INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
Latta, Fletcher Christian
Methods for Internal Reconstruction.
The Ohio State University, Ph.D., 1978

©

Copyright by
F. Christian Latta
1978
METHODS FOR INTERNAL RECONSTRUCTION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

F. Christian Latta, B.A., M.A.

* * * * *

The Ohio State University

1978

Professor Robert Jeffers
Professor Ilse Lehiste
Professor Arnold Zwicky
Professor C. A. Callaghan

Approved By

Robert J. Jeffers
Adviser
Department of Linguistics
To Donna, whose faith and love made my work possible.
ACKNOWLEDGMENTS

Many people have assisted in the preparation of this dissertation and I cannot begin to thank them all here. My adviser and friend, Robert Jeffers, provided hours of helpful discussion, often animated, which served to clarify my thinking on many matters. Professors Lehiste and Zwicky were especially helpful with the formulations for some of the principles. They were also quite patient and carefully pointed out errors in examples and citations.

I owe special thanks to Professors Stampe and Callaghan for their continued encouragement of my study of Austronesian linguistics. Mahalo and aloha to Professor Laurie Reid of the University of Hawaii for his special assistance and encouragement with the Austronesian material.

I would like to thank Dr. John Perkins for kindly reading earlier drafts and for his invaluable assistance proofreading. Dr. Perkins also provided many insightful and substantive comments. Thanks are due to Chris Farrar and Jami Josifek for great discussion and moral support. Special thanks are due Don Churma for his
discussion and valued friendship over the last three years.

Last, but certainly not least, is Mrs. Marlene Payha, whose inexhaustible willingness to help made a difficult job much easier. It is nice to know that there are still good people, like Marlene, in the world.
VITA


1972 . . . . . . . B.A. Philosophy, Ohio University
Athens, Ohio

1972-1973 . . . . Teaching associate, Ohio University
Athens, Ohio

1973 . . . . . . . M.A. Linguistics, Ohio University
Athens, Ohio

Athens, Ohio

1975-1978 . . . . Teaching Associate, The Ohio State
University, Columbus, Ohio

PUBLICATIONS AND PRESENTATIONS

'On stress and vowel harmony in Chamorro.' 1972 Oceanic

'Notes on The Phonology of Reduplication.' 1975 Paper
read at Michigan State University Mini-Conference on

'Nasalization in Sundanese: a closer look.' 1977 Paper
presented at summer meeting of Linguistic Society of
America Honolulu, Hawaii. Summer.

'Nasal infixation in Proto-Austronesian.' 1977 Paper pre-
sentated at the Austronesian Symposium, Honolulu, Hawaii.

'The interaction of nasal assimilation and reduplication
in Tagalog.' 1977 Paper presented at the Annual Meeting
of the Linguistic Society of America, Chicago, Illinois.

FIELDS OF STUDY

Phonetics. Professor Ilse Lehiste
Phonology. Professor David Stampe, Professor Wolfgang Dressler, Professor Arnold Zwicky

Historical Linguistics. Professor Robert Jeffers, Professor Ilse Lehiste, Professor Cathy Callaghan

Psycholinguistics and Language Change. Professor Olga Garnica

Austronesian Linguistics. Professor Lawrence Reid, Professor George Grace, Professor Andrew Pawley, Professor Teresita Ramos

Indonesian Linguistics. Professor John U. Wolff
TABLE OF CONTENTS

DEDICATION ........................................... ii
ACKNOWLEDGMENTS ................................... iii
VITA ................................................... v
LIST OF TABLES ......................................... ix
LIST OF FIGURES ......................................... x

Chapter

I. INTRODUCTION ........................................ 1

II. INTERNAL RECONSTRUCTION: METHODOLOGICAL
    PRINCIPLES AND PROCEDURES .................. 12
    Internal Reconstruction ......................... 13
    Internal Reconstruction and the
    Comparative Method ............................... 22
    Some Methodological Principles .............. 31
    Recovering Sound Changes ..................... 48

III. SAUSSURE'S THEORY OF IE ABLAUT AND
     BENVENISTE'S THEORY OF THE IE ROOT ...... 59
    Laryngeal Theory ................................ 60
    Benveniste's Theory of the Indo-
    European Root .................................... 77
    Argumentation in Saussure (1879)
    and Benveniste (1935) ......................... 83

IV. INTERNAL RECONSTRUCTION AND RELATIONAL
    PATTERNS ........................................... 92
    Regular Relational Patterns .................... 93
    Some Methodological Problems ............... 108
    Some Potential Objections .................... 117
    Principles for Internal
    Reconstruction ................................. 126
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>V. PROTO-AUSTRONESIAN MEDIAL NASAL CLUSTERS</td>
<td>135</td>
</tr>
<tr>
<td>Proto-Austronesian</td>
<td>137</td>
</tr>
<tr>
<td>PAN Canonical Structure</td>
<td>143</td>
</tr>
<tr>
<td>Consonant Clusters in PAN</td>
<td>145</td>
</tr>
<tr>
<td>Dempwolff's Ngadju Article and the Shape of the Infix</td>
<td>158</td>
</tr>
<tr>
<td>Arguments against -VN-</td>
<td>165</td>
</tr>
<tr>
<td>Gonda's Argument and *ng-</td>
<td>174</td>
</tr>
<tr>
<td>Pränasalierung</td>
<td>178</td>
</tr>
<tr>
<td>Determining the Function of *ng-</td>
<td>191</td>
</tr>
<tr>
<td>Evidence from Ngadju Dayak and Tagalog</td>
<td>213</td>
</tr>
<tr>
<td>Some Problems</td>
<td>220</td>
</tr>
<tr>
<td>VI. CONCLUSIONS</td>
<td>231</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>235</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>VL III Instances of PAN CVCVNCVC</td>
<td>153</td>
</tr>
<tr>
<td>2.</td>
<td>PAN CVNCVVC and CVVCVVC Trisyllables</td>
<td>155</td>
</tr>
<tr>
<td>3.</td>
<td>Dempwolff's CVN+CVC+CVCVNCVC Analysis</td>
<td>161</td>
</tr>
<tr>
<td>4.</td>
<td>Exemplary CVNCVVC forms</td>
<td>168</td>
</tr>
<tr>
<td>5.</td>
<td>Tagalog, Toba Batak, and Javanese examples of Pränasalierung</td>
<td>180</td>
</tr>
<tr>
<td>6.</td>
<td>PAN -(N)-C Forms</td>
<td>193</td>
</tr>
<tr>
<td>7.</td>
<td>Relative Frequencies of Cluster Types</td>
<td>205</td>
</tr>
<tr>
<td>8.</td>
<td>Ngadju Dayak OG-NG Pairs</td>
<td>214</td>
</tr>
<tr>
<td>9.</td>
<td>Tagalog pa-, ma-, pan-, and man-</td>
<td>219</td>
</tr>
<tr>
<td>10.</td>
<td>Tagalog, Toba Batak, and Javanese Reflexes of Pränasalierung</td>
<td>223</td>
</tr>
<tr>
<td>11.</td>
<td>Tagalog, Toba Batak, and Javanese Reflexes of Medial Nasal Clusters</td>
<td>224</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Two 'stages' in the Prehistory of Greek</td>
<td>18</td>
</tr>
<tr>
<td>2.</td>
<td>Northern-Southern Correspondences</td>
<td>66</td>
</tr>
<tr>
<td>3.</td>
<td>Greek-Sanskrit Correspondences</td>
<td>68</td>
</tr>
<tr>
<td>4.</td>
<td>Saussure's Analysis of IE Ablaut</td>
<td>69</td>
</tr>
<tr>
<td>5.</td>
<td>Dempwolff's PAN Sound System</td>
<td>139</td>
</tr>
<tr>
<td>6.</td>
<td>Synonymous OG-NG Forms</td>
<td>204</td>
</tr>
<tr>
<td>7.</td>
<td>'That which x-es' OG-NG Relation</td>
<td>208</td>
</tr>
<tr>
<td>8.</td>
<td>Special Case OG-NG Relation</td>
<td>209</td>
</tr>
<tr>
<td>9.</td>
<td>Directionality of Figure 7 OG-NG forms</td>
<td>210</td>
</tr>
<tr>
<td>10.</td>
<td>Directionality of Figure 8 OG-NG forms</td>
<td>210</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

In a very influential article, Chafe (1959) defined internal reconstruction as a 'procedure for inferring part of the history of a language from the material available for a synchronic description of the language, and from that alone.' Indeed, the procedures of internal reconstruction are familiar to most linguists and the method has been well documented in the literature. (See for example, Chafe 1959; Hoenigswald, 1944, 1946, 1960; Kuryłowicz, 1964; and Marchand 1956.) In its most straightforward form the method assumes that paradigmatic allomorphy is not original, but has been introduced by regular sound change. Thus, internal reconstruction yields two results: first, a description of part of the language prior to the introduction of paradigmatic allomorphy; second, it provides a detailed statement of the linguistic changes that are responsible for that allomorphy.

The method has been accurately compared to the comparative method (see for example, Chafe and Hoenigswald, above). However, rather than comparing cognate items from the lexicons of different languages, it is the
different forms of a morpheme within a single language that are compared. Both paradigmatic allomorphy and the assumption that the different forms of a morpheme in a paradigm are cognate\(^1\) are crucial to internal reconstruction. And Chafe has stated this quite succinctly (1959:178-79):

...but with the possible exception of some aspects of dialect geography there is no generally recognized historical technique that does not involve the comparison of cognates. Such comparison seems to me to be the only conceivable basis for internal reconstruction.

The forms which are available for comparison in internal reconstruction are cognate morphs.

Despite this insistence by methodologists that paradigmatic allomorphy and the comparison of cognate morphs are fundamental to internal reconstruction, there are numerous examples of reconstruction which, while treated under the heading 'internal reconstruction,' have nothing to do with either paradigmatic allomorphy or the comparison of cognate morphs, at least in any straightforward way. I have in mind here reconstruction that has its basis in the comparison of structural patterns; that is, skewed figures in the statistical occurrence of linguistic units and structures, asymmetries in structural patterns, and forms which are anomalous with respect to the morphological canons of the language. Indeed, two of the most far-reaching examples of 'internal reconstruction' in Indo-European linguistics--Saussure's Mémoire sur le
système primitif des voyelles dans les langues indo-européennes and Benveniste's theory of the IE root--are examples of reconstruction of this sort. The primary goal of this dissertation is to investigate the type of reconstruction exemplified in the work of Saussure and Benveniste noted immediately above. My main concern will be explicating the principles and procedures of this kind of internal reconstruction. For the moment I will refer to this as pattern reconstruction.

While a methodology for internal reconstruction--as applied to paradigmatic allomorphy--has been worked out explicitly in the literature (see the references cited above), the methodological principles of pattern reconstruction have not yet been elucidated in any systematic way. I will examine Saussure's theory of IE ablaut and Benveniste's theory of the IE root in detail. Through this examination I will identify the methodological principles followed by each. I will then formulate these principles and present a critical analysis of the underlying assumptions, goals and results of pattern reconstruction.

For example, while the method does have a strong comparative element in it, the elements that are compared, structural patterns, are not cognate morphs. And there are those who question the validity of this type of
reconstruction. Hoenigswald (1944:78) has argued:

On the whole, speculation without the benefit of comparative evidence has a reputation for being unsafe, a criticism which is, however, directed more often against certain far-reaching results than against the procedure itself.

It will of course be necessary to review this type of objection as well as others. I will argue that if it can be shown that pattern reconstruction proceeds on the basis of well-motivated linguistic principles, comparative or not, the method must be considered to be 'safe'.

As part of this discussion I will investigate the relationship of the methodological principles of internal reconstruction, and pattern reconstruction, to general principles of linguistic analysis and the philosophy of science.

It is important to note that both internal reconstruction and pattern reconstruction are concerned with paradigmatic relationships in a broad sense. That is, just as different allomorphs can be derived from a common source historically—for example, Latin _genus_ 'kind, type nom. sg.', _generis_ gen. sg. from IE *genes/os—the structural patterns CV and CVC may be derived from a common historical source, namely CVC. In the first case, the paradigm is defined in terms of the constancy of meaning. In the latter case, it is defined in terms of a constancy of shape.
After explicating the methods of internal reconstruction, and pattern reconstruction, I will provide an illustrative example from Proto-Austronesian (PAN). The purpose of this example is to show not only how the method can be applied, but also to reveal the kinds of problems the historical linguist may encounter when reconstructing patterns. I will also examine the results of this kind of work in terms of how it can contribute to a deeper understanding of the history and structure of a language. Thus, the results of pattern reconstruction are often far-reaching and these results often amount to significant gains in knowledge.

The problem in Proto-Austronesian concerns the reconstructed canon *CVCCVC (see Dempwolff, 1934-8). There are only two types of instances of this canon: (A) those which can be analyzed as reduplicated monosyllables, for example *kuðkuð 'rasp', *piipay 'spout', and *tak'tak 'hoe'; and (B), those which contain medial homorganic nasal clusters, for example, *hantaq, 'deliver', *tanduk 'horn', and *d'ambaj 'areka palm'. While Dempwolff reconstructed clusters for (B), Austronesianists have long recognized the severe distributional anomalies this represents for the morphological and phonological systems of PAN. Over the years there have been several proposals to remedy this situation. (See Dahl, 1973 and Blust, 1971 and the references cited there for the most recent suggestions).
I will review these proposals and argue that only CVCCVC (B) represents a distinct canon since the (A) forms can be derived from *CVC and reduplication. Moreover, I will suggest that the distributional limitations on consonant clusters indicate that the (B) canon is the result of linguistic change. After noting that *R (see Dempwolff 1925) and the original nasals in clusters occur where an infix would—that is, -C₁- + CVCVC + CV-C₁-CVC--I suggest, as others have, that the clusters have resulted from infixes, *-ng- and *-R-, and the canon CVCCVC.

I will systematically review the evidence in support of this interpretation and consider the various shapes that have been proposed for the infix. I will review the objections that have been raised and provide an analysis of the function of the hypothetical infix. I will suggest that the weight of the evidence demands an infixation analysis, at least as a working hypothesis.

The infixation hypothesis has some important implications for PAN morphology. Notice that consonant clusters which result from the concatenation of a prefix and a wordbase, CVCₚᵢₑᵣ+CVCₜₑₛₑₑᵢₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑㅔ.xhtml" xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">notice that consonant clusters which result from the concatenation of a prefix and a wordbase, CVC_{prefix}+CVC_{wordbase}, are positionally analogous to the medial clusters. Moreover, the final C's of the prefixes occur exactly where a consonantal infix would. I suggest that it is not coincidental that *R would.
and *-ng- widely occur as the finals of such prefixes, and that those prefixes with final *ng universally participate in morphophonemic nasal assimilation processes with the initials of the following wordbases. The conclusion seems clear: prefixes with final *R and *ng, for example *paR-, *paN- (where 'N' represents the morphophonemic alternants), resulted from a morphological reanalysis of a *CV-prefix, infix *-R- or *-ng- and a base. That is,

$$CV_{\text{pre}} \rightarrow \{ R \}_{\text{ng}} - CVCVC_{\text{w.base}} \geq CV_{\{ R \}_{\text{ng}}{\text{pre}} - CVCVC_{\text{w.base}}$$

Under this analysis, two distinct phenomena, medial nasal clusters and the morphophonemic alternations associated with paN-prefixes, are seen to be historically related.

The purpose of this dissertation, then, is two-fold. On the one hand I hope to clarify the methodological principles of internal reconstruction, particularly those of pattern reconstruction. Second, for the purpose of illustrating those principles and, at the same time, solving a long standing problem in Austronesian linguistics, I will apply the method to the problem of PAN medial nasal clusters. Before beginning, I would like to comment briefly on the organization of the dissertation.

In the next chapter, chapter II, I will review the literature on internal reconstruction with particular attention to Hoenigswald (1944, 1946) and Chafe (1959). My
main concern will be elucidating the method of internal reconstruction as that method has been described in the recent literature. I will discuss the underlying assumptions of the method, its results, and its relationship to the comparative method. I will also formulate certain methodological principles that are frequently appealed to, though often implicitly, in the literature. These principles permit the language historian to draw inferences about the history of a language on the basis of purely internal evidence. Finally, I will discuss general conditions on the recoverability of sound changes of various types. My discussion of the more traditional methods of internal reconstruction in this chapter will serve as a backdrop against which I will evaluate pattern reconstruction.

In chapter III, I will explicate Saussure's theory of IE ablaut and Benveniste's theory of the IE root. My main concern will be to identify the types of evidence upon which Saussure and Benveniste base their theories and to discern the methodological principles to which they appeal. I will argue that the type of evidence they use and the methodological principles they appeal to have not yet been described in a systematic way in the literature of historical linguistics.
In chapter IV I will describe these principles in detail and evaluate their use in making inferences about the history of a language. I will propose a methodology for reconstructing patterns and discuss the different kinds of traces pattern changes leave on linguistic structure. I will also explicate the relationship of pattern reconstruction to the more traditional methods of internal reconstruction described in chapter II. Finally, I will consider potential objections to the proposed method and explicate the relationship of the methodological principles for internal reconstruction—both traditional and pattern reconstruction—to general principles of linguistic analysis and the philosophy of science.

In chapter V I will provide an extended example of the application of pattern reconstruction. I will briefly describe Austronesian studies and focus on the problem of medial homorganic nasal clusters and the infixation hypothesis which attempts to derive the clusters from a CVCVC canon and an infix historically. I will critically review the evidence traditionally cited in support of the hypothesis and argue that it does not really provide support. I will then introduce new evidence and suggest possible alternative hypotheses for the shape of the hypothetical infix. I will also argue that prefixes with final *R or *ng, for example, *paR- or *paN- (see above) result from the morphological reanalysis of a CV- prefix and *-ng- and
*-R- infixes. I will review the objections to the infix hypothesis and provide an initial semantic analysis of *-ng-. My purpose in this chapter is to both illustrate procedures of pattern reconstruction and to provide a systematic evaluation of the problem of medial nasal clusters in Proto-Austronesian. I will demonstrate that the meticulous application of the method to the PAN problem sheds new light on PAN morphology and, thus, significantly advances our knowledge.

In chapter VI I will present a summary of the major conclusions of this dissertation. I will suggest that there is no single unitary method for reconstructing the history of a language on the basis of internal evidence. Rather there are several procedures, all of which are based on sound, well-motivated linguistic principles, which the language historian can appeal to in his work. Thus, it is not a question of the method of internal reconstruction, but of methods for internal reconstruction.
NOTES

1There are of course cases where one or more of the morphemes within a paradigm are not cognate, for example, in a suppletive paradigm. See 2.2 for a complete discussion.
2.0. Introduction. It is the purpose of this chapter to examine in detail the method of internal reconstruction. My fundamental concern will be with the principles and procedures that have been described in the literature on internal reconstruction, particularly in the articles of Hoenigswald (1944, 1946, 1960) and Chafe (1959). After describing these principles and procedures I will review Saussure's reconstruction of the Indo-European ablaut system—what has become, with modification, modern laryngeal theory—and Benveniste's theory of the Indo-European root. I will argue that their hypotheses rest on a different set of methodological principles. In Chapter III I will take a close look at these, heretofore undescribed, principles and explicate their relationship to the more familiar type of internal reconstruction described by Chafe and Hoenigswald.

In the first section of this chapter, I will consider the basic assumption, goals, and results of internal reconstruction. Section 2.2 deals with the relationship of
internal reconstruction to the comparative method. In 2.3 I will describe and formulate three principles that are often appealed to when making inferences about the history of a language on the basis of internal evidence. In 2.4 I will consider the kinds of sound change that are recoverable by internal reconstruction and the types of structural traces that each kind of sound change leaves behind. These four sections form the backdrop against which I will review the hypotheses of Saussure and Benveniste in the next chapter.

2.1. Internal Reconstruction. In his landmark article, 'Internal Reconstruction in Seneca,' Chafe (1959:478) defines 'internal reconstruction' as 'a procedure for inferring part of the history of a language from material available for a synchronic description of the language, and from that alone.' He immediately notes, however, that the bulk of historical inference proceeds from the comparison of cognate forms drawn from either different languages or different periods of the same language as documented by written records. Chafe (478) concludes: 'Such comparison seems to me to be the only conceivable basis for internal reconstruction.' Hoenigswald (1944:80) has also called attention to this fundamental reliance of internal reconstruction on paradigmatic allomorphy, noting: 'It appears essential ... that in all pertinent cases, morphological
alternation in a paradigm be present.' In a later article (1946) Hoenigswald argued (1946:142): 'unambiguous traces ...
[of sound change, FCL] ... necessarily involve the distribution of sounds in morphological paradigms.'

In order to briefly illustrate the method of internal reconstruction, let us consider the familiar example of the alternation of voiced and voiceless obstruents syllable finally\(^1\) in Modern German. Both Hoenigswald (1944, 1946) and Chafe (1959) discuss this example and the presentation here will follow them closely.

Voiced stops and spirants contrast with voiceless stops and spirants in Modern German, but not in all positions; syllable finally, voiced stops and spirants do not occur. Hoenigswald notes (1944:79):

> It is only as a result of this limitation, which is general in the language, that certain paradigms show /d/ before a vowel, but /t/ at the end of a syllable. There is a pair Bund: Bund (/bunde/:/bunt/) with automatic alternation of this kind, while there are also pairs without such alternation, involving the more freely distributed phoneme, e.g., bunte:bunt (/bunte/:/bunt/). This statement involves morphology; it is a morphological statement.

Within this morphological paradigm, then, the following correspondences can be discerned: voiceless stop/spirant: voiceless stop/spirant, for example, tːt in /bunt/: /bunte/, and voiceless stop/spirant: voiced stop/spirant, for example, tːd in /bunt/:/bunde/. Chafe (480) notes that there is no correspondence of voiced stops/spirants
to voiced stops/spirants (for example, dːd) in this position. These correspondence sets, then, are in complementary distribution, and tːd can be grouped with dːd since tːt and tːd represent sets that contrast with each other elsewhere, and consequently, must be assigned to different reconstructed segments.

Internal reconstruction, then, assumes that paradigmatic allomorphy is not original, but has been introduced by a linguistic change or linguistic changes. The method yields two results: (1) a description of a part of the language prior to the introduction of paradigmatic allomorphy; and (2), a detailed statement of the linguistic changes that are responsible for that allomorphy. For the example above, (1) is the statement that voiced stops/spirants enjoyed a wider distribution, occurring syllable finally in the pre-history of Modern German. (2) is, of course, the devoicing of syllable final obstruents.

The Modern German example discussed by Hoenigswald and Chafe illustrates what Hoenigswald calls a compulsory alternation and Chafe, an automatic alternation. That is, the alternation of voiced stops/spirants and voiceless stops/spirants is (Chafe, 1959:481): 'wholly predictable in phonological terms applying throughout the language, not just in a particular grammatical environment'. Chafe argues that automatic alternations (481) 'can be inferred.'
to be the results of the most recent sound changes, since they have not been made unpredictable by subsequent changes.' He concludes (481): 'the forms reconstructed from automatic alternations can be regarded as one stage historically removed from the attested language.'

While an automatic alternation may be the result of a recent sound change, an automatic alternation could be introduced and preserved over a great period of time without interference from subsequent changes. Indeed, the obstruent devoicing change in the history of German was introduced almost 1,000 years ago (King 1969, 1976), but still remains automatic. Moreover, other changes have affected the language, for example, vowel lengthening, without obscuring final obstruent devoicing. An automatic alternation is necessarily, then, the result of a linguistic change that has not been interfered with. This lack of interference, however, is not necessarily an index of how recent a linguistic change was. Automaticity can yield information only about relative chronologies.

In this regard, the use of the term 'stage' in reference to the results of internal reconstruction may be confusing and should be clarified. A closer examination of Chafe's discussion is in order. He argues (481):

Now at this first reconstructed stage it usually happens that certain alternations in the attested language were still automatic. On the basis of these alternations it is possible to reconstruct forms two
stages removed from the attested language, and, barring other complications, the procedure can be continued in this way from each historical stage to one that preceded it.

Chafe illustrates with an example from Greek. He notes (481) that there was 'a stage of Greek' when $s$ alternated with $\emptyset$ intervocalically. Witness his example: *génēsos* 'generation, gen. sg.' (gene-) but *généssí* dat. pl. (genes-).

This alternation, however, was not entirely automatic since some $s$'s do occur intervocalically. Chafe cites *ambrosíā* 'ambrosia'. The $s$'s of forms like *ambrosíā* alternate automatically with $t$; $s$ occurs before $i$ (or $y$), $t$ occurs elsewhere. Beside *ambrosíā*, there is *ámbrótos* 'immortal'.

Consider also *ploúsios* 'wealthy', but *ploutos* 'wealth'; *ergásiā* 'work' but *ergáttēs* 'workman'. Chafe argues that for the $t/s$ alternation we can reconstruct a *$t$* and infer that *$ti>si$* (see below). He concludes (482):

We now find that at the stage which we have reconstructed there are no $s$'s between vowels, so that the alternation exhibited by gene-/*genes*- is automatic. For the correspondence set $\emptyset/s$ we can reconstruct *$s$*, and infer both that *$s>\emptyset$* and that this change preceded the change *$ti>si$*.

It should be apparent by now that when Chafe discusses 'stages' of Greek he is really referring only to the relative chronologies of linguistic changes in the pre-history of a language. Recall that the method of internal reconstruction yields both a description of a part of the
language prior to the introduction of an automatic alternation and a statement of the change(s) responsible for that alternation. A 'stage' in the history of a language, then, is nothing more than a description of part of a language obtained by working backwards from an automatic alternation. Consider Figure 1 below.

Earliest Reconstructed Stage: no $\emptyset$/s alternation--$s$ occurs intervocalically

$g>\emptyset$ intervocally

More Recent Stage; $\emptyset$/s alternation automatic--no $s$'s occur intervocally;
no t/s alternation--t occurs before i

$\emptyset$/s before i

Attested Stage: t/s alternation automatic--no t's before i;
$\emptyset$/s alternation opaque--s's of t/s occur intervocally

Figure 1. Two 'stages' in the pre-history of Greek.

By 'stages in the history of a language' is meant only something like the short descriptions given in Figure 1. It is important to note that no claims are made about a coherent linguistic system. Automatic alternations provide the evidence from which historical inferences are made. Clearly, one cannot make inferences that are not supported by the available evidence. That is, in the above example, one can infer from the t/s alternation only that
t used to occur before i or the antecedent of i. It cannot be inferred that ambrosiā used to be *ambrotia; the alternation does not provide any evidence about the shape of the rest of this word. Etymologically, ambrosiā is from *ŋ-mrot-ia. In the absence of evidence that would indicate the chronologies of the various changes that have given rise to the attested form in Greek, *ambrotiā cannot be established as a form that actually occurred. Likewise, on the basis of the 0/s alternation one can conclude only that 0 used to occur intervocally; that is, gēneos used to have an 0 between the antecedent of e and the antecedent of o. The etymon of this form is *genes-os. The term 'stage', then, as I shall use it, refers to nothing more than a description of a part of a language prior to the introduction of an automatic alternation. Forms like *ambrotiā must be understood as notational conveniences.

Establishing *t as the antecedent of 0 in ambrosiā and *s as the antecedent of 0 in gēneos is not straightforward, by any means. The sounds t and s in the context of those forms which show an alternation are in complementary distribution: 0 occurs before i; t elsewhere. Ignoring the 0/s alternation for the moment, one could assign t/s to either *t or *s. If the assignment is made to *s, one would have to say that *s became t everywhere
except before í. This is clearly phonetically implausible. Moreover, it is incorrect since original *s does not otherwise develop into t. For example, in initial position *s is reflected as spiritus asper (h): witness Gk. hédos, but Lat. sedeo and Skt. sad-, and Gk. hépomai but Lat. sequor and Skt. sac-. An assignment of the t/s alternation to *s would then entail both a phonetically implausible change and a severe complication to the development of *s. For these reasons, such an assignment is prohibited.

It is noteworthy that the assignment of the t/s alternation to the correct segment is complicated by the Ø/s alternation. That is, the astute linguist concerned with the t/s alternation would note that s does not otherwise occur intervocally; when s occurs intervocally, there is always a related form showing a t. If one reconstructs a *t for the t/s alternation and *s for the Ø/s alternation (see below)—if only for illustration—the statement of the linguistic changes that yield the attested forms would be enormously complex: *s>Ø initially in certain morphemes, *s>Ø/V₁V₂ (V₂≠i), but only in certain forms—those that do not show synchronically a related form with t—*s>t, except before í, but only in certain morphemes—those that show a related form with s before í synchronically—and *s>s elsewhere. Needless to say, the assignment of t/s to *t and Ø/s to *s permits the most straightforward
statement of the relevant linguistic changes.

Chafe argues (482, ftnt. 15) that the $\emptyset/s$ alternation must be assigned to *s.

Epenthesis is plausible whenever the environments in which it might have taken place are definable in strictly phonological terms. That is not the case here, so that we are led to assign $\emptyset/s$ to *s.

Clearly any s-epenthesis change would be highly complex since s would be inserted in such a broad range of environments; most likely specific morphemes would have to be overtly mentioned. If $\emptyset/s$ is assigned to *s, then $\emptyset/s$ would be in contrastive distribution with t/s, as Chafe notes (462, ftnt. 14). The reconstruction of an *s for t/s is, therefore, precluded. In short, when one considers both the t/s alternation and the $\emptyset/s$ alternation, it becomes apparent that the most straightforward solution demands the assignment of t/s to *t and $\emptyset/s$ to *s.

In this section I have suggested that the method of internal reconstruction assumes that paradigmatic allomorphy is not original, but has been introduced by regular change. The method proceeds from automatic alternations to a description of the paradigm prior to the introduction of allomorphy. The method also provides a statement of the linguistic changes responsible for that automatic alternation. The method yields information about relative chronologies, the orders of linguistic changes in the pre-history of a language. In the next
section I will consider more carefully the actual procedures of internal reconstruction and the relationship of the method to comparative reconstruction.

2.2 Internal Reconstruction and the Comparative Method

In the preceding section I established that internal reconstruction is essentially concerned with the synchronic alternations in a language. The method assumes that such alternations are not original, but have been introduced by regular linguistic changes. This will be referred to as the Principle of Automaticity. Internal reconstruction has been accurately compared to the comparative method (see for example, Chafe 1959). In this section I will examine the relationship of internal reconstruction to the comparative method. In particular, I will be concerned with the procedural difficulties the use of purely internal evidence may introduce, the assumptions common to both methods, and the applicability of each method to situations of particular types. The discussion here is intended to supplement the exposition of the method begun in section 2.1.

Inferences about the histories of a group of languages--assumed to be related--are made, under the comparative method, by the comparison of cognate forms drawn from the lexicons of the individual languages. Internal reconstruction, since it relies exclusively on material available for a synchronic description of a single
language, compares the different forms of a morpheme within that language. As Chafe notes (479): 'the forms which are available for comparison in internal reconstruction are cognate morphs.'

Internal reconstruction assumes that the various forms of a morpheme within a language are cognate. The only case where this assumption is not borne out is when a paradigm is suppletive. A suppletive paradigm is a paradigm in which one or more of the forms show a different etymology from the other forms. Such forms must be excluded from comparison. Fortunately, since suppletive forms would rarely yield recurrent correspondences, they can usually be spotted quite easily.

Given these forms for comparison—that is, the different forms of a morpheme within a paradigm—the reconstruction follows the procedures of the comparative method. Throughout this discussion I will follow Hoenigswald's 'The Principal Step in Comparative Grammar' (1950). First, recurrent correspondence sets are noted. In general the different forms of a morpheme within a language will yield recurrent correspondences. There are, however, three cases where such correspondences might not obtain: (1) when the paradigm is suppletive; (2) when a form has been borrowed from another dialect (or related language); and (3), when a form has undergone analogical change. The most problematic of these is probably (2) since the form may look
very much like the other forms in the paradigm, and consequently, be very difficult to detect. These forms can be eliminated from comparison only after a detailed examination of the data has yielded the regular correspondences. In most cases, establishing the regular correspondences will require extensive comparisons with forms from other dialects and languages.

Forms that have been influenced analogically will only rarely yield recurrent correspondences. While it is often tempting to dismiss exceptional forms as due to analogy, no real explanation of the exceptional form is offered unless the terms of the analogy can be identified. (See Jeffers 1977 for a more complete discussion of analogical explanations for linguistic changes.) It should be clear that the problems presented by (1)-(3) above are the same whether one is using internal evidence or evidence from different languages. For example, when comparing forms from different languages, borrowings, especially borrowings from a related language, can be extremely difficult to detect. As noted above, extensive comparison is the only method for determining the correct regular correspondences.

