   horse that is like an elephant

(6) Cocompounds

    a.   [[[petti]$_3$[pattaayam]]kal] → pettipattaayaŋŋalə

                          boxes and grainbins

    b.   [[[aana]$_3$[kutira]]kal] → aanakutirakalə

                       elephants and horses

Observe that the stems in (5) and (6) are the same. Therefore, the difference in gemination between subcompounds and cocompounds cannot be attributed to any morphological peculiarity of the stem.

Given that stem final and stem initial gemination takes place in subcompounds but not in cocompounds, how do we encode this in the grammar? One possible way would be to introduce new boundary symbols to distinguish between two kinds of compounds. We may, for example, mark subcompounds with the symbol +#+ , and cocompounds with ## (derivation: + ; inflection: #). The rule of gemination can then be stated as (7):

(7)   onset $\rightarrow \bigwedge \%$ — $+\!\!+\!\!+$ [3]

Such multiplication of boundaries, however, is ad hoc and unrevealing.

Several linguists, Rotenberg (1978), Harris (to appear), Selkirk (1980) among them, have proposed that the use of morphological tree structures allows for the elimination of boundaries. I shall propose that a more efficient way of eliminating boundary symbols is to develop the idea of "level ordering" in Siegel (1974) and Allen (1978) ('stratum ordering' in my terminology). I depart crucially from Siegel and Allen in assuming that phonological rules can apply in the lexicon at various morphological strata.

In order to account for the facts of Malayalam gemination,

Let us assume that the language has the following lexical strata:

(8) Stratum 1 : Derivation

     Stratum 2 : Subcompounding

     Stratum 3 : Cocompounding

     Stratum 4 : Inflection

Let us also assume that gemination applies at the stratum of subcompounding, but not at any other strata. The basic structure of the theory implied by such an approach can be schematised as follows:

(9)   Lexicon

```
                        ⋮
        ┌───────────────────────────────┐
        │         subcompounding         │────→
        └───────────────┬───────────────┘
                        │  ↓
        ┌───────────────────────────────┐
        │         cocompounding          │────→
        └───────────────┬───────────────┘
                        ↓
                        ⋮
```

phonological rules :

1. gemination
   (domain:
   subcompounding)
2. rule ...
   ⋮
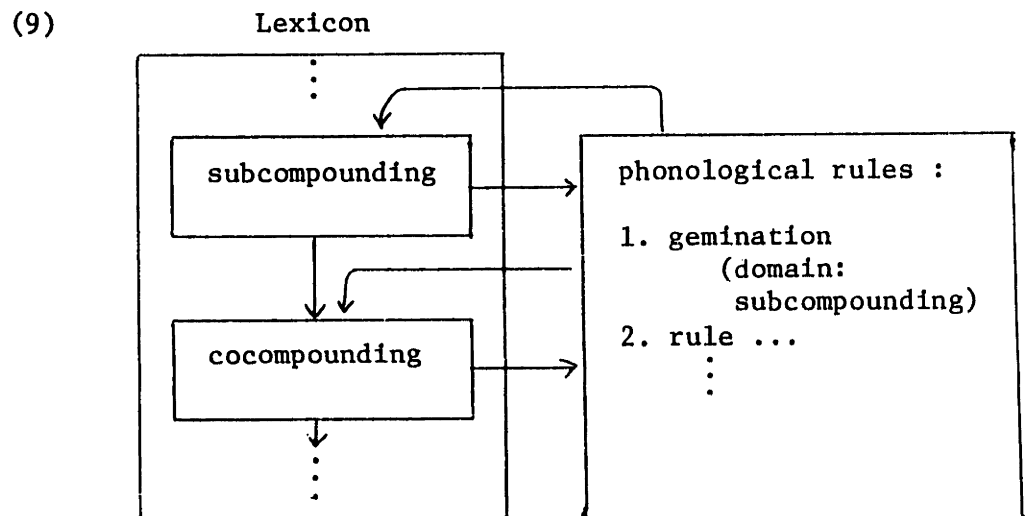
After every morphological operation such as affixation and compounding, the set of phonological rules is scanned for applicability. The domains of application of phonological rules are specified in terms of the strata at which the morphological operations take place, and a phonological rule applies at a stratum just in case the domain of that rule includes that stratum.

Given these assumptions, the rule of gemination can be stated as follows:

(10) Gemination (Domain : Sc) (Sc = subcompound stratum)

Onset → $\bigwedge$ % ―][

The derivation of pa̅sukku̲t̲t̲ikki̲rukkanmaa̅r̲ə 'people crazy about calves' (pa̅su 'cow'; ku̲t̲t̲i 'child'; kirukkan 'mad man'; -maa̅r̲ 'pl.') will be as follows:

(11)

subcompound	[[pa̅su][ku̲t̲t̲i]]	compounding
	[[pa̅su][kku̲t̲t̲i]]	gemination
	[[[pa̅su][kku̲t̲t̲i]][kirukkan]]	compounding
	[[[pa̅su][kku̲t̲t̲i]][kkirukkan]]	gemination
cocompound	————	compounding
	————	phon. rules
inflec-tion	[[pa̅sukku̲t̲t̲ikkirukkan] maa̅r]	affixation
	————	phon. rules

The contrast between kaa̲t̲t̲əmaa̅ram (3a) and aa̲t̲əmaa̲t̲ə (4a) can thus be explained as follows: in kaa̲t̲t̲əmaram, the compounding takes place at the subcompound stratum, where the gemination environment t][is accessible to rule (10). In aa̲t̲əmaa̲t̲ə , on the other hand, the two stems are put together at the cocompound stratum, and t][appears only after the domain of application of gemination is crossed.

2.1.1.2. Nasal Deletion

Malayalam has rules that apply in both subcompounds and cocompounds, but not in a sequence of two words or in a sequence of a stem and an affix. An example of such a process is stem final nasal deletion.

(12) Subcompounds

    a. [[maⁱram]₂ [kutíra]] → marakkutíra
         tree   horse   wooden horse

    b. [[manusyan]₂ [mr̥gam]] → manusyamrə̄gam
         man   animal   beastly man

(13) Cocompounds

    a. [[sukham]₃ [dukkham]] → sukhadukkham
         pleasure   sorrow   pleasure and pain

    b. [[yaksan]₃ [kinnaⁱran]₃ [gandharwan] maaⁱr] →

               yaksakinnaⁱragandharwanmaaⁱrə
               Yaksha's, Kinnara's, and Gandharwa's

(14) Inflections

    a. [[maⁱram] kal] → maⁱrannal   (mk → nn)
         tree pl.

    b. [[manusyan] maaⁱr] → manusyanmaaⁱrə
         man   pl.

    c. [[manusyan] oot] → manusyanootə
         man   dat.

(15) Across words

    a. manusyan mariccu → manusyan mariccu

        man   died   (*manusya mariccu)

    b. yaksan kinnarane nulli → yaksan kinnarane nulli

        Yaksha Kinnara   pinched (*yaksa kinnarane nulli)
          nom,   acc.

        (The Yaksha pinched the Kinnara.)

    c. katam kont mutinnu → katam kontə mutinnu

        debt with destroyed   (*kata kontə mutinnu)

        (We are destroyed by debt.)

There is a process of nasal deletion in derivational morphology that appears to be misleadingly similar to nasal deletion in compounds. This process, which I shall refer to as n-deletion, is illustrated below:

    (16) a. aaroogyam 'health' ; anaaroogyam 'ill health'
         b. aawaranam 'cover' ; anaawaranam 'uncover'
         c. aikyam 'unity' ; anaikyam 'disunity'

    (17) a. sukham 'happiness' ; asukham 'unhappiness'
         b. kramam 'order'   ; akramam 'disorder'
         c. satyam 'truth'   ; asatyam 'falsehood'

The nasal in the negative prefix an- deletes when followed by a consonant. I shall state the rule as follows:

    (18) n-deletion   (Domain : De) (De = Derivational stratum)

           n → ∅ / —— [C

n-deletion differs from the nasal deletion in compounds in

at least two respects. First, n-deletion applies only when followed

by a consonant (cf. (16) and (17)); nasal deletion applies before

vowels as well:

    (19) a. [[wr̥ksam]$_2$[agr̄am]] → wr̩ʔksaagr̄am
          tree   top   tree top

        b. [[mukham]$_2$[aawar̄anam]] → mukhaawar̄anam
           face   cover   cover for the face

        c. [[janam]$_2$[aikyam]] → janaikyam
           peʋple unity   unity of the people

    (20) a. [[[suur̄yan]$_3$[cand̲r̄an]]maar̄] → suur̄yacand̲r̄anmaarɘ
           sun   moon   pl.   the sun and the moon

        b. [[[sukham]$_3$[asukham]]kal̲] → sukhaasukhaŋŋalɘ
           happiness sorrow pl.   happiness and sorrow

        c.[[[jananam]$_3$[mar̄aṇam]]kal̲] → jananamar̄aṇaŋŋalɘ
           birth   death   birth and death

Second, n-deletion applies only to n, while nasal deletion

applies to m as well. Consider the behaviour of <u>sam-</u> 'together':

    (21) a. yoogam 'joining'   ; samyoogam 'joining together'
        b. t̲aapam 'heat'   ; sant̲aapam 'sympathy' (m → n̲)
        c. t̲rɘpt̲i 'satisfaction' ; sant̲rɘpt̲i 'contentment'

Since the a~an alternation in (16) and (17) is not the result

of nasal deletion in compounds, and nasal deletion does not apply in

(21), I shall conclude that nasal deletion does not occur in deriva-

tional morphology. The rule of nasal deletion could be stated as (22)

---

in a theory that makes use of phonological boundaries to encode morphological information:

(22) Nasal → ∅ / — { ## / ## }

Within the stratum theory, on the other hand, the rule is:

(23) Nasal Deletion (Domain : Sc, Cc) (Cc = Cocompound)

　　　　 Nasal → ∅ / —][

2.1.1.3. Vowel Lengthening

Another process that applies in subcompounds and cocompounds, but not in derivations, inflections, and word sequences is stem final vowel lengthening:

(24) Subcompounds

　　a. [[taara]$_2$[kaantan]] → taaraakaantan[4]
　　　　a name　husband　　Tara's husband

　　b. [[rati]$_2$[deewi]] → ratiideewi
　　　　a name　goddess　　the goddess Rati

　　c. [[wadhu]$_2$[gṛham]] → wadhuugṛham
　　　　bride　house　　　bride's home

(25) Cocompounds

　　a. [[[taara]$_3$[kaantan]] maar] → taaraakaantanmaarə
　　　　a name　a name　pl.　　Tara and Kantan

　　b. [[[baalika]$_3$[baalan]]maar] → baalikaabaalanmaarə
　　　　girl　　boy　pl.　　girls and boys

　　c. [[[bhaarya]$_3$[bhartaakkan]] maar] → bhaaryaabhartaakkanmaarə
　　　　wife　　husband　pl.　　couples

(26) Derivations

    a. [ati [kramam]] → atikramam
        extreme order   beyond order

    b. [anu [kāranam]] → anukāranam
        together doing   imitation

    c. [prati [dhwani]] → pratidhwani
        against sound   echo

(27) Inflections

    a. [[taāra] e] → taāraye
        a name acc.

    b. [[taāra] kk'] → taarak'k'ə
               dat.

    c. [[bhaārya] maār] → bhaāryamaārə
        wife   pl.

    d. [[parawa] kal] → parawakalə
        dove   pl.

(28) Across words

    a. taāra parannu → taāra parannu
        a name said   Tara said

    b. bhaārya kaāranam → bhaārya kaāranam
        wife   reason   because of wife

We formulate the rule of vowel lengthening as follows:[5]

(29) Vowel Lengthening (Domain : Sc, Cc)

    Rime → $\wedge$ / —][

2.1.1.4. Vowel Sandhi

We saw that the domain of gemination is the subcompounding

stratum. The rules of nasal deletion and vowel lengthening apply at

the subcompounding as well as the cocompounding strata. I shall now

show that the process of vowel sandhi which merges two adjacent

vowels into a single one applies at the strata of derivation, sub-

compounding, and cocompounding.

In compounding, the vowels fused together undergo further

changes of vowel height, the details of which we shall come to in

4.4.2. Vowel sandhi in compounds is illustrated by the following

examples:

(30) Subcompounds

a. [[mahaa]₂[indr̄an]] → maheendr̄an
    great   Indra   great Indra

b. [[niila]₂[ambar̄am]] → niilaambar̄am
    blue   sky   blue sky

(31) Cocompounds

a. [[[weedam]₃[itihaasam]]kaḷ] → weedeetihaasaŋŋaḷə
    scripture   myth   pl.   scriptures and myths

b. [[[sukham]₃[asukham]]kaḷ] → sukhaasukhaŋŋaḷə
    health   ill health   health and illness

When the infix -a- is inserted in the first syllable of a

noun to form an abstract noun, it merges with the following rime:

(32) a. kumaaṟan 'boy' ; kaumaaṟyam 'boyhood'
                                    ↑
      b. alaṣan 'lazy one'; aalasyam 'laziness'
                              ↑
      c. dhiiṟan 'brave one'; dhaiṟyam 'courage'
         ‗   ‗   ↑

Thus, the domain of vowel sandhi includes the stratum of
derivation.

Vowel sandhi does not apply at the level of inflections, as
shown by (27a).[6] The following examples show that vowel sandhi does
not apply across words:

(33) a. bhaaṟya alari → bhaaṟya y alari
          wife   screamed

      b. paṣu atbhuṭappettu → paṣu w atbhuṭappettu
                        ‥   ‥
          cow   wondered

                 have
In sum, I ʌ shown that morphological information can be made
accessible to phonological rules by allowing the phonological rules
to apply in the lexicon, and specifying their domains in terms of
mo ⱳphological strata. 2.1.1.2., 2.1.1.3., and 2.1.1.4. show that a
phonological rule must be allowed to have multiple strata as its
domain: a theory that restricts domain assignments to single strata
does not account for the facts of Malayalam.

The principle that governs the assignment of domains to
rules is repeated below:

(34) The Stratum Domain Hypothesis

      The domain of a rule is specified as a set of
      continuous st ⱳata.

Notice the stipulation that rule domains should be con-
tinuous. That is to say, if a rule has stratum 2 and stratum 4 as
its domain, then, by (34), we are forced to include stratum 3 also
as its domain. The continuity of strata is by no means a necessary
requirement in the stratum theory. It is possible, for example, to
abandon the continuity requirement and hold a weaker version of the
Stratum Domain Hypothesis, as in (35):

(35) The domain of a rule is specified as a set of strata.

In the absence of empirical evidence against (34), however,
I shall adopt this stronger and more restrictive version of the
hypothesis.

2.1.2. The Mechanics

2.1.2.1. The Stratum Structure

The discussion in 2.1.1.shows that we must assume Malayalam
to have at least four strata of word formation, derivational, sub-
compounding, cocompounding, and inflectional. These strata are ordered
as follows:

(36)   Stratum 1 : derivations
                ↓
          Stratum 2 : subcompounding
                ↓
          Stratum 3 : cocompounding
                ↓
          Stratum 4 : inflections
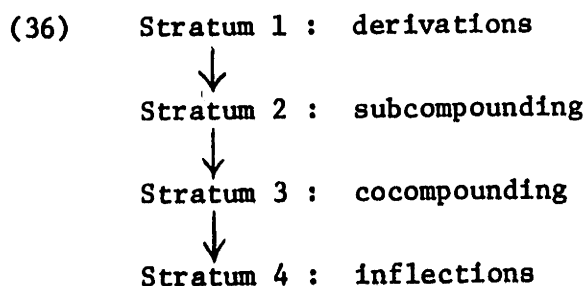
This ordering makes the correct prediction that derivational affixes cannot be attached to compounds (or inflected stems), and that compounds cannot contain inflected stems:

(37) a.   gamanaagamanaŋŋaḷɔ 'goings and comings'

b.   gamanaagamananiyantranam 'control of goings and comings'

c.   *gamanaagamanaŋŋalniyantranam

(gamanam 'going'; aagamanam 'coming';-kal 'pl.'
niyantranam 'control

As (37c) shows, the plural inflection -kaḷ cannot occur inside a compound. This holds for case inflections as well. It is difficult to offer formal evidence for or against the relative ordering of derivations and compounding, as both subcompounds and cocompounds are noun+noun compounds; the semantic relationships, however, show that (36) is not incorrect:

(38) a.   satyam 'truth'   ;   asatyam 'falsehood'
b.   ε ;tyapreemam 'love of truth'; asatyapreemam 'love
of falsehood'

asatyapreemam cannot mean 'hatred of truth', which is what we would expect if the structure were [a [[satyam] [preemam]]]. Arguments based on meaning relationships can sometimes be misleading, but considering the fact that there are no cases in which the meaning of the compound is computed before the meaning of the affix is combined with the stem, I shall assume that (38) is indeed correct, until evidence to the contrary is discovered.

2.1.2.2. The Loop

An interesting problem arises from the fact that a subcompound may contain a cocompound, and a cocompound may contain a subcompound, in both the right and the left members of the compound.

(39) a.


```
                    N₁
               N₂        N₃
            N₄    N₅
        yaksan kinnaran kuuttam  →  yaksakinnarakkuuttam
        Yaksha Kinnara  group         group of Yakshas & Kinnaras
```

b.


```
                                        N₁
                              N₂              plur.
                        N₃          N₄
                  N₅          N₆
              N₇      N₈    N₉      N₁₀
            maatr  sneeham patni widweesam wikaaram kal

            mother  love    wife  hatred    emotion  plur.
```

→ maatrə sneehapatniwidweesawikaaraŋŋələ
   the emotions of mother love and wife hatred

that of cocompounding. The empirical consequences of this ordering
are discussed in 2.1.3.

2.1.2.3. The Opacity Principle

Note that the k in kuuṭṭam geminates, while that in kinnāran

does not, in (39a). This is, in fact, what is expected, because

kuuṭṭam is the constituent of a subcompound, while kinnāran is

the constituent of a cocompound. The derivation of the compound is

as follows:

(41)	deri-vation			
	subcom-pound	————		
	cocom-pound	[[yaksan] [kinnaṛan]]		compounding
		[[yaksa] [kinnāran]]		nasal deletion
				vowel sandhi
		————		
	subcom-pound	[[yaksakinnaran] [kuuttam]]		compounding
		[[yaksakinnara] [kuuttam]]		nasal deletion
		[[yaksakinnaṛa][kkuuttam]]		gemination

Note that the internal bracketing of yakṣakinnāran should

not be seen at the subcompound stratum, if gemination is not to apply to it. This follows from the Opacity Principle, repeated below;

(42) The Opacity Principle

The structure at one stratum is invisible at another stratum.

Observe that (42) imposes a powerful restriction on possible grammars, and is desirable on a priori grounds. The principle also guarantees that the bracketing [[yaksa][kinnaran]] is invisible at the subcompounding stratum.

Pesetsky (1979) proposes a convention that erases internal brackets at the end of a derivation. This convention has similar, but not identical results as the Opacity Principle; it is , therefore, necessary to examine how (42) differs empirically from this convention.

(43) Bracket Erasing Convention

Given the nested constituents

$$_n[\ldots_{n-1}[\ldots]_{n-1}\ldots]_n$$

the last rule of cycle n is: erase n-1.

The following derivation illustrates the application of the Bracket Erasing Convention;

(44)   <u>untheatricality</u>

Stratum 1	[theatre]	
	[[theat re]ic]	-<u>ic</u> affixation
	[theatric]	BEC
	[[theatric] al]	-<u>al</u> affixation
	[theatrical]	BEC
	[[theatrical]ity]	-<u>ity</u> affixation
	[theatricality]	BEC
Stratum 2	[un [theatricality]]	<u>un</u>- affixation
	[untheatricality]	BEC

The effect of BEC is to erase the internal brackets at the end of every morphological operation. The effect of (42), on the other hand, is to erase the brackets at the end of a stratum:

(45)   <u>untheatricality</u>

Stratum 1	[theat re]	
	[[theat re]ic]	-<u>ic</u> affixation
	[[[theatre]ic]al]	-<u>al</u> affixation
	[[[[theat re]ic]al]ity]	-<u>ity</u> affixation
Stratum2	[theatricality]	by (42)
	[un[theatricality]	<u>un</u>- affixation

In 4.5.3., I shall show that the Bracket Erasing Convention is too strong, as there are cases of affixation which are sensitive to the internal structure of the stems.

2.1.3. Stress and Tone

2.1.3.1. P reliminary Facts

The basic facts of word stress in Malayalam are:

(46) a. If the fi rst syllable of a word has a short vowel
        and the second syllable a long one, the primary
        stress falls on the second syllable; in all other
        cases, primary stress is on the first syllable.

     b. Secondary stress (the evidence for which is the
        appearance of a high tone. Discussed below.) falls
        on all the remaining long vowels.

Given below are a few examples:

(47) a. waráam   'will come'
     b. wáram   'boon'
     c. waáram   'week'
     d. waáraàm   'will gather'

(48) a. paráati   'complaint'
     b. paáṇaáyanam   'reading'
     c. máratakam   'eme rald'

(49) a. angáarasàatmiíkaranam   'ca ıbon assimilation'
     b. mr ̃gamad_aadiírasaáyanam 'a medicine'

Malayalam is not a tone language. However, it is necessary to
distinguish between word level intonation and phrase level intonation
in Malayalam. The facts/wo rd level intonation a re:
                    of

(50) a. The primary stressed syllable has a low tone.
     b. The last syllable of the word has a high tone.
     c. All the secondary stressed syllables have high tones.

The assignment of tones is illustrated in (51):

(51) a. paṟaáṭi   b. paṟaṭí   c. paaṟaáyanam

We can account for the rising pitch contour, or the word melody, by assuming that it is of the form LH : the low tone is anchored to the primary stress, while the high tone spreads. In order to specify where it spreads, we propose the following convention for building trees:

(52)   Foot construction

    a. Construct feet on all long vowels.
    b. Mark the first syllable as extrametrical (cf. Hayes (1980)) if it has no foot and is followed by a foot.
    c. Construct feet on initial and final vowels.

(53)   Word tree

    Construct a left branching tree on feet.

(54)   Tone

    H spreads to all tree

Examples for the application of (52)-(54) are given below:

(55)   a. maraṭakam   b. paaṟaayanam   c. paṟaaṭi

$$
\begin{array}{ccc}
\overset{\displaystyle\wedge}{\underset{\text{maratakam}}{\begin{array}{cc} s & w \\ F & F \\ | & | \end{array}}} &
\overset{\displaystyle\overset{s}{\wedge}\wedge}{\underset{\text{paaraayanam}}{\begin{array}{ccc} s & w & w \\ F & F & F \\ \wedge & \wedge & | \end{array}}} &
\overset{\displaystyle\wedge}{\underset{\text{(pa)raati}}{\begin{array}{cc} s & w \\ F & F \\ \wedge & | \end{array}}} \quad (53)
\end{array}
$$

$$
\begin{array}{ccc}
\overset{\displaystyle\wedge}{\underset{\begin{array}{c}\text{ma aatakam}\\ | \\ L\ H \\ * \end{array}}{\begin{array}{cc} s & w \\ F & F \\ | & | \end{array}}} &
\overset{\displaystyle\overset{s}{\wedge}\wedge}{\underset{\begin{array}{c}\text{paaraayanam}\\ |\ / \\ L\ H \\ * \end{array}}{\begin{array}{ccc} s & w & w \\ F & F & F \\ \wedge & \wedge & | \end{array}}} &
\overset{\displaystyle\wedge}{\underset{\begin{array}{c}\text{(pa)raati}\\ | \\ L\ H \\ * \end{array}}{\begin{array}{cc} s & w \\ F & F \\ \wedge & | \end{array}}} \quad (54)^{7}
\end{array}
$$

2.1.3.2. Compounds

The facts of stress and tone dealt with in 2.1.3.1. are of particular interest to us because subcompounds and cocompounds exhibit differences in stress and tone assignment. A subcompound with a single primary stress and a single word melody acts as a unit, irrespective of the number of stems in it, while each stem in a cocompound is a separate unit, with stress and tone assignment taking place on every stem:

(56) Subcompounds

    a. [[[táaraà]₂[kaantan]]maarə]
         L H 2 H H
        a name husband pl. (Tara's husbands)

    b. [[máta]₂[widweèsam]]
       L 2 H H
      religion hatred   (hatred of religion)

(57) Cocompounds

    a. [[[ácchan]₃[amma]]maarə]
       L H 3 L H
      father mother pl.   (parents)

    b. [[[yáksa]₃[kínnara]₃[gándharwan]] maarə]
       L H 3 L H 3 L H
      Yaksha Kinnara Gandha wa pl.

In (56a), the <u>kaa</u> of the second stem is not considered the
first syllable for the purposes of stress assignment, and is not
assigned primary stress. In (56b), <u>ta</u> in the first stem is not con-
sidered the last syllable, and has no high tone on it. Both these
facts follow from the assumption that the two compounds are single
units for the purposes of stress and tone assignment. In contrast,
the first syllable of <u>amma</u> in (57a) is assigned a primary stress,
and the second syllable of <u>acchan</u> receives a high tone, indicating
that the two stems are distinct units for stress and tone assignment.

Thus, noncompound words in subcompounds are treated identically
for the purposes of stress and tone assignment, which takes place only
after all the operations of subcompounding. The domain of stress
and tone assignment is stipulated to be the cocompounding stratum.
Sample derivations are given below:

(58)   a.   mátawidwèeṣam
            L   H H

Sc	[matam] [widweeṣam]	
	[[matam][widweeṣam]]	compounding
	[[mata][widweeṣam]]	nasal deletion
Cc	[mátawidwéeṣam] (F F F / L H)	stress and tone

99

b. acchanammamaar̄ə [8]

    L   H L H

Sc	no rules

Cc	F F F F [acchan] [amma] L H L H	stress and tone
	F F F F [[acchan][amma]] L H L H	compounding

Infl.	F F F F [[acchanamma] maar̄ə] L H L H	affixation

Note that the inflectional ending maar̄ does not receive any secondary stress, and there is no high tone on the syllable. This follows from the stipulations that (i) the domain of stress and tone assignment is the cocompounding stratum, and (ii) inflections are attached after the cocompounding stratum[9] (2.1.2.1.).

The interesting cases of tone assignment are found in those subcompounds which contain cocompounds as one of their constituents. Observe that the sequence mata wiḍweeṣam in (56b) has a single word melody (LHH), while the same sequence has two melodies in (59a) (LH LH):

(59) a. jáati máṭa wiḍwééṣam
         L H L H   L H

b.

```
                    N₁
                   /  \
              stem     N₅
               |        |
              N₂        |
             /  \       |
           N₃    N₄     |
            |     |     |
         jaati matam widweesam
```

N_1 = subcompound

N_2 = cocompound

caste religion hatred   (hatred of caste and religion)

(60)   jaatimatawidweeṣam

De	[jaati] [matam]	
	—— ——	affixation
Sc	—— ——	compounding
Cc	F F F F [jaati] [matam] | | | | L H L H	stress & tone
	F F F F [[jaati][matam]] | | | | L H L H	compounding
Sc	F F F F [[jaatimatam] [widweeṣam]] | | | | L H L H	compounding
	F F F F [[jaatimata][widweeṣam]] | | | | L H L H	nasal deletion
Cc	F F F F F F [jaatimata widweeṣam] | | | | | | L H L H L H	stress & tone
	——	compounding

2.1.3.3. Stress and Tone Assignment in the Boundary Theory

It would be instructive to enquire at this point how the facts of stress and tone described above can be accounted for in a theory that makes use of distinct boundary symbols rather than ordered strata to characterise the domain of phonological rules. In order to distinguish between subcompounds and cocompounds in stress and tone assignment, the boundary theory must stipulate that word tree construction cannot take place across the cocompound boundary ## , the inflectional boundary #, and the word boundary ##.[10] Given the assumptions of metrical theory, such a stipulation would result in an odd projection of feet as well as boundary symbols. This disjunction is unnecessary in the stratum theory.

Even if a projection of boundary symbols and feet were allowed, the boundary theory is incapable of accounting for stress and tone assignment in compounds like (59), in which a subcompound contains a cocompound as one of its members.

Consider the contrast between (61) a and b:

(61) a. [[[jaati][mata]][widweeṣam]] :
         L H   L H   L H

         sub
        /\
      co   N
     /\
    N N
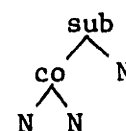
b. [[[kaamuka][[bhartrǝ][sahoodaran]]]maarǝ]:
      L   H   L   H   H

   lover   husband brother   pl.
   (Lovers and husband's brothers)

         co
        /\
       N sub
         /\
        N N
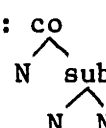
In terms of boundaries, (61) a and b receive identical
representation: X ## Y ## Z :

(62) a.   jaati ## mata ## widweeṣam

     b.   kaamuka ## bhartrə ## sahoodaran # maa r̄ə

The representations in (62) a and b do not make the necessary
distinctions for stress and tone assignment. One may think of exploit-
ing the bracketing difference to account for the distinction. Thus,
in addition to stipulating that ## blocks stress and tone assignment,
one may say that the rules apply cyclically, from smaller to larger
brackets. However, this assumption leads to the incorrect results of
assigning primary stress and a LH melody to every stem in a sub-
compound. Thus, the boundary theory has no account for the contrast
between (61) a and b.

2.1.4.   Summary

The structure of the model of phonology that emerges out of
2.1. is: a subset of phonological rules has the lexicon as the domain
of application. These rules apply prior to lexical insertion, thereby
making a distinction between lexical rule applications and post lexi-
cal rule applications. The lexicon consists of a set of ordered
strata, and the domains of phonological rules are characterised as
sets of continuous strata.

Morphological and phonological operations are interspersed
in a cyclic fashion. Every time a morphological operation takes place
at a stratum, the set of phonological rules which have that stratum
as their domain are scanned for applicability.

(63)   Morphological ——⟶ Phonological
        process   process
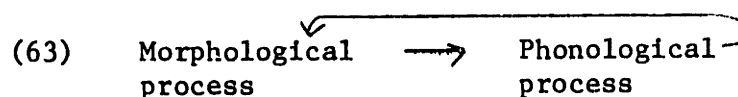
(63) has the effect of making all rule applications in the
lexicon cyclic, without stipulating them to be so (see also Pesetsky
(1979)).

It must be pointed out that the rules we have discussed so
far do not provide any evidence that lexical rules MUST apply cyclic-
ally, after every morphological operation. They could equally well
be accounted for by assuming that the phonological rules apply at
the end of all the morphological operations at every stratum.

(64)   **Stratum i**
        | mo rphological processes | **Phonology**
                                     rule 1
                                     rule 2
        Stratum i + 1   .
        | morphological processes | .
                                     .
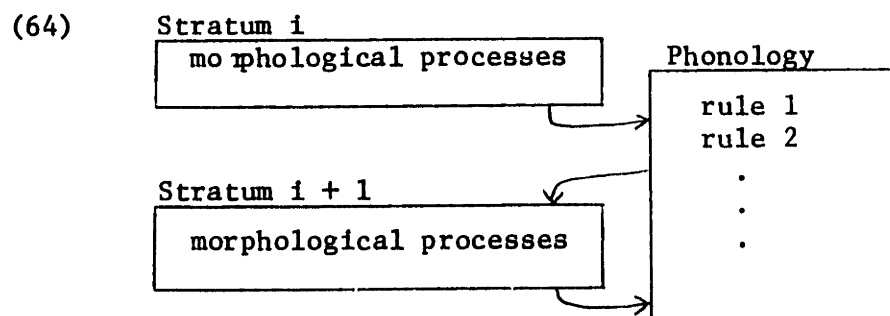
The strongest position would be either that all lexical
rule applications are cyclic, or that all of them are noncyclic. There
is evidence that at least some lexical rule applications must be
cyclic (Kiparsky (1979); Hayes (1980)). Therefore, I shall assume, in

the absence of counterevidence, that all lexical rule applications are cyclic, and adopt (63), which yields this result.