Having noted the recurrent correspondence sets which occur, the next step is to group partially alike sets (see Hoenigswald, 1950). For the purpose of illustration, I will reconsider the example from Modern German (discussed briefly in 2.1). The following recurrent correspondences
can be established: \( t: t \), witness /bun\( \ddot{\text{\textae}} \)/bunte/ 'multi-colored', /rext/:/rext\( e \)/ 'correct' and /kalt/:\( \ddot{\text{k}} \text{al} \text{t}e/ 'cold' all in morpheme final position; \( t: t \) also occurs in other positions, for example, initially: /tak/:/tage/ 'day', /taup/:taube/ 'deaf'; \( t: d \), but only finally; /bund/:bunde/ 'union', /runt/:\( \ddot{\text{\textae}} \text{unde} \)/ 'round', and /vilt/:/vilde/ 'wild'; \( d: d \), for example, /di:p/:/di:be/ 'thief' and /derp/:\( \ddot{\text{\textae}} \text{der} \text{be} \)/ 'coarse'. This last correspondence set does not occur in morpheme final position. All of these recurrent correspondence sets are partially alike. It is to be noted, however, that \( t: d \) and \( d: d \) occur in mutually exclusive environments and that \( t: t \) contrasts with \( d: d \); witness /tozen/ 'roar' but /dozen/ 'boxes, tins (Dosen\( \ddot{\text{\textae}} \text{ff} \text{ner} \)'.

Hoenigswald notes (1950:359): 'partially alike sets occurring in mutually exclusive environments are taken to be continuations of one and the same phoneme of the proto-language.' Following this principle, \( t: d \) and \( d: d \) are grouped together and assigned to the same original segment, \( *d \), distinct from the origin of \( t: t \), \( *t \).

In this example from Modern German the environments for each of the correspondence sets are quite transparent; the final obstruent devoicing change has not been obscured by subsequent changes. It is often the case, however, that the environments for correspondence sets will not be transparent; the alternation of morphs will not be entirely automatic. For example, in the discussion of Greek, it was
noted that while the alternation of \( t \) and \( s \) was automatic, the alternation of \( s \) and \( \emptyset \) was not since some \( s \)'s occurred intervocalically—just those that alternate with \( t \). Working backward from \( t/s \) and reconstructing \(*t*\) for \( t/t \) and \( t/s \), we saw that the \( \emptyset/s \) alternation was entirely automatic at this earlier stage; the environment of \( \emptyset/s \) was no longer opaque. In short, by using automaticity the relative chronologies of the changes \( s>\emptyset/V\_V \) and \( t>s/\_\_i \) were established.

It is this sequential nature of linguistic changes—and the fact that more recent changes may obscure earlier changes—that distinguishes internal reconstruction from the comparative method. As Chafe notes (481),

> For while the latter ... [the comparative method, FCL] ... yields inferences regarding a single proto-language assumed to exist at the particular time when the compared languages diverged, our experience with the history of language shows us that the morphophonemic alternations in a language do not result from simultaneous sound changes, but from a series of independent changes, both regular and irregular, any of which may modify and obscure the results of previous changes. Internal reconstruction thus yields inferences concerning not one but several different historical stages of a language....

One would do well to keep in mind my remarks concerning the usage of the term 'stage' of a language when used in connection with internal reconstruction—that is, a description of only a part of a language prior to the introduction of an automatic alternation. The comparative
method allows one to make historical inferences about an entire linguistic system. For example, on the basis of the comparison of Gk. ambrosia with the cognates in other languages one can infer an original shape, *η-mrot-ia. I noted above (2.1) that internal reconstruction allows us to infer only that *t was the antecedent of s. That is, the method did not permit any inferences about the shape of the rest of the word. In this sense, the results of the comparative method are broader in scope. It is most unfortunate that the term 'stage' is used carelessly to refer to the results of both methods.

It should be noted that the comparative method can be used fruitfully to make inferences about sequences of changes in the histories of languages, albeit in connection with certain methodological assumptions about subgrouping (see, for example, Hoenigswald 1963). But here too, subgroups, as nodes on family trees, represent coherent linguistic systems. A subgroup, as an intermediary between a parent and a subset of the daughters, represents a coherent system distinguished from the proto-language, a coherent system itself, by a set of changes shared exactly and uniquely by all the daughters of the subgroup. Daughters within a subgroup are distinguished from each other and the intermediary by changes unique to each individual's history. Such changes have necessarily occurred since the break-up the intermediary.
That internal reconstruction relies on the procedures of the comparative method entails that the method makes the same assumptions comparative reconstruction makes. The comparative method rests on two basic hypotheses: the relatedness hypothesis and the regularity hypothesis. Since the relatedness hypothesis has to do with assumptions about the relatedness of different languages, that hypothesis will not be of concern here. The regularity hypothesis maintains that sound change is regular. It has its origin in the Neogrammarian tradition of the last century and is implicit in the fundamental assumption of internal reconstruction that synchronic alternations within a language are the result of regular sound changes.

Research into the histories of languages over the last 100 years is supportive of the regularity hypothesis: there is a strong tendency toward regularity. When a sound changes, every instance of that sound in similar environments tends to change in the same way. To be sure, there are exceptions, but frequently these exceptions can be accounted for in terms of general linguistic processes, for example, analogical processes. The mechanism at the base of the tendency is not well-known. Different theories have been proposed--for example, wave theory and lexical diffusion theory--which attempt to account for both the regularity of sound change and the apparent exceptions. For example, in some Chinese dialects (Wang and Chen 1968)
certain sound changes appear to have affected only certain lexical items. Wang and Chen account for this by suggesting that sound changes may spread gradually throughout the lexicon. In the course of this spread, a sound change may be interfered with or even halted before it has affected all the lexical items it could have applied to. The result of such a process is, of course, a situation in which a sound change appears to have been irregular. In my view, the weight of the evidence supports the regularity hypothesis. The hypothesis must be understood as a tendency and there are real exceptions for which there is no explanation. With respect to internal reconstruction, however, automatic alternations—that is, regular alternations—in a language synchronically are a testament to the tendency toward regularity of the changes that produced them historically. The Neogrammrian tenet has withstood the test of time rather well.

In its most straightforward form, then, internal reconstruction proceeds from the following principles and methodological assumptions.

(1.) The cognate forms of a morpheme within a language are compared. Differences in shape are assumed to be the result of regular sound change;
(2.) Recurrent correspondence sets and their environments are noted;
(3.) Partially like correspondences occurring in mutually exclusive environments are considered to be continuations of a common original segment;
(4.) If, as a result of (2) and (3), the environments of other correspondences are seen to be mutually exclusive, (2) and (3) are followed again.

The result of this procedure is an historical account of some automatic alternation in a language. That account, as I have noted above, consists of two statements: (a) a description of part of the language prior to the introduction of the alternation; and (b) a statement of the change or changes responsible for the alternation.

Let me briefly consider the situations where internal reconstruction is typically used. Chafe suggests (478) that internal reconstruction is of crucial importance when:

(1.) a language's relatives are not known and there is no written history;
(2.) the recent details of the history of a language are not accessible from the proto-language;
(3.) when reconstructing the pre-history of a proto-language.

(1) and (3) represent the same basic situation since a proto-language is a special case of a language whose relatives are unknown and which lacks a written history. In the next sections of this chapter I will consider examples of internal reconstruction of this sort.

It should be clear that all historical work with natural languages is constrained by what is known about the structure of these languages. That is, the linguistic changes one posits must be plausible (see 2.1, 2.3, and 5 for examples of plausibility arguments). Moreover, what
the historical linguist reconstructs—descriptions of historical stages of a language—must be descriptions of plausible linguistic systems. Both the comparative method and internal reconstruction, then, are guided by knowledge of the structures natural languages.

2.3. Some Methodological Principles. In the second half of his article on internal reconstruction, Chafe presents a classic example of the workings of the method from Seneca. In part of that reconstruction he uses two arguments—assumption of parallel history and simplicity—to establish relative chronologies. Neither of these arguments rests on automaticity and this is the only principle for establishing relative chronologies I have discussed in the preceding sections. In this section I will examine Chafe's arguments. My purpose is to complete the discussion of the kinds of arguments used in internal reconstruction. In addition to Chafe's Seneca examples, I will provide illustrations from other languages.

In his example from Seneca, Chafe establishes the following changes in the order given below.

\[
(1) \ *n>\emptyset \ / \ \{k\} \rightarrow \{i\} \ \begin{cases} \{t\} \\ \{n\} \end{cases} \ \{y\} \quad \text{(Style B only)}
\]
The evidence for these changes is as follows. The sequence -tn- does not occur in modern Seneca. When a form that ends in t is followed by an initial n, an h appears instead of the t. Witness Chafe's examples (1959:484): for the inclusive person morpheme, -t-, the sequence khn occurs. Chafe argues that the alternation of h and t here must be assigned to *t since .(484)

...there are widely distributed correspondences containing h which contrast with h/t and speak against the assignment of the latter to *h. Assigning it to t, on the other hand, conflicts with no other correspondences and widens the distribution of *t in the pre-language. Thus we can infer that *tn>hn.

Chafe suggests, however, that the history of *tn can be clarified by a consideration of the history of *kn. He points out that t and k are treated similarly in the phonological structure of Seneca.

The sequence k+n (where '+' is a morpheme boundary) does not occur in Seneca. Instead, the sequence khn occurs. Witness Chafe's examples (485): -k- Ist. person, plus -no?se- 'to be uncle to' is hakhnō?seh 'he's my uncle'; and k- 3rd. non-masc. and -ni- dual, in khninše?s
they (non-masc. dual) like it'. Since the here can be predicted phonologically and is clearly devoid of morphological content, Chafe establishes *∅ as its antecedent he considers the alternation in the above forms to be between ∅ and h. Thus, in the prehistory of Seneca, *∅>h / k_n. This is one of the subrules of (2) above.

Chafe then argues (485):

Reconsidering the history of *tn, and assuming a parallel development for t and k, we can infer that *tn>thn, and that there was a second change in which *thn>hn. This hypothesis is supported by the fact that *thn does not occur in Modern Seneca.

I will have more to say about this argument below. For the moment I will consider Chafe's argument in support of (1).

The evidence for (1) stems from a stylistic alternation between the sequences khni and hni (<*thni) in one style, A, and the corresponding sequences ki and ti in another style, B. Both khni and hni derive from *kni and *tni historically, the epenthetic h being introduced by (2). Consider Chafe's examples: khninšeš (A) 'they (non-masc. dual) like it' and kinšeš (B) and hninšeš (A) 'we (incl. dual) like it' and tinšeš (B) (see ftnt 9).

Chafe points out that there are also parallel stylistic alternations before y: knny (A), but ky (B) and hny (A), but ti (B). Chafe argues (485):

Because of the lack of h in the forms of style B, it is most economical to infer that the origin of the stylistic differentiation
preceded the development of the epenthetic h between k or t and n.

It is not entirely clear what Chafe means by 'most economical'. Notice, however, that if h-epenthesis occurred before the introduction of the stylistic variation, then both khni (A) and ki (B), for example, would derive from *khni. In order to arrive at ki (B), both h and n would have to be deleted. (1) would have to be revised since, in its present version, the incorrect khi would result. Moreover, not all hn sequences would delete since n is not deleted in style B between h and i (consider Chafe's formulation of (1)); only h's that have also been inserted would delete. It should be obvious that the most straightforward account of the alternation (and the relevant changes that produced that alternation) can be made only by assuming that the stylistic alternations were introduced before the h-epenthesis change. By economy then Chafe means something like simplicity in the formulation of the sound changes.

Neither of Chafe's chronologies depend on automaticity; the order (2)-(3) is a direct consequence of the assumed parallel history of t and k and the order (1)-(2-(3)) relies on economy. Often these principles are used implicitly and it is to Chafe's credit that he has brought them out into the open. While the principle of economy is fairly obvious—I have provided a rough
formulation below (see also 3.4)—it will be fruitful to take a closer look at Chafe's parallel history argument.\textsuperscript{11}

The Principle of Economy: when there are two alternative hypotheses which attempt to account for the same set of data, the hypothesis which permits the phonetically best motivated formulation of the sound changes is the more highly valued hypothesis.

Chafe notes (484) that 'the two oral stops, \textit{t} and \textit{k}, are distributed in nearly the same way in Seneca phonological structure.' Notice, however, that before \textit{n}, \textit{t} and \textit{k} are not distributed identically: the sequence \textit{t+n} is realized as \textit{hn} while the sequence \textit{k+n} is realized as \textit{khn}. It is noteworthy that both of these automatic alternations—that is, \textit{t/h} and \textit{g/h}\textsuperscript{11} (\textit{k+n + khn})—point to \textit{*tn} and \textit{*kn} where \textit{t} and \textit{k} are distributed similarly. Moreover, both \textit{*tn} and \textit{*kn} appear in Modern Seneca with an \textit{h} before the nasal. It would not be unreasonable to assume that these \textit{h}'s which occur in the same position regardless of whether from \textit{*tn} or \textit{*kn} arose from the same change; that is, both \textit{*kn} and \textit{*tn} underwent \textit{h}-epenthesis. Indeed, the alternative solution would have one believe that the appearance of \textit{h} in the reflexes of both \textit{*kn} and \textit{*tn} is due entirely to coincidence. Such a solution is of course unsatisfactory.
It is important to realize that a parallel history for t and k is assumed only with respect to the shared occurrence of h. The fact that t and k are not exactly parallel in Seneca—t does not occur before hn—is seen to be the result of a recent change, that is, *t>∅ / _hn.

The principle of assumed parallel histories is an extremely useful one in internal reconstruction. In Chafe's Seneca example the principle is used to supplement the reconstruction based on the automatic alternations to provide a sensible account of the distribution of h. Although the principle is widely appealed to in the literature, though often implicitly, it has never been formulated. The rough formulation I offer below is intended to capture only the intuitive force of the principle.

The Principle of Assumed Parallel Histories: when two linguistic elements, x and y, at one stage in the history of a language are reflected as x' and y' respectively at a later stage, and if y' bears the relation to y that x' bears to x, the histories of x' and y' are assumed to be parallel.

My interpretation of element is extremely broad. For example, I consider *tn and *kn to be elements as well as t+n + hn and k+n + khn. *tn and *kn are similar in terms of distribution; hn and khn are similar in terms of the occurrence of h before n. They are different in that t does not occur before hn. By the Principle of Assumed
Parallel Histories one assumes that $h$ was introduced in the same way whether from $*\text{kn}$ or $*\text{tn}$.

It should be evident that this principle rests on general notions of similarity, shared features, pattern, and the like. These notions are difficult, if not impossible, to define precisely. Similarity may manifest itself in phonetic similarity, typological similarity or perhaps something even more abstract. The particulars of each case must be carefully evaluated against the relevant data. The application of the principle to the history of a language will not always be warranted. One must exercise a great deal of care and linguistic sense when appealing to the principle.

Although Chafe uses the principle to supplement what he had inferred on the basis of the automatic alternations in Modern Seneca, that principle can be used by itself. In what follows I will discuss an example from the history of Sanskrit in order to illustrate just how fruitful the principle can be. I will also consider the kinds of evidence one can use to further support the inference of parallel histories for different segments.

On the basis of comparative evidence it is possible to reconstruct a series of palatalized velar stops for Proto-Indo-European, namely, $*\text{k}$, $*\text{g}$ and $*\text{gh}$.13 These segments are reflected in Sanskrit as $\text{g}$, $\text{j}$, and $h$ ($h$ = a voiced glottal spirant). Considering the nature of each
of these changes it would not be unreasonable to assume that *q and *qh paralleled the history of *k and went through an intermediate voiced spirant stage before becoming j and h. That is, that *q>q>j and *qh>q>j just as *k>k. While this hypothesis is quite interesting it lacks attestation.

There are, however, a number of independent observations that suggest voiced sibilants must have occurred as reflexes of certain segments, among them *q and *qh, in the history of Sanskrit. It is well known, for example, that PIE did not have a voiced sibilant phoneme. Nonetheless, there was a rule of regressive voicing assimilation which Lehmann (1955:9) describes as follows: 'voiceless obstruents became voiced before voiced obstruents; voiced obstruents became voiceless before voiceless obstruents.' Witness his examples: Skt. vac plus the instrumental pl. desinence-bhis, vagbhis and pad 'foot' plus loc. pl. -su, patśu. If a PIE root ending in a voiceless sibilant was adjacent to a voiced obstruent, one would expect s+z by this rule. For example, *es- 'to be' plus *-dhi 2nd. imp. should be [ezdhi]. Since Sanskrit has no z, it must be assumed that in Sanskrit words of this type the z of PIE was either lost or changed to something else. The reflex of *ezdhi in Sanskrit is edhi. PIE *e is reflected as a in Sanskrit and PIE *ai as e (witness the treatment of both e and o in sandhi rules, a synchronic
testament to their earlier diphthongal character; see also below). If it is assumed that *ai is the antecedent of *e in edhi, then edhi>aïdhï, hence *z>i.

An examination of the Sanskrit root for 'sit', sad, indicates an identical development of *z. sad is a member of a large class of forms that historically used a reduplicated form of the root to form perfect stems. Witness jïv 'live' perfect stem jïjïv, dã 'give' dadau 'he has given' vïc'speak', uvana 'he has spoken', pac 'cook', papaça 'he has cooked'. The reduplicated normal grade of sad is sasad (<*sesed), but Sanskrit shows sed for the reduplicated zero grade form. Etymologically, sed would have been *se-sd which would have become sëzd by the rule of regressive voicing assimilation. *e>a, and if *z>i, as has been suggested above, then sed has the following history: *sesad [sezd]>sajd>sed.

Consider now Sanskrit nïdas nom. sg. of nïda- 'nest'. The etymon of this form (disregarding the desinence) is *ni 'down' plus *sd, the zero grade of *sed. Notice that the i of *ni is short--the regular reflex of *i is i in Sanskrit--and that the d of the sequence sd is dental--the regular Sanskrit reflex of *d is d. The I and d of nïda are only apparently anomalous. In fact these segments result from quite general sound changes. *ni-sd would become [nïzd] by regressive voicing assimilation. If *z becomes i then niid would result. The sequence i-i
regularly contracts to ī in Sanskrit. Witness the following sandhi examples: ati-iva-ati 'va 'exceedingly'. Although this accounts for the occurrence of the long vowel, the retroflex ḍ has not yet been explained.

There is, however, a well-known retroflexion rule whereby s become ś after r, u, k or i, the so-called ruki rule. For example, the loc. pl. of dhenu- 'cow' is dhenuṣu, although the desinence is *su historically. There is also a second retroflexion rule that retroflexes dentals after s. Witness tiṣṭhāti 3rd. sg. pres. ind. of sthā 'stand'. The comparative evidence, which will not be reviewed here, indicates that both of these retroflexion rules are post- PIE. If it is assumed that ruki applied to voiced sibilants as well as voiceless sibilants then *nizd would have become nizd and then niẓd by cluster retroflexion. If *ί then became ī and contracted with the preceding ī the attested niḍa- results: *ni-ṣd [nizd] niḍ>nizd>niẓd>niḍ>nIḍ.

That voiced sibilants developed into glides is also evident from the history of a number of Sanskrit sandhi formations. For example, the sequence as in final position becomes o before any voiced consonant or short a: devaṣ+ gacchati is devo gacchati 'a god is coming'; vacas plus artham becomes yacortham 'a worthy object'. These examples will be considered separately since they involve slightly different historical developments.
It would not be unreasonable to assume that the o of devo gacchati comes from au since au and o alternate in sandhi; ca plus uvaca is covaca 'and he said', for example. Final voiceless obstruents become voiced when followed by a word that begins with a voiced sound in sandhi. Witness: pattanat plus aggacchati become pattanad aggacchati 'conquest of all regions'. (The sandhi rule is, of course, a generalization of the PIE rule of regressive voicing assimilation discussed above.) The change z\textgreater u completes the history of as to o: as\textgreater az before words with voiced initials, >au (or ai as above, but under different conditions) > o.

Since a is also a voiced sound the as of yacas artham would have become az. Now the sequences aya and ava become e and o in sandhi; for example, te (<tay) plus abruvan becomes te 'bruvan 'they said'. By the same rule yaca\textunderline{u} artham would become yacortham, again indicating the development of *z to a glide.

While the foregoing establishes that voiced sibilants occurred in the prehistory of Sanskrit, it has not yet been established that these voiced sibilants occurred as the reflexes of *\textguillemotright or *\textguillemotleft. Consider, however, a form like m\textasciitilde u\texttilde \textasciitilde ha 'fool'. The etymon of this form is \textasciitilde m\textasciitilde u\textasciitilde g\textasciitilde h\textasciitilde -\textasciitilde ta. There was a general rule of PIE governing the occurrence of aspirates in clusters, Bartholomae's Law, which changed a sequence of a voiced aspirate plus voiceless nonaspirate,
for example –dht–, to a voiced nonaspirate—voiced aspirate sequence, –ddh–. Witness the Sanskrit to- participle of labh 'seize' labāha. By Bartholomae's Law, then, mugh- ta would become mugdha.

It has already been established that long vowels result from the contraction of two short vowels and that retroflexion is introduced secondarily by ruki and cluster retroflexion. Thus, a probable history for mūdha is mūgха<muṣgха<muṣgΔa<muṣdha<mugdha suggesting the change ġ>z.

The methodological procedure in this argument involved first assuming that *ġ and *gh became sibilants by appealing to the Principle of Assumed Parallel Histories. Additional evidence for this hypothesis was then gathered. It was shown that a form like mūgха derives by regular sound changes from muṣhta indicating an intermediate stage with a glide reflex for *gh.

Consider now the following inflectional paradigm from Old Irish.14

<table>
<thead>
<tr>
<th>Noun</th>
<th>Adjective</th>
</tr>
</thead>
<tbody>
<tr>
<td>f'er</td>
<td>g'el</td>
</tr>
<tr>
<td>f'irі</td>
<td>g'ilі</td>
</tr>
<tr>
<td>f'irі</td>
<td>g'ilі</td>
</tr>
<tr>
<td>f'er + Adjective</td>
<td></td>
</tr>
<tr>
<td>f'er g'el</td>
<td>'a bright man, nom. sg.'</td>
</tr>
<tr>
<td>f'irі γ'ilі</td>
<td>gen. sg.</td>
</tr>
<tr>
<td>f'irі γ'ilі</td>
<td>gen. sg.</td>
</tr>
</tbody>
</table>
Both f'er and g'el exhibit morphophonemic alternations within these paradigms. First, the liquids, l and r show palatalized (l⁺, r⁺) and labialized (l⁻, r⁻) alternates. Second, the vowel of these forms appears as e in the nominative sg., but as i in the genitive and dative. Finally, in a noun+adjective syntactic construction, the initial stop, g', of g'el appears as a spirant. While none of these alternations is automatic since they cannot be stated in purely phonological terms (see, 2.1), it would be quite reasonable to assume that the alternations were automatic at some stage in the history of the language, the conditioning factors having been lost. Thus, one might reconstruct a final palatal segment, *[+seg,+pal], for the gen. sg., and a final labial segment, *[+seg,+lab] for the dat. sg. and suggest that l⁺ and r⁺, for example, are to be derived from *l[seg,+pal] and *r[seg,+pal] by a palatalization process; *[+seg,+pal] and *[+seg,+lab] having later been lost. Under this hypothesis the antecedents of the Old Irish gen. and dat. would be as follows:


Turning to the g'/γ' alternation in the gen. and dat. noun+adjective constructions, it would not be unreasonable to assume that *[+seg,+pal] and *[+seg,+lab] are in
part responsible for the spirantization of \( g' \) to \( \gamma' \).

That is, just where it is necessary to reconstruct

\[*[+seg,+pal] \) and \[*[+seg,+lab] to account for the alterations exhibited by the liquids, it is also necessary to reconstruct a segment to account for \( g'/\gamma' \). \( \gamma' \) can be plausibly derived from \(*g'\) by an intervocalic spirantization process. The suggestion, of course, is that

\[*[+seg,+pal] \) and \[*[+seg,+lab] were vocalic. Clearly one could reconstruct a third segment distinct from either

\[*[+seg,+pal] \) or \[*[+seg,+lab] to account for \( g'/\gamma' \). Such a solution, however, would be quite complicated since this mystery segment would have only been present in the noun+adjective construction—there is no spirantization of \( g' \) in the adjective paradigm—and it is difficult to imagine what might condition this. The more elegant solution accounts for both the liquid alterations and \( g'/\gamma' \) on the basis of the same reconstructed segments, \[*[voc,+pal] \) and \[*[voc,+lab], and three sound changes: liquid assimilation, spirantization, loss of \(*[+pal] *[+lab]. The hypothesis relates in a plausible way two apparently unrelated morphophonemic alternations. This hypothesis is therefore more general than any alternative hypothesis that treats these alternations as completely unrelated phenomena.

Finally, consider the vocalic alternations of \( o \) and \( i \). \( i \) occurs instead of \( o \) in the gen. and dat. sg., that
is, in just those cases where it has been necessary to reconstruct \([+\text{voc},+\text{pal}]\) and \([+\text{voc},+\text{lab}]\). This could just be a coincidence. However, \(i\) can be derived from \(*e\) if it is assumed that there was some sort of height harmony rule operative in the prehistory of the language; the mid vowel of a stem like \(\_\text{er}\) raised to agree with the height of the vowel of the desinence. That is, \([+\text{voc},+\text{pal}]\) and \([+\text{voc},+\text{lab}]\) are considered to be high vowels, \(*-i\) and \(*-u\) respectively. Under this hypothesis—which is certainly not the only hypothesis that is consistent with the observed facts—a third phenomenon, \(e/i\), is seen to be related in a plausible way to the liquid alternations and \(g'/y'\). Thus, on the basis of a minimal set of assumptions, the reconstruction of \(*-i\) and \(*-u\) and the related sound changes, it is possible to give a principled account of a wide range of synchronic phenomena; three different synchronic alternations have their origins in the same lost segments and the application of regular sound changes.

In principle, one could account for these alternations in a different way. However, any hypothesis which does not relate the alternations is less general than the hypothesis proposed above. It is a principle of the philosophy of science that the more general hypothesis is the more highly valued hypothesis. In effect, this principle from the philosophy of science is used as a methodological principle in both synchronic and diachronic
linguistic analysis. I will elaborate this claim in a later chapter (4.3). For the moment let me note that as a methodological rule of thumb, if it is possible to plausibly account for an alternation on the basis of what one has reconstructed in order to account for some other alternation, one should do so. I will refer to this principle as the Principle of Optimal Use.

The example from Old Irish represents a slight complication in the application of the principle. On the basis of the alternations shown by the liquids, I reconstructed *[+seg,+pal] and *[+seg,+lab]. On the basis of that evidence alone it was not possible to infer that [+seg,-pal]=*-i or that [+seg,+lab]=*-u. Thus, one cannot apply the Principle of Optimal Use to g'/γ' without modifying the reconstruction of *[+seg,+pal] and *[+seg,+lab] so that they are considered to be vocalic. Recall that I suggested *g'>γ' / V_.V. Such a change is plausible only in so far as it is assumed that *[+seg,+pal] and *[+seg,+lab] were vowels. Thus, applying the principle to g'/γ' entails elaborating the reconstruction made on the basis of the liquid alternations; g'/γ' provides supplementary evidence for determining the nature of *[+voc,+pal] and *[+voc,+lab]. And of course, applying the principle to e/i works in just the same sort of way; *[+voc,+pal] and *[+voc,+lab] are identified with the high vowels, *_-i and *_-u respectively.
By applying the procedures for making historical inferences on the basis of synchronic alternations and the Principle of Optimal Use, it has been possible to recover a great deal of the history of Old Irish inflectional paradigms like those cited above for *f'er and *g'el. Throughout the discussion only internal evidence has been used. The gen. sg. and dat. sg. of these forms historically had desinences, *-_i and *-_u respectively. These desinences conditioned a palatalization and labialization of the preceding stem-final liquids: *₁,l₁>r>₁l₁, r₁ / _i; *₁,l₁>r>₁l₁, r₁ / _u. *-_i and *-_u were also involved in a vowel harmony with the stem vowel: *f'er, *g'el > f'i'r, g'i'l / _v. Finally, in noun+adjective constructions, *g'i spirantized intervocalically. Comparative evidence points to the following desinences:¹⁵ nom.sg. *-_os, gen. sg., *-_i, and dat. sg., *-_u. The Old Irish example stands as a testimony to the power and accuracy of internal reconstruction. The desinences reconstructed above for the gen. and dat. sg. differ from the correct desinences only on the basis of vowel quantity.

In this section I have considered three methodological principles that are often appealed to, though often only implicitly, when making historical inferences on the basis of internal evidence. This section completes my discussion of the methodological procedures and principles of internal reconstruction as that method has been described
in the literature. In the next section I will consider the circumstances under which sound changes of various types can be recovered on the basis of purely internal data.

2.4. Recovering Sound Changes. A sound change in the history of a language will not be recoverable by internal reconstruction unless that change leaves behind some kind of trace on the structure of the language, for example, a morphophonemic alternation. In this section I will be concerned with the nature of the structural traces that the different types of sound change leave behind. While this topic has been discussed to a certain extent by Hoenigswald (1946), he was not primarily concerned with the issue of recoverability. I will review Hoenigswald's article and where necessary supplement his observations on sound changes of various types.

In his article 'Sound change and linguistic structure' (1946), Hoenigswald classifies different sound changes in terms of their effect on linguistic structure. His classification is as follows:

1. non-phonemic change
2. phonemic change without loss of contrast
3. unconditioned merger with loss of contrast
4. conditioned merger with loss of contrast
5. secondary rearrangement induced by a primary loss of contrast
6. borrowed contrast.

My discussion will be organized around Hoenigswald's classification.
Hoenigswald asserts (1946:138) that nonphonemic change—that is, 'pronunciation changes which leave intact existing contrasts between forms of different meaning'—do not leave behind structural traces. He concludes (138): 'Since this kind of change leaves no structural traces, it could not be recovered by internal evidence from the resulting stage.' This assertion is incorrect since under certain circumstances (see below) pronunciation changes may in fact be recovered on the basis of structural evidence in the resulting stage. In such cases, however, it is not the change itself that leaves the structural trace, but something else. For example, in Proto-Austronesian there was a morphophonemic process whereby a base-initial voiceless stop consonant was replaced by a homorganic nasal consonant (see Chapter 5). This process has been inherited in many West Austronesian languages, for example, Javanese and Tagalog. Consider the examples below from Javanese.

<table>
<thead>
<tr>
<th>Javanese</th>
<th>Transliteration</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>payung</td>
<td>'umbrella'</td>
<td>but mayung 'to cover someone with an umbrella'</td>
</tr>
<tr>
<td>tili?</td>
<td>'look at'</td>
<td>nili? 'look for someone'</td>
</tr>
<tr>
<td>t'abut</td>
<td>'expressions'</td>
<td>n'abut 'express something'</td>
</tr>
<tr>
<td>surung</td>
<td>'shove off'</td>
<td>n'urung 'shove something'</td>
</tr>
<tr>
<td>kukup</td>
<td>'take in hand'</td>
<td>ngukup 'take something in hand'</td>
</tr>
</tbody>
</table>

Notice that the g-initial base surung has the prenasalized form n'urung with a palatal nasal, and not the expected
Moreover, all the other bases begin with voiceless stops; ə is the only spirant to participate in this alternation. On the basis of these peculiarities it would not be unreasonable to assume that (a) ə used to have a palatal point of articulation, and (b) that ə used to have an occlusive element in its articulation. Thus, one might reconstruct something like *tʰ or *ऽ as the antecedent of Jv. ə. This would be an example of a pronunciation change that could be recovered from the resulting stage. Comparative evidence (see Chapter 5 and Dahl, 1973) points to a palatal affricate as the Proto-Austro-nesian antecedent of Jv. ə.

The example from Javanese illustrates the two conditions under which a change of this type may be recovered: (1) when a sound conditions an alternation in other segments and then loses the feature responsible for the conditioning; and (2) when a sound is a member of a class of sounds that participate in an alternation but then changes so that it is no longer a member of that class, but still participates in the alternation. In Javanese, (1) is the palatality of *tʰ, and (2) that *tʰ is a member of the class of voiceless stops. Hoenigswald's assertion about the recoverability of nonphonemic change would only be true if the fact that sounds may interact with each other was ignored. The historical linguist cannot ignore what is known to be true of the behavior of sounds in natural
languages. As I have suggested, historical research is necessarily guided by what is known about the structures of natural languages.

Hoenigswald cites Grimm's Law as an example of a phonemic change without loss of contrast. Indo-European voiceless stops became spirants, and voiced stops became voiceless. For example, PIE *t>θ and *d>τ. Hoenigswald notes (138):

But after voiceless consonants, only some kind of [t] occurred in IE. The latter, which was presumably not changed at all at the time of the general shift, was originally most similar to the stop of the old [t]-words; it is now most similar to that of the new [t]-words.

The conditioned change of *t to both t and θ is reflected in Gothic. Hoenigswald cites (139) salbo-b-s 'annointed', but haf-t-s 'restrained' exhibiting the alternation of b and t.

Not all changes of this type will be recoverable on the basis of internal evidence. The necessary condition is that the sound that splits must occur in a morpheme that can itself occur in all the relevant phonological environments. Otherwise, no morphophonemic alternation will be introduced. For example, if *t did not occur initially in morphemes that could follow both voiced and voiceless sounds, there would be no alternation in Gothic and only comparative evidence could reveal the change.
Hoenigswald subclassifies unconditioned mergers with loss of contrast into three general cases: (a) when all the allophones of one phoneme merge with allophones of another phoneme; (b) when a phoneme is lost, that is, merges with zero; and (c) when the allophones of a phoneme merge with the allophones of one or more phonemes and zero. He argues (139) that (a) 'leaves only inconclusive traces in the form of an occasional statistical preponderance of the surviving phoneme.' As examples of this type of change, he cites the merger of IE *dh and *d in Iranian and Slavic, the merger of IE *a, *e, *o to a in Sanskrit, the merger of Latin u and ō in Romance languages, and finally, the multiple sources of Mod. Greek /i/.