The stratum structure of Malayalam and the phonological rules associated with the strata are summarised in (65):

(65)   Strata   Phonological rules

```
    ↓
  stratum 1: derivations      ⇄  vowel lengthening
      ↓                              (domain: 2,3)
  ↘   ↓
  stratum 2: subcompounding   ⇄  nasal deletion
      ↓                              (domain: 2,3)
                                 gemination
      ↓                              (domain: 2)
  stratum 3: cocompounding     ⇄  vowel sandhi
      ↓                              (domain: 1,2,3)
  stratum 4: inflections       ⇄  stress
                                     (domain: 3)
      ↓                          tone (domain: 3)
```
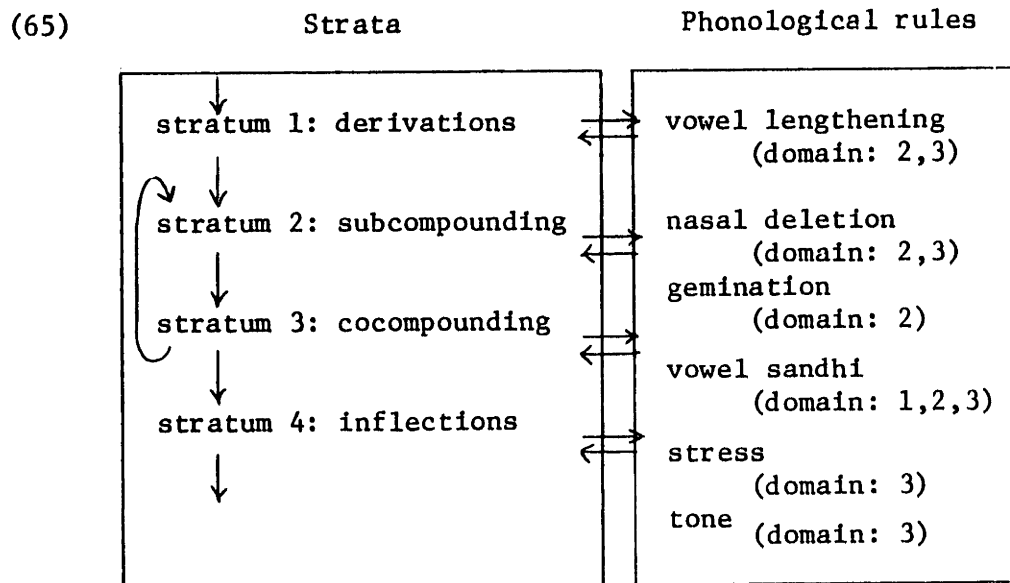
2.2. General Issues

2.2.1. What is a Stratum?

The assumption that lay behind our analysis of subcompounds and cocompounds was that they are generated by rules (66)a and b respectively:

(66) a. N ⟶ N N   (subcompounds)

     b. N ⟶ N*   (cocompounds)

There are no node labels that distinguish between the two

kinds of compounds, such as N^{sub} and N^{co}. Thus, the subcompound

pasukkuṭṭikaḷə 'calves', and the cocompound acchanammamaaṝə 'parents'

will have identical configurational structures:

(67) a.
```
              N
            /   \
          N      plur.
         / \      |
        N   N     |
        |   |     |
      pasu kuṭṭi kaḷ
       cow child pl.
```
b.
```
              N
            /   \
          N      plur.
         / \      |
        N   N     |
        |   |     |
     acchan amma maaṝ
     father mother pl.
```
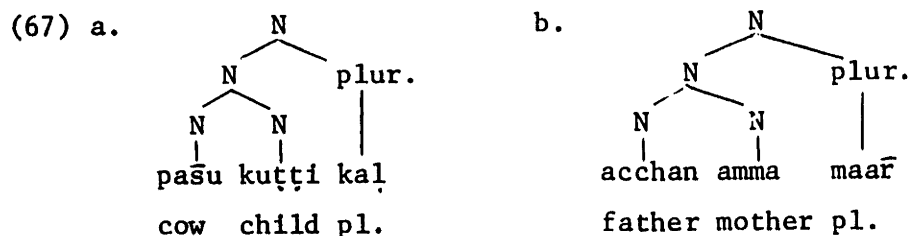
The distinction between subcompounds and cocompounds is made

in terms of the stratum at which the stems are put together. (66a)

has the subcompounding stratum as its domain, and (66b), the cocom-

pounding stratum.

Now, (66a) is not the only morphological rule that applies

at the subcompounding stratum. For instance, as we shall see in 4.5.,

and 4.6., the following rules also apply at this stratum:

(68)   The Causative Rule

$$V \longrightarrow V \text{ -ik'k'} \qquad e.g.$$
```
              V
            /   \
          V      ik'k'
         oot      cause
      'run-intr.'
```
            'run-tr.'
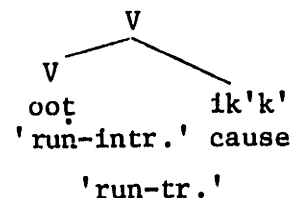
(69) The Verbal Compound Rule

N ⟶ N V -i   e.g.

```
                              N
                          ╱──────╱▲╲
                        N        V   ╲
                        │        │    ╲
      |          ambaṟa(m)   cumb  -i
                    sky        kiss agentive
                        'sky scraper'
```
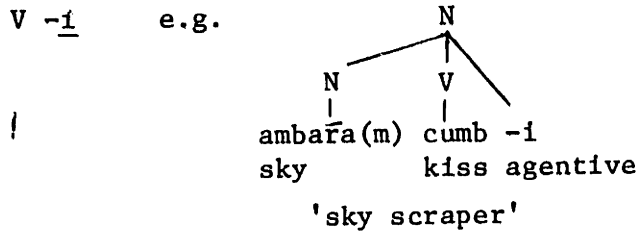
In other words, it is not subcompounds alone which are gene-rated at stratum 2, but causative verbs and verbal compounds as well. For mnemonic purposes, however, I shall continue to refer to stratum 2 as the subcompounding stratum.

Similarly, the derivational stratum is associated with rules such as N → a N ((16a)); N → sam N ((21)); A → prati A ((26c)).The inflectional stratum attaches number and case inflections to nouns: N → N num case; and tense inflections to verbs: V → V tense.

Given this approach to morphology, the rules which make a distinction between two morphological processes must apply at the appropriate stratum, as the morphological distinction is not avail-able in the final output of the word formation component. 2.1. de-monstrates this with respect to phonological rules. Semantic rules follow the same mode of application. Thus, cocompounds have the meaning of 'X and Y and ...', while subcompounds have various mean-ings such as 'X made with Y' (e.g. maṟakkuṭiṟa 'wooden horse'), 'Y's X' (kuṭiṟawaalà 'horse's tail'), 'X that operates with Y' (kuṭiṟawanṭi 'horse cart') etc. These semantic relationships must

be established at the stratum at which each compound is generated, parallelling the application cf phonological rules described in 2.1.

What this shows is that each stratum is associated with sets of morphological, semantic, and phonological rules which have that stratum as their domain. This conception of the lexicon can be schematised as follows:

(70)

strata	morphological component	phonological component	semantic component
1 2	a b c ⋮	X Y Z ⋮	M N O ⋮
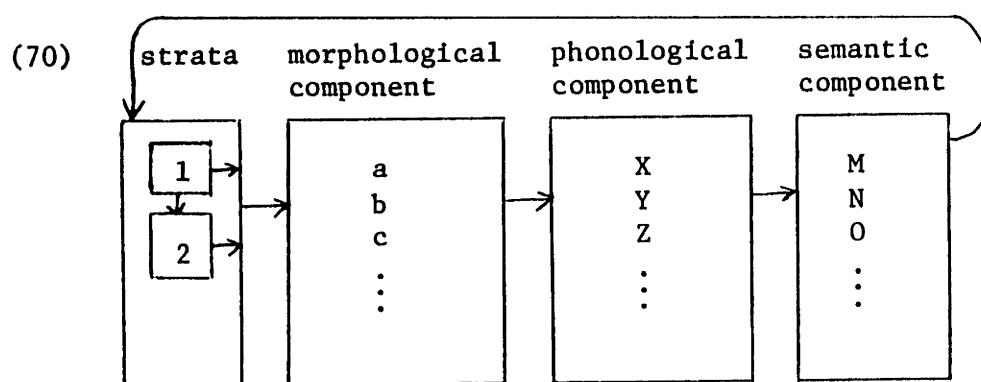
The relative order of the phonological and semantic compo-nents is irrelevant, but it is important that the morphological component precedes the other two.

A stratum, thus, is an abstract domain, at which several morphological, semantic, and phonological properties converge. The only intrinsic property of a stratum is the ordering relation it holds with respect to the other strata. The rest of its properties are derived from the application of various rules which have that stratum as their domain of application.

A stratum may be conceived of as a location, with an index

attached to it. Each location carries the following instructions:

(71) <u>Location i</u>

    a. Search the grammar for rules with index i, and apply them.

    b. Go to location j.

The notion of the stratum is similar to the notion of the component in linguistic theory, with one crucial difference. A component is identified by the processes and types of representations it contains. In other words, a component is uniquely associated with a set of rules and representations. No such unique association is possible with a stratum, as the same rule or representation may be associated with more than one stratum.

2.2.2. Rule Ordering in Lexical Phonology

Does Lexical Phonology require the extrinsic ordering of phonological rules? Supposing that stratum S_i is ordered prior to stratum S_j, and that rules P_i and P_j have S_i and S_j respectively as their sole domains, it would follow that the application of P_i would precede the application of P_j. If so, can we maintain the Unordered Rule Hypothesis such as the one proposed in Koutsoudas, Sanders, & Noll (1974), which says that rules apply whenever they are applicable, subject to universal restrictions on their application? Can the ordering of strata take care of the apparent counterexamples to the Unordered Rule Hypothesis?

A brief look at the facts of Malayalam phonology reveals that extrinsic ordering of phonological rules is, indeed, necessary, independently of stratum ordering. Consider the interaction between nasal deletion (2.1.1.2.) and vowel lengthening (2.1.1.3.). Both nasal deletion and vowel lengthening have the subcompounding and cocompounding strata as their domains; hence the stratum theory says nothing about the relative order of their application. Now, we find that the vowels before deleted nasals do not lengthen, as shown by the contrast between (72) and (73):

(72) a. [[jaaṝan][gṝham]] → jaaṝagrᵊham
         lover   house   lover's house

     b. [[kaalam][saṇḍeešam]] → kaalasaṇḍeešam
         time   message   the message of time

(73) a. [[bʰaaṝya][gṝham]] → bʰaaṝyaagrᵊham
         wife   house   wife's house

     b. [[kala][saṇḍeešam]] → kalaasaṇḍeešam
         art   message   the message of art

The vowel in the last syllable of the first stem lengthens in (73), but not in (72). These examples present a classic case of extrinsic ordering:

(74)   Vowel lengthening ⌐
        Nasal deletion   ⌐

(75)   kaalam saṉḏeešam   kala saṉḏeešam

_____   kalaa saṉḏeešam   Vowel lengthening

kaala   saṉḏeešam   _____   Nasal deletion

Another example that demonstrates the need for extrinsic ordering of phonological rules is the interaction of the process of vowel onglide formation and the processes of post nasal and intervocalic voicing.

(76) Vowel Onglide Rule[11]

$$V \rightarrow {}^{\partial}V \; / \; \begin{bmatrix} +\text{voice} \\ -\text{son.} \end{bmatrix} \underline{\quad}$$

(77) a. t̪alam → [t̪aḷam]   vs   d̪aḷam → [d${}^{\partial}$alam]
                 'a room'   'petal'

b. paalam → [paalam]   vs   baalan → [b${}^{\partial}$aalan]
                 'bridge'   'boy'

c. t̪iiṟam → [t̪iiṟam]   vs   d̪iinam → [d${}^{\partial}$iinam]
                 'shore'   'disease'

The onglide ∂ is inserted after d̪ and b, but not after t̪ and p.

The rule of post nasal voicing is:

(78) Post Nasal Voicing

$$\begin{bmatrix} -\text{son.} \\ -\text{cont.} \end{bmatrix} \rightarrow [+\text{voice}] \; / \; [+\text{nasal}] \underline{\quad}$$

(79) a. kampi → [kambi] 'a metal rod'

    b. kaṇṭu → [kaṇḍu] 'saw'

    c. cem + taamaṟa → [cendaamaṟa]
       red   lotus   red lotus

For the present purposes, I formulate the rule of intervocalic voicing (for details, see 4.1.5.) as follows:

(80) Intervocalic Voicing

$$\begin{bmatrix} -\text{son.} \\ -\text{cont.} \end{bmatrix} \rightarrow [\ +\text{voice}]\ /\ V \text{ --- } V\ ^{12}$$

(81) a. apakaṭam → [abagaḍam] 'danger'

   b. waacakam → [waajagam] 'sentence'

   c. koṭuṭṭu → [koḍuṭṭu] 'gave'
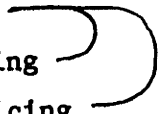
The rules (76), (78), and (80), all apply in the post lexical domain, as shown by the fact that they apply across words:

(82) a. raṣiid̲ ewiṭe → raṣiid̲$^{\partial}$ewiṭe (vowel onglide)
         Rashid where   Where is Rashid?

   b. r̄aaman pooyi → r̄amambooyi (post nasal voicing)
         Rama   went   Rama went

   c. kuṭṭi karaññu → kuṭṭigar̄aññu (intervoc. voicing)
         child cried   child cried

Given that these rules apply in the same domain, the stratum ordering says nothing about their relative order of application. Nevertheless, we find that the vowel onglide rule musc apply prior to the voicing rules. Contrast a and b:

(83) a. manṯan → [mand̲an] 'a fool'

   b. mand̲an → [mand̲$^{\partial}$an]'slow person'

(84) a. waaṯam → [waaḍam] 'rheumatism'

   b. waad̲am → [waad̲$^{\partial}$am]'argument'

In order to account for the fact that the vowel onglide rule does not apply in (83a) and (84a), we stipulate the following order:

(85) Vowel onglide
      Post nasal voicing
      Intervocalic voicing

2.3. Excursions into Other Languages

In this section, I shall present brief illustrations of the application of the stratum theory to English and Dakota, in order to give a clearer picture of the model by examining familiar phenomena, as well as to clarify the kinds of predictions the theory makes.

2.3.1. English

2.3.1.1. How many strata does the lexicon of English require? How are they ordered? The first approximation is provided by (86):

(86) Stratum ordering in English

Stratum 1: Class I affixation (in-,-ity,-ion,-ic...)

Stratum 2: Class II affixation (un-, re-,-dom,-ship...)

Stratum 3: Compounding

Stratum 4: Inflection

It has been pointed out by Selkirk (to appear) that class II derivations can be attached to compound stems, even though class I derivations cannot (Siegel (1974), Allen (1978)). Thus, though *in+self-sufficient and *in+outmoded are impossible, un#self-suffi-cient and ex#frogman are not. In order to account for these facts, we provide a loop from 3 to 2:[13]

(87) Stratum Ordering in English

Stratum 1 : Class I derivations
↓
Stratum 2 : Class II derivations
↓
Stratum 3 : Compounding
↓
Stratum 4 : Inflections
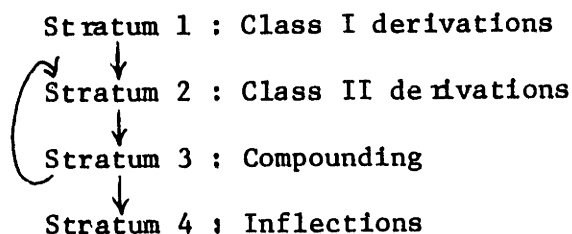
Is there a loop from stratum 4 to 3? Evidence suggests that there cannot be such a loop, even though Selkirk cites compounds like arms conscious and pants suit to show that compounds can contain inflected stems. Alongside these compounds, we also have trouser-leg (*trousers-leg, *trouser, trousers), and scissor-cut (*scissors-cut,*scissor, scissors). Observe that the prohibition on inflected stems extends to the tense inflections as well. Thus, as Selkirk herself notes, *scrubs woman and *scrubbed woman are impossible in English. I shall, therefore, conclude that inflected stems cannot be compounded in English. Compounds such as arms conscious and pants suit are marginal, and some alternative explanation will have to be offered for them.

Note that (87) correctly predicts that derivational affixes cannot be attached to inflected stems: *programser (programmer), *weaponsry (weaponry), *kingsdom (kingdom), *childrenlike (childlike). A loop from 4 to 3 would also amount to a loop from 4 to 2, and we would have no explanation for the impossibility of attaching derivational affixes to inflected stems.

2.3.1.2.   Defining the domain of a rule as stratum 1 has the effect

of allowing the rule to apply across +, but not across # or ##. SPE

notes that there are no rules in English which are blocked by the

presence of +, and stipulates the convention that a rule of type

(88a) is optionally interpretable as (88)b, c, and d:

(88) a.   A → B/ X — Y

     b.   A → B/ X — + Y

     c.   A → B/ X + — Y

     d.   A → B/ X + — + Y

This convention guarantees that + is incapable of blocking

phonological rules, unlike the other boundaries.

In Lexical Phonology, it is unnecessary to have anything

corresponding to this convention. Observe that + is, in fact, the

boundary associated with class I derivations, and is, therefore, the

juncture at stratum 1. The only way for a phonological rule not to

apply across a stratum 1 juncture is for it to be assigned to a

previous stratum, which does not exist. Thus, no phonological rule

can be blocked by the juncture at stratum 1. Generalising, we may say

that defining the domain of a rule as stratum S_i and not stratum S_j

(where S_i is prior to S_j) has the effect of allowing the rule to apply

across the boundaries associated with S_i and the preceding strata,

and blocking the application of the rule across the boundaries asso-

ciated with S_j and the subsequent strata. Thus, the effects of rule

Lexical Phonology accounts for the stress neutrality of class II affixes by assumimg that the domain of foot construction is stratum 1.[14] This domain assignment has the effect of not allowing $\#$ and $\#\#$ inside a foot. Given below are the derivations which illustrate the application of this solution.

(90)

	dialectal	dialectdom	
S_1	F s\ /\ \ sw w dialect	F s\ /\ \ sw w dialect	foot construction
	F s\ /\\ sw w [[dialect] al]	____	affixation
	F F /\ /\ sw s w [[dialect] al]	____	foot construction
S_2	____	F s\ /\ \ sw w [[dialect] aom]]	affixation
	word /\ w s F F /\ /\ sw s w [dialectal]	F word s\ /\ \ sw w [[dialect] dom]]	word tree formation
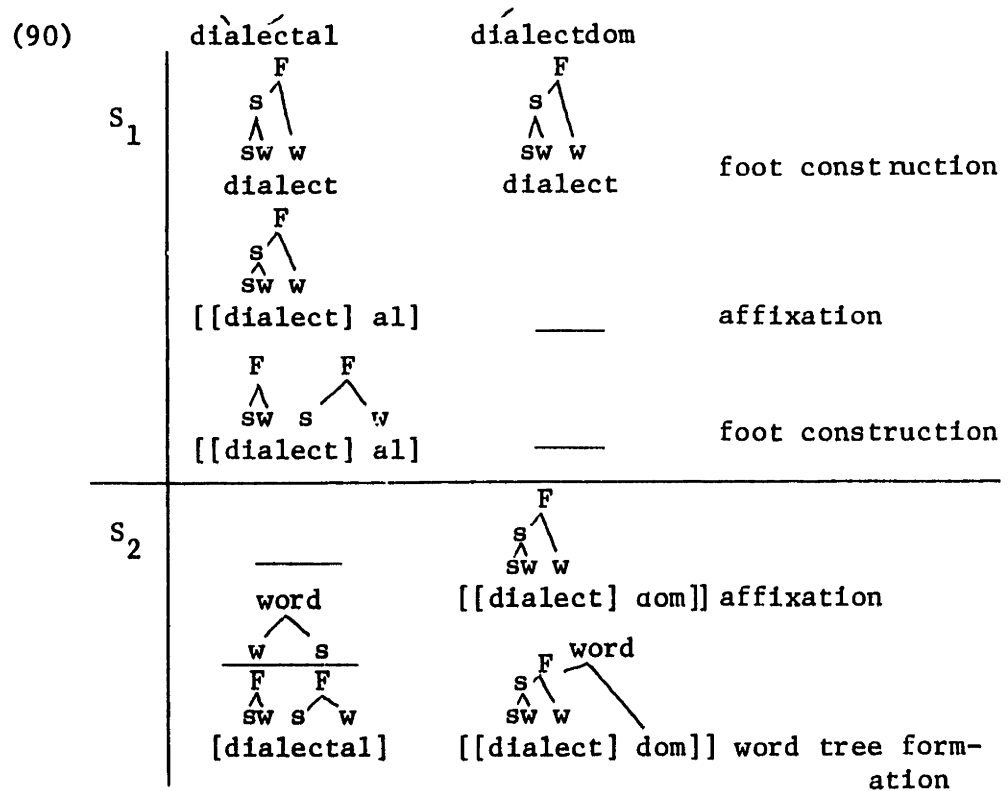
Note that no new feet are formed at stratum 2, and the existing foot trees are undisturbed. The joining together of feet into word trees takes place at stratum 2. Alternately, one may assume that word tree formation can take place at both 1 and 2. I cannot think of any empirical consequence which hinges on these alternatives.

Class II suffixes like -ness, -less, and -dom do not have any stress on the suffixes themselves. On the other hand, some class II suffixes like -ship, -hood, and -like, and all the class II prefixes I can think of (un-, de-, re- ...) are inherently stressed: chìldlìke, kíngshìp, mánhòod, ùnknówn, rèdesígn. In order to account for the contrast, I shall assume that -ship, -hood, un-, etc. have inherent feet.

(91)   kingdom   kingship

S_1 [king]   [king]

           F   F
           [kíng]   [kìng]   foot construction

S_2 F   F   F
           [[kíng] dom]   [[kìng] shìp]   affixation

                                     s   w
           ─────────   ─────────────
           F   F   F
           [[kíng] dom]   [[king] shìp]   word tree

Observe that the rule governing the word tree at stratum 2 is: the stem is assigned S. Thus, we have mánhòod but ùnknówn. One may venture to suggest that this can be accounted for by the following universal principle:

(92) Unless otherwise required by the grammar, the stem
      is dominant.[15]

When a class II affix is attached to a stem, the primary stress is always on the stem. In contrast, most compounds generally

take the primary stress on the first stem. What this means is that

the compound stress rule applies only at stratum 3 (because 3 is

the stratum of compounding), and not at prior or subsequent strata.

We state the compound stress rule as follows:

(93)   Compound Stress Rule   (Domain : Stratum 3)

The right node is strong iff branching.

(Liberman & Prince (1976))

Compare the derivations of <u>unknown</u> and <u>blackbird</u>:

(94)

	ùnknówn	blảckbìrd	
S_1	[known]	[black][bird]	
	F [knówn]	F　　F [blảck][bìrd]	foot construction
	————	————	affixation
S_2	F　　F [ùn[knówn]]	————	affixation
	w⌃s F　　F [ùn[knówn]]	————	principle (92)
S_3	————	F　　F [[blảck][bìrd]]	compounding
	————	s⌃w F　　F [[blảck][bì rd]]	Compound stress rule

To summarise, we find that foot construction takes place in

English at stratum 1. There are no labelling conventions at stratum 2,

and the universally unmarked condition that the stem is dominant

applies. Stratum 3 is the domain of the compound stress rule.

2.3.1.4. There are three types of apparent problems for our analysis of English morphology. The first type is illustrated by words like nuclear physicist. If we adopt the bracketing suggested by the mean-ing of the word ('one who studies nuclear physics', not *'a physicist who is nuclear'), we arrive at [[nuclear physic]ist], which violates the generalisation that class I affixes are not attached to compounds. Note, however, that semantic bracketing does not always correspond to morphological bracketing. English grammarian, for example, can mean 'one who studies English grammar', but we do not for this reason propose the bracketing [[English grammar]$_{NP}$ ian]. Given that some principle of semantic interpretation is necessary to account for English grammarian, we can adopt the bracketing [[nuclear][physicist]], and allow the same principle to apply to it.

The second type of problem is found in words like developmental and governmental. If the -ment in these words is a class II affix, we arrive at the structure [[[stem]class II affix]class I affix], which our analysis of English morphology disallows. We suggest that the -ment in these cases, unlike the -ment in, say, fulfillment, is a class I affix. Some support for this position is offered by the stress shift in these words (development vs developmental). Class II affixes do not cause stress shift.

The third problem is seen in words like ungrammaticality and extrametricality. Assuming that un- is attached to adjectives and not to nouns (unfortunate, but *unfortune), the bracketing

has to be [[un[grammatical]]ity], which involves the attachment of a class I affix to a stem containing a class II affix. At this point, I do not have a solution to this problem.

2.3.2. Compounding in Dakota

2.3.2.1. Basic Facts

Chambers & Shaw (1980; henceforth C&S) show that there are two kinds of compounds in Dakota, which they refer to as lexical and syntactic compounds, illustrated in (95)a and b respectively:[16]

(95) a. hã | wakhã → hãwakhã   (lexical)
          night holy   the northern lights

      b. hã # wakhã → hãwakhã   (syntactic)
                           holy night

C&S use the boundary | for lexical compounds, # for syntactic compounds, and + for inflections. The semantics of syntactic compounds, as illustrated by (95b), is compositional, while that of lexical compounds is almost totally unproductive. There are several phonological processes in which the distinction between the two compounds clearly shows up. As illustrated in (95), for example, lexical compounds have a single stress, while syntactic compounds have two. C&S take care of the stress facts with the following rules:

(96) Dakota Accent Rule (DAR) (p. 325)

$$V \rightarrow \acute{V} \; / \; \#(C_0 V)C_0 \text{---}$$

(97) Compound Accent Rule (CAR) (p. 327)

$$\acute{V} \rightarrow \grave{V} \; / \; [\#\#...\acute{V}...\#... \text{---} \; ...\#\#]_{N,V}$$

The environment $\#$ is met only once in (95a) ($\#\#$hă|wakhă$\#\#$),
but twice in (95b) ($\#\#$hă$\#$wakhă$\#\#$), and therefore, DAR applies to
both stems in (95b). CAR follows, reducing the second $\acute{V}$ to $\grave{V}$.

Dakota has several other processes that demand the distinction
between lexical and syntactic compounds. One of them is the rule of
stem formation, given by C&S as follows:

(98) Stem Formation (p. 330)

$$\emptyset \rightarrow a/ \; C \text{---} \#$$

(98) derives [čepa] 'to be fat' from the underlying $\#\#$čep$\#\#$.
The rule applies word internally in syntactic compounds, but not in
lexical compounds:

(99) a. čheγ $\#$zi $\rightarrow$ čheγazi 'yellow kettle'

b. čheγ| zi $\rightarrow$ čhexzi 'brass kettle'

Similarly, the rule of devoicing applies only in lexical
compounds:

(100)
$$\begin{matrix} z \\ \check{z} \\ \gamma \end{matrix} \rightarrow \begin{matrix} s \\ \check{s} \\ x \end{matrix} \; / \text{---} |$$

(100) changes γ to x in (99b).[17]

2.3.2.2. The Ordering of Strata in Dakota

An interesting feature of Dakota morphology is that the
processes of compounding and derivation apply to inflected stems.
Consider, for example, the behaviour of the inflectional prefixes
wa 'I', ya 'you', and ni 'you':

> (101) Non derived words
>
> > waúpa   'I lay down' (wa + ũpa)
>
> (102) Compound words
>
> > a. čhokhű 'to plan to kill' (čho| khũ)
> > b. čhowákhũ 'I plan to kill (someone)' (čho[wa + khũ])
> > c. čhoyákhũ 'you plan to kill (someone)' (čho[ya+khũ])
>
> (103) Derived words
>
> > a. yeyá 'send' (ye 'go'; ya 'you')
> > b. yewáya 'I send (someone)'
> > c. niyéwasi 'I ordered you to go' (si 'order')

Observe that the inflectional morphemes wa and ya occur inside
compounds ((102)b,c), as well as inside derivations ((103)b,c),
showing that inflectional affixation is an input to both derivation
and compounding. The Dakota facts present a marked case: in most
languages, inflectional affixation is not an input to derivational
affixation, shown by the fact that derivational affixes cannot be

attached to inflected stems. I give below the unmarked universal

stratum ordering and the marked Dakota ordering:

(104) Unmarked Stratum Ordering

        ↓
    stratum i : derivations
        ↓
    stratum j : inflections
        ↓

(105) Dakota

        ↓
    stratum i : inflections
        ↓
    stratum j : derivations
        ↓

Expanding (105) on the basis of the information available

in C&S, we arrive at the following stratum ordering in Dakota:

(106) Dakota stratum ordering

    Stratum 1 : Inflection
         ↓
    Stratum 2 : Derivation
         ↓
    Stratum 3 : 'Lexical' compounding in C&S
         ↓
    Stratum 4 : 'Syntactic' compounding in C&S

In passing, we note that the Dakota facts confirm our assump-

tion that inflectional affixation takes place in the lexicon, rather

than in syntax (see also Lieber (1980), Selkirk (to appear), Bresnan

(in press: a, b)). This assumption is to be contrasted with the

approach taken in , say, the Government Binding Theory (Chomsky (1981))

in which inflections attach to their stems in syntax by means of rules

such as affix hopping. Given Chomsky's assumption that derivational

affixation takes place in the lexicon (Chomsky (1970)), and the fact that inflectional affixation is an input to derivation, it follows that inflectional affixation should also take place in the lexicon.

2.3.2.3. The Domains of Rules

We are now ready to define the dom⁻ins of Dakota rules in terms of the strata in (106). The stem formation rule ((98) for C&S) clearly belongs to stratum 4.

(107) Stem Formation (Domain : Stratum 4)

$$\emptyset \rightarrow a/ \ C \text{ —}]$$

The rule of devoicing ((100)) in the stratum theory would belong to strata 2 and 3:

(108) Devoicing (Domain : Strata 2-3)

$$\begin{matrix} z \\ \check{z} \\ Y \end{matrix} \rightarrow \begin{matrix} s \\ \check{s} \\ x \end{matrix} \ / \text{ —}]$$

The DAR ((96)) should apply at stratum 4, as it applies only when the syntactic compound boundary is present:

(109) Dakota Accent Rule (Domain : Stratum 4)

$$V \rightarrow \acute{V}/ \ [(C_0 V)C_0 \text{ —}]^{18}$$

The CAR ((97)) also belongs to stratum 4:

(110) The Compound Stress Rule (Domain : Stratum 4)

$$\acute{V} \rightarrow \grave{V} \ / \ [[\ldots \ \acute{V} \ \ldots][\ldots \ — \ \ldots]]$$

The derivations of (95) a and b are given below:

(111) hãwákhã 'northern lights'

S$_1$	hã wakhã	
	___ ___	
S$_2$	___ ___	
S$_3$	[[hã] [wakhã]]	compounding
S$_4$	[hãwãkhã]	DAR
	___	CAR

(112) hãwakhã 'holy night'

S$_1$	hã wakhã	
	___ ___	
S$_2$	___ ___	
S$_3$	___ ___	
S$_4$	hã wakhã	DAR
		CAR
	[[hã][wakhã]]	compounding
	___ ___	DAR
	[[hã][wakhã]]	CAR

2.3.2.4. Advantages of the Stratum Treatment

2.3.2.4.1. Assuming that the presence of all boundaries except +
blocks phonological rules unless stated in the structural descrip-
tion of the rule, the formulation of DAR by C&S as in (96) is in-
correct. Observe that the rule applies across | , as shown by (113):

(113) a. cho|k?ũ → chok?ṹ 'to plan to kill' (p. 325, 6)

  b. skal|o+mani → skalómani 'he goes about in order
                                              to play' (p.328)

There can be an optional | both before and after the second
C_0 in the rule. For DAR to apply correctly in (113), C&S will have
to reformulate the rule as:

(114) $V \rightarrow \acute{V}$ / #(C_0V)(|)C_0(|) ——

Rule (114) misses the generalisation that DAR simply ignores
the presence of | anywhere in the string. Why, for example, couldn't
there have been a rule like (115)?

(115) $V \rightarrow \acute{V}$ / #(C_0V)(#)C_0(|) ——

Rules like (115) are unstatable in the stratum framework, as
the Opacity Principle would prevent referring to junctures at diffe-
rent strata. In the boundary theory, however, (114) and (115) have
the same status: there is no reason to prefer one to the other. Now,
as far as I know, rules such as (115) do not exist in natural

languages. If so, the stratum theory allows possible rule types
and correctly disallows impossible rule types, while the boundary
theory allows impossible rule types (e.g. (115)), and sometimes
disallows possible rule types (e.g. stress and tone in Malayalam).

2.3.2.4.2. As seen in 2.3.2.1., Dakota exhibits the following
peculiar properties:

> (116) a. Inflections do not block phonological rules.
>
> b. Inflections can occur inside 표 derivations.

C&S assign the boundary + to inflections. Given the SPE
approach to boundaries, we would expect inflections to be character-
ised in terms of # rather than +. C&S obviously use + instead
because phonological rules apply freely across inflectional junc-
tures in Dakota (e.g. DAR in (113b); (103)), and it is only + which
does not block phonological rules.

C&S's assignment of + to inflections, although it guarantees
the right results, is unprincipled. Their treatment is inadequate
also in that it does not relate (116a) and (116b). In Lexical Phono-
logy, given (116b), (116a) follows. Recall that the juncture asso-
ciated with stratum 1 is incapable of blocking phonological rules.
Given that inflections can occur inside 표 derivations, we are forced
to conclude that the stratum of inflection precedes that of derivation,
and as there is no evidence that anything precedes inflection, the
inflectional stratum becomes stratum 1. It follows, therefore, that

the juncture of inflections cannot block phonological rules.

Given that in English, class I derivations constitute the innermost layer of morphology, our theory would assign it to stratum 1, and thereby make its juncture incapable of blocking rules. Similarly, given that inflections constitute the innermost layer in Dakota, the stratum theory makes the inflectional juncture incapable of blocking rules. In other words, Lexical Phonology explains the association of + with class I derivations in English and with inflections in Dakota.

In sum, we find that the lexical theory allows us to account for the optionality of morphological junctures in a principled manner. This theory disallows those types of optionality of junctures which do not occur in natural languages. The stratum treatment also brings together the two facts of Dakota that inflections occur inside derivations, and that inflectional junctures do not block phonological rules.

2.4. Alternatives to the Stratum Theory

2.4.1. Junctures and Domains

There are two kinds of information that phonological rules require of morphology, which may be characterised as juncture information and domain information. The juncture information, to use Rotenberg's (1978) terminology, consists of edges and junctions.

Thus, when two forms $[a_1, a_2 \ldots a_n]_A$ and $[b_1, b_2 \ldots b_n]_B$ combine together to form $[[\;]_A[\;]_B]$, we can say that a_1 and a_n occur at the edges of A, and b_1 and b_n occur at the edges of B, and that a_n and b_1 occur at the junction between A and B.

All phonological theories require the juncture information. Thus, in SPE, rules of the form (117a) refer to edges, and (117b) to junctions:

     (117) a. $X \rightarrow Y/ \text{---} + $; $X \rightarrow Y/ \#\text{---}$

           b. $X \rightarrow Y/ \text{---} +b$; $X \rightarrow Y/ a\#\text{---}$

In the lexical framework, rules of the form (118a) refer to edges, and (118b) to junctions:

     (118) a. $X \rightarrow Y/ \text{---}]$; $X \rightarrow Y/ \; [\text{---}$

           b. $X \rightarrow Y/ \text{---}][b$; $X \rightarrow Y/ \; a] \text{---}$

In addition to juncture information, phonological rules need to know about the morphological domain in which they apply. One of the primary means of making use of domain information in SPE was to encode it in terms of distinct boundary symbols, thereby packing the domain and juncture information simultaneously into symbols or a par with the linear segmental symbols. Thus, the symbol + would roughly mean 'juncture, class I affixation', and $\#$ would mean 'juncture, class II affixation or inflection'.