Besides statistical traces, changes of this sort may leave other clues as well. Suppose, for example, that a language has a rule whereby k is palatalized to č when it occurs before i. In an inflectional paradigm, stems that end in final k would show č before desinences that begin with i; for example, kak-a but kač-i. If i were to merge with some nonpalatal vowel, say a, the conditions governing the alternation would be extremely opaque. That is one would have forms like kak-a and kač-a (<kak-i). It would not be unreasonable to assume that a at this stage represents more than one source historically, one of which was palatal. The clue in this case is the phonetically anomalous environment for the alternation: why should k>č
before a sometimes, but not always? A very similar change occurred in the prehistory of Sanskrit. \( k \) and \( c \) used to be in complementary distribution: \( c \) only occurred before \( *e, k \) elsewhere. \( *e \) has fallen together with \( *o \) and \( *a \) to \( a \) in Sanskrit obscuring the true nature of the \( k/c \) alternation. Witness, vak 'voice, nom. sag.' vacas nom. pl., vacam acc. sg., vaci loc. sg. The Sanskrit example has been further complicated by a levelling; \( c \) has been generalized to all cases where a vowel or resonant follows. Thus, the \( c \) of the nom. pl., and loc. sg. are etymologically correct, but the acc. sg. should be vakam (*vak-om). (See Kiparsky 1968 and Vennemann 1972 for further discussion of this example.)

Hoenigswald treats secondary phonemic splits (as well as primary phonemic splits) like that above from Sanskrit as (1960:78) 'corollaries of merger processes.' While this makes a certain amount of sense in cases like the above, splits cannot be treated as mergers when an entirely new sound is introduced. For example, in the last section I discussed an example from the history of Old Irish in which it was established that \( *r, *l > r^i, l^i / _i \). The examples cited in that discussion are really only instances of a quite general palatalization process that occurred in the history of Old Irish—for example, the \( f' \) and \( g' \) of \( f'er \) and \( g'el \) are both secondary. The palatalization process has been obscured by subsequent changes, for example the
loss of *-i. In short there has been a secondary phonemic split and a number of new sounds--the palatalized consonants--have been introduced. I do not see how this type of change can be considered to be a merger and such examples would seem to represent a gap in Hoenigswald's taxonomy of sound change. It should also be noted that such secondary splits are recoverable on the basis of internal evidence. Witness my discussion of Old Irish in the last section.

Hoenigswald also asserts (1946:139) that the merger of a phoneme with zero will not be recoverable since (139) 'this process will leave no traces in the subsequent structure.' In the next section of this chapter, I will consider an example that demonstrates convincingly that this assertion is not correct: the loss of laryngeal consonants in PIE left a profound, albeit complicated, trace in Indo-European structure. Hoenigswald does note correctly that changes of type (d)--the merger of the allophones of one phoneme with other phonemes or zero--will be recoverable only if they affect the forms in a paradigm and introduce a regular alternation.

Sound changes that result in conditioned mergers with loss of contrast introduce automatic alternations. As a consequence, changes of this type can be recovered quite easily. This type of change is represented by the example from Greek (2.2) and Mod. German (2.1,2.2). Such changes
may also result in a secondary rearrangement—that is, they may obscure an automatic alternation introduced by an earlier change. As the example from Greek shows, however, the sounds in question will exhibit a particular distributinal pattern and this pattern may be used to discern the original nature of the alternation.

Finally, contrasts may be introduced as the result of borrowing. Hoenigswald (142) notes that in Hindustani [ɾ'] and [ɖ] occur in mutually exclusive environments: [ɖ] initially, in geminates and after [n]; [ɾ'] elsewhere. In loanwords, for example, /reːdiːo:/ [ɖ] enjoys a freer distribution. Hoenigswald argues (142): 'if the foreign vocabulary can be set aside on a descriptive basis ... the old structure can be recovered.'

Chamorro, the language of Guam, provides another example of this sort of change. It is quite difficult to find minimal pairs which contrast the high vowel phonemes /i/ and /u/ with the mid vowel phonemes /e/ and /o/ (see Latta 1972, Topping 1968). Moreover, all such pairs involve at least one loanword from Spanish. Witness: misa 'Mass', but mesa 'table'. If the Spanish loanwords are weeded out, it becomes apparent that i and e and u and o were in complementary distribution: the mid vowels occurred only in closed syllables in pre-Spanish Chamorro. This distributional limitation, however, obtains only in the native vocabulary. Borrowed contrasts, then, may leave
structural traces in the form of observed distributional limitations confined to the native lexicon.

The preceding sections of this chapter represent a survey of the literature. I have attempted to explicate the major assumptions, goals, procedures, and principles of the method of internal reconstruction as that method has been described by methodologists. In the next chapter I will evaluate the theories of Saussure and Benveniste on early Indo-European structure. Their procedures will be evaluated against those discerned in these sections.
NOTES

1Although Hoenigswald claims that devoicing occurs syllable finally, the conditions are quite complicated. See Ralph 1973 and Vennemann 1972. The process does apply without problems word finally. All the examples discussed here are word final.

2Except in certain aorist constructions where the s has been analogically 'rescored'. See Jeffers (1973) for a different interpretation of these aorists.

3To be sure, the comparative method is not without its own set of problems. For example, there is no guarantee that the constituents of *m -mrot-ia actually co-existed in the shapes cited.

4Chafe's discussion of the Seneca material is quite lengthy. I will confine my remarks to those sections that are directly relevant.

5Chafe notes that there is one form which does exceptionally exhibit -tn-.

6The morphological analysis is mine. Chafe does not provide a complete analysis in his discussion.

7The position of the morpheme boundary is problematic, but irrelevant to my discussion. It is also unclear what the phonetic manifestation of this sequence would be.

8Chafe does cite examples of h with morphological content.

9Chafe does not characterize the semantic difference between the two speech styles.

10According to Chafe, non-masc. h in sequences like this shows a slightly different development: hni (A), hi (B), but hny (A), y (B). The loss of h before y will not be considered.

11I prefer the term 'parallel history' to Chafe's 'parallel development' since it cannot be confused with the notion of parallel development as opposed to shared innovation.

12Perhaps this alternation would be better described in terms of an alternation of kh (/n) and k, elsewhere.
Under Kuryłowicz's analysis (1935), the palatal-ized-velars are not original IE, but secondary. His analysis has no consequences for the discussion.

See Jeffers and Lehiste (1978) for a complete discussion of the Old Irish material.

I have not discussed arguments for reconstructing a nominative desinence on the basis of internal evidence. See Jeffers and Lehiste 1978.

Other bases with initial s show nh; such s's do not reflect *tʰ. Only s<*tʰ exhibit nh.
3.0 Introduction. In 1879, Ferdinand de Saussure published his Mémoire sur le système primitif des voyelles dans les langues indo-européennes. This work formed the corner stone for what has come to be called 'laryngeal theory' in modern parlance. The Mémoire is of interest to my analysis of internal reconstruction since Saussure's theory rests entirely on internal evidence. Moreover, many of Saussure's claims about early IE structure were later supported directly with the discovery of what is widely regarded as confirming evidence in Hittite.

Saussure's theory, then, is an excellent example of the power of internal reconstruction.

In the first section of this chapter I will establish the basic arguments and type of evidence that Saussure adduced in support of his theory. In 3.2 I will review Benveniste's theory of the IE root and show that Benveniste followed essentially the same procedures and appealed to the same type of structural evidence as Saussure. In 3.3, however, I will argue that both Saussure's and Benveniste's
theories fall outside the domain of what has been traditionally described as internal reconstruction—that is, the method as I have explicated it in the preceding chapter.

3.1. Laryngeal theory. Laryngeal theory, as Lehmann (1955) has defined it, assumes that the phonological system of early PIE included a number of phonemes, the so-called laryngeals, which have no direct reflexes in the IE languages with the possible exception of Hittite. Lehmann notes (1955:22): 'Evidence for laryngeals in the IE dialects can be deduced only on the basis of phonemes that are also reflexes of other PIE phonemes.' The discovery of Hittite has contributed enormously to the study of PIE. However, Saussure did not have the benefit of the Hittite evidence when he wrote his Mémoire. His conclusions were based entirely on his careful consideration of the internal structure of PIE. Moreover, many of the assumptions held in Saussure's time, for example, that PIE was very similar to Sanskrit, have been revised in light of laryngeal theory and other discoveries. In short, Saussure had a number of very serious obstacles to overcome.

Most modern treatments of laryngeal theory, for example, Polomé's 'Laryngeal theory so far,' Zgusta's 'La théorie laryngale,' or even Lehmann's abridged translation of Saussure (1967) begin with a presentation of Saussure's conclusions and pay relatively little attention to the
particulars of Saussure's arguments. Those arguments will be the fundamental concern of this section.

Saussure began the *Mémoire* by focusing on the vowel a (see note below). He commented (1879:3):

_Aucune matière n'est plus controversée; les opinions sont diverses presque à l'infini, et les différents auteurs ont rarement fait une application parfaitement rigoureuse de leurs idées. À cela s'ajoute que la question de l'a est en connexion avec une série de problèmes de phonétique et de morphologie dont les uns attendent encore leur solution, dont plusiers n'ont même pas été posés._

(Note well: Saussure's notation system is problematic for the modern reader. What Saussure designates with a corresponds sometimes to *e*, sometimes to *e̩*, and sometimes to both *e* and *a* in modern systems of notation. In citations from Saussure I will indicate the modern equivalent in square brackets. In my own prose I will confine myself to using the modern system wherever possible. If it is necessary to use Saussure's, I will indicate the modern equivalent in square brackets.) The morphological problems Saussure alluded to above, primarily concerned IE ablaut, or more correctly, a subpart of that system. In fact, that system provided Saussure with one of the fundamental clues.

It had long been known that IE wordbases showed regular alternations both in the quality and quantity of the wordbase vowels or vocalic segments. For example wordbases with the vowel e regularly showed an alternant
with o in certain morphological contexts. This is qualitative ablaut. IE wordbases also exhibit what are commonly referred to as 'grades'. In the full grade the basic vowel e occurs (but keep in mind that e alternates with o). This vowel may be followed by another vocalic segment, for example i, and the two together form a diphthong, ei (or oi). The vowel e may also be followed by a resonant, for example r, and the two form the sequence er. In the zero grade, the vowel e does not appear. A wordbase with only e in the full grade will lack a vowel in the zero grade. Wordbases with ei in the full grade, show i in the zero grade and wordbases with er show r in the zero grade. The full and zero grades of wordbases occur in different morphological contexts.

Saussure noted that the PIE liquids and nasals existed in two states: as nonsyllabics and syllabics. The liquids and nasals occurred as syllabics in wordbases that had lost the basic vowel e and the liquid or nasal had come to stand alone (that is, in the zero grade). Thus, there were morphophonemic alternations within wordbases of the form el/ı, en/ı and so on. There are also alternations between diphthongs and simple vowels in the same contexts. Saussure pointed out that the alternations involving diphthongs and simple vowels, on the one hand, and the nonsyllabic and syllabic liquids and nasals, on
the other, were fundamentally analogous. Saussure con-
cluded (1879:9):

L'i et l'u de ces racines, ainsi que le
liquid et la nasale des racines telles que
derk, bhendh, peuvent prendre le nom de
'coefficient sonantique.'

Saussure thus argued that much of the ablaut system of
PIE could be understood as an alternation of a root vowel,
e, plus a coefficient sonantique and that coefficient
sonantique alone: e + coefficient sonantique/ coefficient
sonantique. The alternations e₁/₁ and ei/i₁ are both
instances of this general phenomenon. Saussure considered
the phonetic peculiarities of the changes e₁ to ₁ and ei
to i₁, for example, as inherent to the phonetic nature of
the segments involved, but of no consequence to the
overall pattern of alternations.

This analysis does not apply to all wordbases. A num-
ber of wordbases exhibit a different kind of full grade-
zero grade alternation. In the full grade, these word-
bases have a long vowel in the full grade, but o in the
zero grade: *e/ə, *a/ə, and *o/ə (Saussure wrote *a for
*e). Moreover, the development of *e in the IE daughters
was quite problematic and poorly understood. For example,
Saussure noted that *e developed, apparently without con-
ditions, into a, e and o in the Western languages, Italic,
Celtic, and Germanic. These same languages also exhibit
e and o as reflexes of other PIE sounds.
Saussure meticulously evaluated the history of *e in the various IE daughters and noted the following peculiarities (from Saussure, 1879:48-9, §'s 3-7, 9):

a.) The languages of the South—that is, Italic and Greek—show a, e and o as reflexes of *a[ə].

b.) In the southern languages e(<*a[ə]) regularly corresponds to e(<*a[ə]) in the North—that is, Germanic and Slavic.3 a (<*a[ə]) in the Northern languages correspond to both a (<*a[ə]) and o (<*a[ə]) in the South.

c.) While there exist a's which alternate with e in Greek, all such a's reflect an original resonant. a (<*a[ə]) and e (<*e) have been kept distinct.

d.) The IE alternation of e and o is entirely regular in Greek and Latin. Where the languages of the South show o (<*a[ə]), those o's regularly alternate with e (<*e).

e.) When Latin or Greek a (<*a[ə]) corresponds to Northern a (<*a[ə]), the Northern a never alternates with e. Moreover, the Latin and Greek roots show a long vowel in the full grade.

Saussure then asked (1879:49): 'Comment l'a et l'o des langues du sud pourraient-ils donc être sortis d'un seul il même a[ə]?' He concluded (1879:49):

...a et o des langues classiques est originale, et il faut que dans l'a unique du nord deux phonèmes soient confondu.

Saussure's argument is extremely complex—even more so when one considers that Saussure wrote a for both *e and *e. In the interests of clarifying the argument I will review Saussure's (a)-(e) in detail.
(a) simply establishes that there are three reflexes of *ə in the Southern languages and two reflexes in the Northern languages. All things being equal, one would expect to find some Southern a's corresponding to Northern e's, some Southern o's corresponding to Northern e's, and some Southern e's corresponding to Northern a's. (b) demonstrates that this is not the case since Northern and Southern e are uniquely and exhaustively paired as are Southern a and o with Northern a. The conclusion is that the correspondence set Northern e: Southern e represents an original phoneme, *e distinct from Northern a: Southern a, o (see below).

An alternative hypothesis would be to assume that there was in fact only one original phoneme. This phoneme first split into a and e, presumably as the result of a conditioned sound change, the conditioning factors having been lost. At an early stage, then, there would have been an a/e alternation. The Southern languages have further split a into a and o (see below, however). (c) represents counterevidence to this hypothesis since a (*a[e]) and e (*a[a]) have always been distinguished, a fact which supports the original hypothesis.

(d) establishes that in the languages of the South, forms which reflect *ə as o exhibit a morphophonemic alternant with e. (e) establishes that Southern forms with
Saussure’s hypothesis, then, is that these correspondence sets represent three distinct original segments. Saussure used the symbol \( A \) to designate the Northern \( a \): Southern \( o \) correspondence sets and \( A_a \) to designate the Northern \( a \): Southern \( a \) correspondence sets. Consider carefully Figure 2 which summarizes these correspondence sets.

Figure 2. Northern-Southern Correspondences.

Saussure’s hypothesis is that if both sets of correspondences— that is, the Northern \( a \): Southern \( a \) and Northern \( a \): Southern \( o \) pairs—repre- sent the same original segment, the reflexes should participate in the same morphophonemic alternations. One might expect to find, then, some Northern \( a \): Southern \( a \) forms showing alternants with \( e \). Since these expectations are not fulfilled, the conclusion is that two original segments are represented in the correspondence sets.

The hypothesis that more than one original segment is involved, in the normal grade provides further support for this fact that only Northern \( a \): Southern \( a \) forms show a long vowel, never exhibit a morphophonemic alternate with \( e \). The sug-
and that $a_2$ and $A$ have fallen together in the languages of the North.

After discussing $A$ at some length, Saussure points out (1879:66-7):

"C'est pour des raisons toutes pratiques que nous avons jusque'ici considéré l'o gréco-italique comme un tout homogène. En réalité il en existe au contraire deux espèces bien distinctes que nous allons étudier l'une après l'autre... Nous appelons $o_2$ l'espèce qui equivaut à l'ancien $a_2$: le second $o$ recevra la designation $o$."

The facts which support this distinction are as follows (Saussure 1879: pgs. 90ff.):

a.) By Brugmann's Law *o appears in Sanskrit as $a$ in open syllables, but as $a$ in closed syllables. However, some $o$'s in Greek correspond to Sanskrit $a$ in open syllables. Compare the expected Gk. $doru$ 'wood' Sanskrit $dāru$ 'wood' with exceptional Gk. $potis$ 'husband', Sanskrit $patis$ 'husband'.

b.) $o$ (<$a_2[o]$) is limited in its occurrence morphologically (see 1879: pgs. 66ff.). For example, Saussure points out that $o$ (<$a_2[o]$) never occurs in the present of non-derived verbs. Some $o$'s in Greek, however, appear in all the forms of a root: ὅξῳ ὅ 'I smell'.

Saussure argues that it is unlikely that the Gk. $o$'s which correspond to the unexpected $a$'s in Sanskrit open syllables represent the same original phoneme as the Gk. $o$'s which correspond to the expected forms in Sanskrit. Moreover, the latter $o$'s (those from *$o$, Saussure's $o_2$, Brugmann's $a_2$) exhibit restrictions not shared by all $o$'s. Saussure argues (1879:91):
Il est donc inévident que l'o d'un présent comme òdzo, en d'autres termes l'o qui se maintient dans toutes les formes d'un racine, puisse représenter $a_2$.

Figure 3 summarizes this information.

<table>
<thead>
<tr>
<th>Greek</th>
<th>Sanskrit</th>
<th>Morphological Constraints</th>
<th>Original</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>a (open syl.)</td>
<td>no</td>
<td>0 [ə₃]</td>
</tr>
<tr>
<td>o</td>
<td>a (csd. syl.)</td>
<td>yes</td>
<td>a₂[ə]</td>
</tr>
</tbody>
</table>

Figure 3. Greek-Sanskrit Correspondences

The final step in Saussure's argument now follows.

Recall that roots with a reflex of $*A$ [ə₁, ə₂] in zero grade show a lengthened vowel in full grade, ə or ə. Likewise roots with a reflex of $*Q$ [ə₃] in zero grade show ə in the full grade. As I noted above, this ablaut pattern falls outside of the $e + \text{coefficient sonantique}/\text{coefficient sonantique}$ pattern. Saussure argued (1879: 127):

> Le phonème ə₁ [*e*] est le voyelle radicale de toutes les racines. Il peut être seul à former le vocalisme de la racine ou bien être suivi d'une seconde sonante que nous avons appelée 'coefficient sonantique'...

> Les phonèmes $A$ [ə₁, ə₂] et $Q$ [ə₃] sont des coefficients sonantiques. Ils ne pourront apparaître à nu que dans l'état réduit de la racine. A l'état normal de la racine, il faut qu'ils soient précédés de ə₁ [*e*] et c'est des combinaisons $a_1+A$ [e+ə₁, e+ə₂], $a_1+Q$ [e+ə₃] que naissent les longues $A$, $Q$.

> La permutation $a_1:a_2$ [e/o] s'effectue devant $A$ et $Q$ comme ailleurs.

This analysis is represented schematically in Figure 4.
Figure 4. Saussure's Analysis of IE Ablaut (following Saussure, 1879:127).

With the realization that A and Q are coefficients sonantiques it is possible to give an historical account of the history of IE ablaut as a unitary phenomenon. In the full grade, roots contain either the basic vowel *e/o (1 above) or the sequence e + coefficient sonantique (2-8 above). In the zero grade wordbases appear either without a vowel (1), or with just the coefficient sonantique (2-8). Hereafter I will write R for coefficients sonantiques, and the system can be represented eR (full grade)/R (zero grade).

With the loss of *A and *Q, however, the system, in effect, split: (1-6) remain unchanged, but (7-8) became full grade A/ zero grade A (7), and full grade Q/ zero grade Q (8). I will designate the pattern of (7-8) with V/e. The ablaut pattern of (7-8), V/e no longer appeared to be the same kind of alternation as (2-6), eR/R.
There are, then, three major steps in Saussure's argument. The first is the grouping of (2-3) and (4-6) of Figure 4, that is, establishing the general pattern of IE ablaut. The second step involves establishing the existence of A and 0. The final step in the argument is the realization that A and 0 permit the inclusion of (7-8) in the general pattern, eR/R. I will have more to say about each of these steps in 3.3.

Saussure's conclusions were not widely accepted in his day. Many scholars objected to the formulaic character of the reconstruction (see 4.3 for a discussion of this objection) and the apparent lack of concern for establishing the phonetic identity of *A and *0. Polomé notes (1965:10):

In defining his phonemes /A/ and /O/ 'coefficients sonantiques', Saussure obviously visualized them as resonants with either syllabic or non-syllabic allophones, without even broaching the question of the phonetic reality they stand for. This typically algebraic systematization of the PIE vocalism however contributed to a better insight into the structure of the Indo-European 'root'...

Those scholars who accepted the crux of Saussure's theory, for example, Möller and Sweet, attempted to define *A and *0 in phonetic terms. However, Saussure's system exhibited an apparent inconsistency: while *e+Q contracted only to O, *eA contracted to both ð and ð (A). The circumstances governing the occurrence of ð and ð could not
be discerned. Consequently, Möller and Sweet split Saussure's *A into *A [ε₂] and *E [ε₁]. Under this revised analysis, each of the three IE long vowels, *ε₂, *ό and *ǎ in the full grade of wordbases were seen as the result of the contraction of the basic vowel, *ε, and a distinct coefficient sonantique.

This phonetic work added a new dimension to Saussure's theory: *ε₁, *ε₂, and *ε₃ not only contract with *ε, they 'color' it. Saussure's reconstruction of A and *ό asserted only that the full grade long vowels derived historically from a sequence of *ε+A or *ε+ό. The vowel coloring refinement permitted additional predictions. For example, Zgusta (1951:443) notes:

Si nous supposons la voyelle fondamentale ε, ensuite
ε + H₁ > ε
ε + H₂ > ā
ε + H₃ > ẹ.

Or, si par ex. H₂ dans le groupe eH₂ change e en a (et se contracte avec lui), alors, dans ce cas il n'y a pas de raisson pour ne pas supposer que, dans le groupe H₂ε, l'e ne se changerait point aussi en ā.
Donc, nous pouvons supposer:
H₁ + e > e
H₂ + e > a
H₃ + e > 0.

The notion that *ε₁, *ε₂ and *ε₃ (Zgusta's H₁, H₂, and H₃, respectively) color an adjacent vowel has become an important aspect of most modern versions of laryngeal theory. Many of the details, however, have yet to be worked out (see, for example, Zgusta, 1951).
According to Polomé (1965), Hermann Möller was the first to interpret *$e_1$, *$e_2$, and *$e_3$ as laryngeal consonants. Möller later argued that IE *$e_2$ should be identified with Semitic aleph. To a certain extent, Möller's argument lent support to early laryngeal theory. However, Möller's identification is predicated on a genetic relationship between Indo-European and Semitic. Such a relationship is tenuous, at best. Concerning Möller's argument, Lehmann (1955:23) has commented:

The advantage of Möller's argument lay in his ability to cite earlier, though reconstructed, forms with laryngeals; the disadvantage, the forms he cited were so far removed from the various IE developments that the interrelationships seemed nebulous.

In retrospect, the premature marriage of Saussure's theory and various common origin hypotheses—such as Möller's Proto-Euro-Hamito-Semitic—was most unfortunate: significant advances to the theory were lost in the debate over the proposed genetic relationships and were largely ignored.

And there were significant advances. In a note delivered to the Société de Linguistique de Paris, Saussure (1892) suggested that the voiceless aspirates of Sanskrit reflected an original sequence of a voiceless stop and a coefficient sonantique. Thus, a form like Skt. rath- 'chariot' would derive from *rete$H$ ($H$ denotes a laryngeal), zero grade *re$H$ and *$tH > th$. Compare Latin rot$A$ where
$H > \ddot{a}[o_2]$. In this connection it is noteworthy that Skt. kh does not palatalize before IE *e, *i, or *y while all the other velars do (see Lehmann, 1955). As Hoenigswald (1965:93) notes: 'This is most smoothly explained by seeing in the aspiration an earlier consonant which intervened between k and the vowel or semivowel.'

Pedersen had also observed quite early—according to Polomé (1965), as early as 1893—that the ā/ṅ alternation of Skt. IXth class roots could be accounted for in large part under Saussure's assumptions about ablaut. Thus, krīnāmi 'I buy', but krīnīmas 1st. pl. can be seen as reflecting the alternation *neH/ṅH: ā represents the full grade *neH and ṅ the zero grade *ṅH. As Polomé notes (1965:12), however, Pedersen was not aware of Saussure's explanation of the Skt. voiceless aspirates. Pedersen's proposals were unable to account for the aspirate in forms like Skt. pānthāḥ 'path, way, nom. sg.', pathībhyaś dat., abl. pl., pathāḥ gen. sg. These forms are in fact problematic since the aspirate in pānthāḥ is unexpected. Apparently a paradigmatic leveling has occurred and the aspirate has been generalized as a result. Compare, Avestan pantā, and path where the expected forms actually occur.

Saussure had also argued (1879) that Sanskrit set verbs developed from original roots with a final coefficient sonantique. Skt. bhāvī/bhū 'become, be', for example,
represents earlier *bheuH/bhuH. Lehmann (1955) notes that after IXth class roots were analyzed in light of laryngeal theory, Skt. Vth, VIIth and IXth class verbs could be seen as parallel formations in which *n was infixed. Lehmann (1955:24) cites the following examples:

VII rinakti < /ly-n-eky-ty/
V šrnóti < /kl-n-eu-ty/
IX punáti < /pu-n-eH-ty/

He comments (1955:24): 'That a phonological analysis clarified a morphological relationship helped to support the laryngeal theory.'

Recall that Saussure's basic argument for laryngeals rested on those forms which exhibited long vowels in the full grade and which apparently fell outside of the eR/R pattern, and the problematic reflexes of the vowel *e in the zero grade. That additional phenomena, for example, those discussed above, could be accounted for on the basis of laryngeals, made Saussure's hypothesis so much the stronger. These phenomena count as independent evidence in support of the theory.

In spite of the strength of the hypothesis, Saussure's reconstruction was widely ignored apparently because of the absence of any direct evidence for laryngeals in any IE language. In 1927, however, Kuryłowicz argued that in many cases, Hittite (which had recently been deciphered),
exhibited the consonant $\mathfrak{h}$ where a laryngeal had been predicted. Kuryłowicz showed that the missing laryngeals were involved in a number of phonological and morphological problems of IE structure. He expanded the theory and presented a summary of his findings in his *Etudes Indo-européennes* (1935:28-30):

1. All original long vowels represent the contraction of a short vowel and either $\varepsilon_1$, $\varepsilon_2$, or $\varepsilon_3$. That is, $e + \varepsilon_1 > \varepsilon$, $\varepsilon + \varepsilon_2 > \hat{\varepsilon}$, $e + \varepsilon_3 > \hat{\varepsilon}$, and $a$ (from $e$ under ablaut) + $\varepsilon_1 > \hat{\varepsilon}$. These contractions regularly occur when $\varepsilon$ occurs $v_C$.

2. When $\varepsilon$ occurs between two vowels, it is lost and the vowel contracts. But $\hat{\varepsilon}$, $\hat{\varepsilon} + a > a + a$ can be understood as $a\dot{a} + a > a + a\dot{a}$.

3. Between two consonants, $\varepsilon$ disappears completely (except in Greek).

4. Between a consonant and a vowel, $\varepsilon$ disappears. In Indo-Iranian, however, the sequence of a $p$, $t$, or $k$ and $\varepsilon$ becomes $ph$, $th$, and $kh$.

5. Vowel initial IE words began with $\varepsilon$. Some roots with simple initial consonants arise from the sequence of $\varepsilon$ and a simple consonant.

6. $\varepsilon_2$ is reflected in Hittite as $\mathfrak{h}$.

Note that while the Hittite evidence is widely recognized as confirming evidence for Saussure's hypothesis, there are a few serious problems in interpreting the Hittite material. For example, Lehmann has observed (1955: pgs. 25ff.) that the number of forms in which Hittite $\mathfrak{h}$ or $\mathfrak{hh}$ confirms a laryngeal is small. He adds (1955:26):
'This evidence is small, but when one compares it with the total amount of lexical evidence for a Hittite-IE relationship, it is remarkable.' Hittite, however, does not always show an h where one would be predicted and sometimes an h appears where a reflex of a laryngeal is not expected (see Zgusta, 1951, pgs 450ff.). It is also not clear in certain cases which laryngeal should be reconstructed. These problems are beyond the scope of the present work (I refer the reader to Zgusta, 1951 or Lehmann 1955 for excellent discussions of the Hittite material). Suffice it to say here that the discovery of the reflexes of laryngeals in Hittite catapulted Saussure's theory into the fore of Indo-European research. As Zgusta (1951:444-5) has succinctly stated:

Si l'on se trouvait, jusque là, dans le domaine des pures hypothèses plus ou moins probables, en 1927, on a fait une découverte qui semble avoir élevé la théorie laryngale au-dessus du niveau d'une simple supposition.

M. J. Kuryłowicz a montré que dans certains cas, là où la théorie laryngale suppose la présence d'un laryngale originaire, le hittite présente un h. C'était une découverte d'une importance tout à fait extraordinaire, car elle apporte le témoignage extérieur exigé.

It should be evident from the preceding discussion that Saussure's arguments are not based in any straightforward way on the principles of internal reconstruction as I have described them in the preceding chapter. In the
next section, I will review Benveniste's theory of the IE root.

3.2. Benveniste's Theory of the Indo-European Root. The clarification of IE structure that the early work in laryngeal theory permitted culminated in Benveniste's theory of the IE root. In this section I will sketch Benveniste's theory, but pay particular attention to the form of Benveniste's arguments and the types of structural evidence he considered. This section, then, will be quite similar to the preceding section. My discussion in these two sections will serve as groundwork for my analysis of the kind of internal reconstruction exemplified in Saussure's and Benveniste's theories. That analysis will be presented in the next section of this chapter.

Certain limitations on the structure of the IE root had been noted quite early in the history of IE studies. For example, IE roots did not both begin and end with a voiced stop. Nor did roots contain both a voiceless stop and a voiced aspirate or two consecutive resonants. Despite attempts at codifying the structure of the root, for example, Hirt's work with bases, no definite system had been discerned. And like Saussure, Benveniste was concerned with the underlying system. In the opening paragraphs of his 'Esquisse d'une théorie de la racine' Benveniste suggested (1935:148):
Il faut essayer, par de larges comparaisons, de retrouver le système initial sous sa forme la plus simple, puis de voir quels principes en modifient l'économie. C'est ce mécanisme que nous chercherons à définir ici.

Benveniste noted (1935:147 and following) that IE roots of various structures had been reconstructed. He was, however, quick to point out a very peculiar fact of IE root structure. Alongside a reconstructed root like *perk (Lithuanian peršu 'to court') there would very often be another root of a slightly different shape but with a quite similar meaning: *prek (Latin precor 'request').

Kuryłowicz (1935:77-8) made a similar observation in his discussion of heavy and light bases.

In his analysis, Kuryłowicz included forms like *pelH and *pleH beside *perk and *prek. This inclusion of forms with laryngeals is important: without laryngeals there is no obvious parallel between *perk and *prek on the one hand, and *pelH (*pele) and *pleH (ple), on the other.

Benveniste also recognized his debt to Saussure and laryngeal theory arguing (1935:148-150) that it was only as a consequence of Saussure's clarification of the nature of the base vowel *e and ablaut, that one could recognize the significance of certain IE phenomena.

Benveniste argued (1935:149-50), for example:

Dès qu'on a reconnu la véritable nature de e et sa fonction on possède le moyen d'apprécier correctement la classe des
racines 'dissyllabiques' (set). On dénomme ainsi des racines terminées par â, è, ô, en alternance avec e: type *terH/-tre, *gueyH/-guyé- etc. Descriptivement, cette définition est exacte. La tort a été d'admettre qu'il s'agit d'une sorte particulière de racines et de fabriquer au profit de cette prétendue exception des catégories sans fondement.

The first step in Benveniste's argument, then, is the observation that frequently roots of different shapes are systematically paired. Consider the following examples (Benveniste, 1935:151): *werq:*wreg, *perk:*prek, *petr: *pter, *serw:*sreu, *terw:*treu. And with laryngeals: *terH:*treH, *genH:*gneH, *Herg:*Hreq.12 Following the traditional notation, these pairings can be represented by the formula TeRT:TReT, where 'T' is any consonant, 'R' any resonant. R may be a laryngeal and T may include R. (The restrictions noted earlier (see pg. 77) on the phonological structure of the root must be kept in mind here.)

The second step of Benveniste's argument is really nothing more than a careful consideration of the formal relationship that obtains between the members of each pairing. Benveniste argued (1935:150):

En dégageant, au cours d'un chapitre précédent une alternance *déiw/ðyéu, *dérw/ dréu, nous avons implicitement formulé le principe qui domine les alternances radicales. L'élément *-w/*uy est suffixal. Il s'attache à la racine de telle manière que, au degré plein tonique de la racine correspond le degré zéro du suffixe, et réciproquement. Ainsi une racine *der
suffixée par *-w/*-eu offre deux états: *dérw:*drey.