In addition to using boundary symbols, SPE made use of node labels on phrase structure trees to specify the domain information in phonological rules. Thus, to encode the fact that the compound stress rule in English applies in the domain of compounds, and the nuclear stress rule applies in the domain of phrases, SPE labelled the brackets in the structural descriptions of the rules with N, NP, etc.:

(119) a. $...X]_{N,A,V}$ (compound stress rule)

b. $...X]_{NP,VP,PP...}$ (nuclear stress rule)

2.4.2. The Boundary Alternative

SPE has the following mechanism for assigning boundaries:

(120) a. All morphemes are flanked by the boundary +.

b. All major categories and their projections
   (N, V, A, NP, VP...) are flanked by #.

(120a) derives representations like (121):

(121) a.   N   b.   A

            A   A
            | |
       +divine+ +ity+   +un+ +known+

c.

A
N
stem
+tele+ +graph+ +ic+

d.

NP
det.
N
N
A
N
+the+ +black+ +bird+ +pl.+
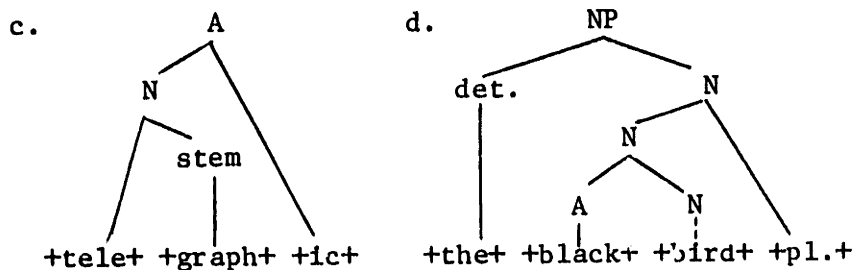
Subsequent application of (120b) yields represenattions like (122):

(122) a.

N
A
##divine# ity#

b.

A
A
#un #known ##

c.

A
N
stem
##tele+graph#ic #

d.

NP
det
N
N
A
N
#the### black## bird## pl.##
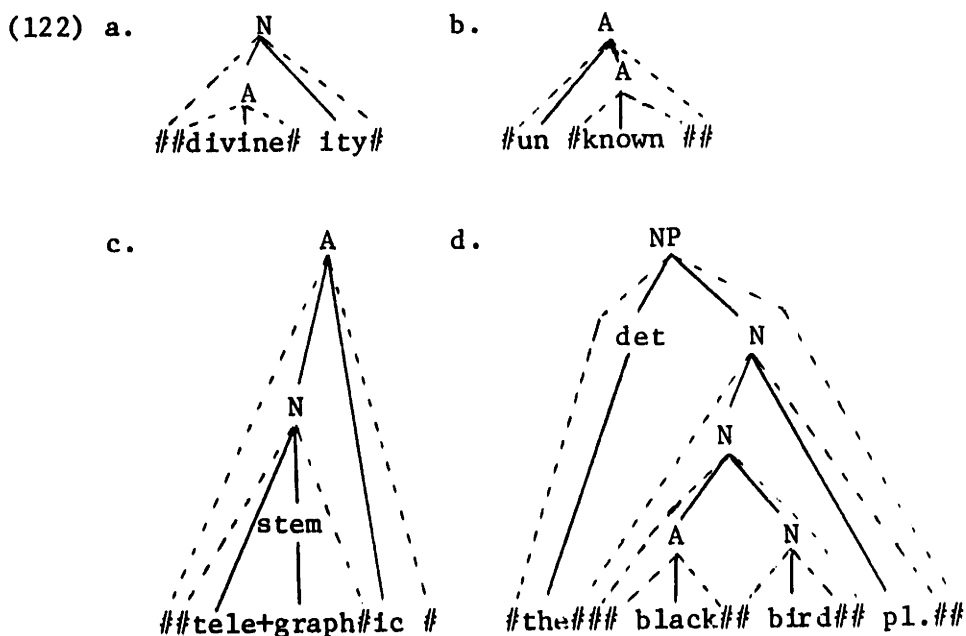
SPE also requires a readjustment rule that changes # to + in the case of affixes like -<u>ity</u> and -<u>ic</u> (class I affixes), thereby deriving <u>divine+ity</u> and <u>tele+graph+ic</u> from (122) a and c. Another readjustment rule changing ## to # derives <u>black##bird#z</u> from (122d), as (120b) incorrectly assigns ## before an inflectional suffix in compounds. This readjustment rule applies to class II affixes as well: <u>ex##frog##man</u> → ex#frog##man.

This discussion shows that there is no principled assignment of boundaries in SPE. The correct boundary is ensured by ad hoc readjustment rules which rectify the incorrect outputs of (120b).

Boundary assignment becomes far more complicated than in English when we take a language like Malayalam. Readjustment rules would be required to change $\#$ to $+$ in derivations, to $\#$ in subcompounds, and to $\#\#$ in cocompounds. What is the environment for these readjustment rules? It cannot be the case that individual morphemes trigger the rules, because the same morpheme sequences can be either subcompounds or cocompounds. One way to insert the correct boundaries would be to assume that subcompounds and cocompounds are distinguished by node labels such as N^{sub} and N^{co}. We may now think of the following mechanism for boundary insertion or for the application of readjustment rules:

(123)   Insert $\#$ between the stems dominated by N^{sub}, and $\#\#$ between the stems dominated by N^{co}.

This approach can be extended to other boundaries as well. Now, if it is the node labels that tell us which boundary symbol to insert, why do we need separate boundary symbols at all? Instead of formulating the rule of gemination in subcompounds as (124), for example, we may formulate it as a rule that applies in N^{sub}, as in (125):

(124)   onset $\rightarrow$ $\wedge$ $\%$ — $\#\#$

(125)   onset $\rightarrow$ $\wedge$ $\%$ —][in N^{sub}

An alternative to (124) would be to adopt the stratum theory for the purposes of morphology, and stipulate that each stratum assigns its own boundary symbol as part of the derivation (cf: Allen (1978)). Once again, the question arises why boundary symbols are necessary in addition to the notion of morphological strata. The point that emerges out of this discussion is that boundary symbols only duplicate the information provided by node labels like N^{sub} or by morphological strata, and are therefore redundant in phonology.

Let us now turn to the consequence of formulating phonological rules in terms of boundaries. The stratum theory crucially assumes that (i) the structure at one stratum is invisible to the rules at another (Opacity Principle), (ii) the domains of rules are continuous strata (Stratum Domain Hypothesis), and (iii) the strata are ordered in relation to each other (Stratum Ordering Hypothesis). These assumptions have empirical consequences which are of relevance to a comparison between the stratum theory and the boundary theory.

Consider, first, a rule of the following type:

(126)   A → B/ X+Y ——#Z

The boundary + belongs to stratum 1, and # to stratum 2 or a subsequent stratum. Reference to the junctions at two different strata in the same rule is disallowed by the Opacity Principle (see also 2.3.2.). Hence, the Opacity Principle forbids the formulation of a rule equivalent to (126). Rules like (126), as far as

I know, do not exist in natural languages. If so, we find that the stratum theory, not the boundary theory, correctly rules out the impossible rule types in natural languages.

One may object that the same result can be derived from the boundary theory by incorporating the Opacity Principle. Note that the Opacity Principle is a necessary component of the stratum theory : without it, there would be no way of distinguishing between the brackets assigned at different strata. In the boundary theory, on the other hand, the principle is an extra stipulation designed for the sole purpose of accounting for the impossible rule types.

Another instance of the restrictiveness of the stratum theory comes from the Stratum Domain Hypothesis, which says that the domains of rules must be continuous strata. Recall that rules like nasal deletion, vowel lengthening, and vowel sandhi in Malayalam have more than one stratum as their domain. The domain of vowel sandhi, for example, consists of strata 1, 2, and 3. In 4.2., we shall see a rule that inserts a glide between two adjacent vowels, with all the lexical strata and the post lexical stratum as its domain. It is logically conceivable that a rule applies at stratum 1, 2, and 4, without applying at 3. The Stratum Domain Hypothesis, however, forbids such a domain assignment, and therefore, such a process is unstatable as a single rule in the lexical theory. Thus, the theory would rule out a dialect of Malayalam in which vowel sandhi applied in derivations and cocompounds, but not in subcompounds. Such

dialects do not, in fact, exist. In the boundary theory, however, nothing prevents rules from applying in the environment of + and ## , and not ##.

One may incorporate the effect of the Stratum Domain Hypothesis in the boundary theory by defining a hierarchy of boundaries such as + ## ## # ##, and stipulating that no rule may mention two nonadjacent boundaries. Such a stipulation, though yielding the desired result, would be entirely ad hoc. Given the independently motivated ordering of domains in the stratum theory, on the other hand, the most natural assumption would be that the domains are not discontinuous.

The third prediction that the stratum theory makes deals with rule blocking, and is derived from the Stratum Ordering Hypothesis. As was illustrated in 2.3.1., the effect of a rule being blocked by a boundary B_j is achieved in the stratum theory by assigning stratum S_i as the domain of the rule, where B_j is associated with a stratum following S_i. It follows from this treatment of rule blocking that a rule which applies across a strong boundary will also apply across a weaker boundary:
(weak $\leftarrow$ + ## ## # ## $\rightarrow$ strong). Conversely, a rule that is blocked by a weak boundary will also be blocked by a stronger boundary. These are powerful empirical predictions, not available to the boundary theory except by stipulation.

The strongest argument against the boundary theory is that it is descriptively inadequate. For the facts for which the stratum theory has a simple account, the boundary theory does not have an adequate description, as seen in 2.1.3.3.

In sum, the lexical model of phonology allows us to eliminate the ad hoc use of boundary symbols by referring to morphological information more directly, in terms of the independently motivated morphological strata and morphological bracketing. This model is explanatorily superior in that it correctly disallows phonological rule types and rule interactions which the boundary theory allows. At the descriptive level, the lexical model can account for all the existing phenomena that the boundary theory can account for, but there exist phenomena in natural languages such as the stress and tone assignment in Malayalam which are easily accounted for within the lexical approach, but not within the boundary approach.

2.4.3. The Node-Domain and Level-Domain Theories

Specifying the domains of phonological rules in terms of morphological strata is not the only alternative to the boundary theory. In the theory of phonology proposed by Selkirk (to appear), for example, domains are characterised in terms of the node labels of phrase structure trees without appealing to the notion of morphological strata (see also Rotenberg (1978)), I shall refer to Selkirk's

theory as the 'node-domain' theory, in contrast to the 'stratum-domain' theory advocated in this thesis. In what follows, I shall give a comparison of the two theories.

In Selkirk's theory, what corresponds to strata are X bar categories like X^0, X^{-1}, X^{-2}, etc., where the negative numerals indicate levels below X^0. In order to account for the distributional properties of class I and class II affixes in English, for example, Selkirk assumes that class I affixes are subcategorised for "roots" (X^{-1}), and class II affixes for "words" (X^0). The structure of theatrical, unbroken, and arm chair would thus be:

(127)

theatre ic al   un break en   arm chair
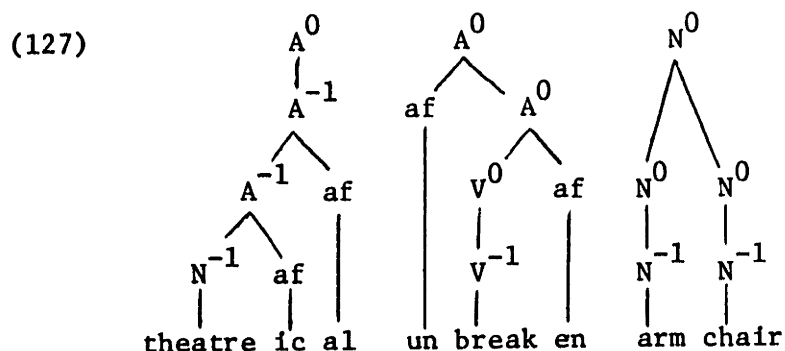
The effects of stratum ordering are built into rewrite rules as follows:

(128) a. $X^{-1} \rightarrow$ affix X^{-1} (X^{-1} affix)
  b. $X^0 \rightarrow X^{-1}$
  c. $X^0 \rightarrow$ affix X^0 (X^0 affix)
  d. $X^0 \rightarrow X^0 X^0$

(128a) defines class I affixation, (128c) defines class II

affixation and inflection, and (128d) defines compounding. As stated earlier, class I affixes are subcategorised for X^{-1}, and class II affixes for X^0. Therefore, X^{-1} can contain monomorphemic forms, and forms derived through class I affixation, but not forms derived through class II affixation or compounding (X^0). The stress neutrality of class II affixes is taken care of by defining the domain of foot construction as X^{-1}. The rule of nasal deletion (in <u>illegal</u>), tri-syllabic laxing, etc., are also restricted to X^{-1}.

This use of X bar categories like X^{-1} and X^{-2} is unmotivated, and does not yield any of the advantages of the X bar theory, such as the characterisation of the rule types, cross categorial generalisa-tions, etc. A more fruitful way of making use of the X bar theory in word formation would be to assume that it is affixes which are X^{-1} categories. For example, one may think of the following representa-tion for words like <u>unpatriotic</u>:

(129)

$$
\begin{array}{c}
A^0 \\
\diagup\ \diagdown \\
A^{-1}\quad A^0 \\
|\qquad N^0\diagup\diagdown A^{-1} \\
\quad\ |\qquad\quad | \\
\text{un}\ \text{patriot}\ \text{ic}
\end{array}
$$
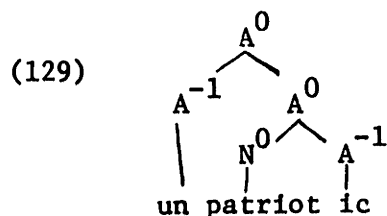
Reinterpreting Williams' (1979) insight, we may assume that the category of the node dominating the affix + stem (or stem + affix) combination is determined by the affix. This assumption need not be stipulated in this theory, since, given the X bar framework, it follows that the affix (X^{-1}) is the head in the rule $X^0 \rightarrow X^0\ X^{-1}$

(or X^{-1} X^0), and the dominating node takes up the category of its head.

In this system, as in Selkirk's, we may assume that affixes are subcategorised for their sisters. Thus, un- and -ic have the following subcategorisation:

(130) a. un- (——— A^0)

b. -ic (N^0———)

This theory allows us to incorporate affixes into the general X bar system. Note that affixes behave like heads in at least two respects. They subcategorise for their sisters, and they determine the categorial features of their mothers. As pointed out earlier, both these properties follow from the X bar theory if affixes are X^{-1}. On the other hand, if words are X^0, and roots are X^{-1}, then roots should be the heads of words, and we would expect roots to subcategorise for affixes rather than vice versa, and for the roots to determine the category of the dominating node. In Selkirk, affixes are special entities that lie outside the X bar. Consequently, the properties of affixes with respect to their sisters and mothers become mere stipulations.

Another advantage of the theory implied in (129) and (130) is that it allows us to specify subcategorisation without violating the assumption that subcategorisations are defined on maximal projections (cf: Emonds (1976)). If we make the perfectly natural assumption

that X^0 is the maximal projection at the level of word formation,
we can say that the condition of the maximality of projection on
subcategorisation applies in word formation as well. In Selkirk's
system, -ity is subcategorised for A^{-1} , and -ness for A^0, and
hence, no such restriction of maximality of projection can be made
on subcategorisation.

Thus, Selkirk's use of the X bar theory in word formation
does not crucially derive any explanatory power from the theory,
such as those involving the notions of head and feature percolation
from the head. The revised system suggested above, on the other hand,
derives natural consequences from the theory for head, feature per-
colation, and subcategorisation. Now, a crucial feature of this
alternative is that the X bar rules do not distinguish between class
I and class II derivation (or between subcompounds and cocompounds)
(e.g. (129)). In order to make the relevant distinction, it would be
necessary in this system to supplement it with stratum ordering.

Thus, Selkirk argues that the use of the X bar theory in mor-
phology eliminates the need for ordered strata, while we conclude
that the maximal exploitation of the X bar theory in word formation
requires the stratum theory to complement it by providing the rele-
vant domain information.

One may still think of replacing the stratum theory with a
node-domain theory, without demanding that the node labels be derived

from the X bar system. Such a theory can be constructed by assigning a numerical index corresponding to a stratum to each root node created at the stratum. One may, for example, p ropose the following grammar of word formation 'for Malayalam:

(131) a. $X_1 \rightarrow X_1$ affix   (affix X_1)   (derivation)

　　　b. $X_2 \rightarrow X_1$

　　　c. $X_2 \rightarrow X_2 \ X_2$ (subcompounding)

　　　d. $X_3 \rightarrow X_2$

　　　e. $X_2 \rightarrow X_3$ (the loop)

　　　f. $X_3 \rightarrow X_3^*$ (cocompounding)

　　　g. $X_4 \rightarrow X_3$

　　　h. $X_4 \rightarrow X_4$ affix   (inflection)

The grammar in (131) yields representations like the following:

(132)   a.   paśukkuṭṭikaḷ 'calves'   b. acchanammamaar 'parents'

paśu kutti kaḷ
cow child pl.
(cf: (67a))

acchan amma maar
father mother pl.
(cf: (67b))
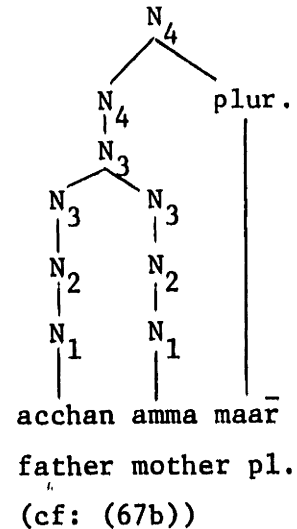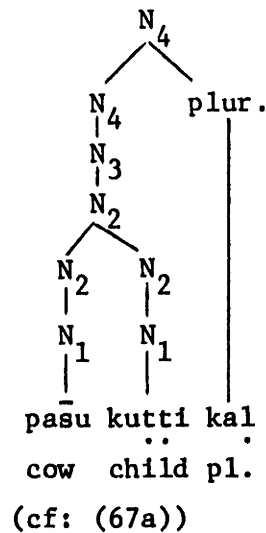
paśukkuṭṭi is a subcompound, acchanamma is a cocompound. We can now specify the domain of gemination as X_2, nasal deletion and

vowel lengthening as X_2 and X_3, and so on. Presented in this fashion, the node-domain theory is a notational variant of the stratum theory, except for one crucial difference. Given the stratum theory, it is imperative that the phonological rules associated with a stratum apply at that stratum in the lexicon, as the stratum information of a construction is not available at subsequent strata (2.2.1.). On the other hand, (131) preserves the stratum information in terms of an index, and therefore, in this theory, it is possible to postpone the application of phonological rules, as in SPE, to the output of morphological and subsequent syntactic operations.

The application of phonological rules in the lexicon in the stratum theory yields the level of lexical representation, which is the output of all the lexical operations. In Chapter 3, I shall show that this level of phonological representation makes important empirical predictions, which are not available to the node-domain theory, in which phonological rules do not apply in the lexicon.

Footnotes for Chapter II

1. Malayalam is a Dravidian language, with massive borrowings from
   Sanskrit. See Appendix 2 for a list of the Dravidian and Sanskrit
   stems used in this thesis.

2. The notation of subcompounds as [[...]$_2$[...]], and of cocompounds
   as [[...]$_3$[...]] is mnemonically related to their stratum struc-
   ture (to be discussed), and have no theoretical significance. I am
   grateful to Douglas Pulleyblank for suggesting this notation.

3. What rule (7) does is to create a branching onset. I assume that
   the t in (3a) and the k in (5b) are onsets. (See 4.1.3.).
   I follow Halle & Vergnaud (1980) in assuming that
   is phonetically realised as [tt], as opposed to , which is
   phonetically [t].

        Rule (7) uses the mirror image notation (%), collapsing (i)
   and (ii), which are disjunctively ordered:
        (i)   onset → Λ / —— ##
        (ii) onset → Λ / ## ——
        I should mention that the rule does not apply to sonorants:
        (i) tii pori → tiippori (tii 'flame'; pori 'fragment')
        (ii) tii naalam → tiinaalam (naalam 'tongue'(?))

4. Gemination does not apply because the stems are not Dravidian,
   but derived from Sanskrit.

5. One may ask why the V~VV alternation cannot be handled in terms
   of a rule that shortens underlying long vowels in the word final
   position instead. The answer is that this move leads to complica-
   tions, as vowels are long in compounds only when the second stem
   is of Sanskrit origin. Compare:

        (i) [[bhaarya][wiita] → bhaaryawiita (wiita 'house' [+Drav])
        (ii) [[bhaarya][grham] → bhaaryaagraham (grham 'house'[+Sans])

6. All inflectional endings in Malayalam are of Dravidian origin, and
   vowel sandhi, like vowel lengthening, applies only when the follow-
   ing element is of Sanskrit origin. Thus, in Dravidian compounds
   like the following, vowel sandhi does not apply, and as a conse-
   quence, a glide is inserted between the two adjacent rimes:
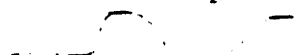
        a. kari ## ila   kariyila (kari 'coal'; ila 'leaf':'dry leaf')
        b. mala ## ooram   malayooram (mala 'hill';ooram 'side':'hill
                                                                 side'
        Given these facts, the nonapplicability of the rule in inflec-
   tional endings may be attributed to the absence of [+Sans.] on in-
   flections.
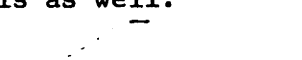
7. Nothing crucially hinges on the analysis proposed here. All that is important for us is the fact that the primary stressed vowel has a low tone, while the last vowel of the word, as well as all remaining long vowels, have high tones.

Observe that the remaining short vowels are unspecified for tone. Thus, the pitch track of paraati shows a low F_0 contour on the first syllable when utterance initial , but a transition from high to low when preceded by another word:

paraati   parannu   vs   kutti paraati parannu
complaint said   child complaint said

The same F_0 contour is observed in other unspecified short vowels as well:

maratakam   vs   paaraayanam

What this means is that the tone bearing unit in Malayalam is the foot, not the syllable. This is characteristic of intonational phenomena, as opposed to tonal phenomena.

It must be noted that final short vowels, unlike initial primary stressed vowels, undergo vowel reduction leading to the disappearance of the vowel in appropriate segmental environments (e.g. raamadaasan $\rightarrow$ raamdaasn). Therefore, the $\underset{F}{\text{v}}$ in (55) is not to be interpreted as secondary stress, but only as marking the location for tone spreading.

8. **acchan** appears to be an exception to stem final nasal deletion.

```
      L  H              L  H                    L  H
9. cf: kutti 'child';  kuttiye 'child-acc.';  kuttiyute 'child-gen.'
   L  H
kuttiyoota 'child-dat.$_2$'.
```

10. The disjunction can be avoided by stipulating a hierarchy of boundary strength such as + ## ## # ##. Rules which cannot apply across ## may also be assumed not to apply across the stronger boundaries # and ##.

11. The raised [$^{\partial}$] in $^{\partial}$V represents a centralised onglide of the following vowel, as opposed to the ∂ in, say, kaata, which is a full fledged central vowel.

12. (78) and (80) cannot be collapsed: cf: sampradaayam $\rightarrow$ sam[b]radaayam 'habit', but aprakaaram $\rightarrow$ *a[b]rakaaram 'likewise'. Also, arkan $\rightarrow$ *ar[g]an 'sun.

13. On the basis of the fact that class II derivation and compounding can be inputs to each other, Selkirk concludes that they belong to the same domain. Selkirk, like Allen, assumes a strict ordering hypothesis which can be stated as: if processes P_i and P_j are inputs to each other, then they define the same morphological domain (level in Allen). In contrast, I assume the looped ordering hypothesis which allows P_i and P_j to be inputs to each other and yet be associated with different domains.

14. This approach to stress is due to Siegel (1974).

15. It is not unlikely that this principle will be applicable in other areas as well (e.g., vowel harmony which proceeds from the stem to the affix).

16. All the Dakota facts are from C&S.

17. For further examples of such phonological processes, see C&S.

18. The rules DAR and CAR can be formulated in a simpler fashion in terms of the metrical framework, but I continue to use the segmental notation so as not to obscure the basic issues involved in the comparison between standard and lexical treatments.

Chapter III :   THE LEXICAL REPRESENTATION

In Chapter II, I demonstrated the advantages of allowing a direct interaction between morphology and phonology, and assuming that this interaction is supervised by ordered strata. The focus of Chapter III is the levels of phonological representation resulting from these assumptions, in particular the level of representation between the underlying and the phonetic.

The principle of specifying the domains of phonological rules in terms of ordered strata yields a theory in which some rule applications take place at various lexical strata, and others outside the lexicon, at the post lexical stratum. In other words, some rule applications take place before lexical insertion, and others after it. I shall argue in this chapter that the stage in the derivation which is the output of the lexicon and the input to lexical insertion is not simply an accidental stage, but is an important level of linguistic representation at which various interesting phonological properties converge. I shall refer to this level as the level of lexical representation.

In 3.1., I shall make explicit the conceptual framework assumed in Chapter II, by answering questions like: what is the nature of the lexicon in Lexical Phonology? What are the levels of phonological representation assumed in this theory, and how are they related to the organisation of the lexicon?

In 3.2., I shall concentrate on lexical rule applications. The issues addressed will be: what are the criteria that separate lexical and post lexical rule applications? What are the predictions made by the assumption that lexical rule applications take place after every morphological operation?

In 3.3., which is the core of this chapter, and of the thesis as well, I identify the lexical representation as a significant level of representation in phonology, and demonstrate its consequences for a wide variety of phenomena such as speaker judgments, effect of pauses on phonological rules, secret code languages, speech errors, acquisition of phonology, speech recognition, and speech production. This section presents a brief glimpse of the potential application of Lexical Phonology for the better understanding of the phonological knowledge that human beings have, and the way they make use of this knowledge.

3.4. compares the level of lexical representation in this theory with two theories that postulate what looks like similar levels, namely, Taxonomic Phonemics and Natural Phonology, with the view of bringing out their empirical differences.

3.1. Underlying, Phonetic, and Lexical Representations

Following the lexicalist trend set by Chomsky (1970), I shall assume that lexical entries are entries of words, not morphemes. The list of unanalysable morphemes in the language constitute what I shall

refer to as the set of morphemic entries. (For the lexicalist approaches to morphology, see Zimmer (1964), Halle (1973), Aronoff (1976), Allen (1978), Siegel (1974), and Selkirk (to appear).) The word formation component which is conceived of as containing lexical redundancy rules maps the set of morphemic entries onto the set of lexical entries.[1]

(1) Morphemes → | Word Formation | → Words
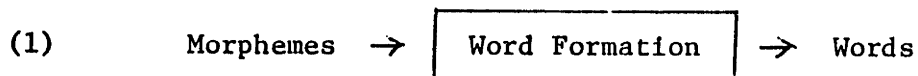
Given the set of morphemes of the language, word formation predicts the potential words of the language, a subset of which constitutes the set of actual words of the language. The component of word formation tells us about the internal structure of words, in terms of the morphemes that constitute them, and the processes that have applied to them. Thus, a morphemic entry specifies the syntactic, semantic, and phonological properties of the morpheme; a lexical entry specifies the syntactic, semantic, and phonological properties of the word. What the component of word formation does is to predict the latter on the basis of the former.

To take an example, the morpheme strong will be listed as an adjective (syntactic property), and the morpheme -ly will be specified as one that makes adjectives adverbs. From these two entries, one may predict that strongly is an adverb. Similarly, the meanings of employer and employee (semantic property) can be predicted from the specifications on the morphemes employ, -er, and -ee: -er picks out

the subject of the verb as its meaning (an <u>employer</u> is an X that employs some Y), and -<u>ee</u> picks out the object (an <u>employee</u> is a Y such that some X employs Y). One can also predict that <u>child</u> <u>employer</u> is ambiguous ('a child who is an employer', or 'one who employs children'), but not <u>child employee</u> ('a child who is an employee', but not *'one whom children employ').[2]

Just as the word formation component predicts the semantic and syntactic properties of words on the basis of the semantic and syntactic properties of their constituent morphemes, it also predicts the phonological prope rties of words on the basis of the phonological properties of their constituent morphemes. Thus, given that the phonological ent ries of the morphemes <u>divine</u> and -<u>ity</u> are /dīvīn/ and /iti/ respectively, we predict that the phonological entry for the word <u>divinity</u> is /dívíniti/. This makes rules like trisyllabic laxing and rules of word stress in English part of the component of word formation.

I use the term Underlying Rep resentation to refer to the phonological representation of morphemes in the set of lexical entries. The phonological representation of words in the set of lex- ical entries is the Lexical Representation. The term Phonetic Repre- sentation is used in the standard sense, to refer to the output of all the phonological rules, given in some universal notation. The diagram given below relates the three levels of representation to the model of grammar:

(2)   exicon

```
+-----------------------------------------+
| morphemes . . . . . . . . .|. Underlying representations
|     |                                   |
|     v                                   |
|  +-----------------------------+        |
|  | lexical rule applications   |        |
|  |     a. syntactic            |        |
|  |     b. semantic             |        |
|  |     c. phonological         |        |
|  +-----------------------------+        |
|     |                                   |
|     v                                   |
| words . . . . . . . . . .|. Lexical Representations
+-----------------------------------------+
      |
      v
 post lexical rule applications
      |
      v
 phrases . . . . . . . . . . Phonetic Representations
```
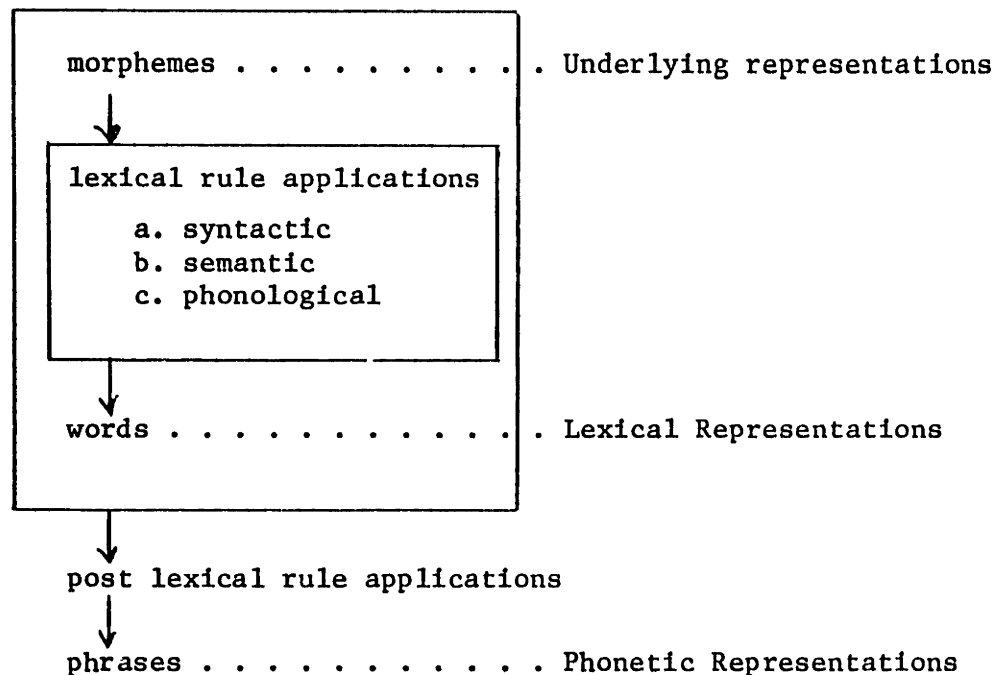
To summarise: it has been recognised in linguistic theory
since Chomsky (1970) that the principles responsible for putting
morphemes together to form words are to be distinguished from the
principles that put words together to form sentences. Most current
theories have accepted this assumption with respect to both the struc-
ture of words and the meanings of words. What Lexical Phonology does
is to complete the set, and say that the principles governing the
structure, meanings, as well as the phonology of words are to be
distinguished from the principles governing the structure, meanings,
and phonology of sentences. My proposal, therefore, is a logical
consequence of the developments in linguistics triggered by Chomsky's
Remarks on Nominalisation.

A note of explanation on the use of the term 'underlying

representation' is in order he ɛ. It is possible to think of a

major difference between the model of phonology proposed in SPE and

the model that I propose as being the addition of an intermediate

level of representation in the latter:

(3)a. SPE

```
┌──────────────────┐      ┌──────────────────┐
│ underlying        │─→ phonological →│ phonetic          │
│ representation     │      rules      │ representation    │
└──────────────────┘      └──────────────────┘
```

b. Lexical Phonology

```
┌──────────────────┐                    ┌──────────────────┐
│ underlying        │                    │ phonetic          │
│ representation     │                    │ representation    │
└──────────────────┘                    └──────────────────┘
         ↓                                        ↑
   phonological  ──→ ┌───────────────┐ ─→ phonological
      rules          │ lexical        │      rules
                     │ representation │
                     └───────────────┘
```
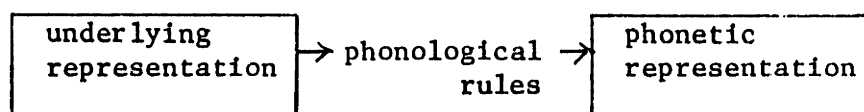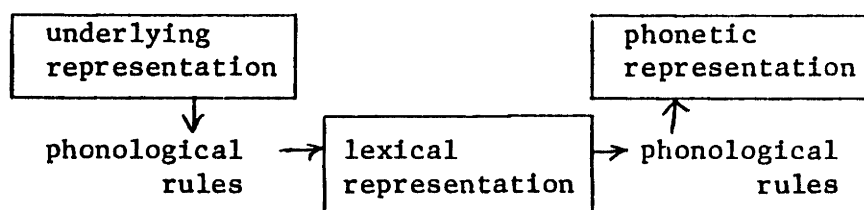
My use of the term 'underlying representation' is not, how-

ever, identical to the SPE use. This can be illustrated with words

like presidential, the SPE underlying representation of which is

/prezīd+ent+y+al/. Consider now the derivation of the word in the

lexical theory (the derivation ignores stress):

(4)   [prezīd]   underlying

          [[prezīd] ent]   suffixation

          [[prezid] ent]   laxing

          [[[prezid] ent] y]   suffixation

          [[[prezid] ens] y]   t → s

          [[[[prezid] ens] y] al]   suffixation

          [[[[prezid] enš] y] al]   s → š

          [[[prezid] enš] al]   y → ∅

                         •
                         •
                         •

I shall discuss at a later point why rules like trisyllabic

laxing, t → s, etc. should be assumed to be lexical rules applying

after every morphological operation. For the present, observe that

the suffix _-y_ is attached not to [[prezīd]ent], but to [[prezid]ent],

which is the output of laxing. Therefore, representations like

[[[[prezīd]ent]y]al] do not exist in Lexical Phonology. Underlying

representations are simply representations of morphemes, and lexical

representations are representations of words, derived through lexical

rules from the representations of morphemes.

3.2. Lexical and Post Lexical Rule Applications

3.2.1. Rules with lexical and Post Lexical Domains

The stratum theory I have developed so far allows for the

application of the same rule in the lexical as well as post lexical

strata. Rules of this kind do exist in natural languages. An example

of such a rule is the rhythm rule in English, first proposed in

Liberman & Prince (1976). The rhythm rule relabels a tree of the form

ŵs s as a tree of the form s w s. It applies at the phrasal level,
　3　2　1　　　　　　2　3　1

changing <u>thirteen men</u> to <u>thirteen men</u>. As shown by Kiparsky (1979),

it also applies at the word level as a cyclic rule, in words like
　　　　　　　　　3　1　　　　3　4　1

<u>artificiality</u> (cf: artificial vs artificiality), <u>anesthesiology</u>,

<u>interchangeability</u>, etc.

The vowel epenthesis in the past tense, plural, etc. is

another rule with lexical and post lexical domains:

$$(5) \quad \emptyset \rightarrow \vartheta \;/\; \begin{bmatrix} +\text{coronal} \\ \alpha\text{strident} \end{bmatrix} \underline{\quad\quad}] \begin{bmatrix} +\text{coronal} \\ \alpha\text{strident} \end{bmatrix}$$

Since the rule has to apply to the past tense ([[pæt]d] →

pætəd <u>patted</u>), present tense ([[wiš]z] → wišəz <u>wishes</u>), and

plural ([[hors]z] → horsəz <u>horses</u>), the inflectional stratum must

be part of the domain of (5). Since the rule also applies in possess-

ives ([[mij]z] → mijəz <u>Midge's</u>), and auxiliary contraction

([[mij]z]gowiŋ → mijəz gowiŋ <u>Midge's going</u>), it must apply at

the post lexical stratum as well.[3]

3.2.2. Criteria for Identification

I shall now try to make explicit the criteria for distinguish-

ing lexical and post lexical rule applications. An obvious property

of lexical rule applications is that they have to be word internal,

as the output of the lexicon is the word, not the phrase. Therefore, no process that applies across words can be a lexical process. An immediate consequence is that external sandhi processes become post lexical operations. To give an example from English, the process that changes s to š in thi[š] year must be post lexical, as the sequence this + year is not available in the lexicon.

The second property of lexical rule applications hinges on their access to lexical information. If a rule is sensitive to word internal morphological information, it must apply in the lexicon. Recall that the Opacity Principle makes the internal structure at one stratum invisible at another stratum. This would mean that the internal structure at any of the lexical strata will be invisible to post lexical processes. It follows that post lexical processes do not have access to word internal structure, and that any rule that is sensitive to word internal structure must apply in the lexicon. We can now identify rules in English like trisyllabic laxing, velar softening, t → s, etc., as rules applying in the lexicon. These rules apply in class I derivations (stratum 1), but not, say, in compounds (stratum 3).

(6) Trisyllabic laxing

line drawing → * l[i]n drawing (cf; l[i]near)

(7) Velar softening

magic eye → * magi[s]eye (cf: magi[s]ian)

(8) t → s

resident employee → * residen[s] employee

(cf: residen[s]y)

In order to guarantee that the rule applications in the environments given above are prevented, these rules must apply at stratum 1, and thereby, in the lexicon. Another example of a lexical rule is the compound stress rule, which applies only in compounds (stratum 3), and not across words.

In sum, we arrive at the following principles:

(9) Lexical processes cannot apply across words.

(10) Post lexical operations are blind to the internal structure of words.

It seems to me that there is a correlation between sensitivity to word internal structure and rule exceptionality. Rules which have lexical exceptions are also rules which are sensitive to word internal structure (cf: Pesetsky (1979)). If so, we may make the following stipulation:

(11) Post lexical operations are exceptionless.[4]

3.2.3. The Issue of Cyclicity

All the rules discussed so far have been shown to apply

cyclically in the lexical domain. The cyclic application follows from the model in which phonological rules apply after every morphological operation (2.1.4.). The prediction that this model makes can be stated as follows:

(12) All lexical rule applications are cyclic.

From the conjunct of (10) and (12), we derive the following principle:

(13) All rules sensitive to word internal morphological information are cyclic rules.

Current approaches to the cyclic application of phonological rules hold that cyclic rules do not apply in the first cycle (Kean (1974) , Mascaro (1976), Halle (1978)). Halle (1978), for example, formulates the principle of the cycle as follows:

(14) "A cyclic rule R applies properly on cycle j only if either a) or b) is satisfied:

   a) R makes specific use of information, part of which is available on a prior pass through the cyclic rules, and part of which becomes first available on cycle j. There are three separate cases subsumed under a):

   R refers specifically to some A and B in:

   i) $[_j$ XAY $...[_{j-1} ...B...] $ Z$]$;
   ii) $[_j$ Z $[_{j-1}...B...]$XAY$]$;
   iii) $[_j$ X $[_{j-1} ...A...]$ Y $[_{j-1}...B...]$ Z$]$.

   b) R makes specific use of information assigned on cycle j by a rule applying before R."

The effect of (14) is to disallow the application of a cyclic rule to an underlying string exhaustively contained in the innermost brackets. It has been argued that certain kinds of phenomena which were previously assumed to be irregular are in fact explained by this view of the cycle (Mascaro (1976), Rubach (1981), Kiparsky (class lectures)). An illustration can be given from Rubach (1981), who gives a detailed analysis of Polish phonology in favour of the cyclic principle cited above. Polish has a rule that tenses high vowels when followed by the Derived Imperfective aj (Yer tensing), and a rule that lowers a high vowel when followed by r (r lowering). Rubach gives the rules as follows:

(15) Yer tensing

$$\begin{bmatrix} +syl \\ +high \\ -tense \end{bmatrix} \rightarrow [+tense] \ / \ \text{---} \ C_0 aj \quad \text{DI}$$

(16) r lowering

$$\begin{bmatrix} +syl \\ +high \\ +tense \\ -round \end{bmatrix} \rightarrow \begin{bmatrix} -high \\ -back \end{bmatrix} \ / \ \text{---} \ r$$

Rubach assumes that these are cyclic rules, and gives the following derivation:

(17) umieraj 'die'

                [u[[mĭr]aj]]

cycle 2	mĭr + aj	
	mir + aj	yer tensing
	mer + aj	r lowering
	mjer+ aj	lab/vel. j insertion

cycle 3	u + mjer + aj	
	no rule applies	

post cyclic	u + m'jer + aj	surface Pal. rule

Observe that the environment for r lowering is strictly con-
tained in the first cycle [mir], and yet the rule applies in the
second cycle, as the environmemt for r lowering is made available by
the application of yer tensing: mir is not an underlying string.
In contrast, r lowering does not take place in words like wirować
'rotate', and wiraż 'turning'. Rubach points out that the underlying
sequence ir is properly contained in the first cycle, and therefore,
the principle of cyclic application in (14) correctly predicts that
r lowering will not apply to these cases. If r lowering were assumed
to be a noncyclic rule, on the other hand, these forms will have to
be listed as exceptions.

Similar accounts have been offered for English as well.
Rubach (1980), for example, argues that (14) coupled with the assump-

tion that trisyllabic laxing is a cyclic rule, will account for the
fact that it does not apply to monomorphemic words like nightingale
(cf: Kiparsky (1974)). /nītingǣl / is an underlying string contained
by the first cycle, and therefore, no cyclic rule can apply to it.

If the formulation of the cycle is as given in (14), (13)
predicts that no lexical rule applications can take place in the
first cycle. There are some problems that this prediction creates.
Consider, for example, the rule of velar softening in SPE:

(18) Velar Softening

$$k \rightarrow s / \text{——} \begin{bmatrix} V \\ -low \\ -back \end{bmatrix} \quad \text{(irrelevant details ignored)}$$

The rule accounts for alternations like criti[k] ~criti[s]ise,
and electri[k] ~electri[s]ity. SPE assumes that the rule applies in
words like receive ([re[kēv]]) as well. The voiceless s in these
words is accounted for by ordering the voicing of s before velar
softening.[5]

(19) [re [kēv]]   [re [sume]]

      ————   [re [zume]]   voicing

      [re [sēv]]   ————   velar softening

Under this account, velar softening applies to an underlying
string contained by the innermost brackets. Therefore, by (13) and (14),
it cannot be a lexical rule. However, contradictorily, velar softening
must be a lexical rule , as it does not apply at the compounding

stratum.

Velar softening is not the only cyclic rule that applies to an underlying string contained by the first cycle. Consider the SPE acco account of the ŋ~ŋg alternations in <u>long</u> [lɔŋ], <u>longer</u> [lɔŋgər], and <u>longing</u> [lɔŋiŋ] in terms of a rule that changes n to ŋ before a vela r consonant, followed by a rule that deletes the g in <u>long</u> and <u>longing</u>, but not in <u>longer</u>. I shall adopt the SPE solution and restate the rules as follows:

$$\left.\begin{array}{l} \text{(20) a. } n \rightarrow \text{ŋ/ } \underline{\quad} \text{ g,k} \\[6pt] \text{b. } g \rightarrow \emptyset / \text{ [+nasal] } \underline{\quad} \text{]} \end{array}\right\} \text{ domain : stratum 2}$$

The derivations are given in (21):

(21)	<u>long</u>	<u>longer</u>	<u>longing</u>	
S_1	[long]	[long]	[long]	underlying
	_____	[[long]er]	_____	affixation
S_2	_____	_____	[[long]ing]	affixation
	[loŋg]	[loŋger]	[[loŋg]iŋg]	(20a)
	[loŋ]	_____	[[loŋ]iŋ]	(20b)

(20b) does not apply to <u>longer</u> because the bracketing [[long]er], by the Opacity Principle, is invisible at stratum 2. Now, it is crucial for this account of the ŋ~ŋg alternation that rule (20b) applies at stratum 2: if the domain of the rule were, say, stratum 3 or 4, the bracketing in [[long]ing] would be invisible

to the rule, and the deletion would fail to apply in <u>longing</u> as well.
Since (20b) must have stratum 2 as its domain, (20b) and the rule
(20a) are lexical rules, and by (13), both are cyclic rules. Now,
observe that (20)a and b apply to an underlying string contained in
the innermost cycle, and are therefore, disallowed by (14).

In order to account for the fact that velar softening applies
to k̄ēv in <u>receive</u>, Rubach assumes that it is a post cyclic rule. Pre-
sumably, a similar solution could be offered to the problem of g-
deletion as well. Given the principle that rules which require word
internal information must be cyclic rules, Rubach's move is not open
to us.

The right solution lies in abandoning (14). Kiparsky (in prepa-
ration), for example, derives the desirable effects of (14) from the
Elsewhere Condition (cf: Kiparsky (1973)). He proposes that all un-
derived lexical entries are identity rules of the form L → L, e.g.

(22) a. <u>nightingale</u> : nītingǣl → nītingǣl

     b. <u>divine</u>   : dīvīn → dīvīn

Given the identity rule approach, (22a) and trisyllabic laxing
are both rules which are applicable to the /ī/ in <u>nightingale</u>, laxing
being the more general of the two. By the Elsewhere Condition, (22a)
applies, and laxing is inapplicable. In <u>divinity</u>, on the other hand,
laxing applies, as there is no identity rule of the form dīvīniti →
<u>dīvīniti</u>.

The identity rule approach to lexical entries correctly predicts that velar softening can apply in receive. -ceive not being a word, there is no identity rule of the form k̄ev → k̄ev, and therefore, nothing prevents the rule from changing k to s in [re[k̄ev]]. The same explanation holds for g deletion as well. Since there is no identity rule of the form long → long, n → ŋ and g deletion are free to apply to long.[6]

3.3. The Lexical Representation

The purpose of this section is to show that the intermediate level of representation that I have been calling Lexical Representation, constitutes a well defined level of linguistic representation. Given the stratum theory, Lexical Representation is the output of the lexicon, and the input to lexical insertion. I shall show that it has a number of other interesting properties as well: it is the input to the assignment of pauses, and to the application of the code in secret code languages, and plays a significant role in speaker judgments, speech errors, language acquisition, etc.

3.3.1. The Effect of Pauses on Phonological Rules

3.3.1.1. It was Rotenberg (1978) who drew the attention of phonologists to the fact that some phonological rules are blocked by intervening pauses, while others are not. As an example of a rule of the

first kind, consider the rule of intervocalic voicing of stops in

Malayalam (2.2.2.):

(23) a. kutti ... (pause)... [k]arannu
        child   cried

     b. kutti[g]arannu

     c. *kutti... (pause)... [g]arannu

The voicing rule does not apply across words if there is an
intervening pause; the same phenomenon is observed word internally
as well. Though speakers do not normally pause word internally, they
do so in careful and slow speech, as when they pause after every
syllable. Thus, in answer to the question, "How many syllables does
population have?", a perfectly natural answer is: "Four. Po..pu..la
..tion." [7] The effect of word internal pauses on the voicing of
stops in Malayalam is illustrated below:

(24) a. a...[p]a...[k]a...[t]am 'danger'

     b. a[b]a[g]a[d]am

     c. *a...[b]a...[g]a...[d]am

We conclude that intervocalic voicing in Malayalam cannot
apply across pauses. In contrast to voicing, the rule of gemination
in subcompounds (2.1.1.) is not blocked by pauses:

(25) a. pasukkutti 'calf' (from pasu 'cow', and kutti 'child')

b. pa...su...kku...tti

c. *pa...su... ku...tti

In order to explain the contrast, all that we need to do is to make the perfectly natural assumption that pauses are assigned not in the lexicon, but after words are put together, i.e., after lexical insertion:

(26)   Syntax   Lexicon

```
   ┌────────────┐      ┌─────────────────┐
   │            │      │  underlying     │
   │            │      │  representation │
   └────────────┘      └─────────────────┘
         ↓                     ↓
   syntactic             lexical
   representation        representation
            ＼          ／
        lexical insertion
        pause assignment
              │
              ↓
        post lexical
        phonological rule
        applications
              └──────→ phonetic
                        representation
```
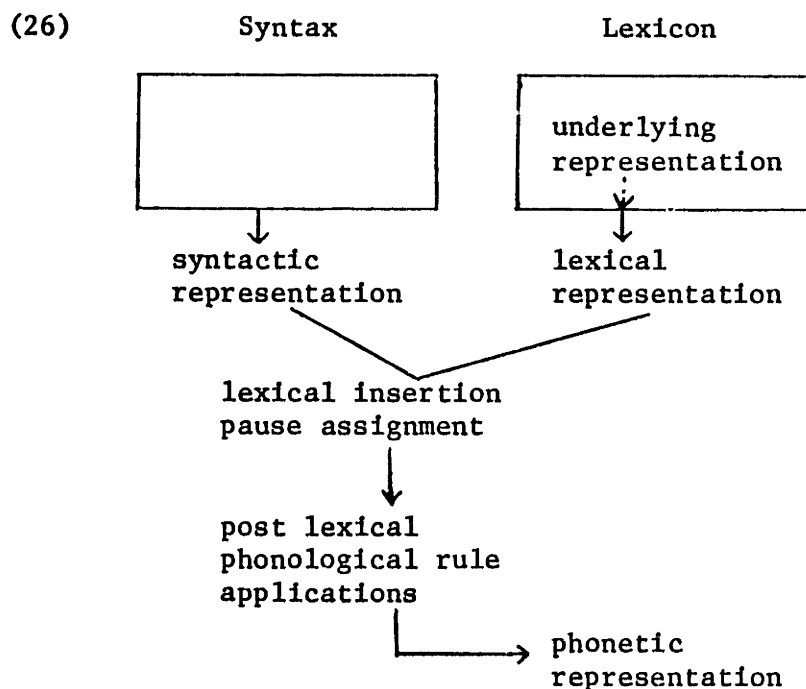
The assignment of pauses creates units which have traditionally been called 'phonological phrases', which are phonological strings bounded by pauses. Just as the rules in the lexicon take place in words, rules outside the lexicon take place in phonological phrases.

Intervocalic voicing applies at the post lexical domain (i.e. after lexical insertion)(2.2.2.), while gemination applies in the lexicon (i.e., before lexical insertion). Each of the syllables <u>a</u>, <u>pa</u>, <u>ka</u>, <u>tam</u> in (24) is a phonological phrase by itself. The environment for voicing (V —— V) is not satisfied by any of the consonants in (24a), and therefore, the rule does not apply. In (25), on the other hand, the rule of gemination applies prior to pause assignment, and therefore, the insertion of the pause after the rule has applied does not affect the rule.

This contrast provides a major confirmation for lexical representation as a significant level of representation in phonological theory. The following properties identify this level:

> (27) a. output of lexical rule applications;
> b. input to lexical insertion;
> c. input to pause assignment, which in turn yields
> d. input to post lexical rule applications.

The model in (26) makes the powerful prediction that no rule which is sensitive to word internal structure is blocked by the presence of an intervening pause. This prediction, to the best of my knowledge, is correct. The following examples from Malayalam show that the lexical rules of nasal deletion (2.1.1.2.), vowel lengthening (2.1.1.3.), and vowel sandhi (2.1.1.4.) are not blocked by pauses:

(28) Nasal Deletion

    a. ma...r̄a...kku...t̲i...r̄a 'wooden horse' (mar̄am, kut̲ir̄a)

    b.*ma..r̄am..(k)ku..t̲i..r̄a (*ma..r̄a..m(k)ku..t̲i..r̄a)

(29) Vowel Lengthening

    a. t̲aa..ˌr̄aa..kaa..n̲tan 'Tara's husband' (taar̄a, kaan̲tan)

    b.*t̲aa..r̄a..kaa..n̲tan

(30) Vowel Sandhi

    a. ma..hee..s̄wa..r̄an 'great god' (mahaa, iis̄war̄an)

    b.*ma..haa..ii..s̄wa..r̄an

In English, lexical rules like trisyllabic laxing and CiV tens-
ing (SPE: 180-81) are blind to pauses:

(31) Trisyllabic Laxing

    a. di..v[i]..ni..ty

    b.*di..v[āy]..ni..ty

(32) CiV Tensing

    a. ma..na..g[īy].. ri..al

    b.*ma..na..g[e]..ri..al

3.3.1.2. Recall that Lexical Phonology allows the same rule to have

both the lexical and post lexical strata as its domain. Given (26), we

would predict the possibility of rules being insensitive to pauses

when they apply in the lexicon, and being blocked by pauses when they

apply post lexically. An interesting example of such a rule is the r deletion in British English, responsible for alternations like the following:[8]

> (33) a. I know the actor.
>   [æktə]
>
>   b. The actor acted hastily.
>   [æktər]
>
>   c. The actor snored.
>   [æktə]

The rule responsible for this alternation can be stated as follows:

(34) r deletion[9]

$$r \rightarrow \emptyset \;/\; \underset{\rule{1em}{0.4pt}}{\overset{R}{|}}$$

Given below are the derivations for (33)a-c:

(35) a. actor   b. actor, acted ... c. actor snored

(phonological tree diagrams)

Resyllabification applies in (35b), linking r to the onset, and saving it from r deletion. No resyllabification is possible in (35) a and c, and hence r is deleted.

Consider now what happens when a pause is inserted:

(36)   The actor ... acted hastily.
     [æktə]
    *[æktər]

Since resyllabification takes place across words, it must be post lexical, and applies after pause assignment:

(37)   The actor ... acted hastily.

                       resyllabification

                       r deletion

The deletion of r applies word internally as well:

(38) a.   bareness   [bɛənəs]

    b.   barest   [bɛərəst][10]

Unlike resyllabification across words, however, word internal resyllabification is blind to pauses, as predicted by (26):

(39) a.   bɛə... rəst

    b.   *bɛə... əst

What is important for us is the fact that a pause makes the r disappear in _the bear is_, but there is no way a pause can make the r in _the barest_ disappear. Resyllabification in the former follows lexical insertion (and hence pause assignment), while in the latter,

precedes it, and therefore, there is nothing that we need to say about the contrast.

It may be pointed out that the contrast between <u>bear is</u> and <u>barest</u> supports the assumption that inflections are attached in the lexicon (cf: Lieber (1980), Selkirk (to appear)). If the attachment of inflections takes place after lexical insertion, the two forms become identical, and we no longer have any explanation for the contrast. Therefore, the facts of r deletion in British English offer crucial evidence against syntactic rules such as affix hopping, which involve attaching the inflectional element to the stem in the syntactic component after lexical insertion. (Note that the behaviour of verbal inflections is identical to <u>-est</u>: e.g. <u>storing</u>: [stɔ ... rɪŋ], *[stɔ ... ɪŋ].)

Yet another rule that is blocked post lexically by pauses, but not lexically, is the rule of palatalisation of s. Following Hayes (1980), I state the essentials of the rule as follows:

(40) Palatalisation

$$s \rightarrow \check{s}/ \underline{\quad\quad} y\check{V} \quad (\check{V} = \text{unstressed vowel})$$

The rule applies word internally, in words like <u>racial</u> (rǣs + yæl → rēyš̌əl), and in some dialects, across words, in sequences like <u>miss you</u> (miss you → mišyə). Speakers of English who do palatalise s across words cannot do so across a pause. Thus,

it is always [mis...yə], never *[miš...yə]. When the pause occurs

word internally, however, it has no effect on palatalisation: it

is always [rēy... šəl] (racial), never, *[ǣys.. əl]. The palata-

lisation in racial is a lexical operation; in miss you, it is post

lexical. Given that pauses are assigned post lexically, this is

precisely what we expect. The pause is assigned after palatalisation

in racial , but before it in miss you.

3.3.1.3. The behaviour of phonological rules with respect to pauses
                  an
provides⟨important empirical distinction between the node-domain

theory and the stratum theory. The essential character of the node-

domain theory is that the components of word formation and syntax

provide labelled trees, and phonological rules apply to these trees,

identifying the domains of rules from the node labels. Observe that

the node-domain theory has no intermediate level of representation

between the underlying and the phonetic: all rules apply after lexi-

cal insertion, at the output stage of the syntactic component. The

only way for accounting for the effects of pauses on phonological

rules in the node-domain theory, as far as I can see, is to stipulate

that the word and the subword domain rules are blind to pauses. Given

a series of node labels X^1, X^2, X^3... X^n, what this stipulation does

is to identify one of the nodes arbitrarily as marking the division bet-

ween pause blind and pause sensitive rules. This arbitrary point, in

principle, could have been, say, the noncompound word node, or an X

node.

The most important distinction between Lexical Phonology and
the node-domain theory is that Lexical Phonology provides an inter-
mediate level of representation between the underlying and the phone-
tic, while the node-domain theory does not. This intermediate level
of representation, as we shall see, provides the input to secret code
languages, to speech errors, and is the basis of speaker judgments.
In the node-domain theory, all these would be unrelated stipulations
on the output of the word domain rules.

3.3.1.4. In addition to predicting that lexical rule applications
are blind to pauses, the model in (26) also makes the prediction that
all post lexical rule applications will be blocked by intervening
pauses. This phenomenon has some apparent counterexamples which fall
into two classes. The first type of apparent exception is that of
the assignment of intonation, which appears to ignore pauses, and
even the segmental material of an appositive type. I give below
Rotenberg's (1978) examples:

(41) a. This is an example sentence.

b. This is... good lord!... an example sentence.

c. This is, as you can see, an example sentence.

If intonation assignment is a rule that applies to segmental

strings after pause assignment, (41) would constitute a counterexam-

ple to our claim that post lexical rule applications will be blocked

by pauses. However, there is no reason to believe that intonation is

handled by rules that operate on segmental material in this fashion

at all. It has been shown by Liberman (1975) and Pierrehumbert (1980)

that the correct approach to intonation is to have a module of grammar

which generates 'intonation words', or tunes. These tunes have interest-

ing internal structures, and are associated with semantic content,

such as uncertainty, surprise, etc. (see Pierrehumbert (1980) for de-

tails). The tunes generated in this manner independently of segmental

strings are then associated with the segments (via stress trees) in

predictable ways.

We may incorporate the Liberman-Pierrehumbert model into (26):

(42)   Intonation   Syntax   Lexicon

                                              underlying
                                              representation
                                                  ⋮
                                                  ↓
         ↓   ↓   ↓
      tunes   syntactic   lexical
                        representation   representation

                        lexical insertion
                        pause assignment

                               ↓

                        post lexical phonolo-
                        gical rule applications

                               |_____→ phonetic
                                                  representation
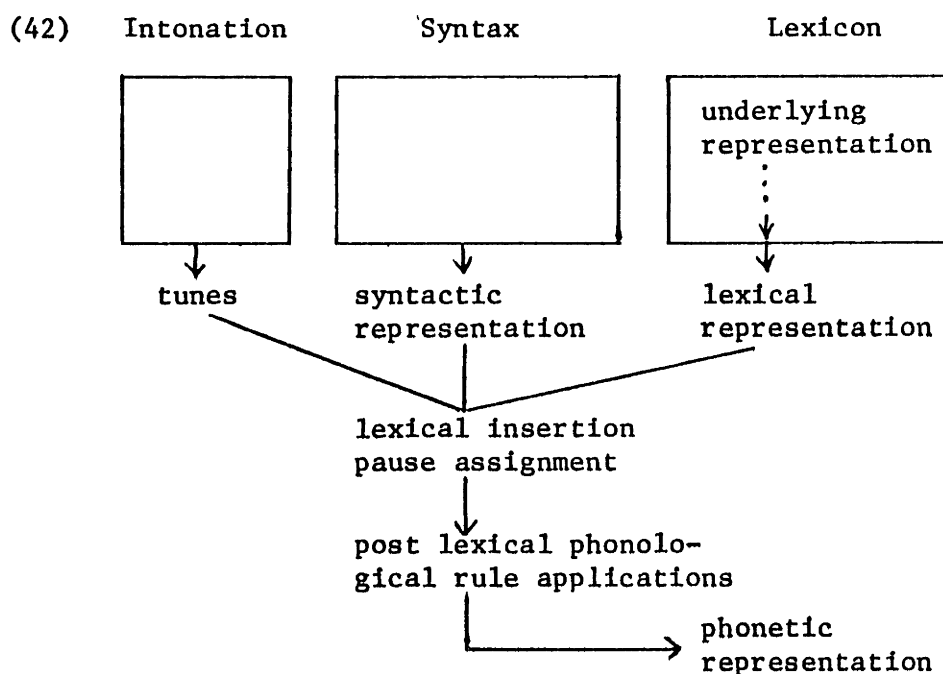
In this approach, lexical insertion is seen as a merging together of the outputs of various components: the lexicon, syntax, and intonation. In this model, cases like (41) cease to be counter-examples.

The second type of apparent exception is seen in words like a and the, which have alternate forms before vowels and consonants: a book vs an old book; th[ə] man vs th[i] old man. Rotenberg (1978) points out that alternations of this kind are blind to pauses:

(43) a. This is an old book.
     b. This is an ... (pause)... old book.
     c. This is a nice book.
     d. This is a ... (pause)... nice book.

If these alternations are handled by post lexical phonological rules, they falsify the claim that post lexical applications will be blocked by pauses. Note, however, that alternations like a~an are restricted to single lexical items. One way to look at them would be to say that a and an are different lexical items related by partial suppletion (cf: child~children) (or perhaps an irregular phonological rule) in the lexicon. Associated with a will be a filter which says that the following word is consonant initial, and the opposite condition will be associated with an. These filters apply during lexical insertion, ruling out *an book and *a old book. Subsequent assignment of pauses would then have no effect on a and an.

Seen in this light, examples like (43) do not present a problem for our model. A true counterexample to the claim that post lexical rule applications cannot cross pauses will be an alternation that is not restricted to a small set of lexical items. If we found, for example, that in some language, all word final obstruents get voiced before a vowel initial word, and that this process applies across pauses, our claim would be falsified. To the best of my knowledge, such languages do not exist.

3.3.2. Speaker Judgments

Given that linguistics is concerned with the properties of the human mind, the intuitions of speakers become data that linguistic theory must account for. Thus, it is not enough for the theory to generate (44)a and b, and block (44c), but also account for the fact that the speaker finds that John and him in (44a), but not in (44b), can refer to the same person:

> (44)   a. John believes that he is a fool.
>   b. John believes him to be a fool.
>   c.*John believes that him to be a fool.

Generative phonology has paid very little attention to speaker intuitions. Judgments on possible and impossible syllables and possible and impossible words have been acknowledged to constitute the data for phonology, but not speaker judgments about sameness and differentness of sounds.

Speakers of English judge the p in pit and spit to be the
same, ignoring the [p] ~ [ph] distinction; the vowels in bit and bid
are judged to be the same, ignoring durational and qualitative diffe-
rences. Obviously, these judgments do not correspond to phonetic
representations. At the other end of the spectrum, speakers judge
the initial vowels of metric and met to be the same, oblivious of
the underlying distinction: /ē/ ~ /e/. Similarly, the first consonant
in sing, the fourth in criticism, and the last in presidency are
judged to be the same, in spite of the underlying s ~ k ~ t distinction.
Underlyingly identical segments are judged to be distinct, such as
the initial vowels of meter and metric, or the second consonant in
race and racial.

Given a theory of phonology with just the underlying and
phonetic levels of representation, these judgments are unexplainable.
Given the theory of Lexical Phonology, what corresponds to the
speaker judgments is the lexical representation:

(45)   underlying   lexical   phonetic

    a. spit   spit   spit   spit
    b. pit   pit   p^hit   pit
    c. bet   bet   bet   bet
    d. bed   bed   be²d   bed
    e. mētr   míytər   míytər   metre
    f. mētr ik   metrik   metr̥ik   metric
    g. kritik ism   krítisism   k^hrítisizm   criticism
    h. prezīd ent y   prézidensi   p^h ̬ézidēnsi   presidency
    i. rǣs   reys   reys   race
    j. rǣs y æl   réyšᴧl   reyšᴧɫ   racial

Judgments of this kind also extend to the results of external

sandhi rules and fast speech phenomena. The speaker who says thi[š]year

thinks that he is saying thi[s]year (see Chapter I, footnote 4). Simi-

larly, speakers think that they are saying divíniti when they actually

say dvíni in fast speech. In short, we find that lexical representations

correspond to the mental representations of the speaker, reflecting

his intuitions of what constitutes his speech. Given that lexical

representations are the representations of lexical entries, this is

hardly surprising: what the speaker has conscious access to is the

set of lexical entries.

3.3.3. Secret Code Languages

A secret code language is one in which one or more rules

(the code) have been added to the grammar of a natural language so

that the result becomes unintelligible to a speaker who doesn't know

the code. A well known example is Pig Latin, in which the first onset

of a word (if any) is moved to the end, and the diphthong ey added to it.

Thus, the words clean (kliyn) and ask (aask) in Pig Latin would be

iynkley and aaskey. SPE gives the following rule for Pig Latin:

(46) $\#\#C_0VX\#\# \longrightarrow \#\#VXC_0ey\#\#$

We may, in fact, think of Pig Latin in terms of the follow-

ing equation:

(47) Pig Latin = Grammar of English + rule (46)

Consider what a secret code does. It creates new lexical items from existing ones so that an outsider does not identify the lexical items. Thus, Pig Latin picks up <u>clean</u> in English and changes it to <u>iynkley</u>:

(48)   Lexical entry$_{Eng.}$ $\rightarrow$ $\boxed{\text{secret code}}$ $\rightarrow$ Lexical entry$_{Pig\ L.}$

Seen in this light, a secret code is an operation on lexical representations. This result gives rise to interesting predictions. To begin with, observe that all secret codes of this kind operate on words, not on morphemes. Thus, rules like (49)a and b are logically conceivable, but nonexistent:

(49)   a. $\#C_0VX\#$ $\rightarrow$ $\#VXC_0ey\#$

b. $+C_0VX+$ $\rightarrow$ $+VXC_0ey+$

What this means is that secret codes are blind to the internal structure of words. Given the lexical theory, this is precisely the expected result. The input to the code is the lexical representation, and therefore, the code cannot be a lexical rule. It follows that the code cannot see the internal structure of words.