Benveniste, then, suggests that members of the pairs TeRT and TReT are to be understood in terms of ablaut. The TeRT forms represent a TER root in full grade and a -eT suffix in zero grade, -T; or TeR+T. The TReT forms, on the other hand, represent the zero grade of the TER root, TR and the full grade of the -eT suffix; TR+eT. The examples cited on the preceding page would then consist of the following roots and suffixes: wérg:wreg, root, wer, suffix eq; pérk:prék, root, per, suffix, ek; pétr:ptér, root, pet, suffix er; sérw:sréu, root ser, suffix, eu; térw:tréu, root, ter, suffix eu; térH:tréH, root, ter, suffix eH; genH:gneH, root, gen, suffix, eH; Herg:Hreg, root, Her, suffix, eg.

Benveniste summarized his theory of the root in the five statements below (1935:170-71):

1. La racine indo-européenne est monosyllabique, trilitère, composée de la voyelle fondamentale e entre deux consonnes différents;

2. Dans ce schéme constant: consonne + e + consonne, les consonnes peuvent être de n’importe quel ordre pourvu qu’elles soient différentes; seule est exclue la coexistence d’une sourde et d’une sonore aspirée.

3. La racine fournit, avec un suffixe deux thèmes alternants: I racine pleine et tonique + suffix zéro; II racine zéro + suffixe plein et tonique.
4. Au suffixe peut se joindre un seul élargissement soit ajouté après le suffixe du thème I, soit inséré entre l'élément radical et le suffixe du thème II (infixation).

5. L'addition supplémentaire d'un élargissement ou d'un suffixe à un thème déjà suffixé et élargi constitue une base exclusivement nominale.

(4) and (5) concern permitted types of formations and are supplementary to (1)-(3) in that they have to do with the types of formal elements that may be added to roots. For this reason, I will not discuss them here (see Lehmann, 1955, pgs. 14ff or Benveniste 1935, chapter four for a complete discussion). (1)-(3) essentially describe a synchronic relationship between TeRT (theme I) and TRêT (theme II). Benveniste does not suggest that something like *TeR-eT was the diachronic antecedent of TeRT and TRêT. In modern terms, Ter-eT would perhaps be interpreted as an underlying form, with TeRT and TRêT alternative surface manifestations. In so far as Benveniste's theory provides an explicit characterization of the etymological relationship of forms like *wérg and *wréq, *téRH and *tréH, and so on, through the entirely abstract system of roots and suffixes, the theory has a diachronic aspect.

As I noted in the last section, Saussure's system was algebraic and abstract. Benveniste's theory of the
root is even more so. As Lehmann notes (1955:16), however,

Roots ...are morphemes found, with phonetic variation, throughout the forms of one paradigm (a paradigm in word-inflection and word-formation). Obviously roots are theoretical units arrived at by linguistic analyses, and we have no evidence, like Hirt's for his bases, that at any stage of pre-IE, roots were words.

It seems to me however that a logical extension of Benveniste's theory would be to reconstruct **Ter-eT as the diachronic antecedent of TReT and TeRT, the latter resulting from the introduction of ablaut into Indo-European at some extremely early period. This is, of course, speculation. The serious analysis such a proposal would entail is beyond the scope of this work.

There are, then, two steps in Benveniste's argument. The first is the observation that there are systematic pairings of reconstructed forms of the shape TeRT with reconstructed forms of the shape TReT. The second step of the argument involves interpreting these forms in terms of ablaut. The conclusion then follows: roots are of the form TeT. To the root may be added a suffix of the shape -eT with either the suffix or the root in zero grade, but not both. In the next section I will take a much closer look at the arguments of Saussure and Benveniste.
3.3. Argumentation in Saussure (1879) and Benveniste (1935). In the last chapter I described in detail the method of internal reconstruction. My emphasis was to elaborate the principles and procedures of the method as they have been described in the recent literature of historical linguistics. In the first two sections of the present chapter I reviewed Saussure's theory of IE ablaut and Benveniste's theory of the IE root. In those sections I was primarily concerned with establishing the principal steps in their arguments. In this section I will review those steps and explicate the principles underlying each step. I will suggest that both Saussure and Benveniste were basing their hypotheses on the same kind of evidence. I will identify that type of evidence and suggest that its use has not been described in the theoretical literature on internal reconstruction. Moreover, I will suggest that the use of this type of evidence falls outside the domain of what some scholars, for example, Chafe, appear to consider to be internal reconstruction.

As I noted in 3.1., there are three steps in Saussure's argument. For ease of reference, these are repeated below (see 3.1 for relevant details):

Step 1: \( \hat{i}, u, \hat{u}, m, \hat{m} \) are all coefficients for sonantiques. There is a general ablaut pattern, \( eR \) (full grade): \( R \) (zero grade). There is a second general pattern, \( \hat{v} \) (full grade): \( e \) (reduced grade).
Step 2: There is evidence that *e represents more than one segment, historically. That is, *A[ŋ₁,ŋ₂] and *Q[ŋ₃].

Step 3. *A and *Q are coefficients sonantiques. The pattern V: e is the result of the loss of *A and *Q. V was originally *eA or *eQ.

Steps (1) and (3) are motivated by the same principle. I will discuss both of these steps together below. For the moment, I will concern myself with Step (2).

Step (2) rests on a fundamental principle of the comparative method: distinct correspondence sets that are not in complementary distribution represent distinct proto-segments. Recall that *e appeared as e and a in the Northern languages, but as a, e, and o in the languages of the South. The conditioning factors surrounding the apparent splits were unknown. Saussure argued repeatedly that if the different correspondence sets all represented the same original phoneme, then one would expect certain phenomena to obtain. For example, if Southerna, e, and o and Northern a and e represent the same original segment, then one would expect a fairly random set of correspondences between the reflexes in the Northern and Southern languages. Saussure showed that this expectation is not fulfilled. In short, Saussure argued that the behaviors exhibited by the reflexes of *e indicated more than one source. That is to say, Saussure showed that there were in fact distinct correspondence sets and that these correspondence sets
represented distinct original segments. Saussure designated these sets by \( *A \) and \( *Q \).

Steps (1) and (3) of Saussure's argument rest on a principle that I will refer to as the Principle of Similarity of Function. This principle can be roughly formulated as follows:

The Principle of Similarity of Function:
linguistic elements that function similarly under similar circumstances are to be analyzed similarly.

In step (1) this principle is used to group both the vowels, \( \text{i, u} \), and the liquids and nasals together as coefficients sonantiques on the basis of their behavior in ablaut. That is to say, Saussure argued that \( \text{i, u, r, m, and n} \) functioned identically in the ablaut system in spite of their phonetic differences. Thus, \( \text{ei/i, eu/u, er/r, em/m, and en/n} \) are all instances of the general formula \( \text{eR/R} \), where \( R \) is a coefficient sonantique.

The Principle of Similarity of Function is appealed to again in step (3). \( *A \) and \( *Q \) are to be considered coefficients sonantiques simply because they behave as coefficients sonantiques. In light of (3), the entire system of IE ablaut is seen to revolve around \( *e \) and the coefficients sonantiques. \( \text{v/e} \), which seemed to represent a different system, now falls within the scope of the \( \text{eR/R} \) pattern historically: \( \text{v} \) is the result of the coalescence of \( *e \) and a coefficient sonantique historically. I will
have more to say about the type of change the loss of laryngeals precipitated in IE ablaut in the next chapter. For the moment I will restrict my discussion to the type of structural evidence Saussure worked with in steps (1) and (3).

In step (1), Saussure established two general patterns of IE ablaut. I have used $eR/R$ and $V/e$ to refer to these. These notations are abbreviations for the systematic pairings of morphemes of different structures. $eR/R$ abbreviates all the morphemes that show a sequence of $e +$ coefficient sonantique in the full grade ($eR$) and just that coefficient sonantique in the zero grade, for example, $*bheidh/bhidh$, $*bheug/bhug$, $*terp/trp$, $*guem/gwm$ and so on. $V/e$ abbreviates all those morphemes that show a long vowel, $\ddot{a}$, $\dddot{e}$, or $\dddot{o}$, in the full grade and $e$ in the zero grade, for example, Doric Gk. histami 'stand, Ist sg.' statos past, perfect participle, tithemi 'place, Ist sg.', thetos p.p.p., and didomi 'give, Ist st.', but dotos p.p.p. The point I wish to make is that $eR/R$ and $V/e$ represent systematic pairings of morphemes of particular general shapes, pairings of morphological canons. It is these two sets of pairings that constitute IE ablaut.

In the next chapter I will explicate the basic assumption Saussure made about the significance of these two sets of pairings. Suffice it to say here that this concern for the systematic pairing of morphemes of different
morphological canons was shared by Benveniste. Recall that the first step of his argument involved establishing the systematic pairing of morphemes of the shape TeRT with morphemes of the shape TReT. The crucial structural evidence for both Saussure and Benveniste, then, can be found in these systematic pairings of morphemes of different canonical forms.

To be sure this kind of evidence is internal to the structure of a language. However, it is a type of internal evidence that has not been described previously in connection with the method of internal reconstruction. Recall that Chafe has argued that the comparison of cognate morphs is the 'only conceivable basis for internal reconstruction' (1959:478; see also 2.1., 2.2). It should be evident at a glance that neither Saussure nor Benveniste were concerned with a comparison of cognate forms, at least in the usual sense of the term 'comparison'. That is to say, while Saussure and Benveniste base their arguments on purely internal evidence—and hence their work can be considered internal reconstruction in a broad sense—they do not appeal to the usual principles or follow the traditional procedures. Notice, however, that at least in Saussure's case, the results obtained are analogous to the results obtained by the more traditional cases of internal reconstruction: just as the different
forms of a morpheme within a paradigm are related historically by regular changes, Saussure accounts for the different form types of morphophonemic alternations in IE ablaut by regular changes, namely the loss of laryngeals. In a very real sense, ER/R and V/ə are 'allopatterns'.

In the next chapter I will propose a methodology for working with the systematic pairings of morphemes of different canonical forms. In so far as no methodology has been described in this area, my explication will be filling an important gap in the literature of historical linguistics. Before turning to that issue it should be noted that internal reconstruction has been defined too narrowly by Chafe, Hoenigswald and others. Particularly, the role of cognate comparison and paradigmatic allo-morphy has been overstated. To be sure, the comparison of the different forms of a morpheme within a paradigm is a very fruitful procedure for making inferences about the history of a language. It is not the only procedure; nor is it the only basis for making accurate inferences about the history of a language. Witness the principles of Economy, Assumed Parallel Histories (discussed in the last chapter) and the Principle of Similarity of Function (discussed above), all of which allow the language historian to make an inference about the history of a language on the basis of internal evidence. The suggestion I wish to
make is that the factor that distinguishes the method of internal reconstruction from other diachronic methods is that it relies exclusively on various sorts of internal evidence. Clearly, paradigmatic allomorphy is but one kind of internal evidence and certainly not the only 'conceivable basis for internal reconstruction.'
Approximately one-third of the Mémoire is directly concerned with establishing Saussure's theory of IE ablaut. The remainder of the book deals with other matters, including certain consequences of the theory for certain reconstructions. Many of these matters will not be considered here.

I use the term 'wordbase' here instead of the more familiar term 'root' in order to avoid any possible confusion with the very special sense of root described in the next section.

The Slavic languages show o instead of a.

'a₂' is Brugmann's designation which Saussure adopts.

'o₂' is the so-called 'a-coloring' laryngeal, 'ə₁', the so-called 'e-coloring' laryngeal. See below for an explanation.

It is noteworthy that some scholars, for example, Kuryłowicz, later argued that phonetic speculations actually hindered the development of laryngeal theory.

H. Möller, 1880, 'A die tönende, E die tonlose kehlkopfspirans?, O das Kehlkopf-r?'

The expected reflex of *H in Sanskrit is short i. The length here is apparently due to analogy.

Saussure, of course, did not refer to his coefficients sonantiques as laryngeals.

For interpretations of the evidence for laryngeals from other languages see Winter (ed.) 1965.

Chapter four of Origines de la formation des noms en Indo-Européen (1935).

See p. 75 for the behavior of laryngeals in positions other than post-vocalically.

Lehmann posits a different shape, TRT, as the antecedent of these forms. See 1955, Chapter 15.
If a root is suffixed, the final T must be R. Benveniste's use of the term 'suffix' is not the normal usage. By suffix he means any ablauting formative added to a root. Élargissements do not participate in ablaut.
CHAPTER IV

INTERNAL RECONSTRUCTION AND
RELATIONAL PATTERNS

4.0. Introduction. In Chapter II I noted that while it was generally agreed that internal reconstruction is fundamentally concerned with paradigmatic allomorphy, there are aspects of internal reconstruction that are only marginally concerned with alternations in a paradigm. Recall, for example, my discussions of the principles of economy and assumed parallel histories. Moreover, in chapter III I described Saussure's theory of IE ablaut and Benveniste's theory of the IE root and suggested that their arguments rested on a type of evidence heretofore undescribed in a systematic way in the literature of internal reconstruction. That evidence was the regular pairing of morphological canons. It is the purpose of the present chapter to propose a methodology for working with this important type of structural evidence.

In 4.1 I will introduce the notion of regular relational pattern and argue that Saussure and Benveniste were concerned with the reconstruction of regular relational patterns. I will show that the basic type of evidence for
relational patterns is the systematic pairing of morphemes of different canonical shapes. In 4.2 I will discuss the methodological procedures for reconstructing regular relational patterns. I will be particularly concerned with noting the special difficulties involved with this type of reconstruction. In 4.3 I will discuss potential objections to the proposed methodology. Finally, in 4.4 I will present a summary statement of the methodological principles and procedures of internal reconstruction.

4.1 Regular Relational Patterns. It is the purpose of this section to specify what is meant by a regular relational pattern. I will consider how this notion is useful in making inferences about the history of a language and argue that while reconstructions that attempt to account for anomalous relational patterns by assuming derivation from a more regular relationship pattern can be favorably compared to reconstruction based on paradigmatic allomorphy, there are important differences. I will also consider the kinds of traces that changes in relational patterns leave on linguistic structure and propose a methodological principle for reconstructing such patterns.

By a relational pattern is meant any of the various formulae by which a language marks grammatical distinctions. For example, in Tagalog instrumental adjectives are formed by prefixing pang- to wordbases. Consider
kamay 'hand', pangkamay 'for the hand'; opisina 'office', pangopisina 'for the office'. In many cases, the final velar nasal of the prefix assimilates to the point of articulation of the initial consonant of the wordbase: simba 'church', pansimba 'for church'; dalawa 'two', pandalawa 'for two'. In still other cases, the initial consonant of the wordbase also disappears: sulat 'write' panulat 'for writing (a pen)'; kuha 'take, get' panguha 'for getting'; takip 'cover' panakip 'for covering'.

There is, then, a regular relational pattern between wordbases and those wordbases prefixed with pang-. The morphophonemic alternations exhibited by the different morphemes that are instances of the relational pattern do not obscure the nature of the pattern. These alternations are inherent in the phonological constituency of the particular morphemes.

In Sanskrit certain root classes form perfect stems by reduplicating a part of the root. In the most straightforward formations the reduplication is a copy of the initial consonant of the root and the root vowel in weak grade. Thus, pug 'nourish', normal grade pug (<paug) shows the reduplication pu-, for example, pupuśīva 'we (two) have nourished' and pac 'cook' shows the reduplication pa-, for example papaca 'he has cooked' (see below). There are other roots in which the reduplication
is not so straightforward. Witness, jīv 'live' perfect stem, jījīv, dā 'give', dadaū 'he has given', vac 'speak', uvāca 'he has spoken', bhid 'split', bibhid 'he has split'. Historically, the reduplication of all of these forms was a copy of the initial consonant of the root and the stem vowel *e. The shape of the reduplication, and in some cases the shape of the root, have undergone linguistic changes. For example, the reduplication of bhid is bi- and not bhi-. On the basis of comparative evidence *bhi-bhid can be reconstructed. The stop, bh, of the reduplication has been de-aspirated as a result of a general sound change, Grassmann's Law, by which the first in a series of two aspirates loses its aspiration. In spite of the application of these changes the shape of the reduplication is predictable on the basis of the shape of the root, though the synchronic rules are quite complicated (see Whitney, 1889 for a complete statement). The important point I wish to make is that the nature of the relational pattern is still transparent: associated with certain verbal roots is a reduplicated form of that root in the perfect.

Benveniste's theory of the IE root is a statement of a regular relational pattern of early IE: to a root of the shape TeT may be suffixed -eT with either the root or the suffix in zero grade. The 'surface manifestation'
of this relational pattern is the systematic pairing of IE roots of the shape \textit{TeRT} with IE roots of the shape \textit{TReT}.

Similarly, Saussure's theory of IE ablaut is the explication of an early relational pattern, namely forms with \textit{eR} (full grade) paired with forms with \textit{R} (zero grade). This relational pattern has been replaced in later IE by two regular relational patterns: (1) forms with \textit{V} (full grade) paired with \textit{*e} (zero grade); (2) forms with \textit{eR} (full grade) paired with forms showing \textit{R} (zero grade). (1) results from \textit{*eH>V} and \textit{H>*e}, and recall that \textit{H} used to be included in \textit{R}. The original relational pattern split (see below).

An inflectional paradigm can be understood as a bundle of relational patterns, for example, base + nom, base + gen, and so on. The paradigm can be represented by the formula 'base + desinence,' which is really nothing more than an abbreviation, a schema for the individual relational patterns that make up the paradigm. Allomorphy within a paradigm is incidental to the relational patterns that define that paradigm; the allomorphy affects instances of the relational patterns, not the general character of the pattern itself.

A relational pattern, then, consists of the systematic pairings of morphemes of different canonical shape.
For example, in the Tagalog instrumental pattern, one finds the following systematic pairings: \texttt{base::pang + base}. (Henceforth I will use '::' to mark systematic pairings.) If Tagalog had only CVCVC wordbases, the pairings could be easily expressed in CVCVC::pang + CVCVC. The structure of Tagalog wordbases is more complex than this, however. I use the abbreviation 'base' for all possible shapes of Tagalog wordbases. Thus, 'base::pang + base' is itself an abbreviation for possible pairings, for example, CVCVC::pang + CVCVC (kamay:: pangkamay), CVCCV::pang + CVCCV (simba::pansimba), CVCVCV ::pang + CVCVCV (dalawa::pandalawa), CVCV::pang + CVCV (kuha::panguha), and so on.

The statement of the pairings for Sanskrit reduplicated perfects is quite complicated synchronically since there have been so many changes that have affected the reduplication system. Originally, the pairings could have been represented by something like CeC::Ce-CeC. At a very abstract level of Sanskrit structure, one could perhaps argue that this is still the pattern, the different forms of the reduplications (and roots) being derived by phonological rules.

The systematic pairings Benveniste worked with are straightforward: \texttt{TeRT::TReT}. As I noted in 3.3 Saussure worked with two sets of systematic pairings, \texttt{eR::R and}
V : e, and hence two regular relational patterns.

In addition to the systematic pairings of morphemes of different canonical shapes, there is usually a functional difference associated with each canon. Thus, base in Tagalog functions differently than pang+base and the IE morphemes of the shape eR (full grade) occur in different morphological contexts than the related morphemes of the shape R (zero grade). However, when reconstructing a relational pattern it is not always possible to identify the functional difference associated with the pairings. It is not known, for example, how TeRT functioned distinctly (if it did at all) from TReT. I will address the issue of identifying the functional difference between paired canons in 4.3. For the moment I would like to consider how it is that patterns change.

There are two types of relational pattern change that I will be concerned with. The first type of change can be accurately described as pattern split. A pattern split occurs when a regular relational pattern at one stage in the history of a language is replaced and two regular relational patterns appear at a later stage of the language. Such a split occurred in the history of IE: the original pattern, eR::R (R includes laryngeals) is replaced by V::e and eR::R (R no longer includes laryngeals since they have been lost). The pattern split has been
induced by regular sound change affecting a subset of the forms participating in the original \( eR::R \) pattern, namely forms containing laryngeals. Thus, while \( eH::H \) may be considered to be a special case of \( eR::R \), the pattern \( \bar{V}::\bar{E} \), which results from the loss of laryngeals, is not readily identified as a special case of the \( eR::R \) pattern. An early regular relational pattern has been replaced by two regular relational patterns.

The second type of relational pattern change I will call pattern reanalysis. When a pattern is reanalyzed essentially either (a), the formal nature of the pairing between canons changes, or (b), the pattern is lost.

Pattern reanalysis of the type (a) is one form of what has been traditionally described as morphological reanalysis. For example, in early IE there was a regular relational pattern for the formation of the locative of verbal nouns: to the verbal root plus a derivative suffix the locative desinence \(*-\bar{I} \) was added. The pairing at this stage could be characterized as base+es::base+es+\( \bar{I} \). This relational pattern is, of course, only one of the many relational patterns that constitute the paradigm for verbal noun inflection in IE. In Latin, the sequence \(*+\text{es}+\bar{I} \) appears as \( e+re \) (\(<*e+si \)) (see Jeffers, 1977, and the references cited there). Thus, \(*\text{douk}-\text{es}+\bar{I} + *\text{douk}-e+si \rightarrow \text{ducere} \). The \( s \) of the derivative suffix
has been reinterpreted as a part of the desinence. Jeffers
(1977:19-20) argues that the reinterpretation has its
basis in related thematic verbal formations. He notes
(1977:20), for example, *douk+e+s > dūcis 2nd sg. pres.,
and *douk+e+t > dūcit 3rd. sg. pres. The new suffix *-si
(>re) has been generalized in Latin as the infinitive
marker. The original pattern, base+es::base+es+i has been
replaced by base+e::base+e+si. There is still a relational
pattern. However, the formal relationship that obtains
between the paired canons has changed.

In the next chapter I will discuss extensively an
example of pattern reanalysis of the type (b), that is,
pattern loss. Suffice it to say here that pattern loss is
a special case of pattern reanalysis: the formal nature of
the relationship between paired canons changes in such a
way that the canons are no longer seen to be systemati-
cally paired. As a result of this loss, members of the
pairing may assume the status of separate, and quite un-
related words. When this happens, it may be difficult to
show that there was in fact a systematic pairing of mor-
phemes of different canonical forms in the history of a
language. That is, it might be quite difficult to show
that there used to be a relational pattern. (See 4.2.
and chapter 5.)
It is important to note that changes to relational patterns are not the same kind of process as sound change. As the examples of pattern split and reanalysis discussed above demonstrate, pattern change reflects reinterpretation and replacement, albeit reinterpretations and replacements that are often precipitated by sound change.

In contrast, sound change is quite different. Although there is a great deal of controversy surrounding the issue, when a sound changes the reflexes of that sound at a later stage may be considered to be continuations of some of the original features of that sound through time. To be sure, the continuation is an abstraction from the observed data, but it is nonetheless assumed to be real.

When a relational pattern changes there is no such continuation of the original features of the pattern. The forms that participate in a relational pattern may, of course, be continued through time. *ple, then, may be considered to be a continuation of its antecedent *pleH. However, the relational pattern that plė participates in, namely Vːːə, is not a continuation of the pattern that *pleH participated in, eRːːR. The loss of laryngeals induced a reinterpretation of the formal relationship that obtained between the normal and reduced grades of forms containing laryngeals. The new pattern Vːːə
replaces \textit{eR::i} (R includes H). It is not a continuation of that earlier pattern in the same way that \textit{ple} may be considered to be a continuation of \textit{pleH}. I will return to this issue briefly below.

As a consequence of the different natures of sound change and relational pattern change, it is not always possible to employ the same methodological procedures in reconstruction. I will consider the particular problems of reconstructing relational patterns and pattern change in the next section. Suffice it to say here that reconstructing a relational pattern as the antecedent of certain structures (perhaps another relational pattern) is quite analogous to the more familiar type of reconstruction discussed in the opening sections of chapter II. In order to illustrate this analogy I turn again to Saussure's theory of IE ablaut.

Saussure noted that IE exhibited two ablaut patterns, \textit{eR::i} and \textit{\ddot{u}:\partial}. It should be noted that both of these are regular relational patterns. \textit{eR} and \textit{\ddot{u}} occur in the same morphological contexts; likewise \textit{R} and \textit{\partial}. The interesting question is: why does IE have two ablaut patterns? To a certain extent this represents a structural anomaly since there is no obvious reason for there being two functionally analogous patterns.
Saussure answered the question above by suggesting that the one pattern, $\bar{V}::\bar{e}$ could be seen as a special case of the pattern $eR::R$, historically, the difference in structure resulting from change. As Lehmann has noted (1955:23):

Saussure's basic assumption was that a similarity...of phonemic variation in morphemes of seemingly different structures pointed to anterior forms of a similar structure. Saussure's assumption has been upheld by the discovery of confirming evidence in Hittite.

Saussure correctly inferred that a pattern split had occurred in the history of IE (see 3.1).

Notice that the occurrence of two patterns, $eR::R$ and $\bar{V}::\bar{e}$, is parallel to the existence of two forms of a morpheme within a language, for example, Greek ambrot- and ambros- (see 2.1 for a discussion of these forms). IE $eR::R$ and $\bar{V}::\bar{e}$ are 'allopatterns' if you will. Just as one accounts for paradigmatic allomorphy by reconstructing an earlier form, *ambrot-, and derives the allomorphs from that form by regular sound change, one can reconstruct an earlier relational pattern and derive the 'allopatterns' by linguistic (sound and pattern) change. Both methods rely entirely on evidence internal to the structure of the language.

Reconstructing a relational pattern involves making a hypothesis to account for certain observed structural anomalies, for example, the existence of two functionally
analogous patterns of ablaut in IE (see also, 4.2). Typically the kinds of evidence used to support the hypothesis will be quite diverse in nature. Recall, for example, the complexity and wide range of phenomena considered by Saussure (3.1). There is, however, one kind of evidence that is fundamental to reconstructing a relational pattern: the systematic pairing of morphemes of different canonical forms. Such pairings are always associated with relational patterns. As a consequence, they may be taken as indicative of the existence of a relational pattern. The pairings may be used to support either a synchronic interpretation, as is the case with Benveniste's theory of the IE root, or a diachronic interpretation, such as Saussure's theory of IE ablaut. When a language historian interprets a systematic pairing of morphemes of different canonical form diachronically, he is appealing to what I shall call the Principle of Systematic Pairings.

The next step in the reconstruction is to establish the formal relationship, that is, the morphophonological relationship, that obtains between the members of pairings that constituted the original pattern. This is not always easy. In Saussure's case, once the multiple sources of *ə were suspected (3.1) the pattern əː could easily be interpreted in light of these lost sources, namely the
laryngeals: eH::H. In other cases, particularly where there has been analogical change, it may be quite difficult to determine the forms of the original pattern. A reconsideration of the Sanskrit reduplication example discussed above will be informative here.

The roots in Sanskrit which show a reduplicated form of the root in the perfect actually have two reduplicated stems: (a) the reduplication, which I will simplistically designate by CV-, (see pages 93-94 above) plus the normal grade of the root; (b) the reduplication, CV-, plus the zero grade of the root. For example, the root puṣ exhibits (a) pu-poṣ-, pu-poṣa (<pu-pauṣa) 'he has nourished' and (b) pu-poṣ-, pu-poṣa 'you (pl.) have nourished'. The (a) stems occur in the singualrs of the active, the (b) stems, elsewhere. For the rest of this discussion I will primarily concerned with the (b) perfect stems. This pattern can be designated by the systematic pairing, C₁V₁C₂:

C₁V₁-C₁V₁C₂, for example, puṣ:pu-puṣ-. For Sanskrit roots of the form C₁a C₂ a different pattern occurs. For example, the (b) perfect stem of sad 'sit' is sed-, sede 'he has sat (middle)' and the (b) perfect stem of pac 'cook' is pec-, pece 'he has cooked (middle)'. However, the zero grade of C₁a C₂ roots is C₁C₂, thus sd and pc. One would therefore expect a pattern like C₁a C₂:

C₁a-C₁C₂, for example sa-sd or pa-pc. The pattern that
actually occurs can be represented $C_1 a C_2 :: C_1 e C_2$. In short, Sanskrit exhibits two functionally analogous regular relational patterns for the formations of the (b) perfect stems.

While this example is similar to the IE ablaut situation with which Saussure was faced, there is one very important difference: the morphophonological relationship between $C_1 V_1 C_2 :: C_1 V_1 - C_1 V_1 C_2$ and $C_1 a C_2 :: C_1 e C_2$ is extremely opaque. A direct comparison of the two sets of pairings is not revealing in the same way that a direct comparison of $e R :: R$ and $\bar{V} :: e w a s$. In order to establish the original nature of the pattern of (b) perfect formation in the history of Sanskrit, it would be necessary to find the form or forms in which the morphophonological relationship is clear. It turns out that this relationship is clear only in the case of one form, the root sad.

Recall that in 2.3 I established that IE *se-sd became sed in Sanskrit as a result of the following series of regular sound changes: *se-sd sezd>sazd>sajd>sed. Thus, the occurrence of sed as the (b) perfect stem of sad is etymologically expected. This is not true, however, in the case of the vowel e of a form like pec, for example. The expected form etymologically is pape, which of course does not occur. sed has apparently become dissociated from the traditional reduplication pattern and
formed the basis for an analogical extension of to roots like pac, pec, then, replaces the expected papc. Since the sequence pc is permitted in Sanskrit one cannot account for the disappearance of papc on phonological grounds.

Earlier in this section I suggested that pattern changes were replacements and reinterpretations, not continuations. The participation of pac in the \( C_1 \ a \ C_2 : \)

\[ \text{sad::sed} \] is, morphophonologically, a special case of \( C_1V_1C_2::C_1V_1-C_1V_1C_2 \). Clearly, \( \text{pac::pec} \) is not; \( \text{pac::pec} \) represents the wholesale replacement of \( C_1V_1C_2::C_1V_1-C_1V_1C_2 \) by \( C_1a \ C_2::C_1e \ C_2 \). It should be evident that there is no continuation of the original pattern in any meaningful sense. Notice that if the form \text{sed} had been lost in the history of Sanskrit, the best one could do would be to suspect, without direct support, that the two (b) perfect stem patterns were descendants of a common pattern. This is, of course, an extremely weak claim. I will have more to say about the problems of recovering the nature of an early relational pattern in the next section.

The reconstruction of a relational pattern, then, involves making a hypothesis in order to account for certain structural anomalies. The crucial evidence in support of a relational pattern is the systematic pairing of morphemes of different canonical forms. The hypothesis consists of
a statement describing the nature of the original pattern, that is, the nature of the original pairing, and a description of the changes that have brought about the pattern change.

Before turning to a discussion of some of the problems involved in reconstructing relational patterns I would like to comment briefly on the reality of relational patterns. Relational patterns are necessarily abstract. However, to the extent that speakers of natural languages can recognize that forms are systematically related, relational patterns are real. When a speaker sees that the different forms of a morpheme within an inflectional paradigm are systematically related forms of the same word, for example, that speaker testifies to the reality of relational patterns.

4.2. Some Methodological Problems. In the last section I argued that sound change and relational pattern change were quite different phenomena and suggested that this difference could have important methodological consequences. In this section I will explore the special problems associated with reconstructing relational patterns.

I have suggested that when a sound changes its reflexes can be considered to be continuations of some of the original features of the sound over time. Consider for example, English foot. This word can be traced
back to IE *ped/pod 'foot'. The shape of the reconstructed word has been established on the basis of the comparison of the cognates of foot in different IE languages. I will confine my remarks to the initial segment of this word.

The formula *pod>foot, particularly *p>f is merely a before-and-after-description of what is assumed to be the same word at different stages in the history of the language. While the formula does not overtly detail the intermediate steps between *p and f those steps are assumed to have occurred. *p>f, then, can be interpreted as reflecting a gradual transition, effected by successive generations of speakers, from *p to f.

To a certain extent the original nature of a sound—the sound's original physical properties—can be inferred by examining the nature of the reflexes of that sound in different languages. Thus, besides English foot, consider Old Norse fōtr, Old High German fuo3, Lat. pedis gen. sg. Avestan pad, Tocharian (A) pe, Attic Gk. podos gen.sg. and so on. In phonetic reconstruction, the decision as to a particular shape is based upon the phonetic properties of the reflexes and a careful consideration of what are known to be plausible sound changes (among other things). Notice that this phonetic reconstruction is possible only in so far as the reflexes of a sound are considered to be
continuations over time of some of the features of that sound.

In the last section I argued that relational pattern change was not like this. I suggested that pattern changes were reinterpretations and replacements and that consequently they could not be considered to be continuations of earlier patterns. Thus, when \( eH > \tilde{V} \), and \( H \geq \), the formal relationship between the normal and reduced grades of forms containing these segments appeared to be outside of the \( eR::R \) pattern. As a result of sound changes affecting the forms of morphemes participating in the pattern, the pattern had to be reinterpreted. The forms that participate in \( eH::H \) (that is, \( eR::R \)) and later \( \tilde{V}::\varnothing \) are continuations through time. The pattern they participate in, is not. It is only as a result of the fact that the morphemes that participate in a pattern are continuations of earlier forms of those morphemes that the original nature of a relational pattern can be recovered. If the sound changes that have affected morphemes that participate in a pattern are known, or can be recovered, the original shape of the pattern can be recovered.