The model of the application of the secret code can be represented as follows:

(50)   Natural Languages   Code Languages

                    Lexicon   Lexicon

underlying representations	underlying representations
↓	↓
lexical rule applications	lexical rule applications

                                                    ↓
                                              lexical entry$_i$
                                                    ↓
                                              secret code
                                                    ↓
                          lexical ent ry$_i$ lexical entry$_j$
                                    ↓   ↓

| post lexical rule applications | post lexical rule applications |

          ↓   ↓
phonetic representations   phonetic representations
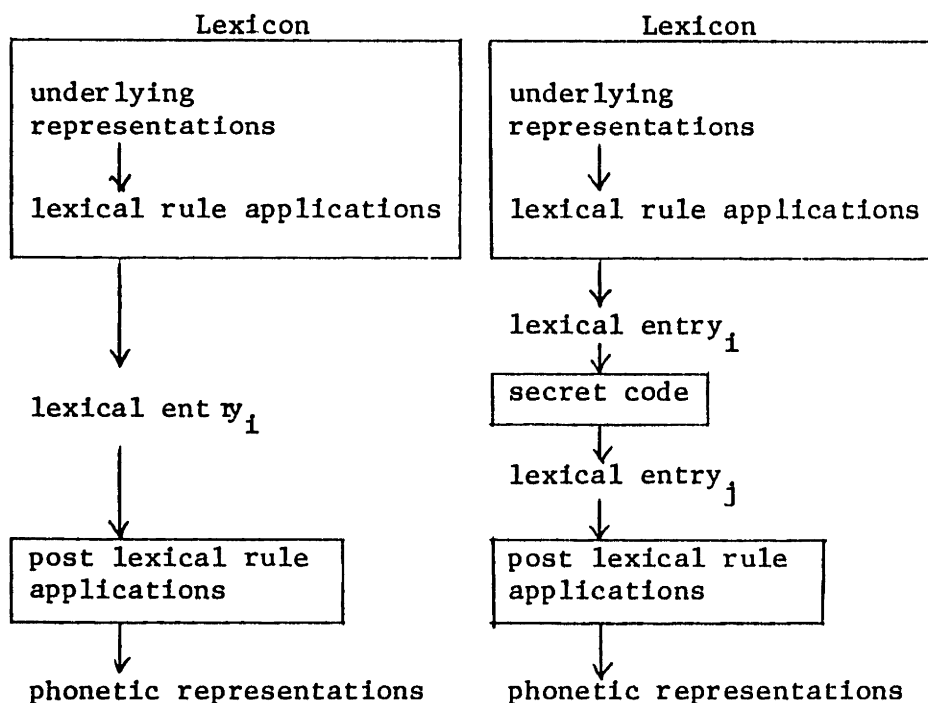
Pig Latin is a language that is perfectly consistent with the model in (50), but since it does not interact with most of the rules in English phonology, it is also consistent with other possible approaches, and does not tell us anything revealing. However, there exist code languages which are more complex than Pig Latin, and show conclusively that the code applies only after the application of lexical rules and before the application of the post lexical rules. One such is the ayb language of English, the code of which is: insert ayb before every rime of the word.[11] Thus, the word America would be aybɔmaybɛ́ɔraybikaybɔ . (50) predicts that ayb insertion takes place after lexical rules like trisyllabic laxing, ClV tensing, velar soften-

ing, etc. The data confirms this prediction. Thus, the ayb versions

of demon (díymən) and demonic (dimáanik) are daybíymaybən and

daybímaybáanaybik. Had trisyllabic laxing applied after ayb insertion,

demon would have been *daybémaybən. Similarly, Canada (kǽnədə) and

Canadian (kənéydiən) in ayb are kaybǽnaybɔdaybɔ and kaybɔnaybéydaybiaybɔn.

Had the CiV tensing applied after ayb insertion, it would have bled

tensing, resulting in *kaybɔnaybǽdaybiaybən for Canadian. The ayb

forms of divider and division are daybivaybáydaybɔr and daybivaybížaybən.

The application of ayb insertion prior to the d → ž rule would have

either blocked the rule in both cases, or allowed it in both, leading

to bad consequences either way. Critical and criticism are kraybítayb-

ikaybɔl and kraybítaybisaybizaybɔm, showing that velar softening also

applies before ayb insertion.

All the above examples show that stress rules must apply before

ayb insertion. If not, the stress would have been on ayb in most cases!

Additional examples like the following show that even secondary stress

patterns remain unaffected by ayb insertion: $\overset{3}{\text{re}}\overset{1}{\text{desi}}$gn → raybĕdaybesayb

$\overset{1}{\text{ign}}$ (examples in English spelling) ; $\overset{3}{\text{art}}\overset{4}{\text{ifici}}\overset{1}{\text{al}}$ity → $\overset{3}{\text{aybar}}$taybifayb$\overset{4}{\text{ic-}}$

$\overset{1}{\text{aybiaybalaybitayby}}$.

There is no lexical rule application that takes place after

ayb insertion in English. It now remains to show that all post lexical

rule applications must follow ayb insertion. The only post lexical

rule that interacts with ayb insertion is the aspiration of voiceless

stops. The ayb version of painter is [paybéyntaybɔr], and not

*[p^{h}aybéyntaybər], which is what the application of aspiration
before <u>ayb</u> insertion would result in.[12]

The prediction made by the model in (50) is fulfilled by
yet another code language I know of, namely, the <u>pa</u> language of
Malayalam, which inserts the syllable <u>pa</u> before every syllable
in the word. The code is illustrated by the following examples:

(51)   <u>Malayalam</u>   <u>pa-language</u>

   a. moohanan 'a name' → pamoopahapanan

   b. kuṭṭi 'child'   → pakupaṭṭi [13]

I shall select two of the rules discussed in 2.1., gemina-
tion and vowel sandhi, to demonstrate that <u>pa</u> insertion cannot pre-
cede lexical rule applications. The rules are repeated below in an
informal fashion for convenient reference:

(52) a. Gemination: Obstruents geminate stem initially
                      in subcompounds.

   b. Vowel Sandhi: The final vowel at the end of the
                      first stem and the vowel at the beg-
                      inning of the second stem fuse to-
                      gether.

(53) <u>Malayalam</u>   <u>pa language</u>

   a. kutira   'horse'   pakupaṭipara

   b. kuṭṭi   'child'   pakupaṭṭi

   c. kutirakkuṭṭi 'baby horse' pakupaṭiparapakkupaṭṭi

(54) a. kala   'art   pakapala

    b. ulsawam   'festival'   payupalsapawam

    c. kaloolsawam 'art festival' pakapaloopalsapawam

Had pa insertion preceded gemination and vowel sandhi, the

result would have been *pakupatiparapakupaṭṭi in (53c), and

*pakapalapayupalsapawam in (54c).

I shall now use the following post lexical processes to show

- that post lexical rule applications must follow pa insertion:

(55) Voicing : Intervocalically, single stops are voiced.
               (see 2.2.2.)

(56) Onset shortening: Branching onsets are shortened after
                   long vowels. (note: all intervocalic
                   consonant sequences form onsets in
                   Malayalam.)

The following examples illustrate the process of onset

shortening:

(57) a. kuṭṭi 'child'   but   kuuṭṭi 'increased'
         L(ong)   S(hort)

    b. kaṣṭam 'pity'   but   kaaṣṭam 'excrement'
         L   S

Consider now the effects of pa insertion on onset shortening

and voicing:

(58) a.　kuutti　→　pa[g]uu[b]atti
　　　　　　　　　　　　　　　　|
　　　　　　　　　　　　　　　　L

　　　b.　kaastam　→　pa[g]aa[b]astam
　　　　　　　　　　　　　　　　　|
　　　　　　　　　　　　　　　　　L

Clearly, if voicing applied before pa insertion, the k in

(58)a and b would fail to voice; if shortening applied before pa

insertion, it would incorrectly shorten the clusters. We must, there-

fore, assume that pa insertion takes place before these rules.

3.3.4.　The Acquisition and Use of Phonological Knowledge

3.3.4.1.　If the model of phonology with a three level system of
　　　　　　　　　　　　　　and
underlying, lexical,/phonetic representations is a more accurate

model of the phonological knowledge of the speaker, it would make

interesting predictions about the acquisition and use of this know-

ledge. In this section, I shall briefly speculate on the possible

directions that theories of the acquisition of phonology, speech re-

cognition and speech production based on Lexical Phonology may take.

It is likely, for example, that what the speaker internalises

during the acquisition of a new word is its lexical representation;

underlying representations of its constituent morphemes are establish-

ed as and when he comes across morphologically related words which

give evidence for these underlying representations. Words are not

presented to learners as sets of related items. A learner who comes across the word presidential may not have come across preside and president, and therefore, may not have arrived at the underlying representation involving /prezīd/ and /ent/. He would store the word in the lexicon as something like /prezidénšæl/ , which is the lexical representation of the word. Similarly, post lexical rules would be acquired by the learner at a relatively early stage, as they do not require morphological evidence involving word internal structure which is hard to come by during the initial stages of learning. The acquisition of lexical processes would follow at a leisurely pace in a random manner, and some speakers may never acquire some of the lexical rules at all,

It must be mentioned that knowledge of related forms is crucial for the acquisition of underlying representations. Thus, it is not enough that the learner guesses the structure of, say, [tinisiti] (nonsense form) in English to have the word structure [tinis+iti], on the basis of his familiarity with the suffix -ity. Even after the suffix has been identified, he will still have to choose between the underlying /tinik/, /tīnik/, /tinit/, /tīnit/, /tinis/, and /tīnis/, and the correct choice is impossible unless he comes across the stem in some other context. It is this aspect of the underlying forms which makes them, and the lexical rules that derive lexical representations from them, slower to acquire.

3.3.4.2. It would not be unreasonable to assume that speech recognition involves the identification of the lexical representation of phonetic strings. The mapping between lexical representation and phonetic representation does not involve morphological information, and therefore, it would be possible to devise a bottom-up model of speech recognition which would arrive at lexical representations from phonetic representations without taking recourse to morphology. ᴛhis would be a significant improvement on the top-to-bottom model proposed by Halle & Stevens (1962), which appears to be the only available model of speech recognition consistent with the assumptions of generative phonology. In the Halle-Stevens model, lexical access is conceived of as applying all the phonological rules to all the underlying forms listed in the lexicon, and comparing the output with the speech signal. Apart from the problem of computational complexity, the Halle-Stevens model suffers from a serious theoretical inadequacy in that it does not account for the fact that speakers are able to 'recognise' nonsense words, the entries for which do not exist in the lexicon. In an experiment conducted by Nakatani and Libeʳman, for example (personal communication), they found that speakers of English were able to assign unique 'phonemic transcriptions' to nonsense words with perfect agreement. If speech recognition is applying phonological rules to existing lexical entries, and matching the output with the speech signal, there is no explanation for the recognition and acquisition of new words, for which lexical entries do not exist,

The SPE model of phonology does not provide a level of representation which can be recovered from the phonetic representation without the use of morphological information. Consequently, it does not yield a bottom-up performance model. The model of Lexical Phonology offers precisely such a level.

3.3.4.3.   One may suggest that speech production also crucially makes use of the lexical level of representation. One may look for evidence for this hypothesis in speech errors which permute segments, as in spin and spac for spic and span. If speech errors of this sort are due to er rors in the input to the speech production system, and the input to speech production is the lexical representation, one would predict that all lexical rule applications would precede the speech error, and all post lexical rule applications would follow it. The error, then, would be analogous to the application of the code in secret code languages (cf: (50)):

(59)   Lexicon

```
      ┌─────────────────────────────────┐
      │  underlying representations     │
      │              │                  │
      │              ↓                  │
      │  lexical rule applications      │
      │                                 │
      └─────────────────────────────────┘
                     │
                     ↓
           correct lexical entry

                     │
                     ↓
           ┌──────────┐
           │  error   │
           └──────────┘
                 │
                 ↓
          incorrect lexical entry
                 │
                 ↓
      ┌────────────────────┐
      │  post lexical rule │
      │  applications      │
      └────────────────────┘
                 │
                 ↓
        phonetic representations
```
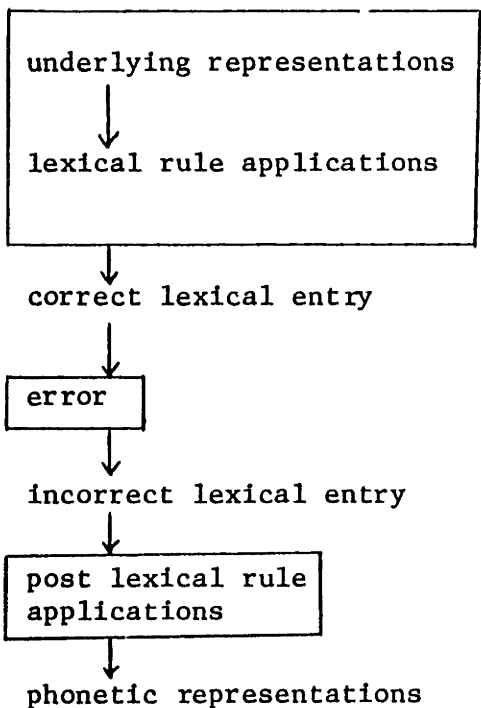
The permutation of the vowels in the first syllable of
metrical grid and linear set would produce the error m[i]trical gr[e]d
and l[e]near s[i]t, never *m[i]trical gr[iy]d and l[e]near s[ay]t. The
correct output is possible only if the lexical process of laxing appli-
es prior to the permutation:

(60)   linear set   linear set

           ī   e   ī   e

           i   ── laxing   e   ī   error

           e   i error   ──   ──   laxing

        l[e]near s[i]t   *l[e]near s[ay]t

Similarly, the permutation of the consonants in electricity

board may result in electri[b]ity [s]oard, but never *electri[b]iti [k]oard. Once again, the desired output demands the application of velar softening before the error.

In contrast, post lexical processes like aspiration apply after the application of the error. Thus, scotch tape becomes s[t]otch [kh]ape, never *s[th]otch [k]ape (example from Stampe (1972)).

(61)   scotch   tape   scotch   tape
        k   t   k   t
        t   k   error   —   th   aspiration
        —   kh aspiration   th   k   error
     s[t]otch [kh]ape   *s[th]otch [k]ape

Bad men [bæd mẽn] may become [bed mæ̃n] , never *[bẽd mæn] : the nasalisation of the vowel, a post lexical process, also applies after the error.

Similarly, I face your miss for I miss your face (mi[š]) may be pronounced as I fa[š] your miss, but never *I face your mi[š]: the external sandhi process (s → š) applies after the error.

3.4. Comparison with Similar Models

3.4.1. Taxonomic Phonemics

By now, the reader must have observed the similarity of lexical representations to taxonomic phonemic representations. A

crucial difference between Lexical Phonology and Taxonomic Phonemics
on the one hand, and Standard Phonology on the other, is that the
latter allows only two levels of phonological representation, the
underlying and the phonetic, while the former incorporate an inter-
mediate level of representation.

<div align="center">

(62)   Taxonomic Phonemics   Lexical Phonology

Morphophonemic. . . I . .Underlying

Phonemic. . . . . .II . .Lexical

Phonetic. . . . . III . .Phonetic

</div>

The similarity ends here. Taxonomic Phonemics focussed on the
restrictions on the mapping between levels II and III, while Lexical
Phonology concentrates on the mapping between levels I and II. While
Taxonomic Phonemics postulates several conditions on the phonemic-
phonetic mapping, such as biuniqueness and invariance, the only con-
dition in Lexical Phonology is that the lexical-phonetic mapping has
no access to word internal morphological information. Therefore, a
number of rules disallowed in Taxonomic Phonemics in the mapping from
II to III are acceptable in Lexical Phonology. A typical example of
such a rule would be one which changes one phoneme to another, like
the voicing rule in Malayalam, which changes voiceless stops to voiced
stops post nasally and intervocalically (2.2.2.), and the palatalisa-
tion rule in English which changes s to š. The lexical representation
for [maṇḍan] in Malayalam is /maṇtan/, and for [ðišyər] (this year)

in English is /ǰis yᴈr/: these are not possible taxonomic phonemic representations. Similarly, rules which delete phonemes would be legitimate post lexical operations, but not legitimate 'allophonic' rules (e.g. list some → lis some). What this means is that lexical representations are more abstract than taxonomic phonemic representations.

All the classical arguments against the taxonomic phonemic level of representation (Chomsky (1964), Halle(1959)) have b. ᴈ essentially arguments against a specific level of representation satisfying conditions like biuniqueness, local determinacy, invariance, and linearity. Since the level of lexical representations does not obey any of these conditions, the objections do not apply to the lexical level.

3.4.2. Natural Phonology

The distinction made by natural phonologists between rules and processes (Stampe (1973), Donegan (1978)) appears to have some correlation with the distinction I have identified as Lexical and Post lexical; it is, therefore, important to examine how the two theories match. A process is defined in Natural Phonology as "a mental operation that applies in speech, to substitute, for a class of sounds or sound sequences presenting a common difficulty to the speech capacity of the individual, an alternate class identical but

lacking the difficulty property" (Stampe (1973:1)). In other

words, a process is motivated by the ease of articulation. One

may say that ty → c in didn't you is a process, but the t → s

change in presidency is a rule, and has no articulatory motivation.

Processes, according to natural phonologists, are universal:

speakers do not have to acquire them. What has to be learnt is the

inhibition of the process in a language that does not exemplify it.

A speaker of English does not have to learn the ty → c change, as

the process is innate, but a speaker of Malayalam has to learn to

inhibit or block the process, as Malayalam does not exhibit the alter-

nation. Rules, on the other hand, have to be learnt.

Another property of processes is that they require special

effort on the part of the adult speaker to violate them. Thus,

velar softening, which is a rule, can be easily violated by a speaker

to produce criti[k]ism, but it is only with great effort that one

can violate the process of aspiration in pin , and say [p]in instead

of [ph]in.

Stampe observes that rules apply before the application of

the code in secret code languages, while processes apply after them.

Thus, the process that changes k to k' before a front vowel applies

after the Pig Latin rule in English, giving the Pig Latin form

ool[k']ey instead of *ool[k]ey for the word cool. On the other hand,

the rule of velar softening must apply before the insertion of ob

in the secret language of ob, since electricity in this language is

obelobectrobi[s]obitoby: the insertion of ob destroys the environment

for velar softening, and therefore, the softening must be assumed

to take place prior to ob insertion (Stampe's examples).

Stampe also points out that processes apply to the result

of slips of the tongue. "Such productions as [staĕkhɛip] for scotch

tape show that the slip /staĕ kɛip/ for /skaĕ tɛip/ arose before the

process aspirating initial stops applied, since otherwise [sthaĕkɛip]

would have been produced" (1973: 44).

The criteria for distinguishing between processes and rules

can be summarised as follows:

(63)   processes   rules

a. motivated by ease of   (?)
   articulation

b. universal   (?)

c. cannot be violated   can be violated

d. follows the applica-   precedes the application
   tion of the code in   of the code in secret
   secret code languages   code languages

e. follows slips of the   precedes slips of the
   tongue   tongue

With respect to (63)d and e, Lexical Phonology and Natural

Phonology make the same distinction, as lexical rule applications

precede the application of the secret code and speech error, while

post lexical rule applications follow them. (63c) is not consistent
with the lexical model, as palatalisation across words (thi[s] yea r)
is a post lexical operation, yet speakers experience no difficulty
in violating it.

The hypothesis that processes are universal ((63b)) and that
they are inhibited in languages which do not have them seems to me
to lack any empirical content. We may find a process that replaces
A by B in exactly one language, but nothing in principle would pre-
vent a natural phonologist from saying that A → B is inhibited in
all human languages except one. I therefore dismiss (63b) as a non-
falsifiable claim.

The claim that processes have articulatory motivation ((63a))
has the potentials of being an empirical claim, but only if it is
accompanied by a theory of phonetics from which we can derive the
notion 'articulatory ease'. Such a theory, given two sounds or sound
sequences, should tell us in unambiguous terms which of them is more
difficult to produce. In the absence of such a phonetic theory, (63a)
is a matter of opinion or intuition, and can hardly be argued with.

There is a crucial difference between the levels identified
by Natural Phonology and Lexical Phonology that has some empirical
consequence: while Natural Phonology makes a distinction between two
types of operations in terms of what they do (i.e. in terms of the
changes they effect), the distinction made in Lexical Phonology is in

terms of where the operations apply (i.e. in terms of their domains). To give an example, the rule of palatalisation in English applies both in the lexical and the post lexical domains. Since the distinction made in Natural Phonology is in terms of the intrinsic phonetic content of the operations, palatalisation should be either a rule or a process, but hardly both.

In sum, the level of lexical representation in Lexical Phonology is not identical to either the taxonomic phonemic level or the level between rules and processes in Natural Phonology.

Footnotes for Chapter III

1. The reason for assuming a morpheme based word formation rather than a word based word formation (cf: Aronoff (1976)) is that languages like Malayalam, Sanskrit, and Latin have several productive and regular forms of affixation on stems which are not independent words. Thus, the tense suffixes in Malayalam (um 'fut.', unnu 'pres', i/tu 'past') are attached to stems which are not independent words: e.g. dukkhikkum 'will grieve', ootum 'will run', but *dukkhikk, *oot. The same observation applies to the formation of nouns, verbs, and adjectives. Thus, dukkhikk- (V), dukkham 'grief', and dukkhitan 'grieving person' are derived from the root dukkh which does not exist as a word. (This is a paradigm case in Malayalam word formation.) Except for the plural suffix, I cannot think of any affix in Malayalam which is attached to 'words'. On the basis of data of this kind common in various languages, I reject the model of word based word formation.

2. This contrast follows from Selkirk's (to appear) principle of semantic interpretation of compounds which forbids the interpretation of the left member as the subject of the predicate of the right member. The meaning "one whom children employ" interprets child as the subject of employ, which the principle disallows. Compounds like Ford employee may appear as counterexamples to this principle, but I would like to suggest that the meaning of Ford employee is "one who is employed at Ford, i.e., at the establishment referred to as Ford", rather than "one whom Ford (the person) employs". cf: *king employee, *president employee, *governor employee.

3. Unlike the past tense, present tense, and plural, the possessive in English is attached to a phrasal category, not a lexical category.

4. It seems intuitively correct to say that only lexical rules can have lexical exceptions (cf: Bresnan (1978), Baker (1979)).

5. Voicing applies only after vowels. cf: re[z]ume, but con[s]ume. (See SPE for the details.)

6. One may ask why there is an identity rule of the form nitingæl → nitingæl, but none of the form long → long. Clearly, this is because lexical rules apply to the underlying form /long/, while no lexical rules apply to /nitingæl/. If so, setting up nitingæl → nitingæl is another way of saying that no lexical rules can apply to /nitingæl/.

7. The discussion in this section is about deliberate pauses, not the kind of accidental pauses created by, say, coughing or surprise.

8. I am grateful to Moira Yip for her help with this rule.

9. The rule is intended to account for the phenomenon of 'linking r', not 'intrusive r' (see Jones (1940), Gimson (1962)). The dialects with linking r have "The car is..." but "The idea is...". The dia-
        [kaar]   [aidiə]
   lects with intrusive r have "The idea is..." with an r at the end
                        [aidiər]
   of idea as well. In the latter dialects, (34) is replaced by a rule that inserts r prevocalically.

10. The transcriptions are from Jones (1977), 14th edition.

11. I am grateful to Joan Bresnan for the ayb data.

12. Ken Hale has pointed out to me that the nasal assimilation in fast speech in words like unclean → u[ŋk]lean applies in Pig Latin after the application of the code. e.g. clean is eanclay which becomes ea[ŋk]lay. Given that fast speech phenomena have to be post lexical, this is the predicted result.

13. Malayalam syllables do not allow codas. Words like kuṭṭi and kampi are syllabified as ku-ṭṭi and ka-mpi (see 4.1.).

14. The rule that inserts a glide between two vowels is discussed in 4.2. (pa u... → payu... (54b)).

Chapter IV: A CASE STUDY : MALAYALAM PHONOLOGY

Having developed a theory of Lexical Phonology, and shown
its consequences for the study of the phonological knowledge of
human beings, I shall proceed to illustrate the application of the
model through an extensive analysis of Malayalam phonology. This
case study has two purposes: first, it provides the proper background
for the phonology of compounding discussed in Chapter II, and high-
lights the details of the theory by examining a set of complex inter-
actions of phonological phenomena; second, it provides additional
supporting evidence for various assumptions made in the preceding
chapters, such as the integration of phonology and morphology, the
Opacity Principle, and the Stratum Ordering Hypothesis.

4.1. Syllable Structure

4.1.1. The Syllable Template

When asked to syllabify the words kampi 'metal rod', and
kaappi 'coffee', native speakers of Malayalam break them up as
ka-mpi and kaa-ppi. Odd though it may seem to those who are used to
the Indo European syllabification, I shall argue that this untutored
syllabification reflects a significant structural property of the
language, namely, that it allows no codas.

Before going on to the details of the syllable in Malayalam,

I should make explicit what I mean by 'coda', and my approach to

syllable structure in general. I follow the theory of the syllable

outlined in Halle & Vergnaud (1980), which sets up a three tier

representation of the syllable, consisting of a tier of distinctive

feature specifications called the 'melody', a skeletal tier of C's

and V's, and a tier of onsets and rimes. (See McCarthy (1979) for the

idea of C/V sequences as the prosodic skeleton.) An example of this

kind of representation is given below:

(1) Whistle

```
      σ     σ
     /\    /\
     O R   O R      onsets and rimes
     | |   | |
     C V   C V      skeleton
     | |   | |
     u i   s l      melody
```
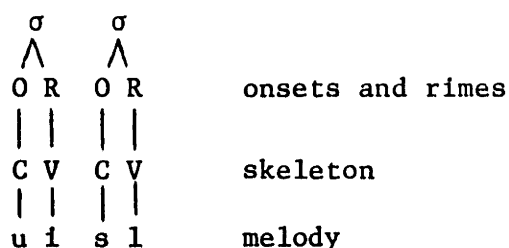
Observe that C's and V's are not to be identified with

'consonants' and 'vowels'. Thus, l in whistle is not a vowel, but

is linked to V element in the skeleton. To use Pike's (1947) termi-

nology, the melody tier consists of the phonetic categories of

vocoids (e.g. u, i) and contoids (e.g. s, l), while the skeletal

tier represents their syllabic (V) and nonsyllabic (C) functions.

Following this approach, I represent the syllable structure

of kampi and kaappi as follows:

159

(4)a. Syllabification: <u>kam-pam</u>; both nasals are codas.

   b. Syllabification: <u>ka-mpam</u>; only the second nasal is
a coda.

   c. Syllabification: <u>ka-mpam</u>; niether of the nasals is
a coda. The first one is part of the second onset,
and the second one is an appendix.

The syllabifications in (4)a-c are:

(5)   a.

```
     σ      σ      b.  σ      σ      c.   σ      σ
     /\    /\         /\    /\          /\    /\
    O R   O R        O R   O  R        O R   O  R A
    | /\  | /\       | |   /\ /\       | |   /\ | |
    C V C C V C      C V C C V C       C V C C V C
    | | | | | |      | | | | | |       | | | | | |
    k a m p a m      k a m p a m       k a m p a m
```
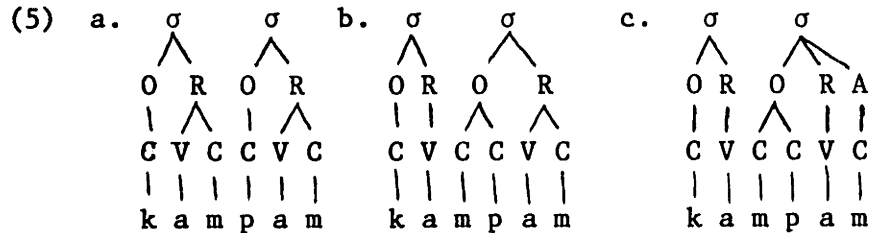
I shall argue for (5c), which illustrates the 'no coda' hypo-
thesis. I propose the following syllable template for Malayalam:

(6) Syllable Template

```
              σ
            /  |  \
    (onset) rime (appendix)
            /\
           V (V)
```
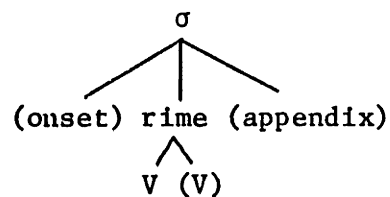
Appendices, it has been observed, have a restricted distri-
bution in natural languages (Halle & Vergnaud (1980), Harris (forth-
coming)). In Malayalam, they occur only in the morpheme final position.

4.1.2. The pa Language

The pa language, a secret code language of Malayalam (3.3.3.),

confirms the V-C_0V analysis of the Malayalam syllable. I formulate

the rule of pa insertion as follows:

(7) Insert pa before every syllable.

Thus, moohanan 'a name', in pa language, is pamoopahapanan.

Versions of the pa language exist in most Indian languages, with

minor dialectal differences. In some versions, the syllable inserted

is ca or sa instead of pa: moohanan → camoocahacanan. In others, the

inserted syllable copies the features of the following vowel:

moohanan → pomoopahapanan; siima 'limit' → piipsiipama. What is

common to all these versions is the insertion of a nonsense syllable

before every syllable in the word.

Consider the pa version of the following words:

(8) a. kañcan 'a name' → pakapañcan
    b. indira 'a name' → payipandipara
    c. susmita 'a name' → pasupasmipata

(9) a. ka-ñcan   b. i-ndi-ra   c. su-smi-ta

If the syllabification of (8) is as given in (9), the pa

versions are correctly accounted for by (7). An alternate syllabifi-

cation, such as the one in (10), fails to account for (8):

(10) a. kan̄-can   b. ī-ndi̱-r̄a   c. sus-mi-ṯa

Given the syllabification in (10), the rule of pa insertion would be unnecessarily complicated:

(11) Insert pa at the beginning of the word, and after every nucleus except the last one.

The syllabification in (10) does exist in the Indo European languages of India, giving the following pa versions:

(12) Hindi, Gujarati

    a. kancan   → cakancacan (or cakancacacan(a)) [1]

    b. indira   → cayincadicara

    c. susmita → casuscamicata

It is well known that languages in India share a number of common properties in vocabulary, phonology, and syntax, in spite of their genetic differences. This has, in fact, led to the notion of 'India as a linguistic area' (Emeneau (1956), Masica (1976)). Given rule (7), the differences between the pa versions of the Dravidial and the Indo European languages of India follow from the assumption that VCCV is syllabified in the former as V-CCV, and in the latter as VC-CV.

4.1.3. Gemination

Another source of evidence for the no coda hypothesis in

Malayalam comes from the rule of gemination (2.1.1.), repeated below:

(13) Gemination (Domain : Stratum 2)

Onset → $\bigwedge$ % ——][

(14) Examples

a. [[kaaṭ][maṟam]] → kaaṭṭəmaṟam (stem final gemi-
nation)

b. [[kuṭiṟa][kuṭṭi]]→ kuṭiṟakkuṭṭi (stem initial
gemination)

The application of (13) in (14b) is illustrated below:

(15)

σ   σ   σ   σ   σ
∧   ∧   ∧   ∧   ∧
O R O R O R O R O R
| | | | | | | | ∧ |
C V C V C V C V C C V
| | | | | | | | ∨ |
k u ṭ i ṟ a k u t i

σ   σ   σ   σ   σ
∧   ∧   ∧   ∧   ∧
O R O R O R O R O R
| | | | | | ∧ | ∧ |
C V C V C V C C V C C V
| | | | | | ∨ | ∨ |
k u ṭ i ṟ a k u t i   by (13)
          (=kk) (=ṭṭ)
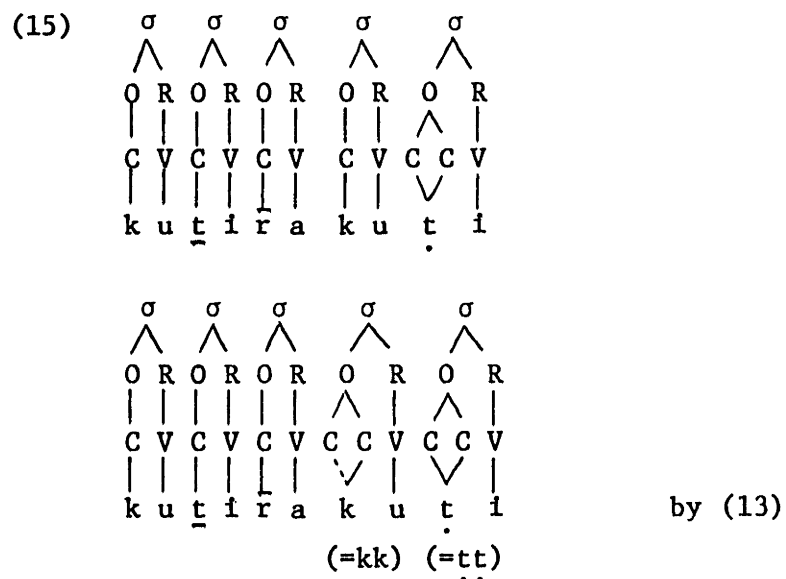
What is interesting is that gemination does not take place
if the onset is already branching, i.e., if the stem final/initial
consonant is a geminate, or in a cluster:

(16) a. [[kutīra][preemam]] → kutīrapreemam

                               *kutīrappreemam

    b. [[paatt] [maas]] → paattɔmaaṣɔ

                               *paatttɔmaaṣɔ

Given the formulation in (13), nothing need be stated about the lack of gemination in (16) a and b. Since the onsets are already branching (ĈC , ĈC), (13) applies vacuously, producing no difference in the output. Crucial to this solution, however, is the assumption that the pr in kutīrapreemam and the tt in paattɔmaaṣɔ constitute onsets, and are not heterosyllabic. If VCCV is syllabified as VC-CV in Malayalam, the formulation of gemination will have to be complicated in undesirable ways.

The rules of stem final and stem initial geminations are disjunctively ordered, stem final gemination applying first. Thus, in kayattɔkaṭṭilɔ 'rope cot' (kayarɔ , kaṭṭilɔ), stem initial k does not geminate, as stem final gemination applies to r (rr → tt). In contrast, the k in marakkaṭṭilɔ 'wooden cot' (maram, kaṭṭilɔ)geminates. What is fascinating is that this disjunctivity is sensitive to the vacuous application of onset branching. If the preceding stem final onset is branching, the stem initial consonant does not geminate:

(17) a. [[cemp][paatram]] → cempɔpaatram *cempɔppaatram
       'copper''vessel'   'copper vessel'

    b. [[swarnam][paatram]]→ swarnappaatram
       'gold' 'vessel'   'golden vessel'

In (17a), (13) applies vacuously to $\overset{\overset{o}{\wedge}}{\underset{mp}{\overset{|\,|}{CC}}}$, and therefore,
does not apply to the stem initial p. In (17b), on the other hand,
stem final gemination does not apply, therefore stem initial gemina-
tion applies to p.

4.1.4. The r̄ ~ r Alternation

The r̄ ~ r alternation in formal Malayalam appears to provide
an argument against the no coda hypothesis.

Underlying r̄ becomes r before consonants and before pauses:

(18)   Formal   Colloquial

    a. awar   awarə   'they'
    b. awar̄e   awar̄e   'they-acc.'
    c. awar̄ootə   awar̄ootə   'they-dat.2'
    d. awarkkə   awar̄əkkə   'they-dat.'
    e. awar̄ ewiṭe   awar̄ ewiṭe   'where are they'
    f. awar... ewiṭe   awarə...ewiṭe   'where are they'
    g. awar wannu   awarɔ wannu   'they came'

If awarkkə is syllabified as a-war-kkə , and the r in awar
and awarkkə is the coda, the rule would be stated as follows:

$$\text{(19)} \quad \bar{r} \rightarrow r/ \underset{\overline{}}{\overset{R}{\overset{|}{}}}$$

The contrast between (18e) and (18f) (formal version) is accounted
for by the syllabification in (20):

```
(20) a.    σ    σ              b.      σ        σ
          /\   /\                     /\       |
         O R  O R                    O  R      R
         | |  | |                    | /\      |
         C V  C V                    C V C     V
         | |  | |                    | | |     |
         a w  a r e wite             a w a r...ewite
                  •                            •
```
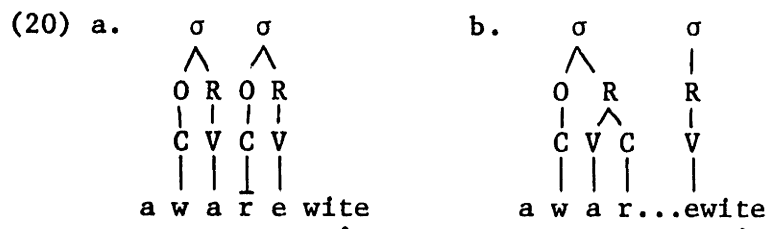
Steriade (1981) argues that the no coda hypothesis fails to
unite the preconsonantal ((18d))and the prepausal ((18f)) r's. [2] I
would like to suggest that the no coda hypothesis can handle the
problem raised by (18) by allowing appendices (r, l, ḷ) in the mor-
pheme final position:

```
(21) a. σ    σ      b. σ   σ    σ    c.  σ   ·σ      σ
        |   /\         |  /\   /\        |  /\      /\
        R  O R A       R  O R  O R       R  O R A  O  R
        |  | | |       |  | |  | |       |  | | |  /\ |
        V  C V C       V  C V  C V       V  C V C  C C V
        |  | | |       |  | |  | |       |  | | |  \/ |
        a  w a r       a  w a  r e       a  w a r  k  ə
```

```
   d. σ   σ    σ    σ    σ
      |  /\   /\   /\   /\
      R  O R  O R  O R  O R
      |  | |  | |  | |  | |
      V  C V  C V  C V  C V
      |  | |  | |  | |  | |
      a  w a  r e  w i  t e
                        •
```
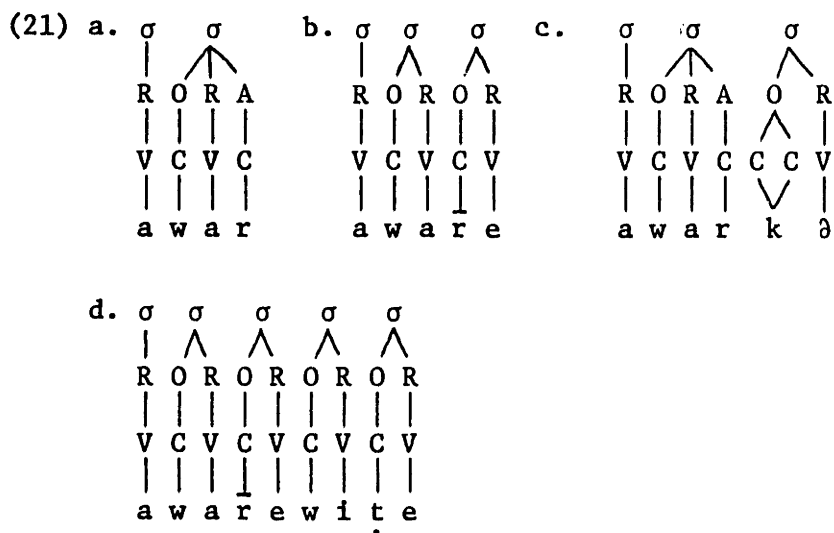
(21) would require an obligatory resyllabification rule such
as (22):

(22) A → 0/ —— R

Given (21), the r̄∼r alternation can now be accounted for as
follows:

167

$$(23) \quad \bar{r} \longrightarrow r / \overset{\overset{A \quad 3}{\displaystyle\perp}}{}$$

4.1.5. The Voicing of Stops

I shall formulate the rule of voicing in Malayalam, tenta-
tively presented in 2.2.2. as intervocalic voicing, in terms of the
syllable structure. What is crucial about the voicing of stops is
that it occurs only in noninitial nonbranching onsets:

(24) a. [kaāṛanam] 'reason' ; [agaāṛanam] 'without reason'
     b. [pawiṯram] 'holy'   ; [abawiṯram] 'tainted'
     c. [paattə]   'song'   ; [paadum]   'will sing'

(25) a. [kramam]   'order' ; [akramam]   'disorder'
     b. [tyaagam] 'sacrifice'; [parityaagam] 'giving up'

(26) a. akalca → [agalca] 'distance'
     b. arkan   → *[argan] 'sun'

Thus, the voicing of stops also takes place when nonbranch-
ing onsets are preceded by appendices in formal Malayalam:[4]

(27) a. awal karaññu → [awalgaraññu]   (NP-V)
         she cried

     b. ṛgam kal tanne → [mrəgaŋŋal danne] (NP-Adv.)
         animal pl. intens.

     c. malar kontə → [malargontə]   (NP-P)
         flower with

     d. awar kal → [awargal]   (N-pl.)
         they pl.   (honorific pronoun)

The rule of voicing can now be stated as follows:

(28) Stop Voicing

   Nonbranching noninitial onsets are voiced.

The sequence rk constitutes a branching onset in (26b), but not in (27c); hence the difference in voicing:

(29) a.
```
        σ         σ
        |        /\
        R     O  R  A
        |    /\  |  |
        V C  C  ¨  C
        | |  |     |
        a r  k     a  n
```

b.
```
    σ      σ      σ      σ
   /\     /\     /\     /\
  O  R   O  R   A O R  O  R
  |  |   |  |   | | |  |  /\ |
  C  V   C  V   C C V  C  C V
  |  |   |  |   | | |  |  | |
  m  a   l  a   r k o  n  t ə
                       .  .
```
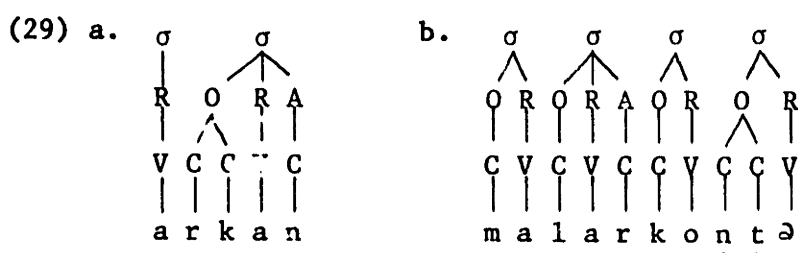
An important fact that must be mentioned at this point is that there are no morpheme internal cases of stop voicing after r, l, and ḷ. Thus, monomorphemic alpam 'a little', kalpaṇa 'order', walkalam 'blouse made of bark', karpuuram 'camphor', wiirpə 'bloating', etc., have voiceless stops: this follows from the assumption that appendices in Malayalam are morpheme final.[5]

4.1.6. Laghu and Guru

Another confirmation for the syllable template in (6) comes from the laghu~guru distinction in Malayalam poetry. 'Laghu' (light syllable) is a syllable with a short vowel not followed by a cluster; 'guru' (heavy syllable) is a syllable with a long vowel or one followed by a cluster. Thus, the initial syllable in kuṭi 'drinking' is a

laghu, while those in kuṭṭi 'child', kuuṭi 'increased-intr.', and kuuṭṭi 'increased-tr.' are guru's.

There are two ways in which the laghu ~ guru distinction can be characterised. Within the no coda hypothesis, the examples given above will be syllabified as in (30), and guru will be defined as in (31):

(30) a. kuṭi   b. kuuṭi   c. kuṭṭi   d. kuuṭṭi
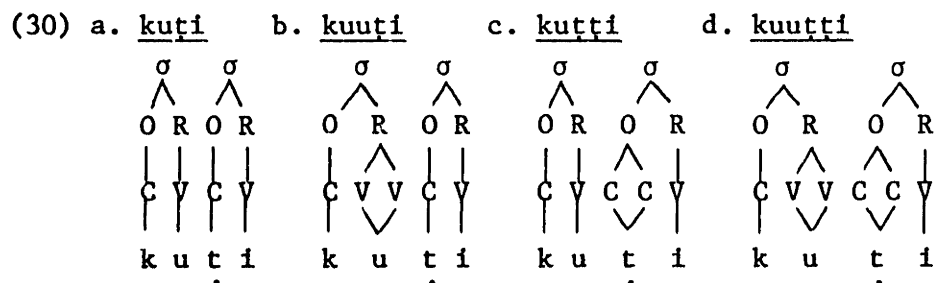
(31) Guru: syllable which has a branching rime, or is
             followed by a branching onset.

Alternatively, one may allow codas, and assume the syllabification in (32), and the definition of guru in (33):

(32) a. kuṭi : as in (30a)

      b. kuuṭi : as in (30b)

      c. kuṭṭi   d. kuuṭṭi

(33) Guru: syllable which has a branching rime.

Inelegant though (31) may seem, I shall demonstrate that (32)

and (33) lead to incorrect predictions, and that (31) is, indeed,

correct.

The first problem which (33) faces is the laghu/guru assign-

ment in the case of word final vowels. The third syllable of kalpana

'order' is a   guru in kalpana prakaaram 'according to the order',

but a laghu in kalpana kaẓiñña 'after the order'. This contrast

follows directly from (31), but for (33) to work, we would require

an additional rule of resyllabification across words that changes

kalpanā prakaaram   to   kalpana prakaaram. This resyllabification will

have to apply every time a word final vowel is followed by a branch-

ing onset.

The second problem is a little more complex. There is an

interesting distinction in Malayalam versification referred tc in

the traditional literature (Rajarajavarma (1904)) as 'laghuprayatnam'

(light effort)and 'tїiwraprayatnam' (intense effort). Thus, one finds

examples like the laghu in the second syllable in malarmaala 'garland

of flowers' (malaṟ 'flowei'; maala 'garland'), in contrast to the

guru in the first syllable of marmaram 'murmur' (monomorphemic),

Observe that the sequence -arm- is taken as a laghu in the former

case (laghuprayatnam), while the same sequence is a guru in the latter

case (tїiwraprayatnam). Our assumptions regarding the syllable struc-

ture of Malayalam allow us to offer a principled account of the tradi-

tional distinction.

What we need to explain is why the second syllable in
malarmaala is a laghu, even though the vowel is followed by a
cluster. This can be accounted for by assuming that the r in this
word is an appendix:

(34)

```
        σ     σ              σ      σ
       /|\   /\            /\     /\
      O R A  O R          O R    O  R
      | | |  | /\         | |    /\ |
      C V C  C V V        C V   C C V
      | | |  | |\/        | |   | | |
     mal a r m a  la  but  m a  r m aram
```
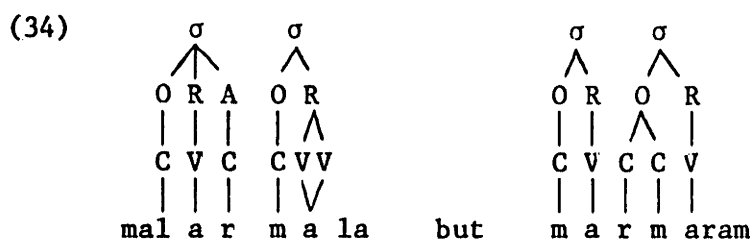
In malarmaala, the second syllable has a nonbranching rime
followed by a nonbranching onset, and is therefore a laghu; in
marmaram, on the other hand, the second syllable has a rime followed
by a branching onset, and is therefore a guru.

Since appendices in Malayalam are restricted to the morpheme
final position, this account correctly predicts that the sequence
VC_iC_j results in a laghu only if C_i is an appendix, i.e., morpheme
final r, l, or ḷ.

(35) a. akalca 'distance' ; muyal cattu 'hare died' (NP V)
           ↑   ↑
           G   L
     b. markaṭan 'monkey' ; malar konṭə 'with flowers'(NP P)
           ↑   ↑
           Ġ   L
     c. arkan 'sun'   ; awarkal 'honor. pronoun'
           ↑   ↑
           G   L

Though the second syllable in malar konṭə 'with flowers' is
a laghu, it is a guru in malar tyajiccu 'sacrificed flowers'. Once
again, the contrast is predicted by (31): ty is a branching onset,

k is not. This contrast becomes clearer with compounds in which

gemination occurs, such as <u>malarppoṭi</u> 'pollen' (<u>poṭi</u> 'powder'). Since

RA   O
ḷḷ   ʌ
ar in this case is followed by pp , the second syllable is a guru,

in contrast to the phrase <u>malar poẓinnu</u> 'flowers fell', in which the

second syllable is a laghu.

Wherever we find laghuprayatnam, we also find the 'irregular'

voicing. Thus, <u>awarkaḷ</u> ((35c)) is phonetically [awargal], while <u>arkan</u>

is [arkan]. It is the nonbranching stop onsets which voice noninitial-

ly, and k in <u>awarkaḷ</u> is a nonbranching onset: the voicing is predict-

ed by the no coda hypothesis.

The behaviour of the <u>pa</u> language confirms our account of

laghuprayatnam. <u>marmaram</u> in the <u>pa</u> language is <u>pamaparmaparam</u>,

while <u>malarmaala</u> is <u>pamapalarpamaapala</u>.

It must be pointed out that nasals never give rise to laghu-

prayatnam. Thus, though the second syllable of <u>malar koṇṭɔ</u> is a laghu,

that of <u>maram koṇṭɔ</u> is a guru. The solution to this problem, I think,

lies in stipulating that nasal appendices are obligatorily attached

to the following onsets:
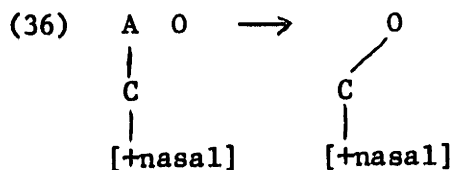
(36)   A   O   ⟶   O
            | ╱
            C   C
            | |
         [+nasal]   [+nasal]

```
                      RA O           R  O
                      || |          _| /|
(36) would change maram konta   to maram konta , giving
```

the desired results.

4.1.7. Malayalam Onsets and Universals of Syllable Structure

A consequence of the assumption that Malayalam has no codas is that all intervocalic consonant sequences, as pointed out earlier, are analysed as onsets: a-rkan, ka-mpi, and sa-mskaa-ram 'culture'. Now, clusters like rk, mp, and msk do not occur word initially, and therefore, it is necessary to stipulate that only a subset of allowable onsets can occur as word initial onsets.

Steriade (1981), who argues for codas in Malayalam, points out that none of the word initial onsets in the language violate the universal principle of sonority hierarchy, given below:

(37) a. Segments belonging to the same onset occur in the order of increasing sonority.

b. Segments belonging to the same rime occur in the order of decreasing sonority.

Language surveys indicate that even though violations of (37) are observed at the beginning and end of words, they are not common word medially. Under the no coda hypothesis, the facts of

Malayalam syllable structure would be inconsistent with these observations. Even though word initial onsets do not violate (37a) (with the exception of onsets like sk, sp, etc. common in other languages as well), word medial onsets do (e.g. rk, mp). Steriade observes that it would be an accident, given the no coda hypothesis, that only word medial onsets violate (37a), and not word initial onsets. On the other hand, if we allow codas (ar-kan, kam-pi), all onsets would obey (37a).

In order to make Malayalam conform to the language survey results, Steriade is forced to complicate the grammar of Malayalam in uninsightful ways. In order to account for the pa language phenomenon, she sets up a late resyllabification rule which groups all consonants into onsets, thereby assuming, in effect, the no coda hypothesis at a late stage in the derivation. She accounts for laghu-prayatnam by setting up another resyllabification rule that changes a nonnasal sonorant into a syllabic consonant morpheme finally just in case it is followed by a single consonant (and not a cluster), and assuming that metrical scansion ignores these syllabic consonants. In short, practically every syllable related phenomenon requires a special rule of resyllabification and ad hoc assumptions.

4.2. Glide Insertion

4.2.1. The Facts and an Account

Malayalam has a rule that inserts a glide between two ad-
jacent vowels, agreeing in roundness and backness with the preceding
vowel :

(38) a.   [[kari̅]$_2$[ila]] → kari̅yila

        coal   leaf   dry leaf

     b.   [[pasu̅]in te]] → pasu̅winte

          cow   poss.   cow's

     c.   taa̅ra alari → taa̅rayalari

         a name roared

Glide insertion can be looked upon as the insertion of an
onset between two adjacent rimes:

(39) RR ⟶ ROR

The principle that is responsible for (39) is:

(40) Rimes and onsets must alternate (in Malayalam).

In order to derive (40), we stipulate that onsets are obli-
gatory in Malayalam except in the phrase initial position:

(41) a.   σ
           /\
          O R

     b.   σ   (condition: σ is initial)[7]
           |
           R

We can now account for glide insertion by the following
general convention:

(42)   Syllable Completion Convention (Domain: Strata 2 - 5,

where 5 is the post lexical stratum)

Complete the syllable structure by providing the
missing rimes and onsets.[8]

The onset node provided by (42) is linked to the melody to
its left:

(43)   R   O ⟶   R   O
         | | | |
         V   C   V   C
         | |
       [-cons.]   [-cons.]
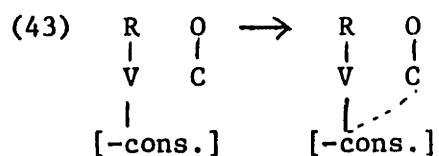
```
C  C      C
|  |      |
i, e, and a are phonetically realised as y by universal
       C    C
conventions, and |  and |  are realised as w.
                 u      o
```

The derivation   of taarayuraṇṇi 'Tara slept' is given below
by way of exemplification:

(44)
```
      σ    σ    σ σ      σ
     /\   /\   |/\     /\
    O  R  O R  R O R   O  R
    |  /\ | |  | | |   | /\ |
    C V V C V  V C V   C C V
    |  V  | |  | | |   | V  |
    t  a  r a  u r a   ŋ    i
```

```
         σ σ
        /\ /\
        O R O R
        | | | |
        C V C V
        | | | |
    t a r a  u r a  ŋ  i        by Syllable Completion
```

```
         σ σ
        /\ /\
        O R O R
        | | | |
        C V C V
        | | | |
    t a r a  u r a  ŋ  i        by (43)
            ↓
            ay
```
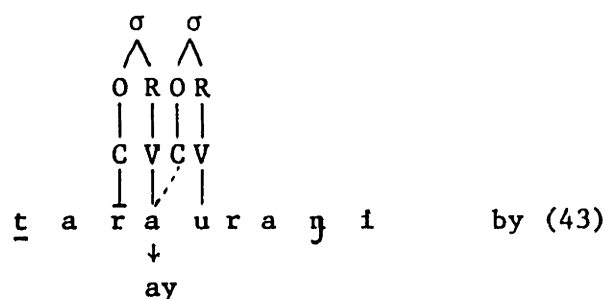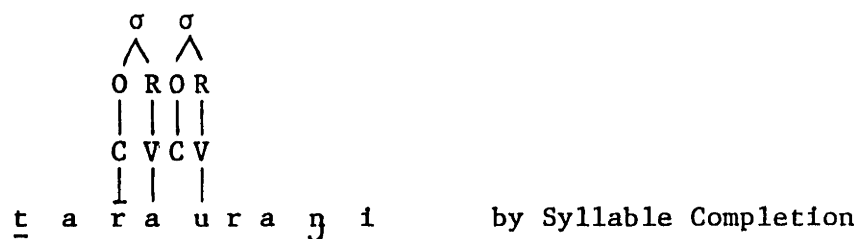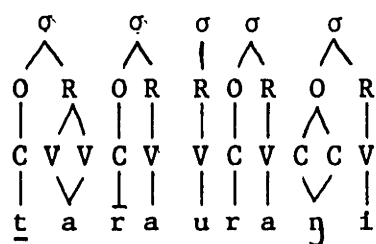
4.2.2. Lexical and Post Lexical Applications

The idea that the same phonological rule can apply both prior to and after lexical insertion makes interesting predictions regarding glide insertion. Recall that pauses are assigned after lexical insertion, and consequently, phonological operations prior to lexical insertion are sensitive to pauses, but not operations after lexical insertion. As predicted, the syllable completion principle is blocked by pauses across words, as (45), but not when it applies word internally, as in (46), (47), and (48):

(45) a. ewite maalu? 'where is Malu?'

b. maaluwewite? (scrambled version of a.)

c. maalu... ewite?

d. *maalu...wewite?

(46) Subcompounds

a. talayina, tala...yina, *tala...ina (tala , ina)
    pillow   head   mate

b. kasuwanti, kasu...wanti, *kasu...anti (kasu , anti)
    cashew nut   cashew   seed

(47) Cocompounds

kutiṛayaanayoṭṭakaŋŋaḷə , kutiṛa...yaana...yoṭṭakaŋŋaḷə
horses, elephants & camels

*kutiṛa...aana...oṭṭakaŋŋaḷə

(kutiṛa , aana , oṭṭakam)
horse elephant camel

(48) Inflection

a. taaṛa 'a name'; taaṛaye 'Tara-acc.'; taaṛa...ye
                                          *taaṛa... e

b. pasu 'cow'; pasuwinte 'cow-gen'; pasu...winte
                                      *pasu...inte

Subcompounds, cocompounds, and inflected words come out
of the lexicon with the glide in them; pause assignment does not
alter the glide. On the other hand, words like maalu and ewite come
out of the lexicon without the glide, and when they are put together,
pause assignment prevents glide insertion.

4.3. ə Insertion

4.3.1. Word Final ə Insertion

While glide insertion is the result of the syllable comple-
tion of unattached rimes, ə insertion is the syllable completion of
unattached onsets. The rule is illustrated by the following examples:

    (49) a. kaṭṭ ayaccu   'sent the letter'
        b. ayaccu kaṭṭə   ((49a) scrambled)
        c. kaṭṭə kiṭṭi   'got the letter'

I assume that the ə in (49)b and c is an inserted one, rather
than assuming that (49a) is a case of ə deletion. Other vowels do not
delete in comparable environments:

    (50) a. meṭṭayayaccu   'sent the bed'
        b. ayaccu meṭṭa   ((50a) scrambled)

    (51) a. kaṭṭiyayaccu   'sent the knife'
        b. ayaccu kaṭṭi   ((51a) scrambled)

The occurrence of ə is restricted to the stem final position,
and after syllabic r (4.3.5.). The solution that inserts the ə
predicts such restrictions.

The generalisation regarding word final ə insertion is: ə
is inserted word finally after any consonant except m and n (in
colloquial Malayalam. In formal Malayalam, m, n, n, r, l, and l)

when not followed by a vowel. The syllabification of m<u>a</u>ram 'tree'

and ka<u>tt</u>ə 'letter' (underlying /ka<u>tt</u>/) are:

```
(52) a.    σ     σ          b.    σ
          /\    /\               /\
         O R   O R A           O R   O
         | |   | | |           | |  /\
         C V   C V C           C V  C C
         | |   | | |           | |   \/
         m a   r a m           k a    t
```
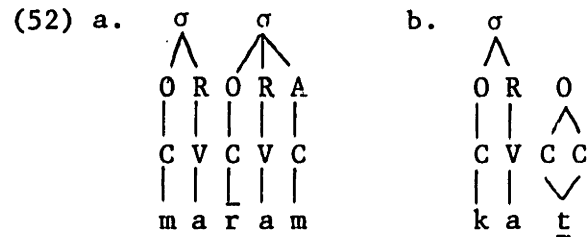
t in (52b) is an unattached onset; by the syllable completion

principle, it is assigned a rime:

```
(53)    σ    σ              σ     σ
       /\    |             /\    /\
      O R    O           O R    O  R
      | |   /\     ──→    | |   /\
      C V   C C           C V   C C
      | |    \/           | |    \/
      k a     t           k a     t
```
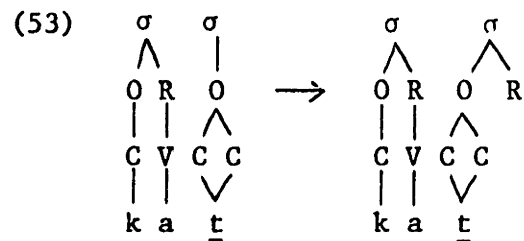
The following rule gives the surface form ka<u>tt</u>ə :

(54) Empty Rime Rule   (Domain : Post Lexical)

   Fill empty rimes with ə .

The derivation of (49)a and c are given below:

```
(55)    σ        σ    σ              σ     σ    σ     σ
       /\       /\   /\             /\    /\   /\    /\
      O R   O  R O  R  O  R        O R   O  R O  R  O  R
      | |  /\  | |  | /\  |   ──→   | |  /\  | |  | /\  |
      C V  C C V C  V C C V        C V  C C V C  V C C V
      | |   \/ | |  | |  \/ |       | |   \/ | |  | |  \/ |
      k a    t a y  a  c   u        k a    t a y  a  c   u
```

(56)

```
        σ               σ     σ
       /\              /\    /\
      O  R  O         O  R  O  R
      |  | /\         |  | /\ |
      C  V C  C       C  V C  C  V
      |  |  \/        |  |  \/  |
      k  a   t        k  i   t  i
             -            .
```

```
       σ    σ      σ    σ
      /\   /\     /\   /\
     O  R O  R   O  R O  R
     |  |    /\  |  |    /\ |
     C  V  C  C  C  V  C  C  V
     |  |   \/   |  |   \/  |
     k  a    t   k  i    t  i
             -            .
```
Syllable Completion
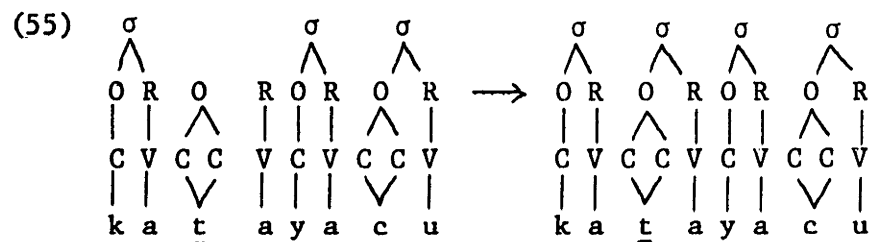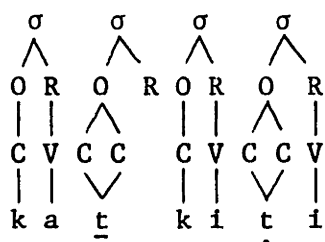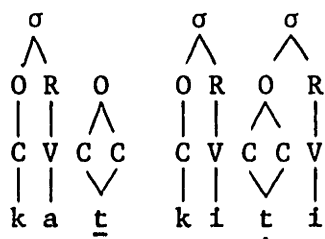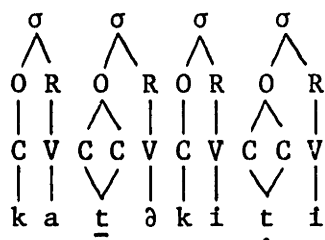
```
      σ    σ      σ     σ
     /\   /\     /\    /\
    O  R O  R   O  R  O  R
    |  | /\ |   |  |  /\ |
    C  V C  C V C  V  C  C V
    |  |  \/  | |  |   \/  |
    k  a   t  ə k  i    t  i
           -             .
```
Empty Rime Rule

4.3.2. ə Insertion in Compounds

Cocompounds behave exactly like phrases as far as ə insertion is concerned, inserting ə before consonants and not before vowels, when a stem ends in a consonant other than m and n:

(57) a. waal ulak'k'a parica kalə 'swords, pounding sticks, &
        .   .   shields'

                            (waalə 'sword'; ulak'k'a
                                                .
                            'pounding stick'; parica
                            'shield')

     b. ulak'k'a waalə parica kalə   (same as (57a))
                    .   .

ə appears after waal in (57b) because the stem final l is
                      .   .

followed by a consonant; in (57a), it is followed by a vowel.

    ə insertion takes place in subcompounds as well, but with
one crucial difference. If the stem final onset of the first stem and
the stem initial onset of the second stem can be combined into a
single onset, ə is not inserted. Compare (58) and (59):

(58) a. waalttala 'sword t' ' (waalə 'sword'; tala 'head'

                              (cf: waalə potti 'sword broke')

     b. aaŋkutti 'boy' (aanə 'male'; kutti 'child'

                              (aanə karaññu 'the man cried')

     c. paalppaatram 'milk vessel' (paalə 'milk'; paatram

                              vessel. cf: paalə pirinñu 'the

                              milk split')

(59) a. kaatʼəpuucca 'Tom cat' (kaatə 'forest'; puucca 'cat')

     b. paattəmaasə 'music tᵉᵃˢher' (paattə 'song'; maasə
                                                'teacher')

     c. cempəpaatram 'copper vessel' (cempə 'copper';
                                                paatram 'vessel')

    The sequence ttm, ttp, mpp (in (59)) are impossible morpheme
internally in Malayalam. It is precisely such sequences which trigger
ə insertion in (59), showing that these sequences are disallowed
across morphemes as well. On the other hand, ltt, ŋk, lpp, etc.,
which are possible morpheme internal sequences, are the ones which
do not require ə insertion (in (58)). What this means is that ttm,

ttp, mpp are not possible onsets in Malayalam, and therefore, ə

insertion must apply in order to avoid nonsyllabifiable strings,

while sequences in (58) are possible onsets, and therefore, no re-

adjustment is necessary.[9]

A clearer contrast is provided by the following examples:

(60) a. paalšakṭi 'strength from milk' (paalə 'milk';

šakṭi 'strength')

b. paaləsooppə 'milk shop' (sooppə 'shop')

lš is a possible onset in Malayalam, while lṣ is not: hence

the ə in (60b).

In order to account for the absence of ə in (58) and (60a),

I assume that adjacent onsets at the subcompounding stratum are

resyllabified into a single onset:

(61) Onset Fusion (Domain : Stratum 2)

O O ⟶ O

(61) obeys the constraints on possible onsets: it applies

only if the result is an allowable onset.