By way of illustrating this point, consider again the Sanskrit reduplicating perfects discussed in the preceding section of this chapter. Recall that there are two patterns for the formation of reduplicated perfect stems
with the zero grade of the root: $C_1V_1C_1::C_1V_1-C_1V_1C_2$ and $C_1a\ C_2::C_1e_2C$. \underline{puš::pu-puš-} is an example of the former pattern and \underline{sad::sed}, an example of the latter. I suggested that the $C_1a\ C_2::C_1e\ C_2$ pattern derives historically from the other pattern by sound changes affecting *se-ad and the analogical extension of the e to forms like pec. If the history of sed were not known, or if the form were absent from the language, the relationship between the two patterns could not be readily recovered. As a consequence, the original shape of the pattern could not be recovered. Knowing the history of sed, there is a connection between the two patterns: \underline{sad::sed} is a special case of $C_1V_1C_2::C_1V_1-C_1V_1C_2$. The peculiarities of sed as a reduplicated perfect stem can be considered to be idiosyncratic to the segments that make up the word. The generalization of e to the forms in which this vowel is not etymologically expected--pec should be papc--establishes $C_1a\ C_2::C_1e\ C_2$ as a distinct pattern in its own right.

This example from the history of Sanskrit and the reconstruction of IE ablaut proposed by Saussure both involve what I have described as pattern split. In general, pattern splits are more easily recovered than pattern reanalyses since splits leave a very conspicuous structural trace: two functionally analogous relational
patterns. Given this kind of clue, it is only necessary to determine the relationship between the two 'allopatterns' and consequently establish the shape of the original relational pattern. When a pattern reanalysis occurs, there may not be any structural evidence indicating that the synchronic pattern had ever been different. For example, in the last section I discussed the origin of the Latin infinitive marker -re and showed how it resulted from a reanalysis of an earlier root plus derivative suffix plus locative desinence pattern: *douk+es+i → *douke+si > dūcere. I am unaware of any evidence internal to the structure of Latin that would allow one to recover this set of changes without resorting to the testimony of other languages.

Pattern losses may be somewhat easier to recover. Recall that when a pattern is lost morphemes that were systematically related are no longer seen to be related; they may become separate words. If enough of these formerly related words are retained, the pattern may be recovered. In such cases, however, it is often difficult to demonstrate that the words used to be systematically related. The kinds of clues the linguist can use to establish such a systematic relationship are quite varied. In the next chapter I will discuss an extensive example from the history of Proto-Austronesian illustrating the
kinds of evidence that can be used. I defer discussion of this point until then.

The first step in reconstructing a relational pattern then is to establish that there is a systematic pairing of morphemes of different canonical forms. Except where a pattern has been lost, this will be quite straightforward. When a pattern split has occurred, there will be two sets of systematic pairings.

The second step in the reconstruction is to determine the shape of the original pattern. That is, one must determine what canons were paired originally. For pattern loss or pattern reanalysis this may be quite difficult (see, for example, chapter five). When a pattern has split, however, it may be possible to determine the shape of the original pattern by comparing the 'allopatterns'. This was the procedure followed by Saussure.

If a direct comparison of the resultant patterns is not fruitful—as in the Sanskrit reduplicated perfect stem example—it is necessary to consider other kinds of structural evidence in order to determine the morphophonological relationship that obtains between the two patterns. If the history of sed in Sanskrit were not known, for example, it would be necessary to compare every form in the $C_1a C_2:C_1e C_2$ pattern to the expected forms in light of the known sound changes. Moreover, as I have
already argued, if the form sed had been lost, the fact
that only one pattern existed at an earlier stage could
not be recovered.

To be sure, the Sanskrit example has been compi-
cated by an analogical extension. However, it is not
inconceivable that a sound change, or even more likely a
series of sound changes, could result in a pattern split
in which the relationship between the two resulting pat-
terns was quite opaque. Consider the following hypothet-
ical example.

Suppose there is a language that was only CVČ and
CVČVC wordbases. Verbal roots may be formed from these
wordbases by suffixing -VC, that is a copy of the final
vowel and consonant, to the wordbase. Consider the forms
below.

túpːːtúp-up
kákːːkák-ak
sítːːsít-it
xátːːxát-at
kárutːːkárut-ut
cákːːcák-ak
pátinːːpátin-in
sálobːːsálob-ob

Now suppose that a general sound change occurs and
the vowel of the syllable immediately following the
stressed syllable is syncopated. The forms above would
appear as follows as a result of this change.

túpːːtúp-p
kákːːkák-k
sítːːsít-t
xátːːxát-t
kártːːkárt-ut
cákːːcák-ak
pátinːːpátin-in
sálobːːsálob-ob
The formal relationship between forms like *tup* paired with *tup-p* and forms like *kart* paired with *kart-ut* is not nearly as transparent as the relationship between IE *eRːR* and *Vːe*, for example. While both *tup* and *kart* show a verb stem with a copy of the final consonant, the *kart* forms also show a vowel of unpredictable quality in the suffix.

Synchronically one could attempt to account for this vowel in a number of ways. Perhaps the most obvious solution would be to establish underlying forms like /karut/ and /cakar/ for the *kart* forms. There would have to be a morphological rule of VC# copying, perhaps restricted to the *kart* forms, that is underlying di-syllables. There would also be a synchronic syncope rule applying after the VC# copying rule. It is problematic whether or not one would want to extend this analysis to include the *tup* forms as well.⁴

Alternatively, one could attempt to predict the vowel on the basis of the root. Since it is not phonologically predictable, it would be necessary to find some morphological answer. It should be obvious that the synchronic rules under such a solution would be enormously complicated and probably quite unnatural. As a third alternative—and one that is probably just as complicated—one
might attempt to derive the tup-p forms from underlying tup-Vp.

While the synchronic situation is complicated under any of these solutions, the sound change that introduced this complexity is quite straightforward, a simple syncope. It is likely that such a change would produce a pattern split since the relationship between the two sets of pairings is so opaque. Thus, single sound changes can produce pattern splits in which the relationship between the two 'allopatterns' is difficult to determine. Nonetheless, I think a clever linguist could recover the original pattern.

The final step in the reconstruction of a regular relational pattern is to explicate the nature of the linguistic changes that the pattern has suffered. Thus, it is necessary to show explicitly how it is that the original pattern changed.

In the next chapter I will provide an extended example of the application of this method to a problem from the history of Proto-Austronesian. It is hoped that my detailed discussion in that chapter will clarify any remaining procedural difficulties. The principal steps of the method of reconstructing regular relational patterns are repeated below.
1.) establish the existence of systematic pairing of morphemes of different canonical structure.

2.) infer the original shape of the pattern on the basis of (a) a direct comparison of the pairings (in the case of pattern split); or (b) the careful consideration of the available structural evidence.

3.) provide a detailed account of the linguistic changes that have affected the original pattern and resulted in the pattern change.

In the next section I will review some of the potential objections one might have to working with systematic pairings of morphemes of different shape in reconstruction.

4.3. Some Potential Objections. In this section I will consider the three objections that seem to me to be the most likely to be raised against the method outlined in the first two sections of this chapter. These objections are sketched below.

(1) The method does not rely on the comparison of cognate morphs.

(2) The results of the method are highly formulaic in nature.

(3) It is often impossible to determine the function of the reconstructed relational pattern.

I will deal with these in turn and argue that each objection is ill-founded.

In his discussion of internal reconstruction, Hoenigswald (1944:78) made the following observation:
On the whole, speculation without the benefit of comparative evidence has a reputation for being unsafe, a criticism which is, however, directed more often against certain far-reaching results, than against the procedure itself.

It is important to re-emphasize here the importance Hoeningswald and Chafe ascribe to the comparison of cognate morphs in internal reconstruction. Consider, for example, Chafe's claim (1959:478):

Most of historical linguistics is fundamentally dependent upon the technique of comparing cognate forms. These forms may be drawn from different languages, as in the comparative method proper, or from different stages of the same language, as in the use of written records; but with the possible exception of some aspects of dialect geography, there is no generally recognized historical technique that does not involve the comparison of cognates. Such comparison seems to me to be the only conceivable basis for internal reconstruction.

It should be obvious from my discussions of various methodological principles that such comparison is not the only conceivable basis for internal reconstruction. (See especially 2.3 2.4 and 4.1, 4.2.) Indeed, Automaticity is the only principle that has been described in this work that is concerned directly with paradigmatic allomorphy, and consequently, cognate morphs.

For example, I discussed at some length the Principle of Assumed Parallel Histories. This principle is based, as all principles must be, on the experience of linguists
with the histories of natural languages. And, there is no element of the comparison of cognates associated with the use of the principle. It would be most unreasonable to reject the principle simply because it does not depend on the comparison of cognate morphs. To do so would be to blindly ignore a good deal of what has been learned about the ways in which languages develop.

One cannot object to a methodological principle simply because it is independent of cognate comparison. In the next section I will have more to say about the principles formulated above and argue that most are instances of more general principles of inference, not even necessarily specific to the discipline of linguistics. It is sufficient to note here that the reconstruction of regular relational patterns is firmly rooted in empirical observations of the ways languages change. Indeed, the crowning glory to this kind of reconstruction can be seen in the discovery of the reflexes of laryngeals in Hittite. Recall that Saussure (see 3.1) did not have the Hittite evidence when he reconstructed *A and *O on the basis of his careful analysis of the structure of IE. That the method of reconstructing regular relational patterns is not based on the comparison of cognate morphs, then, is not a legitimate objection to the proposed method.
Recall that Hoenigswald has noted (see citation above) that the criticism against reconstruction without the benefit of the testimony of other languages—that it is unsafe—has to do more often with the far-reaching results than with the actual method. It is not entirely clear what Hoenigswald means by 'far-reaching,' though he does refer explicitly (1944:78) to the works of Kuryłowicz and Benveniste:

Even the most conservative scholars, for instance, have held that there is some relationship between the IE accent and ablaut, and most of them would not hesitate to say that the zero grade must have developed, at some time in IE history, from full grade under weak stress. The difference between this example and the more sweeping reconstructions of 'IE prehistorical grammar' by Hirt, Benveniste, Kuryłowicz, and others, is merely one of degree.

In 3.1 I noted that some scholars objected to Saussure's findings because of their formulaic character and Saussure's disregard for the phonetic identity of *A and *O. In light of this and Hoenigswald's reference to Benveniste and Kuryłowicz immediately above, I will interpret 'far-reaching' as formulaic.

Clearly there is nothing inherently objectionable to formulaic reconstruction. Notice that the comparative method is traditionally formulaic: a proto-segment is technically an abbreviation of the particular correspondence set (or sets) that support the reconstruction of a
proto-segment distinct from all other proto-segments. For example, *p in a reconstructed form like *ped/pod 'foot' is really nothing more than an abbreviation of the correspondence set, English f (foot), Old Norse f (frôtr), Old High German f (fuô3), Latin p (pedis), Avestan p (pad), Tocharian (B), p (pe) and Attic Greek p (podos), for example (see 3.1 above). It seems to me, then, that the objection cannot be directed entirely against the algebraic character of the reconstructions. There must be something more. I submit that this kind of objection is directed against the systematicity or regularity—not so much the algebraic character—imposed on the structure of a language. Importantly, the regularity imposed is not transparent. Thus, it seems to me that the objection is that reconstructed relational patterns impose a system or regularity that is not really there; languages are artificially being made to look more regular than they really are. This view is mistaken, however.

I pointed out in 2.2 that internal reconstruction, to an even greater extent than the comparative method, provides a description of only a part of a language at some earlier point in its history. Internal reconstruction, whether concerned with alternations in a paradigm or canonical forms, does not permit inferences about a complete, coherent system. A regular relational pattern
is, by definition (4.1) a formula. However, this formula does not, by any means, represent the whole language. In reconstructing a relational pattern, one asserts that a sub-part of a language was systematic and regular. One is not asserting that every aspect of that language at that point in time historically, was entirely systematic and regular.

Such assertions are entirely consistent with what is known about natural languages. Every known language exhibits regular relational patterns. Indeed, much of synchronic linguistic analysis is concerned with discovering those patterns. And experience has shown that much of what is irregular in a language synchronically is often the result of change to a historically more regular system. It seems to me then that the objection to the formulaic character of reconstructed relational patterns, manifests a confusion between reconstructing a subsystem of a language which is quite regular and reconstructing a complete linguistic system (language) that is entirely regular. While the latter is fundamentally inconsistent with the experience of linguists with the history and structures of natural languages, the former is wholly consistent with that experience. I conclude that this objection is unfounded.
The final objection I will address concerns the frequent inability of the language historian to characterize the function of a regular relational pattern that he has reconstructed. For example, Benveniste established a TeT root which, when suffixed, forms two themes, TeRT and TReT. The functional difference between these two themes is unknown. Presumably there was some sort of functional distinction. The problem here is not one that is specific to reconstructing relational patterns. Rather, it is a more general problem, that of reconstructing morphological function. In order to illustrate this point, I will consider a hypothetical example where the same problem appears in connection with comparative reconstruction.

Suppose one is comparing a group of related languages, all of which exhibit various case endings. In certain instances, the endings correspond regularly both in form and function. In other cases, there is only a correspondence of function, and in still others, there is only a correspondence of form. Thus, each language might show -log (or something phonetically similar) for the genitive and one would, of course, reconstruct *-log. But suppose for the accusative some languages show -xat while the others show -mat and one language has -ko marking the dative where others have this desinence marking the locative.
Both of these latter cases are problematic. For the first one would probably be forced to reconstruct both *xat and *mat for the accusative. For the second, the one directly relevant for the point at hand, one could only reconstruct the form of the desinence, *-ko, but not its function. Unless there is an exact identity between the functions of cognate morphemes, it is extremely difficult to determine the original function to be associated with the original morpheme. The language historian is forced to make an intelligent guess based on the characteristic functions of each of the cognate morphs in the daughter languages.

The hypothetical example described above is not unusual. In IE, for example, there are some languages which indicate *bhi/bho for the instrumental, for example, Sanskrit, others indicate *mi/mo, for example Slavic languages. Moreover, the desinence is used to make dative, locative, instrumental and ablative. There is no tried and true method for determining the original function marked by this desinence.

On the basis of the preceding remarks, this final objection can be placed in the correct perspective; the problem is one of reconstructing morphological function in general. This objection is not specific, therefore, to reconstructing relational patterns. It would be most
unreasonable to fault the method proposed above because of this short-coming of linguistic analysis without also faulting other methods, for example, the comparative method. Unfortunately, the solution to the problem of reconstructing morphological function has not been solved. But, notice that the inability to assign the correct function to a reconstructed relational pattern does not preclude the existence of that pattern. Frequently, the language historian is dealing with enormous time depths. It is not unreasonable, then, that problems of this sort are encountered.

None of the potential objections to the method of reconstructing relational patterns discussed in this section is well-founded. I have argued that it is unreasonable to require that all principles of internal reconstruction be based on the comparison of cognate morphs--there are quite legitimate principles that have nothing to do with such comparison. Moreover, I have argued that the systematicity associated with relational patterns is entirely consistent with expectations about reconstructed and natural languages. Finally, I have suggested that there is a general problem in reconstructing morphological function, a problem that may be encountered when working with relational patterns. This short-coming is not, however, specific to such patterns and cannot,
therefore, be legitimately posed as an objection to the method proposed here. In the next section, I will take a closer look at the principles for internal reconstruction proposed in this work.

4.4. Principles for Internal Reconstruction. In the course of this dissertation I have referred to and proposed formulations of some six methodological principles that can be appealed to when making an inference about the history of a language on the basis of internal structural evidence. In my view, these principles jointly constitute the method of internal reconstruction. It should be evident, however, that there really is no unitary method of internal reconstruction; hence the title of this work, methods for internal reconstruction. In this section I will review these six principles. The point I wish to make is that all of these principles have either correlates in synchronic analysis or they are instances of more general principles of argumentation in the philosophy of science and are therefore not specific to diachronic analysis. I think this is an important point to make in light of the neglect for the principles of diachronic analysis evident in the literature of theoretical linguistics. The methodological principles discussed in this dissertation are listed below for ease of reference. I will discuss each in turn.
1. Automaticity
2. Economy
3. Assumed Parallel Histories
4. Optimal Use
5. Similarity of Function
6. Systematic Pairings

The Principle of Automaticity asserts that alternations in a morphological paradigm are not original but have been introduced by linguistic change. In order to recover the original form of the morpheme and the changes that have introduced the synchronic allomorphs one follows the procedure of comparing the cognate morphs. Automaticity is the only principle discussed in this dissertation that has to do with cognate morphs directly. This is, of course, contrary to what has been generally asserted about internal reconstruction (see 2.1, 2.2 and 4.3).

Both the Principle of Automaticity and the procedure for working with alternations in a paradigm have correlates in synchronic analysis. In modern methods of phonological analysis, for example, different forms of the same morpheme at the surface are derived from a common underlying form by the application of regular phonological rules. This is quite analogous to deriving the different forms of a morpheme from a common form historically. To my knowledge the principle evident in this type of synchronic
analysis has not been given an explicit formulation, though it is evident in the theoretical discussions of many scholars, for example, Sapir, Bloomfield, and Chomsky. And of course, the procedures of cognate comparison have a direct synchronic correlate in structuralist principles of phonemic analysis. Since Hoenigswald has already discussed this parallel, I will not have anything more to say about this subject.

The Principle of Economy essentially guarantees the simplicity of the reconstruction: the preferred hypothesis is the one that permits the most straightforward and natural account of the changes involved in terms of phonetic motivation. It should be apparent that this principle is quite similar to the Principle of Optimal Use, which is also concerned with simplicity. To date, there has been no generally accepted systematic account of simplicity as it relates to linguistic analysis, though there have been attempts (see Halle 1964 for one example). And witness the concern for this issue in the theoretical literature of synchronic analysis, particularly generative techniques.

That hypotheses and descriptions should be simple is a basic tenet of the philosophy of science. By their very nature, descriptions and hypotheses are simplifications. Simplicity, however, manifests itself in many different
ways, for example, formal simplicity as opposed to say, conceptual simplicity. It seems to me that the discussion of linguistic simplicity in the literature is directly concerned with formal simplicity. A thorough discussion of this issue is beyond the scope of the present work. It is sufficient to point out here that both the Principle of Economy and the Principle of Optimal Use are also concerned with formal simplicity.

With respect to the former principle it is widely assumed that the simplest hypothesis and the simplest description will also be the most natural and straightforward (see, for example, Chomsky and Halle, 1968 particularly their discussion in chapter nine). The assumption is warranted for both synchronic and diachronic hypotheses. Unfortunately, the exact connection between naturalness and simplicity has not been elaborated successfully.

It is noteworthy that much of the controversy surrounding the notion of linguistic simplicity has to do with the best way to measure formal simplicity, that is, the best form of the so-called evaluation metric. That linguistic hypotheses and descriptions should be simple is not questioned.

The Principle of Assumed Parallel Histories has a synchronic correlate in modern phonological analysis.
When two or more sounds in a language exhibit partially similar behaviors, it is quite customary to account for that similarity by positing parallel derivations. For example, suppose a language weakens velars intervocally; /g/ appears as ɣ and /k/ as ʰ. A typical solution to this kind of problem would be to posit a spirantization rule which weakens both /g/ and /k/ to the corresponding voiced or voiceless velar spirant intervocally, that is to ɣ and ʰ, respectively. A second rule would apply only to ʰ and weaken this sound further to ɦ. Indeed, one would have to have a very good reason for not proposing this solution.

Diachronically, if *g and *k changed to ɣ and ʰ, one would posit a set of changes exactly parallel to the synchronic rules proposed above. In this case the diachronic procedure is exactly parallel to the synchronic procedures. This is a testament to the tightly interrelated behavior of sounds in languages synchronically and diachronically.

I noted above that the Principle of Optimal Use is essentially concerned with simplicity. The synchronic correlate of this principle is to be found in the practice of making synchronic hypotheses as general as possible. For example, it is quite common to derive nasalized vowels at the surface from an underlying sequence of an oral vowel
plus a nasal consonant. Suppose some other alternation, for example, the nasalization of the initial consonant of a following word, can be plausibly derived by appealing to the hypothetical nasal consonant. One would have to have a very good reason for not optimally using that hypothetical nasal nasal. And the parallel to diachronic procedure is clear. If one can plausibly account for an alternation on the basis of what it has already been necessary to reconstruct, one should do so. The result is a simpler hypothesis.

The Principle of Similarity of Function is not really a diachronic principle; it is much more general and permeates much of modern linguistic analysis. For example, it can be found at the base not only of phonological argumentation, but at the base of syntactic argumentation as well. For example, this principle underlies much of the work that has been done towards reducing the basic syntactic categories. Such arguments are usually of the form: X and Y are distinct categories that function similarly in important ways; therefore they should not be analyzed as separate categories. The disagreement over these hypotheses frequently is directed at the issues of whether or not X and Y do behave similarly, whether those ways are in fact important, and so on. The principle is not called into question, as indeed it should not be.
I have suggested in 4.1 and 4.2 that the systematic pairing of morphemes of different canonical forms can be taken as indicative of the existence of a regular relational pattern historically. One then constructs a hypothesis describing the nature of the original pattern and the changes the pattern has suffered. Much of modern linguistic analysis rests on the same procedures synchronically. For example, in syntactic analysis, when one syntactic pattern of a language is systematically paired with another sentence pattern, one can establish an abstract underlying pattern and derive the two surface patterns by transformational rules. At base of this type of analysis is the assumption that systematically paired linguistic elements are indicative of a 'deeper' unity. This unity can be expressed in a number of ways, for example, in an abstract underlying sentence pattern, in the underlying form of a morpheme, or diachronically, in a common antecedent pattern for two synchronic relational patterns.

It should be clear, then, that the six principles discussed in this work are not specific to internal reconstruction or even diachronic analysis. I have shown that all have parallels in linguistic analysis in general or the principles of the philosophy of science. In short all the principles represent the diachronic versions of
tried and tested principles of argumentation and analysis. As a consequence, inferences made about the history of a language on the basis of these principles and internal evidence--assuming, of course, that the principles have been carefully applied and the data correctly interpreted--are of the highest integrity.

While I have illustrated the application of each of these principles to actual linguistic data, I have not dealt with any particular reconstruction problem extensively. Chafe has provided an excellent illustration of the application of the more traditional principles of internal reconstruction in his (1959) analysis of Seneca. However, the method of reconstructing relational patterns proposed above has not been illustrated in such a way. In the next chapter I will provide a quite extensive example of relational pattern reconstruction to better illustrate the application of the method. It is hoped that any remaining uncertainties surrounding the method will be clarified by that exposition.
NOTES

1 See Schachter and Otanes, 1972, for a thorough discussion of the pang-formation in Tagalog.

2 The two aspirates are always separated by at least one sonorant. Thus, the disimilation is a distant one. The second aspirate is always a member of a root.

3 Other important considerations would be the relative antiquity of the attestations from the various daughter languages, the probability that certain developments are due to areal phenomena, and the possibility of language contact in the histories of the daughters.

4 One could propose a derivation like /tup/ + tup-up + tup-p. However, the only motivation for doing so is the resulting symmetry of the analysis. It is not clear that this consideration is that important. One might also suggest that the vowel of the suffix, for example u of -ut should be analyzed as part of the base; kart would then be /karty/ and a rule of final vowel deletion would yield the surface form.

5 It is important to note that I am talking about what has been called phonemic reconstruction here, and not phonetic reconstruction. See 108-109 above.

6 I cannot even begin to give a systematic account of simplicity here. I use 'simplicity' in its pre-theoretic sense.

7 Of course it is not always possible to derive nasalized vowels from this source.
5.0 Introduction. In the preceding chapters I have discussed methods for internal reconstruction in quite general terms. The purpose of this present chapter is to amplify that discussion by way of an extended example. My main concern will be to illustrate the method of relational pattern reconstruction explicated in chapter 4. The focus of my example is the problem of Proto-Austronesian (PAN) medial nasal clusters and the issue of whether or not they should be derived (historically) from an earlier canon without such clusters and a nasal infix. As I noted in the introduction, Dempwolff's reconstructed nasal clusters represent severe distributional anomalies for PAN morphological and phonological systems. However, to date, no serious systematic attempt at a solution to these problems has appeared. Besides illustrating relational pattern reconstruction, then, this chapter provides a systematic analysis of the problem of medial nasal clusters in PAN.

In 5.1 I will briefly discuss Austronesian studies and Dempwolff's reconstruction of the PAN sound system. I will describe the reconstructed morphological canons in
5.2, and the reconstructed consonant clusters in 5.3. These three sections provide the necessary background information, define the problem the nasal clusters represent, and introduce the nasal infix hypothesis. After reviewing this evidence, I will discuss Dempwolff's solution and the problems it entails in 5.4. I will consider general arguments against proposing an infix of the shape -VN- (as Dempwolff did) in 5.5 and argue that *-ŋ- is the most likely shape for the nasal infix. In 5.6 I will review Gonda's objection to reconstructing an infix on the grounds that its function has not been determined and propose the assignment of the infix to *-ŋ-. In 5.7, I will discuss the morphological process Dempwolff termed Pränasalierung and suggest that the assumption of an infix, *-ŋ-, allows one to draw an explicit connection between the development of medial nasal clusters and Pränasalierung *pən-, for example, is the result of a morphological reanalysis of a CV- prefix (*pə-) and *-ŋ-. In 5.8 I provide a tentative semantic analysis of the function of *-ŋ- surveying the forms Dempwolff reconstructed. This analysis is then considered in light of the semantic analysis of *-ŋ- in Ngadju Dayak and Tagalog presented in 5.9. Finally I will discuss some of the problems the nasal infix hypothesis entails in 5.10. I will suggest that the weight of the evidence demands adopting *-ŋ-, at least as a working hypothesis.
5.1 Proto-Austronesian. The Austronesian language family consists of 500-800 languages spoken throughout the Pacific Ocean and in certain mainland Southeast Asian areas.\(^1\) Geographically the family spreads from Madagascar off the south east coast of Africa in the west, to Formosa in the north, to the Hawaiian Islands in the north east, and to Easter Island in the south east. There are also non-Austronesian languages spoken within this area in New Guinea and Melanesia. These languages are commonly referred to as Papuan languages although more than one language family is thought to be represented by them.

While the early European explorers remarked on the apparent similarity of Austronesian languages—particularly the Austronesian languages of Polynesia—and some early scholars suggested that all Pacific languages were the remnants of an older common parent (see for example W. Schmidt (1906), W. von Humboldt (1836-9), and Dyen (1965) for an excellent review of early writings on Austronesian) the first comprehensive study of this question was published by Otto Dempwolff in his Vergleichende Lautlehre des Austronesischen Wortschatzes (three volumes, 1934-38).

In the first volume—Induktiver Aufbau einer indonesischen Ursprache (1934)—Dempwolff compared Tagalog (Tg.) Toba Batak (TB.) and Javanese (Jv.) and established a sound system for proto-Indonesian (Urindonesisch). In
volume two—Deduktive Anwendung des Urindonesischen auf austronesische Einzelsprache (1937)—Dempwolff showed that the sound system he established for proto-Indonesian was sufficient to account for the languages of Melanesia and Polynesia. He then felt justified in identifying his proto-Indonesian with the parent language of all Pacific Island languages, proto-Austronesian (Uraustronesisch). The final volume of his study was a lexicon for proto-Austronesian.

While Dempwolff's method has received much criticism, it is widely acknowledged that his reconstruction was very accurate. Although Dempwolff refers only to three languages in volume one (VLI), his earlier publications (see references) attest to his familiarity with some thirty or more Austronesian languages. As Dahl (1973:5-10) notes, Dempwolff used only Tg., TB., and Jv. in VLI because he knew that these three languages would attest all the reconstructed sounds he considered necessary for proto-Austronesian.

Figure 5 (p. 139) shows the sound system Dempwolff reconstructed for proto-Austronesian (PAN). The reconstructed sounds p, b, t, d, t̪, d̪, k, g, l, y, m, n, n', ń, mp, mb, nt, nd, n̂t, n̂d, nk, ng have their usual values and warrant no further discussion here. Dempwolff considered t̪ and d̪ to represent (pre-)palatal stops,² voiceless and
Stops
  p  t  t'  k'  k
  b  d  d'  g'  g

Spirants
  v  l  l'  j  γ

Nasals
  m  n  n'  n

Nasal
  mp  nt  n't  n'k  nk
Clusters
  mb  nd  n'd  n'g  ng

Vowels
  i  u  e
    a

Laryngeals
  '  h

Figure 5. Dempwolff's PAN Sound System
voiced respectively. However, the reflexes of *t' and *d' are radically different. Tg., TB., and Jv. all show s for *t' (as do most Austronesian languages). Tg., however, merges *d' and *d to d where TB. and Jv. show d' for *d'.

Dempwolff argued that *t' must be reconstructed even though the reflexes unanimously indicate *s because the morphological process of **Pränasalierung**—a process whereby bases with an initial voiced stop appear with a nasal homorganic to the stop (for example, a base that begins with b appears as mb) and voiceless stop initial bases appear with a nasal homorganic to the original stop substituted for that stop (for example a p initial base appears as m)—applied only to stop consonants in all other cases (1934:39):

Aus #22,d geht hervor, dass die **Pränasalierung** 10 Verschlusslauten verändert, aber nur 2 Reibelaute, Jv. w und Tg., TB., Jv. s. Von diesen geht Jv. ſ, wie in #29 gezeigt werden wird, auf UIN b zurück, also auf einen Verschlusslaut. Indem also ſ als einziger Reibelaut übrig bleibt, der durch **Pränasalierung** verändert wird, drängt sich der Analogie-Schluss auf, das auch ſ ursprünglich ein Verschlusslaut gewesen ist.

It should also be noted that while *t' was reconstructed for initial, medial, and final positions, *d' was reconstructed only for initial and medial positions.

Dempwolff considered *k' and *g' to be palatal stops made with the front part of the back of the tongue. Here too, however, the reflexes of each sound are different.
Tg. and TB. show $k$ for *$k'$, while Jv. shows $t'$. For *$q'$, Tg. exhibits $l$ medially and $d$ finally. TB. shows $g$ ($k$ finally as the result of a general process of final obstruent devoicing) and Jv. shows $r$. Dempwolff reconstructed *$k'$ for only initial and medial positions, but *$g'$ for only medial and final positions.

While Dempwolff considered *$v$ to be a voiced bilabial fricative, the sound is reflected as a semivowel, $w$, or zero almost universally. Likewise, Dempwolff interpreted *$j$ as a voiced palatal fricative even though the sound is most often reflected as the semivowel $y$ or as zero. *$j$ is also extremely rare in initial position. Finally, Dempwolff considered *$l$ to be a voiced retroflexed fricative.

In discussing the phonetic values associated with these sounds, it is worth noting that Dempwolff actually considered his reconstructed segments to have the phonetic values he described for them. That is, he did not consider his reconstructions to be mere formulas, abbreviations of correspondence sets (see 4.3). However, in several cases Dempwolff found that he had already 'used up' a sound in a prior reconstruction and had to make an alternative assignment. For example, the correspondence set Tg. *$l$, Jv. and TB. *$r$ would indicate an *$r$-like sound. However, Dempwolff noted (1934:52):
Es läge nahe, für diesen Laut ein alveolares r in die Ursprache aufzunehmen. Aber im Hinblick darauf, dass im TB. noch ein r anderer Ursprungs vorkommt (#36), und dass im Jv. r auf mehrere andere Laute zurückgehen kann (#37 und #50, g und h) wird hier vorgezogen, einen neuen Laut für die Ursprache zu wählen, den stimmhaften, retroflexen Reibe­laut, ˌ...

It should also be noted that not all sounds were reconstructed for all three positions; that is, initially, medially, and finally. The retroflex *t was reconstructed only for initial and medial positions, as were the palatals *k', *d', and *n'. The homorganic nasal clusters occur only medially between two vowels and never initially or finally within a wordbase.

The vowels *i, *u, *a and *ə as well as the laryngeal *h all have their usual values. *i is, however, problematic. Dempwolff used this symbol to denote smooth vowel onset—weicher Vokal-Einsatz, -Zwischensatz, and -Absatz. He did not permit reconstructed items to begin or end in a vowel. Instead he inserted *i. Likewise, he irregularly inserted i into the forms of extant languages to regularize their canonical form to CVCVC. Dahl devotes a chapter of his book, Proto-Austronesian, to a discussion of Dempwolff's use of this symbol. Dahl argues (1973:22):

'Thus the only meaning the symbol i has is C zero at a syllable limit' (italics Dahl's). Most modern Austronesian scholars eliminate the symbol from reconstructions and I
will follow that practice here. (See below for other changes in Dempwolff's notation adopted here.)

In the discussion of Dempwolff's assignment of Tg., TB. and Jv. s to *t', that is a stop articulation rather than some sort of fricative, mention was made of Pränasalierung, a morphological process affecting the initials of wordbases. The term Pränasalierung is actually a cover term for two distinct phonological processes, nasal substitution (nasaler Ersatz) and nasal accretion (nasaler Zuwachs). Nasal substitution as a rule occurs only with voiceless stop initial wordbases. In conjunction with certain prefixes, p, t, and k, for example, are replaced by m, n, and ng respectively. Nasal accretion, on the other hand, affects only initial voiced stops; b, d, and g, for example, in conjunction with the same prefixes occur as mb, nd, and ng. Since Pränasalierung will be discussed in much greater detail in later sections of this chapter, I will not say any more about the process here.

Throughout the remainder of this chapter I will use the following substitute symbols for Dempwolff's: for the retroflexes *t, *d, *l, *nt, and *nd I write *T, *D, *L, *NT, and *ND respectively. For *γ I write *R. For *ŋ I write *ng. For *e I write *e, and *ι has been omitted.

5.2 PAN Canonical Structure. In the last section I briefly described the history of Austronesian studies and
Dempewolf's reconstructions of the PAN sound system. In this section I will review the kinds of wordbases to be found in VLIII. The majority of wordbases reconstructed by Dempewolf are disyllabic. Monosyllabic wordbases are all of the form CVC: for example *bak 'split off', *bun 'to be abundant', *Duk 'dwelling, habitation', *d'eng 'standing', *gang 'gape, split open', *gat 'asunder, in two', and *tet 'tear, rip up'. Disyllabic wordbases are predominantly CVCVC, but there are others of the shape CVCCVC. Examples of this latter type fall into two general classes: (A) those which are composed of two identical syllables (commonly abbreviated CVC\textsuperscript{2}); and (B), those which contain a medial homorganic nasal cluster. Consider, *
kuDkuD 'rasp', *piRpiR 'spurt, splash' *TakTak 'crash, clatter' (A), and *hanteD 'deliver', *tanduk 'horn' and *d'a(m)baj 'to hang down' (the parentheses surrounding the nasal in forms like *d'ambaj will be discussed below). The (B) class will be designated by the formula CVNCVC and it is this canon that will be the major concern of this chapter.

Trisyllabic wordbases have the following shapes: CVCVCVC, CVNCVCVC, CVCVNCVC (where, as above, 'NC' represents a homorganic nasal cluster) or CV+CVC\textsuperscript{2}, that is a reduplicated monosyllable plus a CV- prefix. There are only four examples of tetrasyllabic wordbases in VLIII and
according to Dahl (1973:10) these consist of 'two normal disyllabic wordbases.'

PAN exhibits prefixes of the shape CV-, the most common, and CVC-. Infixedes are all of the form -VC- and infixed before the first vowel of a wordbase (see below for a slight modification of this statement). Suffixes are most commonly -VC, but -CVC suffixes occur as well. Dahl (1973: 10) notes that affixing with CVC- or -CVC will not produce any new consonant clusters; that is, 'clusters that do not already exist in CVC\textsuperscript{2} wordbases.' Thus, the structure of PAN words does not differ significantly from the structure of PAN wordbases.

5.3 Consonant Clusters in PAN. Having considered the segments and canons reconstructed for PAN, it is now possible to focus on the problem of PAN medial nasal clusters. Notice that the examination of PAN morphological canons above indicates that the occurrence of consonant clusters in PAN was extremely limited. While CC clusters were permitted only across a morpheme boundary, NC clusters could also occur morpheme internally. The consonant clusters that occur in CVC\textsuperscript{2} wordbases at least originally contained a morpheme boundary since they are clearly reduplicated forms. Dempwolff considered the monosyllabic wordbases to be the historical antecedents of the CVC\textsuperscript{2} forms. In VLI he argued (pp. 118-19):
Man wird nicht umhin können, diese einfachen Wurzelwörter als Vorläufer der iterierten anzusprechen, und man kann sie daher, obwohl sie nur in einer der drei Sprachen lebendig sind, als UIN... ansetzen.

It is not clear that the reduplication process which yielded the CVC\(^2\) canon in PAN was still productive in PAN times. There are many pairs CVC and CVC\(^2\), for example, *bak and *bakbak 'to take off the bark of a tree' and *bun and *bunbun 'to amass, accumulate'. But there are also CVC monosyllables that lack a corresponding CVC\(^2\) form, and likewise, CVC\(^2\) forms that lack a CVC counterpart. The missing monosyllables are assumed to have been lost\(^4\) and the missing CVC\(^2\) forms either have been lost or never existed.

Regardless of the productivity of this type of reduplication in PAN, the important point is that all CC clusters are derived--either synchronically or diachronically--and include a morpheme boundary (at least originally).

Homorganic nasal clusters could occur across a morpheme boundary when the nasal was the final consonant of a prefix; that is, as a result of Pränasalierung. However, the second member of all such clusters was always a voiced stop. When prefixes with final nasals were affixed to wordbases beginning with a voiceless stop, nasal substitution (nasaler Ersatz) occurred. Thus, CVng- plus bVCVC (a schematic example of a voiced initial wordbase) will yield CVm-bVCVC, while CVng- plus pVCVC will yield
CVmVCVC. I will have more to say about Pränasalierung in later sections of this chapter and I only mention it here to illustrate the conditions under which NC clusters occurred with an intervening morpheme boundary.

Non-homorganic nasal clusters occur in CVC² forms when the final consonant of the monosyllable is a nasal. Under these circumstances the nasal did not assimilate and the tendency in the development of the daughter languages is to preserve the original point of articulation of the nasal. For *bunbun, for example, Tg. shows bunbon, TB. bunbun, but compare Malay bumbun. There are no examples in VLIII that indicate the final nasal of a reduplicated monosyllable has changed and assimilated to the point of articulation of the initial of the following (identical) monosyllable. That is to say, any homorganic nasal clusters that occur in CVC² forms are due to the coincidence of the points of articulation of the initial and final consonants of the original monosyllable, for example *gang and *ganggang 'gape'.

The only clusters that do not obviously contain a morpheme boundary, then, are morpheme internal homorganic nasal clusters; that is, CVNCVC, CVNCVCVC and CVCVNCVC forms. However, a close look at these forms reveals that there are cases where the daughter languages do not unambiguously attest a nasal cluster. In these cases, CVNCVC
in one daughter may correspond to CVCVC in another. There are also cases where a daughter exhibits both CVCVC and CVNCVC corresponding to CVNCVC or CVCVC in other daughters.

Where a nasal cluster was not unambiguously indicated, Dempwolff reconstructed an optional nasal considering the nasal to be a special case of nasaler Zuwachs. He argued (1934:109):

Aus #53 und #54 wird hier der Schluss gezogen, dass die Lauterscheinung des nasalen Zuwachses beim Inlaut ursprünglich ein Lautvorgang gewesen ist, ebenso wie beim Anlaut, dass aber der erste früh erstarrt ist, während der letzte lebendig blieb. Man darf wohl vermuten, dass es sich beim nasalen Zuwachs des Inlauts um ein Inflix gehandelt hat, aber dessen ursprünglichen Lautbestand und seine grammatische Funktion sind nicht aufzuklären.

If this optional (fakultative) nasal does represent an infix, I find it extremely unlikely that its occurrence was truly optional. Undoubtedly, morphological function played a determining role in the occurrence of the nasal. Moreover, it is not clear from Dempwolff's statement whether he thought all homorganic nasal clusters reflected an earlier infix or just these optional cases. It is also not clear what Dempwolff thought the original shape of the infix was. Since all other reconstructed infixes are -VC-one might expect the infix here to have this shape as well. However, with the evidence available a narrower assignment cannot be made. -um- and -in- can be eliminated as potential
shapes since *-um- and *-in- have already been reconstructed and there is no evidence linking Dempwolff's hypothetical infix with either of these (see below 5.4). However, even if it were possible to establish the original identity of the vowel and nasal, there are serious problems with a -VN-infixation hypothesis. Most notably, affixation of a -VN-infix into a wordbase will yield C₁-V₁N₁-V₁C₂V₂C₃, and in a later section (5.4) I will show that there is no straightforward way to account for the loss of V₁, a necessary syncope if the nasal of the infix (N₁) is to become adjacent to C₂ of the wordbase.

I will refer to the hypothesis that PAN medial nasal clusters—either -(N)C- or -NC- or both—exhibit (the remnants of) a nasal infix as the Nasal Infix Hypothesis. Many Austronesianists have accepted some version or other of this hypothesis. However, this acceptance has not been accompanied by an abundance of evidence for the hypothesis nor has a systematic description of the details of the hypothesis appeared. In this regard, R. Blust has captured nicely the working assumptions underlying the acceptance of the Nasal Infix Hypothesis by many scholars. He argues (1971:76):

...the theory that Proto-Austronesian had only homorganic nasal clusters suffers from another serious drawback. By this theory one is forced to regard the agreement in point of articulation of the nasal and following consonant as merely a fortuitous
fact of Proto-Austronesian canonical shape. But a distributional limitation of this kind would most naturally be explained as the outcome of an assimilation. Whether the unity of the nasal infix is attributed to Proto-Austronesian, or to some earlier stage, it seems clear that its phonetic diversity arose ultimately through conditioned change.

Before turning to a detailed analysis of the Nasal Infix Hypothesis, one would do well to consider just what evidence there is for the hypothesis. I have already shown that PAN had some disyllables and trisyllables that exhibited an ambiguously attested nasal cluster. It should be pointed out that these forms do not represent scattered examples. Rather there are 196 cases in VLIII where Demp-wolff reconstructed an optional nasal (see Table 6, pp. 193-203). This number represents 42 percent of the total number of forms reconstructed with nasal clusters—both -(N)C- and -NC-. In addition to this apparent alternation between -C- and -NC-, that is, -(N)C- forms, there is the distributional limitation referred to by Blust. However, this putative limitation should be looked at in more detail.

The distributional limitation Blust refers to above is no doubt the occurrence of only one type of morpheme internal consonant cluster in PAN—in—homorganic nasal clusters—and the restriction on the occurrence of this type of cluster to medial position. However, one would do well to ask whether 'a distributional limitation of this
kind would most naturally be explained at the outcome of an assimilation'? Suppose that PAN medial nasal clusters are not derived from an infix. Does this represent a real distributional anomaly for PAN?

The answer based on the evidence presented so far would have to be in the negative. There is nothing phonologically anomalous in a language exhibiting only one kind of consonant cluster, particularly homorganic nasal clusters. There are many languages that permit only this one type of cluster. R. Herbert (1977a) has suggested that the reason for this preference of NC clusters has to do with the ability of nasals to function as syllable coda. He notes that it is not infrequent for a predominantly CV language—that is, a language that permits only open syllables—to permit CVN syllables. Herbert reports that when a language allows only homorganic nasal clusters, the clusters are usually found word internally as opposed to initial or final positions. He has also found (personal communication) that there is a tendency against having more than one cluster within a word, particularly in successive syllables. Thus, nasal cluster languages—as I shall refer to languages that permit only -NC- clusters—often do not permit CVNCVNCVC forms even though such forms may be consistent in other respects with the morpheme structure constraints of the language.
It is clear at a glance that PAN is consistent with Herbert's findings concerning nasal cluster languages. PAN has no word initial or word final nasal clusters. Moreover, with the expected exception of CVNCVNCVC, PAN permits nasal clusters between any two syllables word internally. That is, CVNCVC, CVNCVCVC, and CVCVNCVC all occur.

It is especially noteworthy that PAN exhibits a CVCVNCVC canon since this canon would apparently count as counterevidence to the Nasal Infix Hypothesis. No matter what the shape of the infix may have been, it is fairly certain that it was infixed into the first syllable of a wordbase and not the second. Notice that there are no CVCVCVC canons which one would expect if the infix occurred in the second syllable. However, while both PAN CVNCVC and CVNCVCVC are expected, no morphologically simplex CVCVNCVC forms should be found.

There are twelve instances of this canon in VLIII. These forms are listed below in Table 1.
**Xbalan'd'a** and **Xpalangka** will not be considered any further here since they are loanwords. With the exception of (g) and (i) all the remaining forms show *1 or *L (Dyen’s *r) for C₂. Moreover, with the exceptions of (f), (g) and (l) all forms show *a for V₁. These -al-, -aL- sequences immediately suggest the old *-al/r- infix referred to frequently by Dempwolff (VL, see also 1925) and others. Notice also that Dempwolff cross-referenced some of the forms to shorter, disyllabic forms. For example, he cross-references *ka[L]an'd'ang to *[l]an'd'ang and *makempu to *empu. I would add here *galinggang and *galunggang to each other as well as to *galinggang 'name of a plant'.

TABLE 1. VIILIII Instances of PAN CVCVNCVC

- a.) Xbalan'd'a expenses Skrt. loanword
- b.) galinggang name of a plant
- c.) galunggang name of a tree
- d.) galumbang to roll
- d.) ka[L]an'd'ang wicker basket c.f. [l]an'd'a
- f.) kulambu curtain, mosquito net
- g.) Lima(n)[t]ek small leech
- h.) ma[L]an[t]i name of a tree
- i.) makempu grandchild; add. to *empu
- j.) Xpalangka sedan chair Skrt. loanword
- k.) t'alu(m)bung covering veil
- l.) t'e[L]ampang scion, tine
The suggestion here is that these forms are morphologically complex, containing either an infix or a prefix or suffix.\(^7\) If these forms are in fact morphologically complex, PAN has no CVCVNCVC forms. That is, there are no PAN trisyllables with a homorganic nasal cluster between the second and third syllables. But before one can claim that this constitutes a phonological anomaly—such a gap would be without explanation in terms of Herbert's observations—it is necessary to show that CVCVVCVC and CVNCVVCVC trisyllables are not also morphologically complex in the same ways. If they are, the best one can say is that PAN did not permit trisyllabic wordbases. Table 2 (p. 155) lists the trisyllables other than CVCVNCVC found in VLI.

Even a quick glance at the trisyllables of Table 2 indicates that a considerable number of morphologically simple trisyllables must be considered PAN. It is noteworthy that the sequences -al- and -al- are common in these forms, but the gross frequency is less, 40 percent, than for the forms of Table 1, 60 percent. Moreover, while some of the forms of Table 2 are cross-referenced, there is no reason to suspect that all, or even most of the forms are morphologically complex. Indeed, fully one-half of the forms are above suspicion.

If one interprets PAN as a nasal cluster language, then, PAN exhibits a canonical peculiarity. There is no
# TABLE 2. PAN CVNCVCVC and CVCVCVC Trisyllables

| a.) | alimet' to be visible |
| b.) | ali[n]u shade |
| c.) | anitu departed soul |
| d.) | apu[L]a reconciliation |
| e.) | bangkudu name of a tree |
| f.) | balakang back |
| g.) | balanak name of a fish |
| h.) | bali[dD]a name of a fish |
| i.) | balig'a lathe of the loom |
| j.) | bali[j]ung planishing axe |
| k.) | balikat shoulder blade |
| l.) | baliLang sulphur |
| m.) | ba[l]ingin name of a tree |
| n.) | ba[L]uvang white spot on the face of animals |
| o.) | banu[v]a land, settlement |
| p.) | banu[v]lang name of a tree |
| q.) | bi[t]uhen star |
| r.) | bi[t]uka intestines |
| s.) | bulajlaj trunk, nozzle |
| t.) | daRami straw, blade, stubble |
| u.) | Daluvang cloth made from bast |
| v.) | galumat deck of a (water) vessel |
| w.) | habut'a accumulation, massing |
| x.) | halaman outer court, porch |
| y.) | haluvan bow of a ship |
| z.) | haLimav beast of prey |
| aa.) | hampelat name of a tree of which the leaves are used for polishing |
Table 2 (continued)

bb.) kamudi  helm  
c.  katapang  name of a tree  
dd.) kulapu  name of a fish or aquatic animal  
 ee.) ku[v]ala  estuary  
 ff.) k'aLemin  crystal glass, mirror  
gg.) lampujang  name of a plant  
 hh.) la[t]'u[n]a[h]  onion  
 ii.) Lumbi[j]a  sago palm  
 jj.) muva[L]a  mouth of a river  
 kk.) paLahu  water-craft, boat  
 ll.) pa[n]iki  flying fox  
m.) pan'enget  wasp  
 nn.) [t]ambilang  spade  
oo.) [t]a(m)buni  afterbirth  
 pp.) talinga  ear  
 qq.) [t]ali[t']aj  name of a tree  
 rr.) ta[L]ima  accept  
 ss.) [t]amanang  to be'barren  
 tt.) tangiRi  name of a fish  
uu.) [t]ingaDah  look over head  
vv.) t'angkalan  chopping block  
w.) t'uligi  spear  
 xx.) t'umanget  spirit, mind, intellect
reason to limit the occurrence of nasal clusters to a position between only the first and second syllables of trisyllabic wordbases. Thus, the lack of a morphologically simplex CVCVNCVC canon is crucial since such a canon would (a) provide counter evidence to the Nasal Infix Hypothesis, and (b), provide evidence for the hypothesis that PAN was a nasal cluster language.

In summary, there are two pieces of evidence that support the Nasal Infix Hypothesis: (1) the widespread occurrence of -(N)C- forms (nasaler Zuwachs); and (2) the conspicuous absence of CVCVNCVC forms that are not also morphologically complex. The Nasal Infix Hypothesis accounts for both of these: (1) merely represents the historic alternation between uninfixed di- and trisyllables—that is, CVCV \sim CVNCVC and CVCVVC \sim CVNCVVCVC; and (2) obtains as a consequence of the infix occurring between the first and second syllables of wordbases. (1) and (2), then, can be explained as the remnants of a relational pattern in the prehistory of AN. The systematic pairings of morphological canons that constitute this relational pattern can be represented as CVCVC::CVNCVC and CVCVVC::CVNCVVCVC for the di- and trisyllabic wordbases respectively. The task remains to construct a hypothesis of sufficient detail. It is not sufficient to simply assert that there was a nasal infix (of some shape or other) that gave rise
(through some set of linguistic changes or other) to PAN medial nasal clusters. Rather, the parenthetical material above is of paramount importance.

In terms of the method of relational pattern reconstruction proposed in the last chapter, only the first of three principal steps has been taken: evidence in the systematic pairing of morphemes of different canonical structure has been presented. It is now necessary to determine what the original shape of the infix was (step 2) and consider what changes induced the pattern change (step 3). Since it is unlikely that nasal infixation was productive in PAN times (see 5.10) the pattern change under discussion is an example of the loss of a relational pattern.

5.4 Dempwolff's Ngadju Article and the Shape of the Infix.

In 1922 Dempwolff's *Entstehung von Nasalen und Nasalverbindungen im Ngadju* (Dajak) appeared. This work was a detailed report of his investigations of nasals, nasal clusters and Pränasalierung in Ngadju. Although the investigation was primarily concerned with the history of a particular language, Dempwolff included cognates from other Austronesian languages and in his conclusions, Dempwolff made claims about PAN and pre-Ngadju. Dempwolff analysis was quite thorough and he made some rather interesting suggestions concerning the origins of nasal clusters. As
a consequence, I will devote a considerable amount of
attention to discussing this article.

Dempwolff (1922:203) argued that there were two
sources for nasal clusters:

1. Der grauen Vorzeit der austronesischen "Ursprache" gehören die Nasalverbindung an, die durch den Zusammentritt zweier Wurzelwörter zu einem Stammwort entstanden sind...

2. Die Nasalverbindungen, die durch Infigierung eines (hypothetischen) m entstanden sind, sind einer alten Schicht des eigentlichen Ngadju zuzuschreiben.

According to Dempwolff then, nasal clusters arose from (1) the concatenation of two distinct monosyllabic wordbases where the final nasal of the first wordbase assimilates to the point of articulation of the initial of the second wordbase, and (2), the infixation of an -m- into a CVCVC wordbase. While Dempwolff apparently abandoned (1) by the publication of VL—there is no mention of this kind of word formation process in VL—and hesitated in assigning the infix to -m-—recall that in VL Dempwolff claimed that the form of the infix was not clear—both of these analyses will be reviewed here. It is noteworthy that Dempwolff never discussed the reasoning behind his decision to omit (1) from VL or his hesitation concerning (2).

The basic argument of (1), which I will refer to as the compound analysis, is that when monosyllabic basewords
(Wurzelwörter) of the shape CVN (where 'N' is any nasal) and CVC are placed together, the nasal assimilates to the point of articulation of the following consonant and a form, CVNCVC, results. The scheme is actually a little more restricted than this. C₁ of the monosyllable C₁V₁C₂ must be a stop. Table 3 lists the formulae cited by Dempwolff (1922:178-9, 190-1). The number to the right of each example indicates the number of instances Dempwolff had found. In some cases, the example he cites is a trisyllable. While he often indicates that the trisyllable contains some affix, others are apparently basic and cited under the formula for convenience. The affixes of morphologically complex trisyllables are enclosed in parentheses to the right of the gloss.

The sequences CVn'+CVC, CVn+CVC and CVN+CVC are conspicuously absent. Recall however that *n' did not occur finally in PAN. Ngadju (ND) merely reflects this situation. Moreover, the retroflex segments of PAN have merged with their nonretroflex counterparts in ND. Since the basis of Dempwolff's discussion is ND, it is not surprising that these clusters are absent. However, for PAN it might be expected that some -NT- and -ND- clusters would arise through the same process. Recall that *T and *D do not occur initially. Consequently these segments will never occur as C₁ of a C₁V₁C₂ monosyllable. Since the
TABLE 3. Dempwolff's CVN+CVC > CVNCVC Analysis

I. Voiceless initials (CVC)

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Morphology</th>
<th>Meaning</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVm+pVC</td>
<td>&gt;CVmpVC</td>
<td>kalumpit to be pinched (-al-)</td>
<td>9</td>
</tr>
<tr>
<td>CVm+tVC</td>
<td>&gt;CVntVC</td>
<td>lantai field</td>
<td>9</td>
</tr>
<tr>
<td>CVm+t'V</td>
<td>&gt;CVn't'V</td>
<td>kun't'i palace</td>
<td>2</td>
</tr>
<tr>
<td>CVn+pVC</td>
<td>&gt;CVmpVC</td>
<td>pampang to be plain, smooth</td>
<td>2</td>
</tr>
<tr>
<td>CVn+tVC</td>
<td>&gt;CVntVC</td>
<td>tantan summit</td>
<td>10</td>
</tr>
<tr>
<td>CVn+kVC</td>
<td>&gt;CVngkVC</td>
<td>bungkor cover, coverlet</td>
<td>8</td>
</tr>
<tr>
<td>CVng+pVC</td>
<td>&gt;CVmpVC</td>
<td>pampang prong</td>
<td>16</td>
</tr>
<tr>
<td>CVng+tVC</td>
<td>&gt;CVntVC</td>
<td>garantong kettle drum (ga-)</td>
<td>32</td>
</tr>
<tr>
<td>CVng+t'VC</td>
<td>&gt;CVn't'VC</td>
<td>karin't'in chime (ka-)</td>
<td>2</td>
</tr>
<tr>
<td>CVng+kVC</td>
<td>&gt;CVngkVC</td>
<td>rangkum suddenly</td>
<td>44</td>
</tr>
</tbody>
</table>

II. Voiced initials (CVC)

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Morphology</th>
<th>Meaning</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVm+bVC</td>
<td>&gt;CVmbVC</td>
<td>lambon to choke (as weeds)</td>
<td>18</td>
</tr>
<tr>
<td>CVm+dVC</td>
<td>&gt;CVndVC</td>
<td>kandong be hidden</td>
<td>13</td>
</tr>
<tr>
<td>CVm+d'VC</td>
<td>&gt;CVn'd'VC</td>
<td>batan'd'ak to shove off</td>
<td>4</td>
</tr>
<tr>
<td>CVm+gVC</td>
<td>&gt;CVnggVC</td>
<td>gonggom sea snail</td>
<td>2</td>
</tr>
<tr>
<td>CVn+bVC</td>
<td>&gt;CVmbVC</td>
<td>d'amban footpath</td>
<td>7</td>
</tr>
<tr>
<td>CVn+dVC</td>
<td>&gt;CVndVC</td>
<td>endeng place</td>
<td>10</td>
</tr>
<tr>
<td>CVn+d'VC</td>
<td>&gt;CVn'd'VC</td>
<td>d'und'un on</td>
<td>8</td>
</tr>
<tr>
<td>CVn+gVC</td>
<td>&gt;CVnggVC</td>
<td>sanggop confession</td>
<td>3</td>
</tr>
<tr>
<td>CVng+bVC</td>
<td>&gt;CVmbVC</td>
<td>bumbong buds, unopened leaf</td>
<td>19</td>
</tr>
<tr>
<td>CVng+dVC</td>
<td>&gt;CVndVC</td>
<td>landok list</td>
<td>28</td>
</tr>
<tr>
<td>CVng+d'VC</td>
<td>&gt;CVn'd'VC</td>
<td>tan'd'ang growth</td>
<td>8</td>
</tr>
<tr>
<td>CVng+gVC</td>
<td>&gt;CVnggVC</td>
<td>karanggam handful</td>
<td>14</td>
</tr>
</tbody>
</table>
final nasal of CVN assimilates to the point of articulation of C₁, no N's (retroflex nasals) could result in this fashion. And, even though a CVN monosyllable could occur in PAN, the final N would assimilate to some other point of articulation. Thus, CVN+CVC will not yield -NT- or -ND-clusters. The origins of these clusters will have to be found elsewhere. All the other combinations of nasal and stop that do occur in PAN could have their origins in formulae like those of Table 3.

There are at least three serious problems with Dempwolf's monosyllabic compound analysis. First, although based on Brandstetter's Wurzel und Wort in den indonesischen Sprachen (1910) where a number of monosyllabic root words and word formation rules are posited to derive disyllabic words, Brandstetter's work is not widely accepted. While his analysis is suggestive and a large number of CVC sequences can be found across words with similar meanings—for example, Tg., Bisayan, Gayo nipis 'thin', Tontemboan apis 'small', tompis 'sunken (of cheeks)', Old Jv. tapis 'small', and Makassar nipisiq 'thin', root pis (*pit')—the second part of the theory has not been worked out. That is, other than reduplication, there are no constant CV- prefixes or -VC suffixes to form words with. Second, supposing that the necessary word formation analysis were forthcoming, I find it extremely odd that only monosyllables that end in a nasal can be the first
member of a compound. That is, only CVN+CVC compounds occur, and no CVC_{1}+CVC_{2} (where CVC_{1}≠CVC_{2}). One certainly could not claim that a tendency to avoid clusters was involved in this restriction since CVC_{2} forms permit a wide range of clusters and CVC_{2} forms may be regarded as a special case of monosyllabic root compounding. Besides this unnatural restriction on what can be the first element of a compound (CVN) there would be an additional restriction on the second element of the compound: the initial of the second monosyllable must be a stop. Since neither of these restrictions is apparently motivated, I think this hypothesis must be regarded as weak at best.

Third, note that some of Dempwolff's examples are reduplicated monosyllables. Examples like tantan are not problematic because the homorganic cluster is the result of the coincidence of the points of articulation of the initial and final consonants. However, examples like pampang are problematic. Recall that the general rule is for the nasal to remain unassimilated in reduplicated monosyllables. While ND. is one of the few languages that does exhibit an assimilation, this is not the case for PAN or the daughter languages in general.

On the basis of these three objections I conclude that monosyllabic root compounding cannot account for the origins of the medial homorganic nasal clusters of PAN,
except, of course, for certain obvious cases like *ganggang.

There are also problems with Dempwolff's *-m- analysis. Concerning this hypothetical infix Dempwolff argued (1922-3:179):

Ein Formelement, das als m oder als m mit einem vorangehenden Vokal, meist als um, in vielen austronesischen Sprachen als Infix nachgewiesen ist, lässt sich im ND. als -um- in vereinzellen Fällen aufspüren.

But, concerning the function of the infix Dempwolff (1922-23:180) only said 'Auf die Grundbedeutung dieses austronesischen Formelemente, -um-, -m-, kann hier nicht einge­gangen werden.' By considering the infix to be essentially consonantal—that is, -m- --yet connected with -um-, Dempwolff attempts to have the best of two worlds. On the one hand, he can skirt the problem of introducing a new canonical shape for infixes by connecting consonantal -m- with the more regular -um-. Presumably, there would also be a connection in morphological function, though Dempwolff does not mention this. On the other hand, Dempwolff can also skirt the problem of syncope mentioned earlier (see below 5.5 as well). However, his success is only apparent: Dempwolff still must motivate the relationship between -m- and -um-.
Austronesian languages generally do not exhibit -m- as a variant of -um-. Dempwolff noted (1922-3:180):

But, the connection being made here is of -m- with ma- or ma- and Pränasalierung, not -um-. While some languages exhibit a similarity in function of ma- and -um-, I fail to see the significance of this similarity for the -um/-m- relationship. Moreover, when -um- is prefixed, it usually appears as um- or m- and never as ma-. Dempwolff proposed to connect -um- with ma- through the following argument (1922-23:180): 'Es liegt die Vermutung nahe, dass -um- durch Umstellung von *mu- entstanden ist, und dass *mu- eine Variante zu ma- ist'. However, in light of the facts presented above concerning the behavior of -um- as a prefix, Dempwolff's suppositions must be rejected. There is simply no evidence supporting a relationship between a hypothetical mu- and the prefix ma-.

5.5 Arguments Against a -VN- Shape. In the last section I argued against Dempwolff's assignment of the hypothetical infix to *-um-. Recall that I mentioned earlier that there were several problems connected with positing a -VN- infix
to account for the PAN nasal clusters. I suggested that the main problem concerned the necessary change from $C_1-V_1N_1-V_1C_2V_2C_3$ to $C_1V_1/V_1N_1C_2V_2C_3$. -VC- infixes in PAN are infixed before the first vowel of the root. There is no evidence indicating that pré-AN infixes behaved differently. Thus, it is safe to assume that the CVNCVC forms (for example) of PAN would derive from something like $C_1-V_1N_1-V_1C_2V_2C_3$. If a syncope of $V_1$ occurred, the resulting canon would be $C_1V_1N_1C_2V_2C_3$, with the nasal of the infix adjacent to the medial consonant of the wordbase. If $N_1$ then assimilated to the point of articulation of $C_2$, CVNCVC would have resulted. It is the purpose of this section to evaluate this hypothesis and others which posit a -VN- shape for the pre-AN infix. I will argue that the weight of the evidence is against positing such an infix.

The set of changes $C_1-V_1N_1-V_1C_2V_2C_3 > C_1-V_1N_1-C_2V_2C_3 > C_1V_1NCV_2C$ is, perhaps, the first set of changes that comes to mind if we adopt a -VN- infix. However, there are at least two reasons for suspecting that these changes never in fact occurred. First, many Austronesian languages show a marked tendency for penultimate stress. While we know very little about the accentual system of PAN, it has been suggested that PAN words were stressed on the penultimate syllable (see for example, Brandstetter 1916, Essays II and IV). If penultimate stress was the rule, then the set
of changes under discussion would involve the syncope of a stressed vowel, namely \( V^1 \). The syncope of a stressed vowel is phonetically implausible.

But even stress aside, notice that after the syncope of \( V^1 \), \( V_1 \) is left as the vowel of the initial syllable. This trace should be fairly easy to spot. That is to say, the initial vowel of CVNCVC forms should be identical, or at least show a marked statistical disproportion. Table 4 (p. 168) lists representative CVNCVC forms. Notice that any of the four vowels reconstructed for PAN can occur as the initial vowel of the wordbase. Moreover, C. Douglas Chrétien's study of the statistical structure of PAN wordbases (1965) does not indicate any disproportion in the frequencies of any of the vowels in this position.

One might attempt to account for this lack of a constant initial vowel in CVNCVC forms by claiming that the vowel of the infix participated in some morphophonemic alternation or underwent subsequent linguistic changes. However, unless these changes were unconditioned, this alternative hypothesis lacks support. Chrétien's study does not indicate any positive or negative correlations between the initial vowels of CVNCVC forms and the other consonants and vowels of the forms that would suggest that there was any sort of conditioning of this vowel by consonants or vowels. In light of this, the only other
alternative is to claim that there was conditioning but that subsequent linguistic changes have been so severe as to have obscured the original relationship. This alternative is, however, clearly ad hoc. Its only purpose seems to be to preserve the necessary syncope in the face of a complete lack of evidence.

TABLE 4. Exemplary PAN CVNCVC Forms

where $V_1 = *i$

*bintih to hit with the shin bone
*in[Di]u mother
*ti(ng)ka to move rhythmically

where $V_1 = *u$

*umbang carried by the water, float
*puNTuk elevation
*dungkug to bend

where $V_1 = *a$

*an'd'ang to be long
*ha(n')t'ah grind, polish
*ta(ng)kuL rake with the hand

where $V_1 = *e$

*enggan rhinoceros bird
*keNTung bird-clapper
*t'embah reverence, respect
As an alternative to the syncope of $V_1$, one might argue for the deletion of $V_1$ and a subsequent metathesis of $N_1$ and $V_1$. That is, $C_1-V_1N_1-V_1C_2V_2C_3 > C_1N_1V_1C_2V_2C_3 > C_1V_1N_1C_2V_2C_3$. The motivation for the metathesis would presumably be the avoidance of the $C_1N_1$ cluster. But, the syncope of $V_1$ could be blocked to begin with and the cluster avoided. This is a weak objection since it is not uncommon for a sound change to introduce a sequence that later undergoes further change, as the metathesis hypothesis suggests. However, the following sequence of changes has been documented for many languages: $CN > Cn > nC$ (see for example, Herbert 1977a). Since PAN permitted prenasalized stops in initial position, albeit in conjunction with Pränasalierung, I think it likely that this set of changes would have occurred had there been a syncope of $V_1$. But of course this set of changes will not result in $N_1$ being adjacent to $C_2$. It should be noted that the set of changes $C-V_1N_1-X-CN_1-X-N_1X$ ('X' is any permissible phonological string) is well documented for PAN (see for example Wolff, 1973). Thus, from the prefix $ma$- we derive $na$- through infixation with -in-: $m-in-a > mna- > na$-. Since this set of changes involves the syncope of the vowel of an infix and results in a sequence, $C_1N_1$, analogous to the sequence to which the metathesis should apply (but clearly does not apply) I submit that there is very little
support for the metathesis hypothesis.