The derivation of (60) is given below:

```
(62) a.   σ        σ    σ      b.  σ          σ
         /\       /\   /\         /\         /\
        O  R     O O  R  O R     O  R  O    O  R  O
        | /\     | | |  | /\      | /\ |    | /\ /\
        C V V   C C V  C C V     C V V C    C V V C C
        | \/     | | |  | | |     | \/ |    | \/  \/
        p  a    l s a  k t i     p  a  l    s  o  p
                  ⌣       _                 •

         d̚       σ    σ
        /\       /\   /\
        O  R    O  R  O  R
        | /\    /\ |  | /\ |                  _____   onset fusion
        C V V  C C V  C C V
        | \/    | |_|  | | |
        p  a   l s a  k t i
                 ⌣      _

                            σ      σ  σ        σ    syll. com-
                           /\     /\ /\        /\      pletion
         _____          O  R   O R O  R    O  R
                           | /\   | | | /\     /\
                           C V V  C  C V V C C
                           | \/   |   | \/  \/
                           p  a   l   s  o  p
                                       •

                            σ      σ  σ        σ    empty rime
                           /\     /\ /\        /\        rule
         _____          O  R   O R O  R    O  R
                           | /\   | | | | /\   /\ |
                           C V V  C V C V V C C V
                           | \/   | | | | \/  \/ |
                           p  a   l ə s o  p  ə
                                      •
```
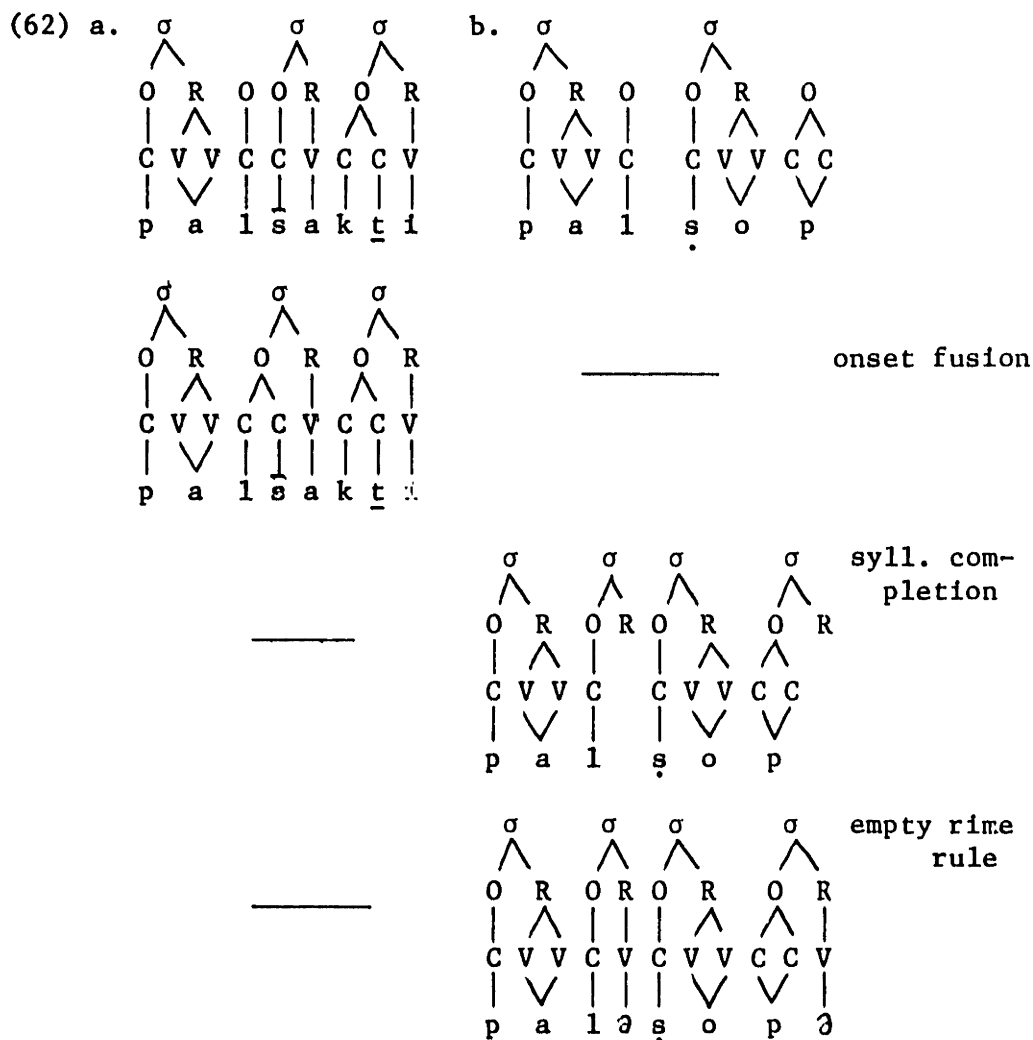
Observe that onset fusion does not take place in cocompounds and phrases because the domain of the rule is restricted to the sub-compound stratum. Hence the contrast between, say, (57b) and (58a).[10]

4.3.3. Lexical and Post Lexical Applications

The contrast between word internal and word external glide insertion is parallel to the contrast between word internal and word external ə insertion. Though ə is not inserted before a vowel, it

is, before a pause, word finally:

(63) kaatewite 'where is the forest?' vs kaatə...ewite

ə is not, however, inserted before a pause, word internally:

(64)a. Subcompounds

kaattaana 'wild elephant'; but *kaattə...aana

(cf: kaa...ttaana)

b. Cocompounds

waalulak'k'akal 'swords & pounding sticks'; but

*waalə...ulak'k'akal

c. Inflected Words

rasiidinte 'Rashid's'; but *rasiidə...inte

Under the assumption that syllabification is cyclic (cf:
Kiparsky (1979)), these results follow from Lexical Phonology.

4.3.4. ə in Datives

My account of ə essentially says that it appears after un-
attached onsets. Since m and n appear as appendices in both formal
and colloquial Malayalam, they do not create unattached onsets, and
the theory predicts that ə will not appear after m and n. This predic-
tion is violated in a special instance, that of the dative case of
nouns ending in n. A sample list of case forms in Malayalam is:

(65)

	nom.	acc.	dat.	gloss
a.	kuṭṭi	kuṭṭiye	kuṭṭik'k'ə	child
b.	makal(ə)	makale	makal(ə)kkə	daughter
c.	maram	maraṭṭine	maraṭṭinə	tree
d.	pasu	pasuwine	pasuwinə	cow
e.	waalə	waaline	waalinə	sword
f.	baalan	baalane	ɬaalanə	boy

It is clear that the stem for the nonnominative form in
(65) c-f are mara-ṭṭ-in, pasu-in, waal-in, and baalan. In the dative
case, ə is attached after n to these stems. If n is an appendix, how
is the ə accounted for?

I suggest that the dative is a suffix of the form $-\overset{R}{\underset{V}{|}}$,
with no segmental melody:

(66)

Dative suffixation

Resyllabification
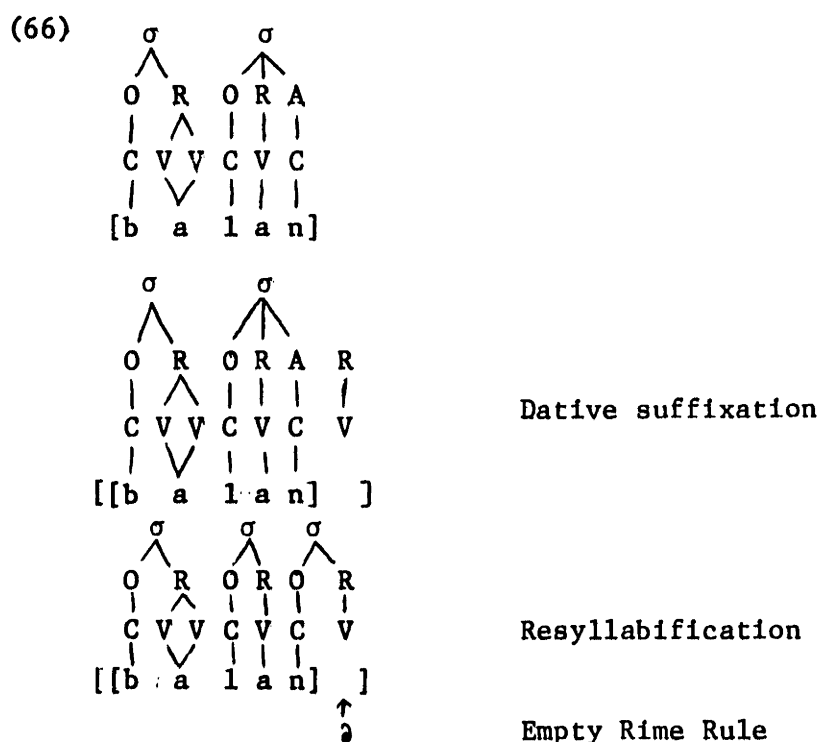
Empty Rime Rule

4.3.5. Syllabic r

Another context in which empty rime filling applies is after syllabic r. That it is necessary to postulate a distinction between syllabic and nonsyllabic r in Malayalam, though the distinction is neutralised on the surface, has been amply argued for in Warrier (1976). Her arguments are summarised as follows:

(67) a. Morpheme internally, ə can occur only after r, in words like <u>rəṣi</u> 'sage'. This fact can be accounted for by assuming an underlying /r̩ṣi/ (to distinguish it from examples like /raaṇi/ 'queen'), and a rule that inserts ə after r̩.[11]

   b. The rule of intervocalic voicing of stops (4.1.5.) applies in the environment of V — r̩, but not V — r: [sugrədi] 'good deed' (from /sukr̩ti/), but [sukriya] 'good deed' (from /sukriya/).

   c. The postulation of syllabic r, along with the independently required vowel sandhi rule (2.1.1.4.), explains the contrast between a and aa in <u>maharṣi</u> 'great sage' (from mahaa 'great', and r̩ṣi 'sage') and mahaaraaṇi 'great queen').

One can account for the insertion in <u>rəṣi</u> by adding the following rule to the grammar:

(68) Desyllabification (Domain : Post Lexical Stratum)

$$V \rightarrow C/ \underset{r}{\top}$$

(68) detaches r from the rime node and attaches it to an existing or newly formed onset node. Consider the derivations of rəṣi and maharṣi:

(69)

```
        σ  σ           σ   σ      σ   σ
        |  ∧          ∧   ∧      |   ∧
        R  O R        O R O  R    R  O R
        |  | |        | | |  ∧    |  | |
        V  C V        C V C V V   V  C V
        |  | |        | | | V     |  | |
        r  ṣ i        m a h a     r  ṣ i
```

                    [[m a h a][r ṣ i]]   Compounding

_____   _____   Onset fusion

_____   _____   Syll. completion

_____   Vowel Sandhi

```
                σ    σ          σ
               ∧    ∧          ∧
               O R  O  R       O R
               | |  |  ∧       | |
               C V  C V      V C V
               | |  | |      | | |
              [[m a h a]  [r ṣ i]]
```

                          ⋮

```
        σ   σ           σ   σ      σ
       ∧   ∧          ∧   ∧      ∧
       O R O R        O R O R     O R
       ⋮ | | |        | | | |     / | |
       C   C V        C V C V C C V
       |   | |        | | | | | | |
       r   ṣ i        m a h a r ṣ i
```
                                          Desyllabification

r ə ṣ i   _____   Empty Rime Rule

Subcompounding stratum
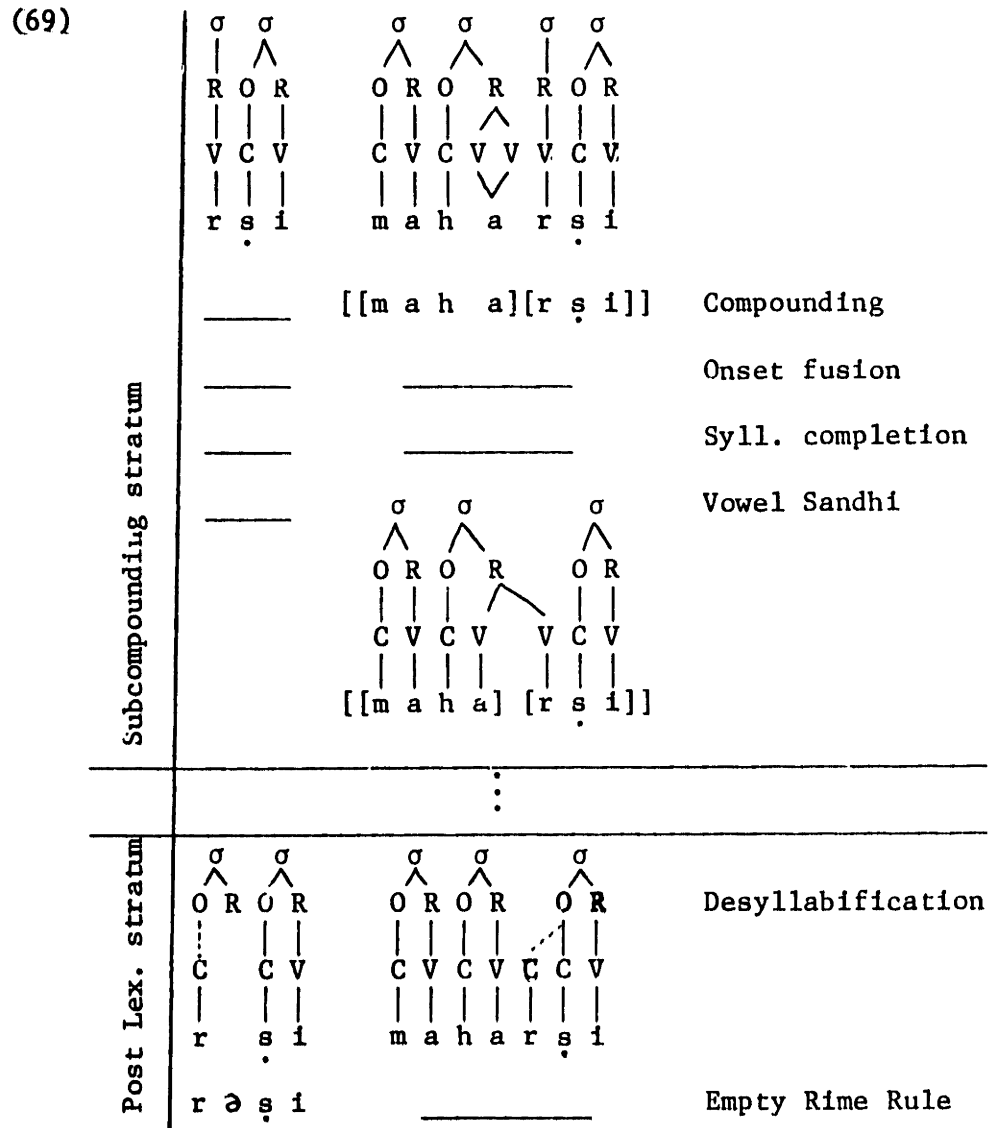
Post Lex. stratum

ə appears in rəṣi but not in maharṣi because at the appropriate

stage in the derivation (the post lexical stratum), the former has
an empty rime but not the latter.

It is important to note that desyllabification does not
apply at stratum 2. If it did, and the application of vowel sandhi
followed desyllabification, we would have unacceptable forms like
*mahaarǝsi. Evidence from a different area forces us , in fact, to
conclude that desyllabification is a post lexical rule. Observe that,
in order to account for forms like [sugrǝdi] (/sukr̥ti/), wǝ must
assume that desyllabification takes place after stop voicing (2.2.2.).
If not, the contrast between /sukr̥ti/ and /sukriya/ would be lost,
and voicing would not take place. We know that voicing is a post lex-
ical rule, and since desyllabification follows voicing, it also must
be post lexical.

An advantage of the metrical treatment of ǝ insertion is that
it provides a unified account of the insertion at the end of words,
at the end of stems in subcompounds and cocompounds, and after sylla-
bic r. Within a segmental treatment, the fact that it is the same
vowel that is inserted in these environments would be an accident.

Other Dravidian languages like Tamil and Telugu offer strong
support for the assumption that the vowel insertion after syllabic r
and word finally are, in fact, the same process. In Telugu, the English
word belt is pronounced [belṭu] as a consequence of syllable completion
and empty rime filling: the empty rime filling vowel in Telugu is u. It
is the same vowel that appears after syllabic r also: /r̥si/ in Telugu is

[ruṣi]. In Tamil, <u>belt</u> and /rṣi/ are <u>belti</u> and [rɨṣi].

4.4. Vowel Sandhi

4.4.1. Derivations

Recall that vowel sandhi has the derivational, subcompound-
ing, and cocompounding strata as its domain. There is a derivational
process (2.1.4.) in Malayalam, restricted to words of Sanskrit origin,
which inserts <u>a</u> to the left of the first rime in the word:

(70) a. alasam 'lazy'   ; aalasyam 'laziness'   : a ∼ aa
     b. kumaaran 'boy'   ; kaumaaryam 'boyhood'   : u ∼ au
     c. wikalam 'distorted' ; waikalyam 'distortion' : i ∼ ai
     d. diinam 'illness'   ; dainyam 'ill-looking' : ii ∼ ai
     e. kruuram 'cruel'   ; krauryam 'cruelty'   : uu ∼ au
     f. krəṣi 'farming'   ; karsakan 'farmer'   : r̩ ∼ ar

What is inserted, I suggest, is the melody <u>a</u> at the melodic
tier: given the assumptions regarding the linking of the melody to
the skeleton, the desired consequences follow. In the case of nonbran-
ching rimes, it creates a new branch ((70)a-c):

(71)   σ   σ
           /\   /\
          O R   O R
          | | | /\
          C V   C V V
          | | | | |
          u i   →   u a i
        wikalam   waikalyam
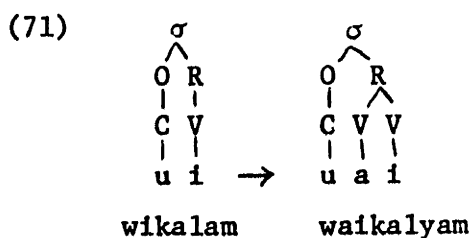
In the case of rimes which are already branching ((70)d, e),

the addition of the new melody does not create a new branch, since

a rime can have at the most two branches in Malayalam. Consequently,

the linking of the a to the V results in a delinking, once again,

following from general principles:

(72)

```
      σ              σ
     /\             /\
    O   R          O   R
    |   /\         |   /\
    C  V  V        C  V  V
    |   \/         |  |  |
    d    i         d  a  i
```

d̲iinam   d̲ainyam
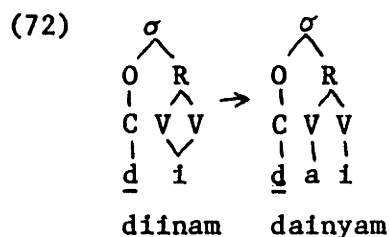
The solution sketched above accounts for (70f) as well:

(73)

```
     σ    σ            σ    σ
    /\   /\           /\   /\
   O R  O R          O  R  O  R
   | |  | |          |  /\ |  |
   C V  C V    →     C V V C  V
   | |  | |          | | | |  |
   k r  s i          k a r s  a  kan
        •                  •
```
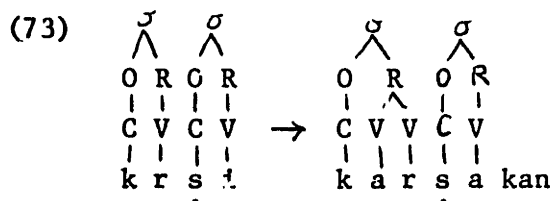
The insertion of ə in kr̩si is a consequence of the desylla-

bification of r̩, which creates an empty rime node, as in r̩əsi ((69)).

The absence of ə in karṣakan is similar to that in maharṣi: the desy-

llabification does not result in an empty rime node, and hence there

is no ə insertion.

4.4.2. Compounds

The kinds of alternations in derivational morphology sketched

in 4.4.1. are found in compounding as well, with a few additional com-

plications. Given below are examples of subcompounds which illustrate

the phenomena (cocompounds exhibit identical alternations):

(74)a. a][a → aa : niila][ambaram → niilaambaram
                    blue   sky

   b. a][aa → aa : niila][aakaasam → niilaakaasam
                     sky

   c. aa][a → aa : mahaa][atbhutam → mahaatbhutam
                    great   wonder

   d. aa][aa → aa : mahaa][aanandam → mahaanandam
                     pleasure

(75)a. i][i → ii : rawi][indran → rawiindran
                    sun   god   sun god

   b. i][ii → ii : rati][iiswaran → ratiiswaran
                     a name god   Cupid

(76)a. a][au → au : wrksam][ausadham → wrksausadham
                     tree   medicine

   b. aa][au → au : mahaa][ausadham → mahausadham

   c. a][ai → ai : janam][aikyam → janaikyam
                    people unity   people's unity

   d. aa][ai → ai : mahaa][aiswaryam → mahaiswaryam
                     godliness

(77)a. a][r → ar : saptam][rsi → saptarsi
                    seven   sage the seven sages

   b. aa][r → ar   mahaa][rsi → maharsi

Consider what happens to the syllable trees:

(78) mahaatbhutam ((=74c))
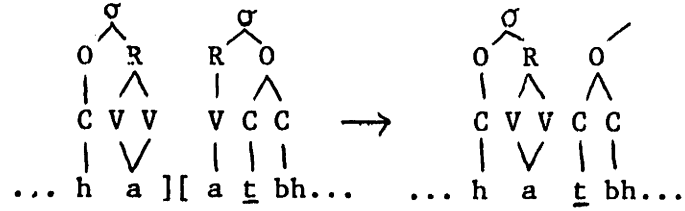
If we look at the change in (78) as one involving the erasure of the rime node, leaving the melody tier intact, the facts illustrated in (74) - (77) become identical to (70).

I propose a rule that erases a rime node when followed by another in Sanskrit compounds:

(79)   Rime Erasure   (Domain : Strata 2, 3)

$$R \rightarrow \emptyset / \text{---} R$$
$$+skt.$$

Since a rime node can have at the most two V's, the left over melodies which do not get linked do not surface.[12] The derivations for (74a), (76)a, d, and (77b) are given below:

(80)a.   niilaambaram
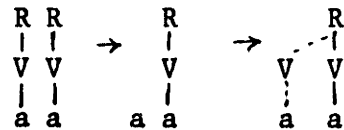
b.   wrksausadham

c. mahaíswaryam

```
R     R            R            R
/\    /\           /\           /\
V V   V V    →    V V    →    V V
 \/   | |         | |          | |
 a    a i         a a i        a i
```

d. maharsi
       .

```
R   R           R            R
/\  |      →    |      →    /\
V V V          V          V V
 \/ |          |           | |
 a  r          a r         a r
```
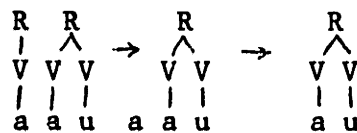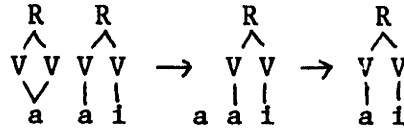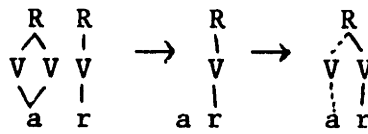
Now, both rime erasure and glide insertion apply to adjacent
rimes, and one bleeds the other. Rime erasure is a less general rule,
applying at strata 2-3 to words of Sanskrit origin, while glide in-
sertion (resulting from the principle of syllable completion) applies
at all strata, and it applies to both Sanskrit and Dravidian words.
I shall assume that rime erasure is a rule, and that syllable comple-
tion is a general convention. In n̲ii̲laambar̄am ((74a)), rime erasure
reduces the two rimes to a single rime, eliminating the need for syl-
lable completion. In the Dravidian compound panayoola 'palm leaf'
(from pana 'palm'; oola 'leaf'), rime erasure does not apply, and
syllable completion, therefore, does apply, resulting in glide in-
sertion.

We are now ready to tackle vowel combinations of a more
complicated sort. To begin with, i and u become glides when followed
by a vowel.

(81)a. i][ai → yai : r̄ati̱][aiśwar̄yam → r̄aty̱aiśwar̄yam
a name   godliness   godliness of sex

b. i][au → yau : r̄ati̱][auṣadham → r̄aty̱auṣadham
medicine   erotic medicine

c. u][ai → wai : śiśu][aiśwar̄yam → śiśwaiśwar̄yam
child

d. u][au → wau : śiśu][auṣadham → śiśwauṣadham

e. i][a → ya : pati̱][ambar̄am → paty̱ambar̄am
husband   sky   a sky of husbands

f. u][a → wa : śiśu][ambar̄am → śiśwambar̄am
a sky of children

These examples can be accounted for by assuming that ⌐ high
vowel becomes nonsyllabic (i.e. linked to a C element in the skeleton)
when followed by another vowel.

(82) Glide Formation

$$V \longrightarrow C / \underset{[+\text{high}]}{\underline{}} \quad [-\text{cons.}]$$

(82) creates a new C element and links the vowel to it. The
following derivation illustrates the application of (82):

(83) r̄at̲yausa̲dham

```
        O R      R
        | |      /\
        C V      V V
        | |      \ |
.... t̲ i      a u ...
```

```
        O    R    R
        /\   |    /\
       C C   V   V V
       | |   X   | |
       t̲ i       a u        Glide Formation
```

```
        O         R
        /\        /\
       C C       V V
       | |       | |
       t̲ i       a u        Rime Erasure
        ↓
        y
```

It is crucial that t̲y in (83) is seen as a branching onset
(C̰ C), rather than as a nonbranching onset(C̰). Recall that nonbran-
ching onsets voice noninitially in Malayalam, while branching onsets
do not (4.1.5.). Since voicing does not apply to t̲y in (83)
(*r̄a[d̲]yausa̲dham), we must conclude that t̲y is a branching onset.

As stated earlier, i is realised phonetically as y when linked
to a C element. The linking of C̰ to an appropriate onset node
follows from general principles.

I now turn to the phenomenon of mid formation. If the second
element in vowel sandhi is a high vowel, the result is a mid vowel
agreeing in other features with this vowel:

(84)a. a][i → ee : kala][indrān → kaleendrān
art god

b. aa][i → ee : mahaa][indrān → maheendrān
great

c. a][ii → ee : kala][iiswaran → kaleeswaran
god

d. aa][ii → ee : mahaa][iiswaran → maheeswaran

(85)a. a][u → oo : kala][utsawam → kalootsawam
festival

b. aa][u → oo : mahaa][utsawam → mahootsawam

c. a][uu → oo : looham][uurjam → loohoorjam
metal energy energy from metal

d. aa][uu → oo : mahaa][uusmalam → mahoosmalam
warmth

The process of mid formation can be characterised as a rule that applies at the melody tier, changing a sequence of a vowel melody and another high vowel melody into a single mid melody. The rule is stated as follows:

(86) Mid Formation (Domain: Strata 7,3)

$$[-\text{cons.}] \begin{bmatrix} +\text{high} \\ \alpha\text{back} \end{bmatrix} \longrightarrow \begin{bmatrix} -\text{high} \\ -\text{low} \\ \alpha\text{back} \end{bmatrix}$$

The application of the rule is illustrated in (87):

(87) kaleend̲r̄an

```
     R   R
     |   |
     V   V
     |   |
    kala  i̲n̲d̄ran
```

_____   Glide Formation

```
        R
      .-᷄|
    V   V
    |   |
    a   i
```
Rime Erasure

```
       R
      ᷄|
    V  V
     \_/
  kal  e  n̲d̄ran
```
Mid Formation

It is important to note that the skeletal tier remains un-
disturbed by the application of the rule, and the rime structure

remains $\overset{R}{\underset{V\ V}{\bigwedge}}$, independent of the changes at the melody tier.

Recall that sequences like i][a and u][a result in ya and wa:
the vowel after a glide is a short one, which is an automatic result
of the desyllabification of i and u. The surface forms of sequences
like i][u, u][i, and u][u, however, are yoo and woo, with a long vowel
after the glide:

(88)a. i][u → yoo : pat̲i][u̲tsawam → pat̲yoot̲sawam
                      husband festival   festival of husbands

b. u][u → woo : s̄is̄u][u̲tsawam → s̄is̄woot̲sawam
                  child

c. u][i → wee : s̄is̄u][i̲nd̄ran → s̄is̄weend̲r̄an
                  god

d. u][ii → wee : s̄īs̄u][īis̄waran → s̄īswēēs̄waran
god

e. u][uu → woo : s̄īs̄u][uuṣmaḷam → s̄īswōōṣmaḷam
warmth   warmth of children

We saw that vowel sandhi applies only to words of Sanskrit origin. Now, it is a fact of Sanskrit phonology that it has no short mid vowels, even though it has short as well as long high and low vowels. Clearly, it is this property of Sanskrit which carries over to (88). Thus, there are no short e and o in words of Sanskrit origin in Malayalam; we need to stipulate a rule that lengthens mid vowels in these words.

(89) $\begin{bmatrix} -\text{high} \\ -\text{low} \end{bmatrix} \longrightarrow \overset{V\ V}{V}$ in Skt. words

The derivation of (88c) is:

(90) s̄īs̄wēēnḏran

```
      O R   R
      |  |   |
      C  V   V
    _ |  |   |   _
    s̄īs̄ u  īndr̄an

      O     R
     ⌐·.    |
      C  Ċ  V
      |  |  |
      s̄  u  i          Glide Formation

            R
            |
            V
            |
            e            Mid Formation

       O      R
      ⌒      ⌒
      C  C   V  V
      |  |   |  |
    s̄īs̄ u   e  ḗnḏr̄an    (89)
```

Mid formation and glide formation du not apply in sequences involving i][i as in (75), in contrast to the sequences in (88)b and e. In order to prevent the application of mid formation, it is necessary to stipulate that i][i is an exception to mid formation. i][i does not undergo glide formation because there are no sequences of the form <u>Cyi</u> or <u>Cwu</u> in Malayalam. In order to account for the absence of these sequences, I set up the following filter:

(91)   * C   C   V
          | |
       [α back] [+high]
                [α back]

4.5. Causatives

4.5.1. The Phonology of Causativisation

A causativised verb in Malayalam can be causativised further, thereby creating several 'layers' of causativisation:

(92)a. t̲innum   ; t̲iittum   ; t̲iittik'k'um
        will eat   will feed   will cause to feed

   b. ootum   ; ootik'k'um ; ootippik'k'um
       will run   will run   will cause X to run Y
         intr.   tr.

The devices used to express causation fall into two categories:

(93) Type A

    a. Denasalisation

        (i) uraŋŋum 'will sleep' ; urakkum 'will put to
                                                    sleep'

        (ii) kuumpum 'will fold ; kuuppum 'will fold
                            intr.'   tr.'

      (iii) unnum 'will eat'   ; uuttum 'will feed'
                ..   ..

    b. Gemination

        (i) aatum 'will swing   ; aattum 'will swing
            •   intr.' · ·ν ··   tr.'

        (ii) murukum 'will tighten; murukkum 'will tighten
                            intr.'   tr.'

      (iii) maarum 'will change ; maattum 'will change
                            intr.'   tr.'

    c. Suffixation of -u<u>tt</u>

        (i) wa̅rum 'will come'   ; wa̅ru<u>tt</u>um 'will make X
                                                    come'

        (ii) coo̅rum 'will leak'   ; coor<u>tt</u>um 'will cause to
                                                    leak'

      (iii) ketum 'will go off' ; ketu<u>tt</u>um 'will put off'
              •   • --

(94) Type B

    a. Suffixation of -<u>ikk</u>

        (i) ootum 'will run   ; ootik'k'um 'will run
            •   intr.'   •   tr.'

        (ii) parayum 'will say'   ; parayik'k'um 'will make
                                                    X say'

      (iii) cee̅rum 'will join   ; ceerkkum 'will join
                            intr.'   tr.'

    b. Infixation of -<u>ipp</u>-

        (i) ootik'k'um 'will run ; ootippik'k'um 'will make
            •   tr.'   •   X run Y'

(ii) pathik'k'um 'will study' ; pathippik'k'um 'will
                                                teach'
(iii) dukkhik'k'um 'will grieve'; dukkhippik'k'um 'will
                                            cause X to grieve'

How do we formally relate the processes of gemination and

nasalisation to causativisation?

I assume that a word formation process can consist of a set

of operations which specify semantic, syntactic, and phonological

changes. This idea is inherent in the treatments of lexical rules

like passivisation in the Lexical Functional Grammar. Thus, Bresnan

(in press: a) formulates the passive rule for English as follows:

(95) Passive

a. (SUBJ) → ∅ / (BY OBJ)
    (OBJ) → (SUBJ)

b. V → V part.

(95a) gives the syntactic changes in the lexical entry of

the word; (95b) gives the morphological change. This format can be

extended to derivational processes as well. Thus, the effect of -er

attachment (employ ~ employer) can be characterised as:

(96) The Agent Rule

a. $]_V$ → $]_V$ er$]_N$

b. Pred 'F $\langle$(SUBJ)..$\rangle$' → Pred 'X such that
                                    F$\langle$(SUBJ)..$\rangle$ is true'
                                       X

(96a) specifies the morphological and syntactic change,
(96b) specifies the semantic change.

We can now formalise the causativisation in (93)a-c as:

(97) Type A Causativisation

    a. (SUBJ) → (OBJ)

        ∅ → (SUBJ)

              agent

    b. Pred 'F' → Pred 'cause F'

    c. Onset → $\wedge$ / —— $]_V$ (gemination)
     or

       Onset → [-nasal]/ —— $]_V$ (denasalisation)
     or

       $]_V$ → $]_V$ utt$]_V$ (-utt suffixation)

(97a) specifies the syntactic change: the existing subject
becomes an object, and an agent subject is inserted into the lexical
form. (97b) specifies the semantic change. (97c) corresponds to the
phonological and morphological changes, the choice of which is deter-
mined idiosyncratically by the verb.

Alternately, one may assume that processes like gemination
and denasalisation are triggered by an abstract 'cause' morpheme
suffixed to the verb:

(98) a. Onset → $\wedge$ / —— $]_V$ 'cause'$]_V$

    b. Onset → [-nasal]/ —— $]_V$ 'cause'$]_V$

(98) is strictly within the assumption that all morphological operations are cases of affixation or compounding (McCarthy (1979), Lieber (1980)). That this assumption is too strong, and that the richer mechanism illustrated by (97) is needed in morphology, is shown by the a...y derivation in alternations like anukuulam 'supporting -Adj.' ~ aanukuulyam 'support -N', and cancalam 'fickle' ~ caancalyam 'fickleness' (for more examples, see 4.4.1.:(70)). The derivational change from A to N is accompanied by two phonological operations: a is inserted at the beginning of the first rime, and y is inserted at the end of the last onset:

(99)   A → N   Rule

   a. $]_A$ → $]_N$

   b. $\emptyset$ → a / [σ
                          R
                       V

      $\emptyset$ → i /   σ]
                        O
                         C

If the categorial change from A to N, and the phonological changes of a and y insertion were not directly associated, and the phonological rules applied to the output of the syntactic or morphological component, the facts would become less amenable to description. We would have to postulate an abstract morpheme morphemd which changes A to N, and triggers the rules of a and y insertion. The disadvantage of this solution is that we would have to allow the abstract

morpheme $\emptyset_k$ in (100), attached at one end of the stem, to effect a phonological change at the other end:

$$(100) \quad R \rightarrow \underset{\substack{\text{V} \\ \mid \\ \text{a}}}{R} / [[\sigma \ Q]_A \ \emptyset_k]_N$$

The only way to formulate such a rule is by resorting to the use of variables (Q), which is unmotivated in phonology.

4.5.2. The Two Strata of Causativisation

Type A causativisation and type B causativisation take place at two different strata. Stratum A (=type A) causativisation is less productive, and applies to intransitive verbs only. Among the intransitive verbs, it applies only to some; the choice between gemination, nasalisation, and -utt suffixation is determined by the verb idiosyncratically. In contrast, stratum B (=type B) causativisation is quite general, applies to both transitive and intransitive verbs, and the choice between the two kinds of suffixation is not arbitrary. This difference in productivity is not unlike the difference in productivity between stratum I affixation and stratum II affixation in English (Allen (1978)).