David Stampe (personal communication) has suggested another hypothesis based on his experience with Munda languages.\(^\text{10}\) He suggests that the shape of the infix could have been \(-V_iN_i-\) with \(V_i\) assimilating to \(V_1\). If \(V_i\) copies the qualities of \(V_1\), and \(V_1\) is subsequently lost, there would not be any trace. That is, \(C_1-V_iN_i-V_1C_2V_2C_3 \rightarrow C_1-V_1N_i-V_1C_2V_2C_3 \rightarrow C_1-V_1N_i-C_2V_2C_3\). While Stampe's analysis does avoid the problem of the diverse qualities of \(V_1 < V_i\), the deleted vowel will still be the stressed vowel. This rather weak objection is supplemented by another, stronger argument: while there is comparative evidence in Munda languages for this vowel copy analysis, no such comparative evidence exists for Austronesian linguistics. In the absence of such evidence, we can only consider Stampe's analysis to be suggestive.

At this point we would do well to question the reasons for positing a \(-VN-\) infix, rather than just \(-N-\). After all, if a \(-N-\) infix were to be affixed after the first vowel of the root, the nasal would be adjacent to \(C_2\) at the outset. The only linguistic change that would be necessary would be for the nasal to assume the point of articulation of \(C_2\). The only apparent drawback to this analysis is that it represents a new canon, namely \(-N-\).
It is not clear, however, just what kind of an argument one can adduce against establishing a new canon. It would not be correct to argue that since PAN had no -N-infices, pre-AN did not have them either. Pre-AN could have used consonantal infixes extensively and lost them through regular linguistic changes. Indeed, the hypothesis that pre-AN $C_1V_1-N_1-C_2V_2C_3$ became PAN CVNCVC makes a suggestion in that direction. There is also some independent evidence for the new canon. In Dempwolff (1925), both the word for 'canoe paddle' and the word for 'bundle' were reconstructed with a medial consonant cluster---*bert'ay 'canoe paddle', *beRket 'bundle'---namely -RC-. Blust (1971:13) has recently drawn our attention to this fact. He notes:

This segment was later omitted from reconstructions, however, even when it was reflected in all of the languages compared, as in Tg. bi-g-kis, TB. bo-r-hos, Ml. be-r-kes < *bekes 'bundle'.

Blust considers the preconsonantal *R to be an infix, as Dempwolff (1925) apparently did. Blust then argues:

...the assumption of an infix -ng- brings out the parallelism between prenasalization and preconsonantal *R, and allows a more general statement to be made about Proto-Austronesian canonical shape than is otherwise possible: Proto-Austronesian permitted no consonant clusters within a morpheme, but allowed the preconsonantal infixation with *R and *ng.
In other words, Blust uses the independently attested pre-consonantal *R to argue for a -N- shape (see below for Blust's assignment to *-ng-).

In terms of economy (here, the number of discrete morphological canons it is necessary to reconstruct for PAN), establishing an -N- infix does not alter the number of canons; PAN CVNCVC would be derived and one would then be trading PAN CVNCVC for pre-AN -N-. The ultimate decision for a trade such as this would rest on the further predictions made by each hypothesis. That is, if one hypothesis yields insights that the other hypothesis does not, the one which advances our knowledge of AN structure the most would be the more highly valued hypothesis.

In section 5.7 I will consider this question in detail. I will argue that establishing a *-N- infix allows one to make a specific and plausible connection between the PAN medial nasal clusters and the Pränasalierung of initials. This connection cannot be made under any other hypothesis. Consequently, the additional insights offered by the -N- analysis make it the more highly valued hypothesis.

At this point in the investigation there are five competing hypotheses concerning the origins of PAN medial nasal clusters. These are presented in summary form below. (2)-(5) are all different versions of the Nasal Infix Hypothesis.
1.) The CVNCVC canon is not derived synchronically or diachronically. It is original. PAN is a nasal cluster language. (NCL)

2.) The CVNCVC canon is derived from a pre-AN *-V1N1- and the syncope of V1 of the baseword. That is, C1-V1N1-V1C2V2C3 > C1V1N1C2V2C3. (SYN)

3.) The CVNCVC canon is derived from a pre-AN *-V1N1- and the syncope of V1 with a consequent metathesis of N1 and V1. That is, C1-V1N1-V1C2V2C3 > C1N1V1C2V2C3 > C1V1N1C2V2C3. (MET)

4.) The CVNCVC canon is derived from a pre-AN *-V1N1- and the assimilation of V1 to V1 and subsequent deletion of V1. That is, C1-V1N1-V1C2V2C3 > C1-V1N1-V1C2V2C3 > C1-V1N1-C2V2C3. (VCO)

5.) The CVNCVC canon is derived from a pre-AN *-N- infix that occurred after the first vowel of the wordbase.

In light of the arguments presented above, hypotheses (1) and (5) are the least problematic. (2) and (4) entail the deletion of a stressed vowel. SYN also fails to account for the different qualities of V1 < V1 and VCO lacks comparative evidence. (3) involves an unlikely metathesis. Following the Principle of Economy discussed in 2.3 and 4.3 (1) and (5) are the more highly valued hypotheses. Most of the literature since Dempwolff assumes that the shape of the infix was as in (5). The discussion in Dahl's Proto-Austronesian (1973) is limited entirely to an -N- infix, but even this is just assumed, and not argued for. Dahl does not even consider the possibility of a -VN- infix (see below).
5.6 Gonda's Argument and *ng-. In the last section I discussed different hypotheses concerning the shape of nasal infix and suggested that \(-N-\) (as opposed to \(-VN-\)) seemed the more plausible of the two shapes. In the present section I will further refine the \(-N-\) hypothesis and argue that \(-N\) should be identified with \(*-ng-\). Before turning to that discussion, however, I would like to briefly consider a general objection to any treatment of the nasal clusters as deriving from an infix and a CVCVC or CVCVCCV canon. J. Gonda (1942, 1943) has suggested that it is not correct to consider the nasal clusters as the historic result of a CVCVC baseword and a hypothetical infix. His argument (see especially 1943) is essentially methodological. Gonda notes that affixes always have some defineable grammatical function. If the function of a hypothetical affix cannot be defined, Gonda argues there is no justification for considering it an affix at all. He also argues that there are other reasons for lengthening words. For example, he discusses what he calls lasvormen, forms that are lengthened without any apparent grammatical function associated with the lengthening, and argues (1942) that many languages—not just Austronesian languages—lengthen words with liquids and nasals to make them more intensive or emphatic. Gonda considered the nasals which appear in PAN nasal clusters to be of this nature.
There are several problems with Gonda's reasoning. First, while affixes most certainly have some sort of grammatical function, this does not mean that a historical linguist will always be able to reconstruct that function. Indeed, there are serious problems involved in reconstructing meaning in general. I do not take an inability to reconstruct the function of an historical affix as a serious objection.

Moreover, if the function of the nasal infix was 'intensity' or 'emphasis', it seems to me that Gonda has identified the necessary grammatical function (unless Gonda does not want to count these as legitimate grammatical functions (see Gonda 1952)). However, I think he would be hard pressed to support such a position. I would suggest that if there are sporadic lengthenings without apparent function, this lack of function is due to the analyst's inability to discern it, rather than to a genuine lack of function. In my experience languages do not lengthen words arbitrarily. It is possible that the nasal was ideophonic. However, it is not clear to me how one would go about demonstrating this. (See 5.8.)

Dahl (1973: Chapter 18) discusses Gonda's position at length and apparently adopts the lasvormen analysis. However, Dahl considers Pränasalierung to be the same type of process. He argues (1973:98):
As long as nobody has been able to assign a grammatical function to it, we have no right to call it an infix. The only thing we can establish is that the prenasalized forms have a nasal element in addition to other forms with similar or identical meaning.

And later (p. 99):

The only difference from the nasal accretion of the medial consonant is that in initial position the ng may be placed before any consonant, but the nasal substitution takes place only when the initial consonant is a voiceless stop or *b, i.e. part of those which may be prenasalized in medial position.

While Dahl does not provide any arguments to support his assignment of *-N- to the velar nasal, his assignment is no doubt due to the fact that it is the velar nasal which appears in initial position (Pränasalieru̱g) when no assimilation takes place. That is, the velar nasal appears before vowels and in most languages w, h, y, and ʔ. Moreover, Javenese unexpectedly shows -ngs- for *-n't'. Both Blust (1971) and Dahl (1973) account for this problematic reflex by arguing that the assimilation of *-ng- to the point of articulation of the following stop followed the change of PAN *t' to ʔ. Dahl (1973:100) argues:

But where the original stops and affricates remained as stops, the assimilation took place. Thereafter, the natural way to express an 'emphatic' nasal was in the form of a homorganic prenasalization of the following stop.
In an earlier section (5.3) I left open the possibility that the -NC- (as opposed to -(N)C-, nasaler Zuwachs) clusters might not be derived from the nasal infix. That is, the Nasal Infix Hypothesis might have only been concerned with -(N)C- forms where the -C-\-NC- alternation is evident. However, if Blust and Dahl are correct in their interpretation of the Javanese -ngs- reflex, then all PAN medial nasal clusters will have to be derived from *-ng-. Let us consider why this must be the case.

If only -(N)C- forms derive from a pre-AN disyllable or trisyllable and *-ng-, then Jv. -ngs- for Dempwolff's *-(n')t'-' will follow Blust and Dahl's explanation. What Dempwolff reconstructed as *-(n')t'-' should be -(ng)t'-' (and *t'>s yielding -ngs-). But, the forms that Dempwolff reconstructed without an optional nasal, that is *-n't'-', would presumably derive from pre-AN *-n't'-. That is to say, the nasal of these forms would never have been ng, but n'. Thus, the Javanese -ngs- reflex of PAN *-n't'-' would still be problematic. There is no reason to suppose that *-n't'-' > -ngs-. If one derives both -(N)C- and -NC- forms from pre-AN *-ng-, one can readily account for Jv. -ngs- from Dempwolff's *-n't'-' since this cluster too would derive from pre-AN *-ngt'-' (and *t' > s). For the moment then, I will accept Blust and Dahl's assignment of the infix to *-ng-, but turn to the question
again in later sections.

Having identified -N- as -ng-, it is now possible to describe the original nature of the relational pattern, that is, to specify the original nature of the pairings: CVCVC:CVngCVC for disyllables and CVCVCVC:CVngCVCVC for trisyllables. The only sound change that it is necessary to posit to derive PAN CVNCVC and CVNCVCVC is regressive nasal assimilation. The plausibility of such a change is beyond question. Nonetheless, the reason for the pattern loss is unclear and would not seem to be a direct consequence of the nasal assimilation. I will take up the issue of the pattern loss in 5.10. For the moment I would like to consider the relationship between Pränasalierung and the origin of the medial nasal clusters.

5.7 Pränasalierung and Medial Nasal Clusters. In 5.1 I briefly described the morphological process of Pränasalierung. Recall that Pränasalierung effects the initial consonants of wordbases: if the initial stop is voiceless, the initial is replaced by a nasal of the same point of articulation—for example, initial p appears as m, initial t as n, and so on. If the initial is voiced, the stop remains but is prenasalized or preceded by a homorganic nasal—for example, initial b appears as mb, initial d as nd and so on.
In this section I will argue that if one adopts the nasal infix hypothesis, it becomes possible to relate Pränasalierung to the origins of the medial nasal clusters; namely, both the prefixes which participate in Pränasalierung and the nasal clusters reflect an infix. Moreover, I will suggest that the prefixes provide support albeit indirectly, for identifying the infix as *-ng-.

Consider now the examples from Tagalog, Toba Batak, and Javanese Pränasalierung in Table 5 below (p. 180).

In addition to the regular alternations of Table 5, Tg. and TB. show some initial b's appearing as m—for example, Tg. bili 'to buy', namili 'have shopped' and TB bunu 'kill', mamunu 'kill someone'—and Javanese shows w appearing as m—for example, warah 'announcement', marah 'announce something'. Jv. w, however, is a regular reflex of PAN *b (see §5.1). Tagalog also exhibits some d's appearing as n, for example, dikit 'to strike', nanikit 'has struck'. Pränasalierung is fully productive in most West Austronesian languages. However, Dempwolff notes (1934:31) '...[Pränasalierung]...kommt aber auch zuweilen als erstarrte Lauterscheinung vor, z.B. in Jv. nipis-tipis 'thinness'.'
TABLE 5. Tagalog, Toba Batak, and Javanese Examples of Pränasalierung

<table>
<thead>
<tr>
<th>Voiceless Initials</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>p to m:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pili</td>
<td>'to choose'</td>
<td>namili</td>
<td>'has chosen' Tg.</td>
</tr>
<tr>
<td>pulung</td>
<td>'collected'</td>
<td>mamulung</td>
<td>'collect something' TB.</td>
</tr>
<tr>
<td>payung</td>
<td>'umbrella'</td>
<td>mayung</td>
<td>'cover someone with an umbrella' Jv.</td>
</tr>
<tr>
<td>t to n:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>takot</td>
<td>'fear'</td>
<td>nanakot</td>
<td>'have others 'fearful' Tg.</td>
</tr>
<tr>
<td>tambun</td>
<td>'mound'</td>
<td>manambun</td>
<td>'make a mound' TB.</td>
</tr>
<tr>
<td>T to N:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TiTi</td>
<td>'tick'</td>
<td>NiTi</td>
<td>'strike fire' Jv.</td>
</tr>
<tr>
<td>t' to n':</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t'abut</td>
<td>'expressions'</td>
<td>n'abut</td>
<td>'express something' Jv.</td>
</tr>
<tr>
<td>s to n:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sulat</td>
<td>'write'</td>
<td>panulat</td>
<td>'for writing= pen' Tg.</td>
</tr>
<tr>
<td>surat</td>
<td>'write'</td>
<td>manurat</td>
<td>'write something' TB.</td>
</tr>
<tr>
<td>s to n':</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>surung</td>
<td>'shove off'</td>
<td>n'urung</td>
<td>'shove something' Jv.</td>
</tr>
<tr>
<td>k to ng:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kita</td>
<td>'see'</td>
<td>nangita</td>
<td>'has sought' Tg.</td>
</tr>
<tr>
<td>kukup</td>
<td>'take in hand'</td>
<td>ngukup</td>
<td>'take something in hand' Jv.</td>
</tr>
</tbody>
</table>
TABLE 5 (continued)

Voiced Initials

<table>
<thead>
<tr>
<th>b to mb:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>balot</td>
<td>'wrapped up'</td>
<td>pambalot</td>
</tr>
<tr>
<td>butbut</td>
<td>'plucked out'</td>
<td>mambutbut</td>
</tr>
<tr>
<td>bubut</td>
<td>'pluck out'</td>
<td>mbubut</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>d to nd:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>dipa</td>
<td>'fathom'</td>
<td>nandipa</td>
</tr>
<tr>
<td>dolos</td>
<td>'glide'</td>
<td>mandolos</td>
</tr>
<tr>
<td>duga</td>
<td>'opinion'</td>
<td>nduga</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D to ND:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dukun</td>
<td>'indigenous doctor'</td>
<td>NDukun</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>d' to n'd':</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>d'arum</td>
<td>'needle'</td>
<td>man'd'arum</td>
</tr>
<tr>
<td>d'undung</td>
<td>'raise up high'</td>
<td>ndundung</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>g to ngg:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>gawai</td>
<td>'cast a spell'</td>
<td>nanggawai</td>
</tr>
<tr>
<td>gomgom</td>
<td>'to rule'</td>
<td>manggomgom</td>
</tr>
<tr>
<td>gantung</td>
<td>'hang up'</td>
<td>nanggantung</td>
</tr>
</tbody>
</table>
In most languages the process is connected with certain prefixes. Dempwolff, and other early scholars identified these as ma-, na- and pa- for example. Most modern treatments attach a nasal to these prefixes. Wolff (1973) for example, cites them as maN-, naN- and paN- where 'N' denotes the morphophonemic alternations the initial of the base participates in. In Tagalog N is most straightforwardly analyzed as ng since it is this segment which appears when nasal assimilation does not apply (see below).

While very little has been written concerning the form or function of *Pränasalierung* in the proto language, there is general agreement that (a) the process only affected initial stop consonants, and (b) that substitution (nasaler Ersatz) only occurred if the initial stop was voiceless. Recall that Dahl (1973:99) argued '...the nasal substitution takes place only when the initial consonant is a voiceless stop or *b....' It is not clear to me that *b* should be included with the sounds that undergo nasaler Ersatz. While Tagalog (and most other Philippine languages), Javanese and Toba Batak all indicate that at least some b's underwent nasaler Ersatz other languages suggest that only the voiceless initials were susceptible to the process. For example, in Dempwolff's Ngadju article (1922-23) he reports that b appears as mb. He does not cite a single example where b appears as m. Moreover,
the daughter languages do not agree in which words that begin with b undergo nasaler Ersatz. That is, since there is no agreement, the b-substitutions may very well be parallel developments.

According to Wolff (1973), m substituted for b and p in connection with the infix *-um-, which Wolff reconstructs as the 'non-past independent active verbal affix.' This infix was inserted before the first vowel of the root unless the root began with a bilabial. In this case, *p and *b were changed to *m (Wolff, 1973:72-4). Wolff claims that Mongondow preserves the rule intact and cites the following examples (1973:84).

-um- + baya? 'go' maya? 'will go'
-um- + posiyug 'sleep' mosiyug 'will sleep'
-um- + kosing 'laugh' kumosing 'will laugh'

He also suggests that Atayalic and Tsouic have preserved the process although he notes that there are no Tsouic examples which reflect *-um- added to a PAN *b- initial base.

That nasaler Ersatz from *-um- is distinct from nasaler Ersatz in Pränasalierung is evident from the Old Javanese forms. Wolff notes (1973:84):

In contemporary Jv. there is no way to distinguish reflexes of *-um- and reflexes of *maN- for bases which begin with /p/ or /w/ (from *b). In Old Jv., however, when -um- is added to bases with p- or w-, the morphophonemics are different from when maN- is added. With
the addition of *-um-, the productive pattern is for the /p/ or /w/ of the base to be dropped, eg. umangun 'be up' (formed by adding *um- to a base wangun). This is an innovation on the original pattern, which is still retained in forms of high frequency in Old Jv.: initial /w/ or /p/ is changed to /m/ and the initial consonant is dropped: e.g. Old Jv. maturu 'sleep' (formed by adding *um- to a base paturu, as exemplified by the imperative form in Old Jv.). These forms umangun and maturu contrast with forms like mamangun 'erect something' which consists of maa- added to a base wangun.

In these examples, *-um- apparently occurred as a prefix. Old Jv. maturu and umangun would then derive from *um-paturu and um-wangun. Wolff also argues that prefixes like *maR- are derived from underlying *-um- + *paR- in a similar manner. We are apparently faced then, with two distinct cases of nasaler Ersatz in connection with different affixes in the proto language.

R. Herbert (1977b) has suggested, however, that nasal substitution might very well be a secondary development. He begins his argument by noting that the phonetics of the process are quite peculiar. That is, more distant (phonetically) segments apparently totally assimilate (m+p+m) while more similar segments do not (m+b+mb). He argues that this violates a universal proposed by Hutcheson (1973) which claims that the total assimilation of less similar segments implies the total assimilation of like segments. Herbert then draws on his own work with
phonetic processes affecting nasals (1977a) and notes that the sequence mp is highly unstable. He argues that the most common process affecting obstruents following nasals is progressive voicing. However, if mp + mb, the resultant sequence would be easily confused with /b/ initial bases. Moreover, he argues that if the nasal were to regressively devoice before the following voiceless stop, mp + mp, the resultant sequence would be almost indistinguishable from simple voiceless initial (as opposed to prenasalized initials). Thus, if either voicing process applies, the result will be perceptual difficulty. Since neither process could affect an mb cluster, these sequences would remain unchanged. Herbert then suggests that there was only a nasal assimilation process in the proto language and that the perceptually motivated deletion of the voiceless stop was a West Austronesian innovation. The basis for this claim is the apparent lack of nasal substitution in Oceanic languages.

Herbert's analysis of nasaler Ersatz is essentially correct. There is no real substitution at all, but a perceptually motivated deletion. His work represents an important advance in our understanding of phonetic process in Austronesian linguistics. However, it is not clear that his claim that the deletion is a West Austronesian innovation is correct since there are Oceanic forms which
apparently exhibit the remnants of nasaler Ersatz, though not productively (George Grace; Andrew Pawley, Lawrence Reid, personal communications). Moreover, recent research in New Guinea has turned up languages which are Oceanic and which exhibit something quite similar to Pränasalierung productively. However, the details are not clear enough at present to be certain that there are in fact Oceanic retentions. Nonetheless, we can be certain that nasaler Ersatz is a development from an earlier system where there was only nasal assimilation. Thus, at some point in the history of Austronesian languages, mp, nt, etc. existed along side of mb, nd, etc.

Several scholars have implied that PAN medial nasal clusters and Pränasalierung have the same source historically. Dempwolff hinted (1922-23, 1934) at this and Dahl (1973:99) argues:

Among the values of a lengthened form mentioned by Gonda we found the emphatic, expressive, or intensifying character. If this was the original value of the insertion of ng before the medial consonant, there is a parallel to the same procedure before the initial consonant. Adriani regards this as an intensive form...

The only difference from the nasal accretion of the medial consonant is that in initial position the ng may be placed before any consonant...

However, the details of this connection have not been established.
It has already been pointed out that the Pränasalierung of initials often occurs in connection with certain prefixes, for example maN-. Let us reconsider a few of the examples cited in Table 5. Tg. sulat goes back to PAN *t'ulat (TB surat). pang- in Tagalog is a general instrumental prefix and pang- + sulat → panulat 'for writing, a pen'. Tg. dipa 'fathom' goes back to PAN *depa. nang- is the completed aspect form of mang-, a general verbal affix; nang- + dipa → nandipa. In PAN, panulat would presumably have been *pan't'ulat and nandipa, *nandepa. Taking these forms back even further (following Herbert's suggestions) we can reconstruct pre-AN *pan't'ulat and *nandepa.

It is important to note here that every prefix with a final N, that is, a nasal that participates in Pränasalierung, has a corresponding prefix with final *R. Even more importantly, there is always a corresponding prefix without a final consonant. Wolff (1973), for example, reconstructs the following triplets:

a) pa- paR- paN-
b) ma- maR- maN-
c) na- naR- naN-

pa-, ma-, and na- and likewise the other triplets are related to each other. Recall that Wolff suggests -um- + pa- → ma-, and I have pointed out that -in- + ma- → na-,
albeit by a different route. These prefixes are also found in larger prefixes, for example ipaq-, ipanag-. Finally, let us note that N appears as ng before vowels, /l/, /w/, /y/, and /h/ and elsewhere when assimilation does not take place (see Schachter and Otanes 1972 and 5.9 below).

It is quite interesting that the only two consonants that appear as the finals of these prefixes, and consequently appear preconsonantally in affixed forms, are exactly those two consonants that can appear preconsonantally wordbase internally; that is, *ng and *R. Moreover, ng and R in these affixed forms occur in positions that are analogous to the wordbase internal preconsonantal positions. That is, there is no difference in kind between the pairs of structures below.

\[
\begin{align*}
(1a) & \quad [CV-ng-CVX]_\text{wordbase} \\
(1b) & \quad [CV-R-CVX]_\text{wordbase} \\
(2a) & \quad [CV-ng-]_\text{wordbase} \quad [CVCVX]_\text{wordbase} \\
(2b) & \quad [CV-R-]_\text{wordbase} \quad [CVCVX]_\text{wordbase}
\end{align*}
\]

Let us consider *depa, and prefix this wordbase with *na-. The resultant form is *nadepa. If we now infix *-ng- or *-R-, (3) below will result.
It is clear that (3a-b) are simply instances of (2a-b) respectively. The suggestion here is that the final consonants of prefixes like *paN- and *paR- are the remnants of earlier infixes, the same infixes that give rise to medial clusters in PAN wordbases, namely, *-ng- and *-R-.

The only additional assumption one needs to make is that infixation followed prefixation. But this assumption is not at all unwarranted. Notice that na- is derived from m-in-a- (perhaps -um-p-in-a, if Wolff is correct) which implies that infixation follows prefixation. Indeed the ordering of infixation and prefixation is well attested in Austronesian languages. Under this analysis then, PAN *paR-, *paN-, etc. result from the morphological reanalysis of a CV prefix, an infix, and a baseword as in (4) below.

\[
(4) \quad CV_{pre} - \text{(ng)}_{R} - CV_{CVCX} > CV(\text{ng})_{R} - CV_{CVCX}
\]

The motivation for the proposed reanalysis would have been the loss of productivity of *-ng- (and presumably *-R-). I will argue below that the function of *-ng- became opaque and that wordbase internally it ceased to mark
any clear grammatical distinction. In conjunction with the CV- prefixes discussed above, the function of *-ng- became fused with the function of the prefix (see 5.10).

It should be noted that version 5 (see 5.5) of the Nasal Infix Hypothesis is the only version consistent with the proposed morphological reanalysis. If one assumes that PAN was a nasal cluster language (see 5.3, 5.5) there could not be a reanalysis since there would not be an infix. If the shape of the infix were -VN- (see 5.4-5.5), then paN, for example would have to derive from something p-VN-a. Such a sequence is analogous to *m-in-a, which develops into na- (see 5.5), that is a CV prefix and not a CVN- prefix.

Version 5, then, provides a principled account of the relationship between Pränasalierung and the medial nasal clusters. Both are instances of an early relational pattern of *-ng- infixation. The prefixes which participate in Pränasalierung are the result of a pattern change, pattern reanalysis. The medial clusters, however, are the result of a pattern loss. (See 4.1 for a discussion of these types of pattern change). In so far as this is the only version of the nasal infix hypothesis that can relate Pränasalierung and the medial clusters, it is to be considered the more highly valued on the basis of the Principle of Optimal Use (see 2.3). Hereafter I use
'Nasal Infix Hypothesis' to refer exclusively to version 5.

5.6 Determining the Function of *-ng-. The only aspect of the Nasal Infix Hypothesis that I have not considered in the preceding sections of this chapter concerns the function of *-ng- and I find it quite surprising that no systematic study has appeared to date which attempts to determine the function of the infix. The absence of such a study is particularly striking in light of the widespread acceptance of the hypothesis and Gonda's argument (5.6) that it is not proper to speak of an affix unless its grammatical function can be identified. It is the purpose of this section to provide a preliminary analysis of the function of *ng-. I will confine my analysis to the forms reconstructed with an optional medial nasal since it seems probably that such forms might still reflect the function of *ng-.

Recall that Dempwolff reconstructed an optional nasal when the daughter languages did not unambiguously indicate a medial nasal cluster. I will refer to the reflexes without a nasal as the 'oral grade' (OG). The reflexes that indicate a cluster will be referred to as the 'nasal grade' (NG). The 196 forms that Dempwolff reconstructed with the optional nasal are listed below in Table 6 (p. 193). The first column of the table lists Dempwolff's reconstruction. The second column contains Dempwolff's gloss for the.
reconstructed form. The third column lists the NG reflex
glosses and the final column, the oral grade reflex glosses.
The general strategy here is to compare the 'general sense'
of the NG with the 'general sense' of the OG as indicated
by the glosses of the reflexes of each grade. The goal, of
course, is to ascertain any systematic relationship that
might obtain between the OG and NG. Since I am interested
in the 'general sense' of the NG and OG I have not indicated
the language particular references for individual glosses
unless a language exhibited both an OG and an NG. In such
cases the language reference is indicated in order to show
that such doublets are not limited to one or two languages,
but found throughout Austronesia.

As an example of how I adapt Dempwolff's entries to
fit the format of the table, consider the first item of
Table 6, *ambaj. Dempwolff's entry for this form is listed
below (section references deleted).

'to move to and fro'...TB. m/abeabe 'to wave
to and fro', ambi 'to wave with the arms';
Jv. awe 'to wink', Ml. ambai 'basket to catch
crabs', ND. awei 'temporary wall=screen'.