There is also a semantic distinction between A and B. Causatives derived through A are unambiguously of the 'direct causative' type, while those belonging to B could have either the direct or the indirect reading of causation. A direct causative is one in which the

causer is the agent of the action (e.g. eat~feed, in English), while

an indirect causative is one in which the causer is not the agent,

and gets the action performed in an indirect fashion (e.g. eat ~

cause to eat, in English) (see Masica (1976) for a discussion of

direct and indirect causatives in Indian languages).

Causatives derived through A can undergo further causativi-

sation through B, but the reverse is not true:

(101) a. tinn ; tiitt   ; tiittik'k'
        eat   feed   cause to feed

      b. war   ; warutt   ; waruttik'k'
        come   make X come cause to make X come

(102) a. tinn ; tinnik'k'   ;*tinnik'k'utt
        eat   cause to eat

      b. oot   ; ootik'k'   ;*ootik'k'utt
        run   run -tr.

These results would follow from the ordering of stratum A

before stratum B.

In sum, differences in morphological productivity, semantic

properties, and distributional criteria distinguish type A causatives

from type B causatives, and justify the decision to assign the two

processes to two different strata.

4.5.3. Opacity vs Bracket Erasure

The facts of causativisation are of particular interest because

they provide crucial evidence in favour of the Opacity Principle.
A characteristic of -ikk is that it can be attached to any verb
which does not have a branching structure ([[X] Y]). Thus,
from the simple verb wiḻakk 'to weld', and from the derived verb
mayakk 'to hypnotise' (from mayaṇṇ 'to doze'), neither of which
contain any suffixes, we get wiḻakkik'k' 'to cause to weld', and
mayakkik'k' 'to cause to hypnotise'. However, -ikk causativisation
is impossible in (103)-(106)b :

(103) a. [[ḏukkh]ik'k'] 'to grieve' (cf: ḏukkham 'grief'
                                         ḏukkhiṯan 'one who is
                                                 grieving')

      b. * ḏukkhik'k'ikk

      c.   ḏukkhippik'k' 'to make X grieve'

(104) a. [[paṯh]ik'k'] 'to study' (cf: paaṯham 'lesson')
      b. * paṯhik'k'ikk

      c.   paṯhippik'k' 'to teach'

(105) a. [[mara]kk] 'to forget' (cf: marawi 'forgetfulness')

      b. * marakkik'k'

      c.   marappik'k' 'to cause to forget'

(106) a. [[ooṯ]ik'k'] 'to make X run'

      b. * ooṯik'k'ikk

      c.   ooṯippik'k' 'to cause X to run Y'

What (103)-(106) show is that -ikk suffixation must necessarily
see the internal structure of the stem. The -ikk in (103)-(105)a is

the stem forming -ikk, and that in (106a) is the causative -ikk.
The stipulation that -ikk cannot be attached to a branching struc-
ture rules out the unacceptable examples above, while allowing
wilakkik'k' and mayakkik'k', which do not have branching structures.

This solution has interesting consequences for the choice
between the Bracket Erasure Convention (2.1.7.) and the Opacity
Principle. According to the Bracket Erasure Convention, the last rule
in the cycle of [[mara]kk] ((105)) will erase the internal brackets,
giving [marakk], thereby making it identical to mayakk in structure.
If these are the forms to which -ikk causativisation is applicable,
there is no reason why it should apply to one and not to the other.

The Bracket Erasure Convention thus destroys valuable morpho-
logical information, and is too strong an empirical hypothesis. Given
the principle of Opacity, on the other hand, the right consequences
follow. All that the Opacity Principle demands is that the internal
brackets at one stratum be invisible to the processes at another stra-
tum, not to the processes at the same stratum. If the attachment of
the verb forming -ikk and the causativising -ikk take place at the
same stratum, -ikk suffixation can see the internal structure of
both dukkhik'k' and ootik'k'.

It is interesting that -ikk causativisation can apply to verb
forms which contain the causative suffix -utt: [[war utt]ikk] 'cause

X to make Y come'; [[walar̄ tt]ikk] 'cause X to grow Y'. Note that

these are branching forms, and yet -ikk can be attached to them,

because -utt is attached at a stratum prior to the attachment of

-ikk. Therefore, the internal structure of, say, war̄utt, is, by

Opacity, invisible to -ikk attachment, which treats it as a nonbran-

ching form.

Similar conclusions can be drawn from the facts of -ipp-

infixation, the restriction regarding which can be stated as: -ipp-

is inserted only in branching structures, preceding the last branch

(as shown by (103)-(106)). The following examples show that (i) -ipp-

cannot be attached at the end of a stem, and (ii) it cannot be insert-

ed into a nonbranching structure.

(107) a. *dukkhipp (cf: (103c))

b. *pathipp (cf: (104c))

c. *marapp (cf: (105c))

d. *ootipp (cf: (106c))

(108) a. wilakk 'weld' ; *wilappik'k'

b. mayakk 'hypnotise' ; *mayappik'k'

c. murukk 'tighten-tr,' ; *muruppik'k'

(107) shows that -ipp- is an infix, not a suffix. (108) shows

that it cannot be infixed into underived structures ((108a)), struc-

tures derived through denasalisation ((108b)), or structures derived

through gemination ((108c)): none of these are branching structures,

Since -utt is attached at stratum A, and -ipp- at stratum B, the bracketing assigned by -utt suffixation is, by the Opacity Principle, invisible to -ipp- infixation, and thus, -ipp- cannot be inserted in -utt stems:

(109) a. * waripputt   (cf: war ; warutt)

b. * talaripputt (cf: talar ; talartt)

Thus, the facts of -ikk and -ipp affixations show that these affixes must be sensitive to the presence or absence of prior suffixations at that stratum, but not affixation at a previous stratum. These results are consistent with the Opacity Principle, but not the Bracket Erasure Convention.

4.5.4. The Place of Causatives in the Grammar

The question that now arises is: where do the two strata of causativisation, which I called strata A and B, fit in with the rest of the grammar?

We can either assume that the two strata of causativisation constitute two strata of derivational morphology, similar to the two strata of derivation in English, or assume that type A causativisation belongs to stratum 1 (Derivational stratum), and type B causativisation belongs to stratum 2 (Subcompounding stratum). The first

solution would entail adding an extra stratum to the set we already

have:

      (110)   Stratum 1 : Derivations (A)
                 Stratum 2 : Derivations (B)
                 Stratum 3 : Subcompounding
                 Stratum 4 : Cocompounding
                 Stratum 5 : Inflections
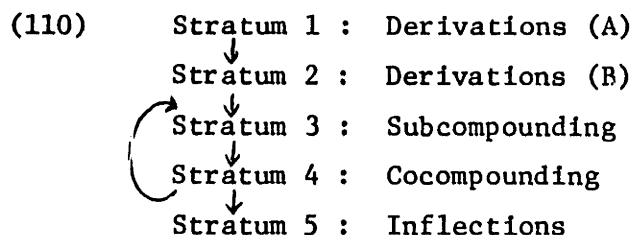
The second solution is less elegant in that it puts affix-

ation and compounding in the same stratum:

      (111)   Stratum 1 : Type A Causativisation and other
                               Derivations
                 Stratum 2 : Type B Causativisation and Subcom-
                               pounding
                 Stratum 3 : Cocompounding
                 Stratum 4 : Inflections
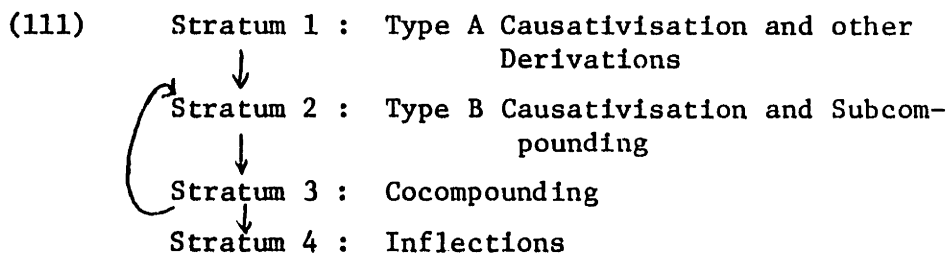
I shall reserve the choice between the two solutions for the

next section.

Type B causativisation cannot be associated with either

the cocompounding or the inflectional strata, because of the facts

of tone assignment. The word melody LH is assigned to the output of

the subcompounding stratum (2.1.3.). If cocompounding and type B cau-

sativisation belonged to the same stratum, the high tone would in-

correctly occur at the end of the stem to which the suffix is attached,

not at the end of the suffix itself.

```
            L  H      L  H
(112) a.  ootum  ;  ootunnu    (run + fut.;pres.)
          L ˙ H      L ˙  ̄ ̄H
      b.  ootikkum ;  ootikkunnu  (run+cause + fut.;pres.)
          L ˙   H   ˙L      H
      c.  ootippikkum ;  ootippikkunnu  (run+cause+cause+ fut.;
          •            •        ̄ ̄                         pres.)
```

Tone assignment ignores inflectional endings, but is sensitive

to causative endings. This contrast follows from the assumption that

type B causativisation takes place prior to tone assignment, and in-

flectional suffixation, after it. Thus, type B causativisation must

be associated with a stratum prior to cocompounding and inflection.

4.6. Verbal Compounds

4.6.1. The Morphology of Verbal Compounds

Malayalam has a verbal compound construction similar to

constructions like slave driver, movie goer in English:


```
(113) a.  ambaracumbi ' skyscraper '  (ambaram;  cumbikk;  cumbanam)
                                        sky        to kiss   kiss
      b.  kutirapreemi 'horse lover'(kutira;  preemikk;  preemam)
                                      horse    to love    love
      c.  gaganacaari 'sky traveller' (gaganam;  carikk  )
                                        sky        to travel
```


I shall assume that these words have the structure in (114),

and are generated by (115):

(114)

```
              N
        N⎯⎯⎯⎯⎯V
        |     |
    ambaṟam  cumb  -i
```
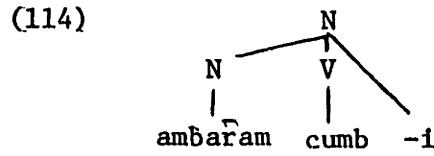
(115) N → N V -i̲

The ternary branching construction is chosen in preference
to the possible binary branching structures in (116), because neither
[N V] nor [V -i] ((117)) are possible words in Malayalam:

(116)

```
         N                        N
      ⎯⎯⎯⎯N                   V⎯⎯⎯⎯⎯
     N    V                  N    V
     |    |                  |    |
 ambaṟam cumb -i        ambaṟam cumb -i
```
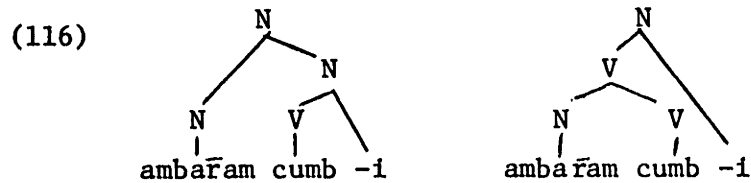
(117) a. * ambaṟacumb (ikk) ; * cumbi

b. * kuṯiṟapreem (ikk); * preemi

c. * gaganaca(a)r̄(ikk); * caar̄i

4.6.2. Nasal Deletion

Verbal compounds crucially differ from subcompounds and co-
compounds. In verbal compounds ([N+V+-i]), the stem final nasal of
the first stem deletes only if the second stem is 6f Sanskrit origin;
in subcompounds and cocompounds ([N+N]), the nasal deletes irrespec-
tive of the etymology of the second stem. Nasal deletion takes place
in ambaṟacumbi ((113a)), because cumb is of Sanskrit origin, but not
in maṟancaati 'tree swinger' (maṟam 'tree'; caaṭ 'jump'), because

caat̪ is a Dravidian verb. In contrast, deletion takes place in the

[N+N] compound maraccaṇṇala 'wooden chain', even though caṇṇala 'rope'

is Dravidian. Given below are more examples which bring out the Drav-

idian vs Sanskrit contrast in verbal compounds:

(118) Dravidian second element

    a. r̄asaŋkolli 'kill joy' (r̄asam ; koll)
                                      interest to kill

    b. wair̄amwiẓuṇṇi 'diamond swallower' (wair̄am ; wiẓuṇṇ)
                                             diamond swallow

    c. maanamṇookki 'sky gazer' (maanam ; ṇookk)
                                    sky   look

(119) Sanskrit second element

    a. janad̄roohi 'enemy of the people' (janam ; d̲r̄oohikk)
                                      people harm

    b. maargad̲ar̄si 'path shower' (maargam ; d̲ar̄sikk)
                                    path   see

    c. mat̲awid̲weesi 'religion hater' (mat̲am ; wid̲weesikk)
                                  religion hat̄e ·

The nasal deletion in verbal compounds can be stated as foll

follows:

(120) [+nasal] → ∅ / ——][+ skt.

Whether the nasal deletion in [N+N] compounds and in verbal

compounds should be collapsed into a single rule is an issue that I

will not go into. If necessary, it can be done as follows:

(121) [+nasal] → ∅ / ——] ⟨+v⟩⟨+skt⟩

4.6.3. The Branchingness Condition

An interesting property of verbal compounds is that the verb
has to be nonbranching. Thus, i can be attached to d̅r̅ooh and cumb,
but not to [[d̅r̅ooh]ikk] and [[cumb]ikk]: *janad̅r̅oohikki,
*ambar̅acumbikki. The verb can be a causativised one as long as the
process does not create new brackets:

(122) a. s̅akunammuṭakki 'ill omen' (s̅akunam ; muṭaŋŋ)
                                    omen   be prevented

      b. pir̅imurukki 'thread tightener' (pir̅i ; muruk)
                                         thread tighten-intr.

      c.*kuppiyuṭakki   (kuppi ; uṭa ;   uṭakk)
                          bottle   break-intr. break-tr.

The restriction on the branchingness of the stem in verbal
compounds parallels that in -ikk and -ipp- affixation. As in the
case of -i, -ikk cannot be attached to branching structures. Now, if
the attachment of -i is to treat [[d̅r̅ooh]ikk] and [[uṭa]kk] as
branching structures, it follows that verbal compounding must belong
to the same stratum as type B causativisation. At other strata, the
internal structure of type B causativised verbs will, by the Opacity
Principle, be invisible to verbal compounding, and hence, d̅r̅oohikk
will not be seen as a branching structure. Given that verbal compound-
ing and type B causativisation belong to the same stratum, our theory
predicts that the brackets created at a previous stratum, such as that
of type A causativisation, will be invisible to verbal compounding.
This prediction is fulfilled by compounds such as kaṭamwar̅utti

'debt bringer', in which the verb is derived through -uṭṭ suffix-
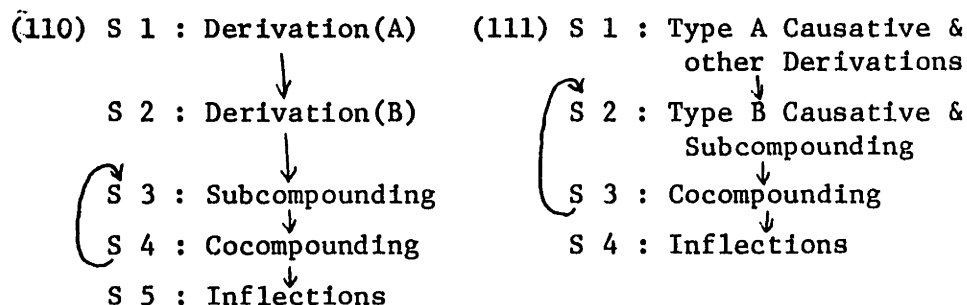ation: war̄, war̄uṭṭ.

Recall that we found that type B causativisation must precede
tone assignment. Since verbal compounding and type B causativisation
belong to the same stratum, verbal compounding must necessarily pre-
cede tone assignment. The facts confirm this prediction:

<pre>
            L   H
(123) a. ambar̄acumbi
            L   H H
       b. kuṭir̄apreemi
            L   H
       c. wair̄amwizuŋŋi
</pre>

If verbal compounding, like cocompounding, followed tone
                                              L   H L H
assignment, we would derive the unacceptable results *ambar̄acumbi,
   L H L H   L H L   H
*kuṭir̄apreemi, and *wair̄amwizuŋŋi . There is no dialect of Malayalam

in which these forms are acceptable. If there were such a dialect,

Lexical Phonology would not be able to account for the facts. Thus,

the actual facts are the only facts allowed by Lexical Phonology,

and therefore, they confirm the validity of the theory, particularly

that of the Opacity Principle.

4.6.4. The Stratum of Verbal Compounding

As observed earlier (4.5.4.), the data from type B causativisa-
tion is consistent with either (110) or (111):

(110) S 1 : Derivation(A)   (111) S 1 : Type A Causative &
                                              other Derivations
     S 2 : Derivation(B)   S 2 : Type B Causative &
                                              Subcompounding
     S 3 : Subcompounding   S 3 : Cocompounding
     S 4 : Cocompounding   S 4 : Inflections
     S 5 : Inflections

Given that type B causativisation and verbal compounding must
belong to the same stratum, (111) predicts that the N in [N+V+i] can
be a subcompound. By transitivity, subcompounding and verbal compound-
ing must belong to the same stratum, which would predict the possibi-
lity of one being inside the other. On the other hand, (110) predicts
that subcompounds cannot occur inside verbal compounds, as verbal
compounds are formed at a previous stratum.

The examples in (124) show that (111) makes the right predic-
tions for Malayalam:

(124) a. [[[paśu]₂[kkuṭṭi]][preem]i] 'lover of calves'
         cow   child   love -er

      b. [[[mahaa]₂[nagaṛa]][waas]i] 'dweller of a great city'
         great city   dwell-er

If the N in a verbal compound can be a subcompound, it follows
that it can also be a cocompound, as subcompounds and cocompounds are
inputs to each other:

(125) a. [[[jaaṯi]₃[maṯa]][p æem]i] 'lover of caste and reli-
caste   religion love er   gion'

b. [[[sṯrii]₃[puⁿuṣa]][wiḏweeṣ]i] 'hater of men and
woman   man   hate   -er   women'

On the basis of these facts, I conclude that Malayalam has

the following stratum organisation:

(126)   Stratum 1 : Derivations, including type A Causativisa-
                                                        tion

        Stratum 2 : Subcompounding, Verbal Compounding, and
                                    type B Causativisation

        Stratum 3 : Cocompounding

        Stratum 4 : Inflections

Footnotes for Chapter IV

1. In some versions of the pa̱ language, a vowel is added to the word final consonant. Thus, bhaaskar 'a name', in this dialect, is cabhaas caka cara.

2. It must be pointed out that the r̄ → r rule cannot be reversed as a r → r̄ rule which applies prevocalically: underlying r never becomes r̄.

formal	colloquial	
kayar	kayarɜ	'rope'
kayarine	kayarine	'rope-acc.'
kayarewi̱te	kayarewi̱te	'where is the rope?'

3. An inadequacy of this solution is that it does not explain why, in single morphemes, r̄C is possible only when the C is y. Thus, arkan 'sun', and bhaar̄ya 'wife' are possible, but not *ar̄kan and *bhaarya. One may think that this constitutes an argument against the no coda hypothesis, as (19), and not (23), accounts for both morpheme final and morpheme internal r. I do not take this as a serious objection, as there exist independent restrictions on the distribution of r̄ and r: dr̄/*dr, tr/*tr̄, jr/*jr̄, ḍr/*ḍr̄, etc. require special morpheme structure or onset structure conditions.

4. Examples with morpheme final r, l, and ḷ are used because stops voice in Malayalam when they are preceded by nasals whether they are appendices or not (2.2.2.).

5. Observe that the correct stipulation on these appendices is that they are morpheme final, not word final or phrase final. The r in awarkaḷ ((27d)) is morpheme final, but not word or phrase final.

6. Words like n̠alki 'gave' and pulki 'embraced' pose a problem. The l in these examples cannot be an appendix, if appendices occur only morpheme finally ([[n̠alk] i] ; [[pulk] i]). As expected, the initial syllable in these words constitutes a guru; yet, k gets voiced. One way of explaining this would be to say that the segment is underlyingly voiced. However, the final i in these examples does not exhibit the ɜ onglide which follows underlying voiced stops (see 2.2.2.). Examples like these are extremely rare, while examples in which the voicing does not apply are numerous: alpam 'a little', kalpana 'order', t̠aalpar̄yam 'interest'. Since voicing is a post lexical rule, n̠alki and pulki cannot be exceptions: post lexical rules do not allow exceptions. I reserve this problem for future research.

7. Appendices are ignored here for convenience of presentation.

8. See Vergnaud & Halle (1979), for the original proposal to treat epenthetic segments in Berber and Hanan in terms of the conventions of syllabification.

9. I have not done a thorough dictionary search to check my statements about possible and impossible morpheme internal consonant sequences in Malayalam, but have depended on my intuitions.

10. Since Syllable Completion would destroy the environment for Onset Fusion, it must be ordered after Onset Fusion.

11. Since syllabic r is restricted to the environment $\left\{ \begin{matrix} C \\ \# \end{matrix} \right\}$── C, it would not be necessary to set up an underlying distinction between /r/ and /ṛ/. r is assigned to the rime node just in case it occurs in the environment mentioned above.

   It may be pointed out that the ə in ragi is phonetically [ɨ], a high central vowel. This is a consequence of a late rule that raises ə when followed by a consonant: cf: [kaadə] 'forest', but [kaattɨtii] 'forest fire'; [kaarə] 'car', [kaarɨwannu]'the car came', [kaarə... wannu].

12. Compare the treatment of Luganda in Vergnaud & Halle (1979). See also the erasure convention in Harris (forthcoming).

EPILOGUE

The early stages of generative linguistics had no provision
for morphology. The grammar was conceived of as a device that maps
a finite set of morphemes onto an infinite set of sentences (Chomsky
(1957)). Morphemes were directly mapped onto sentences: the tradi-
tional notion 'word' played no role in linguistic theory, except by
way of an accidental stage in the derivation of the sentence. The
basic assumption was that the principles that determine the way mor-
phemes are put together to form words were in no way distinct from
the principles that determine the way words are put together to form
sentences. Consequently, word structure and sentence structure were
handled by the same module of the grammar.

During this period, the lexicon was viewed as an unstructured
collection of whatever was idiosyncratic and unpredictable. All pheno-
mena which were regular and were worthy of the linguist's attention
were the burden of the nonlexical components of syntax and phonology.
As a result, little attention was paid to the nature of the lexicon.

With Chomsky's seminal paper, 'Remarks on Nominalization'
(1970), the traditional notion 'word' came back to linguistics. Chom-
sky proposed that certain regular relationships between words could
be expressed in terms of 'lexical rules', and that these rules were
of a different nature from the syntactic rules which determined

sentence structure. A lexical rule was a redundancy rule which captured the regularities in the lexical entries, such as the relation between destroy and destruction. This was the beginning of the recognition that word structure and phrasal structure were not governed by the same set of principles, and that they belonged to different modules of grammar. The lexical entry was no longer the morpheme; it was the word.

The second significant work in this direction was Halle's 'Prologomena to a Theory of Word Formation' (1973). Halle undertook to investigate the principles governing word structure in depth, and added a new module to the grammar as part of the lexicon, namely, the word formation component. These works attributed a far richer structure to the lexicon, and focussed on the nature of lexical operations. Ever since, there has been an increasing awareness of the role of the lexicon in linguistic theory. Halle was followed by several linguists who undertook to study the organisation of the lexicon and the rules which characterise word structure: Aronoff (1974), Siegel (1974), Allen (1978), Hust (1978), Amritavalli (1980), Lieber (1980), and Selkirk (to appear), to mention a few. Several linguists recognised that the mechanism of lexical rules could be extended to handle phenomena which were previously handled in terms of syntactic transformations, such as passivisation and raising (Bresnan (1978), Brame (1978), Baker (1979)). As a result, the lexicon, and the lexical

rules, became the appropriate device for the characterisation of
a great number of syntactic and semantic regularities, thereby re-
ducing the power of the nonlexical syntactic component to a minimum.

This thesis may be viewed as an extension of the lexical
approach to phonology as well, almost as a logical step in the
course of the developments sparked off by Chomsky's Remarks.

Subham

Append$x 1 : PHONETIC CHART FOR MALAYALAM

Pure Vowels

	Front	Central	Back
High	ii 　i		uu 　u
Mid	ee 　e	ə	oo 　o
Low			aa 　a

Diphthongs

ei

oi

ai　au

Consonants

	bilab.	dental	den.alv.	alveolar	pal.alv.	retrofl. or post alv.	palato-alveolar	palatal	velar	glottal
Stop	p　b ph　bh	t̪　d̪ t̪h　d̪h		t		ṭ　ḍ ṭh　ḍh	c　j ch　jh	k'　g' k'h　g'h	k　g kh　gh	
Fricative				s	s̄	ṣ				h
Lateral				l		ḷ				
Tap			r̄	r						
Nasal	m	n̪		n		ṇ	ñ	ɲ	ŋ	
Friction-less continuant						ẓ				
Glide	w							y		

Appendix 2 : INDEX OF DRAVIDIAN AND SANSKRIT STEMS

aadi S	janam S	ooram D
aagamanam S	jananam S	parawa D
aana D	kaalam S	pasu S
aaroogyam S	kaantan S	pati S
aata D	kaaranam S	patni S
aawaranam S	kaata D	pattaayam D
acchan ?	kala S	petti D
agram S	kampi D	pori D
aikyam S	kanna D	preemam S
alasan S	kari D	rati S
ambaram S	katam D	sandeesam S
amma ?	kattila D	satyam S
asukham S	kayara D	sneeham S
baalan S	kinnaran S	sukham S
baalika S	kirukkan D	suuryan S
bhaarya S	kramam S	taamara ?
bhartaawa S	kumaaran S	taapam S
candran S	kutira D	taara S
cannala D	kutti D	tii S
deewi S	kuuttam D	trpti S
dhiiran S	maata D	uma S
dhwani S	mahaa S	waala D
dukkham S	mala D	wadhu S
gandharwan S	manusyan S	wanti D
grham S	maram D	wargam S
gamanam S	maranam S	weedam S
ila D	mrgam S	wiita D
indran S	mukham S	yaksan S
itihaasam S	naalam S	yoogam S
jaaran S	niyantranam L	

REFERENCES

Allen, Margaret (1978). Morphological Investigations, Ph.D.
        Thesis, University of Connecticut.

Amritavalli, R. (1980). Aspects of the Organization of Redundancy
        Rules in the Lexicon, Ph.D. Thesis, Simon Fraser University,
        Canada.

Aronoff, Mark (1976). Word Formation in Generative Grammar , MIT
        Press, Cambridge, Mass.

Baker, C.L. (1979). 'Syntactic Theory and the Projection Principle',
        Linguistic Inquiry 10: 553-81.

Brame, M. (1978). Base Generated Syntax, Noit Amrofer, Seatt'e,
        Washington.

Bresnan, Joan (1971).'Sentence St ress and Syntactic Transformations',
        Language 47: 257-87.

_____ (1978). 'A Realistic Model of Transformational Grammar', in
        M.Halle, J.Bresnan, & G.Miller (eds) Linguistic Theory and
        Psychological Reality, MIT Press, Cambridge, Mass.

_____ (in press: a), 'The Passive in Lexical Theory', in J.Bresnan
        (ed) The Mental Representation of Grammatical Relations, MIT
        Press, Cambridge, Mass.

_____ (in press: b). 'Control and Complementation', in J.Bresnan
        (ed) The Mental Representation of Grammatical Relations, MIT
        Press, Cambridge, Mass.

Chambers, J.K. & P.A.Shaw (1980). 'Systematic Obfuscation of Morpho-
        logy in Dakota' Linguistic Inquiry 11: 325-36.

Chomsky, Noam (1957). Syntactic Structures , Mouton, The Hague.

_____ (1964). Current Issues in Linguistic Theory, Mouton, The Hague.

_____ (1965). Aspects of the Theory of Syntax , MIT Press, Cambridge,
        Mass.

_____ (1970). 'Remarks on Nominalization', in Studies on Semantics
        in Generative Grammar, Mouton, The Hague.

_____ (1981), Lectures on Government and Binding , Foris
        Publications, Dordrecht.

Chomsky, N. & M.Halle (1968). The Sound Pattern of English, Harper & Row, New York.

Donegan, P. (1978). On the Natural Phonology of Vowels, Ph.D. Thesis, Ohio State University.

Emeneau, M.B. (1956).'India as a Linguistic Area', Language 3.2.1.:3-16.

Emonds, J. (1976). A Transformational Approach to English Syntax, Academic Press, New York.

Gimson, A.C. (1962). Introduction to the Pronunciation of English, St. Martin's Press, New York.

Halle, Morris (1959). The Sound Pattern of Russian, Mouton, The Hague.

_____ (1973).'Prologomena to a Theory of Word Formation', Linguistic Inquiry 4: 3-16.

_____ (1975). Formal vs. Functional Considerations in Phonology, Indiana University Linguistics Club.

Halle, M. & K.Stevens (1962). 'Speech Recognition: A Model and a Program for Research', in IRE Transactions of the Professional Group on Information Theory, Vol. 11-8, Number 8.

Halle, M. & J-R. Vergnaud (1980). 'Three Dimensional Phonology', Journal of Linguistic Research, 1.1.: 83-105.

Harris, James (forthcoming). 'Spanish Syllable Structure'.

Hayes, Bruce (1980). A Metrical Theory of Stress Rules, Ph.D. Thesis, MIT.

Hust, J.R. (1978). 'Lexical Redundancy Rules and the Unpassive Con-struction', Linguistic Analysis 4: 61-89.

Jackendoff, R. (1975) 'Morphological and Semantic Regularities in the Lexicon', Language 51: 639-71.

_____ (1977). $\bar{X}$ Syntax: A Study of Phrase Structure, MIT Press, Cambridge, Mass.

Jones, D. (1940). An Outline of English Phonetics, Heffer, Cambridge.

Kahn, D. (1976). Syllable Based Generalizations in English Phonology, Ph.D. Thesis, MIT.

Kean, M-L. (1974).'The Strict Cycle in Phonology', Linguistic
        Inquiry 5: 179-203.

Kiparsky, Paul (1972).'Explanation in Phonology', in S.Peters (ed)
        Goals of Linguistic Theory, Prentice Hall, Englewood Cliffs,NJ.

_____ (1973). 'Elsewhere in Phonology', in S.R.Anderson & P.Kiparsky
        (eds) A Fest schrift for Morris Halle, Holt, Rinehart & Winston
        New York.

_____ (1974). 'Phonological Representations', in O. Fujimura (ed)
        Three Dimensions of Linguistic Theory, Paishukan, Tokyo.

_____ (1979). 'Metrical Structure Assignment is Cyclic', Linguistic
        Inquiry 10: 421-441.

_____ (in preparation). Lexical Phonology.

Koutsoudas, A., G.Sanders, & C.Noll (1974). 'On the Application of
        Phonological Rules', Language 50: 1-28.

Liberman, Mark (1975). The Intonational System of English, Ph.D.
        Thesis, MIT.

Liberman, M. & A.Prince (1977). 'On Stress and Linguistic Rhythm',
        Linguistic Inquiry 8: 249-336.

Lieber, Rochelle (1980). On the Organisation of the Lexicon, Ph.D.
        Thesis, MIT.

McCarthy, John (1979). Formal Problems in Semitic Phonology and
        Morphology, Ph.D. Thesis, MIT.

Mascaro, J. (1978). Catalan Phonology and the Phonological Cycle,
        Ph.D. Thesis, MIT. Reproduced by IULC.

Masica, C. (1976). Defining a Linguistic Area: South Asia, The Univer-
        sity of Chicago Press, Chicago.

Pierrehumbert, Janet (1980). The Phonology and Phonetics of English
        Intonation, Ph.D. Thesis, MIT.

Pesetsky, David (1979).'Russian Morphology and Lexical Theory',
        Unpublished ms, MIT.

Pike, K.L. (1947). Phonemics, University of Michigan, Ann Arbor.

Rajarajawarma, A.R. (1904). wr̥ttamañjari, NBS, Kottayam, India.

Rotenberg, Joel (1978). The Syntax of Phonology, Ph.D. Thesis, MIT.

Rubach, J.(1980).'Segmental Rules of English and Cyclic Phonology',
        Unpublished ms.

_____ (1981). Cyclic Phonology and Palatalisation in Polish and
        English, Warsaw University Press, Warszawa.

Selkirk, Elizabeth (1980). 'Prosodic Domains in Phonology: Sanskrit
        Revisited', in M.Aronoff & M-L.Kean (eds) Juncture, Anma
        Libri , Sarasota, California.

_____ (to appear). Phonology and Syntax: The Relation between
        Sound and Meaning, MIT Press, Cambridge, Mass.

Siegel, D. (1974). Topics in English Morphology, Ph.D. Thesis, MIT.

Stampe, D. (1973). How I Spent my Summer Vacation, Ph.D.Thesis,
        Ohio State University.

Steriade, Donca (1981).'On the Structure of Malayalam Onsets', Un-
        published ms, MIT.

Sweet, H. (1871). A New English Grammar: Logical and Historical,
        Clarendon Press, Oxford.

Vergnaud, J-R. & M.Halle (1978). 'Metrical Structures in Phonology',
        Unpublished ms, MIT.

Warrier, Tara (1976). The Phonetics and Phonology of Malayalam,
        M.Litt. Thesis, CIEFL, India.

Williams, Edwin (1979). 'On the Notions Lexically Related and Head
        of a Word', Unpublished ms, U.Mass.

Zimmer, K.E. (1964). Affixal Negation in English and Other Languages,
        Monograph No. 5, Word 20.2.