The first two columns of the table for this entry are
straightforward. Under the nasal grade column I list: '(TB)
to wave with the arms'. For the oral grade I list: '(TB)
to wave to and fro'. I have omitted the Ngadju Dayak and
Malay reflexes from comparison since the semantics are
questionable.
<table>
<thead>
<tr>
<th>Gloss</th>
<th>NG</th>
<th>OG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ambaj</td>
<td>to move to and fro</td>
<td>(TB) to wave with the arms</td>
</tr>
<tr>
<td>2. ambat</td>
<td>to hinder, impede</td>
<td>(TB) to hinder; to stand</td>
</tr>
<tr>
<td></td>
<td></td>
<td>in the way; speech defect</td>
</tr>
<tr>
<td>3. andav</td>
<td>sun, day</td>
<td>(ND) day; day; daily</td>
</tr>
<tr>
<td>4. ang'g'i</td>
<td>tribe, relatives (young)</td>
<td>brother, sister</td>
</tr>
<tr>
<td>5. angken</td>
<td>appropriate, acquire, (the appropriated = mine)</td>
<td>mine; (Jv) acknowledge as mine (with prefix)</td>
</tr>
<tr>
<td>6. angkup</td>
<td>hold together</td>
<td>(TB) hindrance; to stop; check</td>
</tr>
<tr>
<td>7. ang'k'am</td>
<td>plan, intention</td>
<td>wish</td>
</tr>
<tr>
<td>8. antah</td>
<td>to be raw, unripe</td>
<td>to eat raw meat</td>
</tr>
<tr>
<td>9. antat'</td>
<td>to be up</td>
<td>(Tg) rank</td>
</tr>
<tr>
<td>10. bengkah</td>
<td>discord, dissension</td>
<td>(Tg) commission; to be up; obvious; on</td>
</tr>
<tr>
<td>11. benteng</td>
<td>abdomen</td>
<td>wary; to be wavering</td>
</tr>
<tr>
<td>12. bengkung</td>
<td>arch, vault</td>
<td>belly; to be pregnant</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ankle bond</td>
</tr>
<tr>
<td>13. bentit'</td>
<td>shin bone</td>
<td>shin bone; calf</td>
</tr>
<tr>
<td>14. bindang</td>
<td>to be broad</td>
<td>cloth; breadth; to be broad</td>
</tr>
<tr>
<td></td>
<td></td>
<td>spread</td>
</tr>
<tr>
<td>15. binting</td>
<td>bulwark; trench</td>
<td>(Jv) bulwark</td>
</tr>
<tr>
<td>16. bungkal</td>
<td>to turn up the ground; agitate</td>
<td>(TB) destroy a village</td>
</tr>
<tr>
<td></td>
<td>Gloss</td>
<td>NG</td>
</tr>
<tr>
<td>---</td>
<td>------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>17.</td>
<td>bungkul</td>
<td>bulb</td>
</tr>
<tr>
<td>18.</td>
<td>buntun I</td>
<td>growing boil</td>
</tr>
<tr>
<td>19.</td>
<td>buntun II</td>
<td>name of a tree</td>
</tr>
<tr>
<td>20.</td>
<td>bun't'u</td>
<td>youngest child</td>
</tr>
<tr>
<td>21.</td>
<td>dandang</td>
<td>to heat</td>
</tr>
<tr>
<td>22.</td>
<td>DampaR</td>
<td>to be flat</td>
</tr>
<tr>
<td>23.</td>
<td>Dampi</td>
<td>protruding</td>
</tr>
<tr>
<td>24.</td>
<td>d'ambaj</td>
<td>to hang down</td>
</tr>
<tr>
<td>25.</td>
<td>d'ambat</td>
<td>to take hold of; grab</td>
</tr>
<tr>
<td>26.</td>
<td>d'[ae]mbut</td>
<td>hair</td>
</tr>
<tr>
<td>27.</td>
<td>d'engkal</td>
<td>span</td>
</tr>
<tr>
<td>28.</td>
<td>d'ungkit</td>
<td>incise</td>
</tr>
<tr>
<td>29.</td>
<td>d'ungkung</td>
<td>small water craft; canoe</td>
</tr>
<tr>
<td>30.</td>
<td>d'umput</td>
<td>to take with the fingers</td>
</tr>
<tr>
<td>31.</td>
<td>embun</td>
<td>rainfall</td>
</tr>
<tr>
<td>32.</td>
<td>empat</td>
<td>four</td>
</tr>
<tr>
<td>33.</td>
<td>empu</td>
<td>ancestor; grandchild</td>
</tr>
<tr>
<td>34.</td>
<td>entut</td>
<td>flatus ventris</td>
</tr>
<tr>
<td>35.</td>
<td>entik</td>
<td>to be small</td>
</tr>
<tr>
<td>36.</td>
<td>gambal</td>
<td>portrait</td>
</tr>
<tr>
<td>Gloss</td>
<td>NG</td>
<td>OG</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>37. gan'd'il</td>
<td>to be strange odd number; to be strange; troublesome; difficult to compute</td>
<td>incorrect pronunciation</td>
</tr>
<tr>
<td>38. gampit</td>
<td>to hold together pinch</td>
<td>pliers</td>
</tr>
<tr>
<td>39. gan't'ang</td>
<td>inflame easily inflamed</td>
<td>easily inflamed, burnt</td>
</tr>
<tr>
<td>40. gan't'a</td>
<td>odd number odd number, five odd number</td>
<td>to be fragile; full of cracks; broken through; gap; break</td>
</tr>
<tr>
<td>41. gentut'</td>
<td>to break broken off, rent</td>
<td></td>
</tr>
<tr>
<td>42. gung'k'i</td>
<td>glazed clay pot clay jar</td>
<td>glazed clay pot</td>
</tr>
<tr>
<td>43. Rambung</td>
<td>atmospheric turbidity dusk, twilight; hail</td>
<td>heaven; sky</td>
</tr>
<tr>
<td>44. Rangkit</td>
<td>tie together; raft</td>
<td>horse drawn vehicle; binding; raft</td>
</tr>
<tr>
<td>45. Rampat'</td>
<td>to tear off torn away, off; harvested; to rob</td>
<td>harvest</td>
</tr>
<tr>
<td>46. Rempuh</td>
<td>to pound; reduce to powder to crush</td>
<td>powdered; to reduce to powder tension</td>
</tr>
<tr>
<td>47. Ren[t]eng</td>
<td>to stretch; span; tighten to be tightened powder</td>
<td></td>
</tr>
<tr>
<td>48. Rimbaj</td>
<td>to hang woven bast curtain</td>
<td>suspended strap; to hang to be high</td>
</tr>
<tr>
<td>49. hambav</td>
<td>to be high; up; on top to surface; drawn high; to be high</td>
<td>to scatter wooden puppets magically with rice to be high</td>
</tr>
<tr>
<td>50. hambuR</td>
<td>to strew; spread to scatter; sprinkle</td>
<td></td>
</tr>
<tr>
<td>51. handang</td>
<td>to prepare an attack lie in wait</td>
<td>lie in wait; await; watch for front</td>
</tr>
<tr>
<td>52. handep</td>
<td>facade; front forehead</td>
<td>sovereign; excellent; admirable; prince</td>
</tr>
<tr>
<td>53. handi</td>
<td>ruler; sovereign ruler</td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Word(s)</td>
<td>Gloss</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>54.</td>
<td>hang'g'av</td>
<td>day, sun</td>
</tr>
<tr>
<td>55.</td>
<td>hang'g'i</td>
<td>tribe; relatives (of other family)</td>
</tr>
<tr>
<td>56.</td>
<td>hampit</td>
<td>hold together</td>
</tr>
<tr>
<td>57.</td>
<td>hanted</td>
<td>to convey; accompany</td>
</tr>
<tr>
<td>58.</td>
<td>han't'ah</td>
<td>grind; polish</td>
</tr>
<tr>
<td>59.</td>
<td>han't'ang</td>
<td>gill of fish</td>
</tr>
<tr>
<td>60.</td>
<td>hentah</td>
<td>to be raw</td>
</tr>
<tr>
<td>61.</td>
<td>hindam</td>
<td>loan</td>
</tr>
<tr>
<td>62.</td>
<td>hiNTaj</td>
<td>look at</td>
</tr>
<tr>
<td>63.</td>
<td>huNDang</td>
<td>crustaceans</td>
</tr>
<tr>
<td>64.</td>
<td>indah</td>
<td>to be beautiful</td>
</tr>
<tr>
<td>65.</td>
<td>in'd'uk</td>
<td>to step; dance</td>
</tr>
<tr>
<td>66.</td>
<td>impaL</td>
<td>beyond; across</td>
</tr>
<tr>
<td>67.</td>
<td>impi</td>
<td>to dream</td>
</tr>
<tr>
<td>68.</td>
<td>impun</td>
<td>to collect</td>
</tr>
<tr>
<td>69.</td>
<td>intem</td>
<td>to be black</td>
</tr>
<tr>
<td>70.</td>
<td>in't'ang</td>
<td>jaw</td>
</tr>
<tr>
<td>71.</td>
<td>kambang</td>
<td>inflate; blossom; unfold</td>
</tr>
<tr>
<td>72.</td>
<td>kaNDung</td>
<td>womb</td>
</tr>
<tr>
<td>73.</td>
<td>kang'k'ang</td>
<td>beans</td>
</tr>
<tr>
<td>74.</td>
<td>kang'k'ing</td>
<td>means for fastening</td>
</tr>
<tr>
<td>75.</td>
<td>kampak</td>
<td>to slam (door)</td>
</tr>
<tr>
<td>76.</td>
<td>kampit</td>
<td>hold together</td>
</tr>
<tr>
<td>Gloss</td>
<td>NG</td>
<td>OG</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>KembaL</td>
<td>twin</td>
<td>(Jv) twin; neighbor</td>
</tr>
<tr>
<td>kembung</td>
<td>inflate</td>
<td>big box; distended;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>inflated</td>
</tr>
<tr>
<td>keng'k'ing</td>
<td>smell of urine</td>
<td>urinate; cadaverous</td>
</tr>
<tr>
<td></td>
<td></td>
<td>smell of urine</td>
</tr>
<tr>
<td>kempit</td>
<td>to hold together;</td>
<td>to carry under the arm;</td>
</tr>
<tr>
<td></td>
<td>cramp</td>
<td>(ML) cramp</td>
</tr>
<tr>
<td>kentut</td>
<td>flatus ventris</td>
<td>flatus ventris</td>
</tr>
<tr>
<td>keNTeL</td>
<td>to tremble, vibrate</td>
<td>(TB) tremble; (Jv)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>stretch</td>
</tr>
<tr>
<td>kiNDat</td>
<td>open the eye</td>
<td>blink; wink</td>
</tr>
<tr>
<td>kin'd'ak</td>
<td>to step; dance</td>
<td>dance</td>
</tr>
<tr>
<td>king'k'u</td>
<td>deception; deceit</td>
<td>swindle; impostor</td>
</tr>
<tr>
<td>k'anggap</td>
<td>skim; scoop off</td>
<td>draw water</td>
</tr>
<tr>
<td>k'angkup</td>
<td>close with a snap;</td>
<td>to chew</td>
</tr>
<tr>
<td></td>
<td>snap at</td>
<td></td>
</tr>
<tr>
<td>k'ampuL</td>
<td>to mix</td>
<td>mixed</td>
</tr>
<tr>
<td>k'iNDuk</td>
<td>ladle; scoop out</td>
<td>(Tg) small pot; pick up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>morsel</td>
</tr>
<tr>
<td>k'impi[I]</td>
<td>flat dish; platter</td>
<td>(TB) flat plate</td>
</tr>
<tr>
<td>k'ungkil</td>
<td>poke</td>
<td>(Jv) engraving tool;</td>
</tr>
<tr>
<td>k'ungku[L]</td>
<td>to shear; clip</td>
<td>jack</td>
</tr>
<tr>
<td>lambat'</td>
<td>to be open</td>
<td>(TB) to be wide</td>
</tr>
<tr>
<td>lang'k'ung</td>
<td>falsify</td>
<td>falsify</td>
</tr>
<tr>
<td>lampit'</td>
<td>layer</td>
<td>layer</td>
</tr>
<tr>
<td>lan't'at</td>
<td>name of tree, fruit</td>
<td>Lansium Domesticus</td>
</tr>
<tr>
<td>lembung</td>
<td>ditch, grave</td>
<td>wheel rut</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gloss</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>-------</td>
</tr>
<tr>
<td>98.</td>
<td>lengkung</td>
<td>to be concave</td>
</tr>
<tr>
<td>99.</td>
<td>lempit</td>
<td>to fold</td>
</tr>
<tr>
<td>100.</td>
<td>liNDung</td>
<td>protection; refuse; shelter</td>
</tr>
<tr>
<td>101.</td>
<td>lingkEL</td>
<td>to wring; revolve</td>
</tr>
<tr>
<td>102.</td>
<td>lingkung</td>
<td>to be concave</td>
</tr>
<tr>
<td>103.</td>
<td>lingkup</td>
<td>encircle</td>
</tr>
<tr>
<td>104.</td>
<td>ling'k'ad</td>
<td>to slip</td>
</tr>
<tr>
<td>105.</td>
<td>liman[t]ek</td>
<td>small leech</td>
</tr>
<tr>
<td>106.</td>
<td>limpung</td>
<td>javelin</td>
</tr>
<tr>
<td>107.</td>
<td>lungka</td>
<td>to injure; hurt; wound</td>
</tr>
<tr>
<td>108.</td>
<td>lungkat</td>
<td>remain open</td>
</tr>
<tr>
<td>109.</td>
<td>lun[t]uh</td>
<td>to cook</td>
</tr>
<tr>
<td>110.</td>
<td>[l]amba</td>
<td>wilderness</td>
</tr>
<tr>
<td>111.</td>
<td>Lambak</td>
<td>extend</td>
</tr>
<tr>
<td>112.</td>
<td>[L]ambut</td>
<td>to be hairy</td>
</tr>
<tr>
<td>113.</td>
<td>Lambut'</td>
<td>snatch away</td>
</tr>
<tr>
<td>114.</td>
<td>Lampak</td>
<td>crumble to pieces</td>
</tr>
<tr>
<td>115.</td>
<td>Langgat'</td>
<td>defoliated</td>
</tr>
<tr>
<td>116.</td>
<td>Limpun</td>
<td>gather; collect</td>
</tr>
<tr>
<td>Gloss</td>
<td>NG</td>
<td>OG</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>117. Lungkub</td>
<td>protection shaded; (MI) protecting hanging over something</td>
<td>protection; (MI) same as protection; (MI) same as</td>
</tr>
<tr>
<td>118. mun't'uh</td>
<td>enemy</td>
<td>enemy</td>
</tr>
<tr>
<td>119. pandang</td>
<td>grassy plain</td>
<td>grassy plain</td>
</tr>
<tr>
<td>120. paNDung</td>
<td>head piece</td>
<td>(Tg) ornament of the head</td>
</tr>
<tr>
<td>121. pangku</td>
<td>fondle; hug</td>
<td>(ND) to take in one's lap</td>
</tr>
<tr>
<td>122. pang'k'ang</td>
<td>piece of wood</td>
<td>post; pole</td>
</tr>
<tr>
<td>123. pantik</td>
<td>to be pointed</td>
<td>invaded; (Jv) nail</td>
</tr>
<tr>
<td>124. pantuk</td>
<td>to pick</td>
<td>to nail</td>
</tr>
<tr>
<td>125. peNDem</td>
<td>close the eyes</td>
<td>(TB) grave; (Jv) to bury</td>
</tr>
<tr>
<td>126. pundu1</td>
<td>blunt, dull</td>
<td>blunted</td>
</tr>
<tr>
<td>127. puNDuk</td>
<td>hut</td>
<td>hut, dwelling</td>
</tr>
<tr>
<td>128. pungkul</td>
<td>to throw; beat</td>
<td>beaten</td>
</tr>
<tr>
<td>129. pung'k'ak</td>
<td>summit, top</td>
<td>offshoot; beginning and end; (Jv) top</td>
</tr>
<tr>
<td>130. pung'k'u</td>
<td>top, peak</td>
<td>top; point; (Jv) corner</td>
</tr>
<tr>
<td>131. puntung</td>
<td>break off; demolish</td>
<td>(MI) remnant; broken; (ND) cut off</td>
</tr>
<tr>
<td>132. pun't'u</td>
<td>termite hill</td>
<td>termite hill</td>
</tr>
<tr>
<td>133. tambaRI</td>
<td>without savour</td>
<td>(Jv) insipid; insipid; sweet water</td>
</tr>
<tr>
<td>134. tambaRII</td>
<td>antidote; remedy</td>
<td>(TB) remedy; recover</td>
</tr>
<tr>
<td>135. tambaRIII</td>
<td>to proffer</td>
<td>wedding present</td>
</tr>
<tr>
<td>136. [t]ambat</td>
<td>to fasten</td>
<td>fastened (prefix)</td>
</tr>
<tr>
<td>137. tambeng</td>
<td>to be veiled; wrapped up</td>
<td>mine (redup.); to be veiled (suffix)</td>
</tr>
<tr>
<td>138. tambIL</td>
<td>rim; flat receptacle</td>
<td>brim; flat bowl; platter; annex; flat basket</td>
</tr>
<tr>
<td>Gloss</td>
<td>NG</td>
<td>OG</td>
</tr>
<tr>
<td>-------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>tambuh</td>
<td>drumstick</td>
<td>pipe</td>
</tr>
<tr>
<td>tambun</td>
<td>heap</td>
<td>mound; heap; (ML dam; (ND) heaped</td>
</tr>
<tr>
<td>[t]ambuni</td>
<td>afterbirth</td>
<td>afterbirth</td>
</tr>
<tr>
<td>tan'Da</td>
<td>sign; symbol</td>
<td>(Tg) seal; stamp; symbol</td>
</tr>
<tr>
<td>tan'd'ak</td>
<td>name of boat rope</td>
<td>name of boat rope</td>
</tr>
<tr>
<td>ten'd'uk</td>
<td>projecting</td>
<td>to stumble</td>
</tr>
<tr>
<td>tanggal</td>
<td>loosen; detach</td>
<td>detached; (ML) detached</td>
</tr>
<tr>
<td>tanggih</td>
<td>remind someone of debts</td>
<td>same as Gloss</td>
</tr>
<tr>
<td>tangkav</td>
<td>theft</td>
<td>theft</td>
</tr>
<tr>
<td>tangkit'</td>
<td>defense</td>
<td>defense; (ND) to parry</td>
</tr>
<tr>
<td>tangkub</td>
<td>cover over with; cover; lid</td>
<td>agree; fittingly close (Jv) fit, be suited</td>
</tr>
<tr>
<td>tampi</td>
<td>remove dust and chaff</td>
<td>winnow; to pound rice</td>
</tr>
<tr>
<td>[t]embak</td>
<td>clear a woods</td>
<td>clear a woods</td>
</tr>
<tr>
<td>tembing</td>
<td>edge; bank; shore</td>
<td>edge; border</td>
</tr>
<tr>
<td>tembuk</td>
<td>perforate</td>
<td>perforated; (ML) perforated</td>
</tr>
<tr>
<td>tenduh</td>
<td>calm</td>
<td>calm</td>
</tr>
<tr>
<td>tempa</td>
<td>form</td>
<td>(TB) form</td>
</tr>
<tr>
<td>[t]empap</td>
<td>hand's breadth; span</td>
<td>span</td>
</tr>
<tr>
<td>tentu</td>
<td>certainly</td>
<td>certainly</td>
</tr>
<tr>
<td>[t]imbav</td>
<td>height; to be high</td>
<td>(Tg) height; elevated</td>
</tr>
<tr>
<td></td>
<td>Gloss</td>
<td>NG</td>
</tr>
<tr>
<td>---</td>
<td>-------</td>
<td>----</td>
</tr>
<tr>
<td>159.</td>
<td>[t]in'd'ak</td>
<td>tread; step</td>
</tr>
<tr>
<td>160.</td>
<td>tingkah</td>
<td>move rhythmically</td>
</tr>
<tr>
<td>161.</td>
<td>[t]impun</td>
<td>collect; gather</td>
</tr>
<tr>
<td>162.</td>
<td>tumbuh</td>
<td>thrive; grow</td>
</tr>
<tr>
<td>163.</td>
<td>tunduh</td>
<td>refer to</td>
</tr>
<tr>
<td>164.</td>
<td>tungkang</td>
<td>artisan</td>
</tr>
<tr>
<td>165.</td>
<td>[t]ungked</td>
<td>support; stick</td>
</tr>
<tr>
<td>166.</td>
<td>[t]ungkeh</td>
<td>piece; tie; point</td>
</tr>
<tr>
<td>167.</td>
<td>tungkup</td>
<td>to cover</td>
</tr>
<tr>
<td>168.</td>
<td>TangkuL</td>
<td>rake with the hand</td>
</tr>
<tr>
<td>169.</td>
<td>TungkuL</td>
<td>germinate</td>
</tr>
<tr>
<td>170.</td>
<td>t'ambuR</td>
<td>to strew</td>
</tr>
<tr>
<td>171.</td>
<td>t'aNDeR</td>
<td>lean against</td>
</tr>
<tr>
<td>172.</td>
<td>t'anggu</td>
<td>plant pulp; sago</td>
</tr>
<tr>
<td>173.</td>
<td>t'alumbung</td>
<td>covering; veil</td>
</tr>
</tbody>
</table>

201
<table>
<thead>
<tr>
<th>Table 6 (continued)</th>
<th>Gloss</th>
<th>NG</th>
<th>OG</th>
</tr>
</thead>
<tbody>
<tr>
<td>174. t'embuh</td>
<td>saturated</td>
<td>satisfied; recover;</td>
<td>bite; mouthful</td>
</tr>
<tr>
<td>175. t'empah</td>
<td>chew; masticate</td>
<td>(ND) chew betel</td>
<td>(ND) betel bite; chewed</td>
</tr>
<tr>
<td>176. t'empit</td>
<td>to be tight; narrow</td>
<td>narrow</td>
<td>narrow; pliers</td>
</tr>
<tr>
<td>177. t'entagi</td>
<td>band</td>
<td>broad strap; string;</td>
<td>same as NG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>collar</td>
<td></td>
</tr>
<tr>
<td>178. t'entul</td>
<td>name of a tree</td>
<td>Sandricum Indicum; name</td>
<td>same as NG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>of a tree</td>
<td></td>
</tr>
<tr>
<td>179. t'INDil</td>
<td>ajar; oblique</td>
<td>ajar</td>
<td>oblique; slanting</td>
</tr>
<tr>
<td>180. t'ingkap</td>
<td>grasp</td>
<td>grasp</td>
<td>name of bird of prey</td>
</tr>
<tr>
<td>181. t'ingkep</td>
<td>to be finished</td>
<td>sufficient</td>
<td>to be equipped</td>
</tr>
<tr>
<td>182. t'impang</td>
<td>side road; cross</td>
<td>step aside; side road;</td>
<td>way to next world (infix)</td>
</tr>
<tr>
<td></td>
<td>road</td>
<td>road</td>
<td></td>
</tr>
<tr>
<td>183. t'umbing</td>
<td>jagged; notched</td>
<td>fragment; jagged</td>
<td>jagged</td>
</tr>
<tr>
<td>184. t'undu</td>
<td>spoon</td>
<td>spoon</td>
<td>spoon</td>
</tr>
<tr>
<td>185. t'ungkaL</td>
<td>toil, labor</td>
<td>difficult, hard</td>
<td>toil</td>
</tr>
<tr>
<td>186. t'ung'k'i</td>
<td>to be clean</td>
<td>wash clothes</td>
<td>clean</td>
</tr>
<tr>
<td>187. t'umpit</td>
<td>to be narrow; blowpipe</td>
<td>blowpipe; narrow; boring</td>
<td>pliers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hole</td>
<td></td>
</tr>
<tr>
<td>188. umbu[dg']</td>
<td>leaf bud; heart of the palm</td>
<td>heart of the palm</td>
<td>heart of the palm</td>
</tr>
<tr>
<td>189. undu[L]</td>
<td>go to the rear</td>
<td>(TB) go together; return; recede</td>
<td>(TB) in ranks one after another</td>
</tr>
<tr>
<td>190. unda[Dang]</td>
<td>crustacean</td>
<td>crab and other crustaceans</td>
<td>crustaceans</td>
</tr>
<tr>
<td>191. umpak</td>
<td>bark</td>
<td>waste wood; (ND) gold plated</td>
<td>(ND) bark; bark</td>
</tr>
<tr>
<td>Gloss</td>
<td>NG</td>
<td>OG</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>---------------------------</td>
<td></td>
</tr>
<tr>
<td>192. untah</td>
<td>to vomit</td>
<td>to vomit</td>
<td></td>
</tr>
<tr>
<td>193. un[t]aj</td>
<td>name of a fruit</td>
<td>orange</td>
<td></td>
</tr>
<tr>
<td>194. untek</td>
<td>brain</td>
<td>brain</td>
<td></td>
</tr>
<tr>
<td>195. untung</td>
<td>gain; profit</td>
<td>(TB) gain; gain; lottery</td>
<td></td>
</tr>
<tr>
<td>196. untut'</td>
<td>commission; order</td>
<td>ambassador</td>
<td></td>
</tr>
</tbody>
</table>

|                          |             | (TB) wealthy through      |
|                          |             |fortunate circumstances    |
|                          |             |commissioned; sent         |
I have also recorded the relative frequency of each kind of cluster, that is, for example, \(-nt-\) as opposed to \(-ngk-\). These figures are reported in Table 7 (p. 205). The relative frequency figure is the percent of the total \((N)C\) clusters. These figures have been checked against the relative frequency figures cited by Chrétien (1965). The figures cited in Table 7 seem to be consistent with his findings. Thus, the clusters do not exhibit any statistical oddities, but seem to pattern in the same way that other PAN segments pattern.12

Of the 196 forms of Table 6, 72 may be considered to exhibit essentially synonymous reflexes in both the nasal and oral grades. Figure 6 below lists the Table 6 items by number (and language reference where relevant) that I consider to be synonymous in the OG and NG.


Figure 6. Synonymous OG and NG Forms
<table>
<thead>
<tr>
<th>Cluster Type</th>
<th>Fre</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>mp</td>
<td>33</td>
<td>16.8</td>
</tr>
<tr>
<td>mb</td>
<td>37</td>
<td>18.8</td>
</tr>
<tr>
<td>nt</td>
<td>28</td>
<td>14.2</td>
</tr>
<tr>
<td>nd</td>
<td>14</td>
<td>7.1</td>
</tr>
<tr>
<td>n't'</td>
<td>10</td>
<td>5.1</td>
</tr>
<tr>
<td>n'd'</td>
<td>05</td>
<td>2.5</td>
</tr>
<tr>
<td>NT</td>
<td>02</td>
<td>1.0</td>
</tr>
<tr>
<td>ND</td>
<td>12</td>
<td>6.1</td>
</tr>
<tr>
<td>ng'k'</td>
<td>13</td>
<td>6.6</td>
</tr>
<tr>
<td>ng'g'</td>
<td>03</td>
<td>1.5</td>
</tr>
<tr>
<td>ngk</td>
<td>34</td>
<td>17.3</td>
</tr>
<tr>
<td>ngg</td>
<td>05</td>
<td>2.5</td>
</tr>
</tbody>
</table>
It is interesting to note that while certain authors have claimed that the function of *-ng- was emphasis (see for example, Dahl 1973, Chapter 18; Gongda 1942, 1943) the items in Table 6 do not support this claim. Under my interpretation of the forms, only three items seem to exhibit a OG-NG relationship that I would characterize as 'emphasis': (41) *gentat', OG 'to be fragile, full of cracks, broken through, gap, break', NG 'broken off, rent'; (65) *in'd'uk OG 'to step', NG 'to skip, trample down'; and (69) *intem OG 'black', NG 'very black'. There may be a few more forms that someone else may interpret as exhibiting an 'emphasis' relation. However, unless one defines 'emphasis' in such a way as to make it meaningless, there are not enough forms in Table 6 to support the claim that the function of *-ng- was 'emphasis'.

A careful inspection of Table 6 indicates that there are a large number of synonymous forms. Slightly over 33 percent of the PAN -(N)C- forms show no change of meaning between the oral grade and the nasal grade. However, if *-ng- was a productive affix in pre-AN, it is not inconceivable that the original distinction was lost in this many forms by PAN since the time depth is so great. I suggested earlier that whatever the function of *-ng- was in pre-AN times, it had lost its productivity by PAN. Indeed, I think you could use this 33 percent figure to argue that the infix was no longer productive.
Certain forms in Table 6 stand out prominently since they seem to exhibit a derivational relationship between the OG gloss and the NG gloss. Consider, for example, (2) \*ambat, OG ' (TB) hindrance', NG ' (TB) to hinder'; (46) \*Rempuh, OG 'powdered', NG 'to pound, to reduce to powder'; and (101) \*lungka, OG 'injured, wound', NG 'to pierce'. These forms suggest that OG nouns and adjectives can be made into verbs by the insertion of \*-ng-. There is also some independent—albeit indirect—evidence for such a relationship. Hamp (1974) has argued that it is necessary to reconstruct Proto Oceanic \*susu 'breast', but \*nsusu 'to suck' in order to account for certain problematic reflexes in Fijian. The POC forms would derive from PAN \*\text{t'ut'}\text{u} and \*ng-\text{t'ut'}\text{u} and we would have an instance of the same derivational relationship in initial position under Pränasalierung. I will discuss the significance of the link between medial clusters and Pränasalierung for determining the function of \*-ng- below (5.9). Suffice it to say here that there are very few cases in Table 6 where the relationship is as clear as it is in the few forms cited above.

A significant number of forms in Table 6 exhibit a reverse, derived verb relationship. That is, the OG can be interpreted as 'that which has the property of x' (where 'x' is the NG gloss) or 'that which x-es' or 'that
which does \( x' \). Consider, for example, (35) *entik, OG \\
'(Jv) to make small', NG '(Jv) little chap'; (44) *Rangkit, \\
OG 'horsedrawn vehicle, binding, raft', NG 'raft', and \\
(77) *kembaL, OG '(Jv) double', NG '(Jv) twin'. Figure 7 \\
lists the other forms in Table 6 that I think fit into \\
this same general scheme.

2, 3, 8, 10, 11, 12, 14, 21, 22, 23, 25, 35, 37, 38, 44, \\
45, 46, 47, 48, 52, 60, 66, 72, 75, 77, 79, 80, 82, 84, \\
87, 89, 90, 94, 95, 98, 101, 103, 104, 107, 111, 112, \\
114, 115, 120, 121, 123, 124, 129, 130, 131, 134, 135, \\
136, 138, 144, 151, 153, 155, 162, 163, 169, 170, 171, \\
175, 176, 179, 180, 185, 191, 195, 196.

Figure 7. 'That which \( x-es \)' OG-NG Relation.

Perhaps related to this relationship are the forms \\
of Figure 8. The forms of this figure seem to exhibit a 
relationship where the NG is a special case of the OG. 
Notice that there is some overlap between Figure 7 and 
Figure 8. The semantics of 'special case of' and 'that 
which \( x-es/ \) does \( x' \)' are very close. It may very well be 
that both of these relationships are manifestations of a 
common, more basic relationship.
4, 9, 10, 11, 16, 17, 18, 23, 26, 28, 36, 39, 41, 42, 43, 50, 52, 65, 83, 85, 86, 90, 91, 94, 97, 103, 104, 113, 114, 117, 119, 120, 125, 134, 137, 138, 145, 148, 152, 158, 166, 172, 175, 182, 186, 189, 190, 191.

Figure 8. 'Special Case' OG-NG Relation

Figure 7 has 71 entries (36.2%) and Figure 8, 48 entries (24.4%). If the two figures can be put together then more than 50 percent of the PAN-(N)C- entries exhibit some sort of systematic relationship between the OG and the NG.

Ideally it would always be the NG that was the derived case, with the OG interpreted as basic. However, in a sense we have defined Figure 7 as if the NG was basic. I will turn to this issue shortly. Even so, the NG is derived a significant number of times (see Figure 9). Likewise, the OG is derived almost as often as it is basic in Figure 8. That is, there are many cases where the OG has to be interpreted as a special case of the NG, and not the other way around. Figures 9 and 10 list the directionality of the relationships of Figures 7 and 8, respectively. It is noteworthy that for some of the forms it is almost impossible to determine which form is basic and which form is derived. Consider, for example, (38) *gampit, OG 'pliers' and NG 'pinches'. This can be
A. Oral Grade Basic: 14, 35, 44, 48, 60, 66, 77, 84,
90, 112, 114, 115, 129, 131, 144, 153, 163, 169, 170,
179, 191, 196.
B. Nasal Grade basic: 2, 3, 8, 10, 11, 12, 21, 22, 23,
25, 37, 38, 45, 46, 47, 50, 52, 72, 75, 77, 79, 80,
82, 87, 94, 95, 98, 101, 103, 104, 107, 111, 120, 121
123, 124, 130, 135, 136, 138, 151, 155, 162, 166,
175, 176, 180, 185, 195.

Figure 9. Directionality of Figure 7 OG-NG Relation.

interpreted either as (a) 'pinching as a property of pliers'
in which case the OG is basic and the NG derived, or (b)
as 'pliers pinch' in which case the OG is derived and the
NG basic.

A. Oral Grade basic: 9, 10, 16, 28, 36, 39, 41, 42, 43,
53, 65, 83, 85, 86, 90, 91, 97, 114, 117, 125, 137,
148, 166, 172, 175, 186.
B. Nasal Grade basic: 4, 11, 17, 18, 23, 26, 50, 94,
103, 104, 113, 119, 120, 134, 138, 145, 152, 158, 182,
189, 190, 191.

Figure 10. Directionality of Figure 8 OG-NG Relation.
The consequence of interpreting the OG-NG relationship as we have in Figure 7 is that the NG appears to be basic more often than the OG. There are two alternatives to this problem. On the one hand we can accept the directionality figures of Figure 9 and assume that the majority of forms in the nasal grade have crossed categories—that is, ceased to be distinct from the oral grade grammatically. On the other hand we could assume that we have characterized the OG-NG relationship incorrectly and reinterpret the relation in a way that is more consistent with the directionality figures of Figure 9. If the OG reflexes represent something like 'that which x-ea' or 'that which does x' where the NG is 'x', then the OG apparently represents something that is essentially nominal in function. The NG, along the same lines, is essentially verbal. Consider, (38) *gampit, OG 'pliers', NG 'pinch', (46) *Rempuh, OG 'powder, to reduce to powder', NG 'to crush', (94) *langkung, OG 'bracelet made from false gold', NG 'to falsify', (121) *pangku, OG '(ND) lap', NG '(ND) to take in one's lap', (151) *[t]embak, OG 'cutter instrument', NG 'to clear a woods', and (180) *tingkap OG 'name of bird of prey', NG 'to grasp'. I submit, then, that *-ng- was a derивational affix used to derive verbs from nouns.
Along side of this relationship we have the 'special case' relationship of Figure 8, where the NG represents a special case of the OG. As we mentioned earlier, it may be the case that both of these relationships—Figure 7 and Figure 8—are different manifestations of a more basic, as yet uncharacterized, relationship. The connection between the two relationships is suggested by the overlap between Figures 7 and 8.

The methodology used above to determine the morphological function of *-ng- suffers from a few drawbacks. First, by using 'general senses,' inferred from the glosses of the OG and NG reflexes, the study became subject to the investigator's judgment and bias. It is hoped that these were kept sufficiently under control so as to not invalidate the findings presented above. Second, 'general senses' are tantamount to reconstructed meanings. In 4.3 I argued that there are general problems associated with semantic reconstruction. These problems have not been solved, and the results presented here must be regarded as tentative at best. Finally, the daughters taken collectively might not reflect the original pre-AN state of affairs as clearly as particular languages taken individually may. That is, there may be a particular daughter that has preserved the original OG-NG relationship enough to give us a detailed picture of the original situation. I turn to this question in the next section.
5.9 Evidence from Ngadju Dayak and Tagalog. There are no studies of individual languages that address the issue of the function of *-ng-. However, Dempwolff's Ngadju study provides a ready source of material.

In this section I will briefly consider the material from Dempwolff's article in order to identify any relevant semantic material. I will also discuss the Tagalog prefixes mang- and nang- in juxtaposition to na- and na- in order to isolate the function of *-ng-. Recall that in 5.7 I argued that prefixes like nan- derive from a CV- prefix, *na- and *-ng-. The analysis presented here is not intended to be conclusive, but merely suggestive of fruitful directions in future research. The forms of Table 8 point to the same relationship discussed in Section 5.8. At the very least, the ND forms are consistent with the findings presented there. In many cases the OG and NG glosses are identical. In others, for example tepe and tempe, the NG functions as a derived verb. In still others, for example, upak and umpak, takis and tangkis, the NG can be interpreted as a special case of the OG.

Unfortunately, the ND data does not offer any new insights into the functions of *-ng-. For the moment we will have to be content with the available material. Ideally, there would be several studies to draw from and compare. This work must remain, however, for future studies.
TABLE 8. Ngadju Dayak OG-NG Pairs

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>mepan</td>
<td>thump, thud (prefix me-, root pan)</td>
</tr>
<tr>
<td>mempan</td>
<td>heaviness, shapelessness</td>
</tr>
<tr>
<td>tepe</td>
<td>to be crushed</td>
</tr>
<tr>
<td>tempe</td>
<td>to crush rice</td>
</tr>
<tr>
<td>d'upang</td>
<td>oscillation</td>
</tr>
<tr>
<td>d'umpang</td>
<td>oscillation</td>
</tr>
<tr>
<td>upak</td>
<td>rind, bark of a tree</td>
</tr>
<tr>
<td>umpak</td>
<td>engraving</td>
</tr>
<tr>
<td>atoh</td>
<td>reflection, consideration</td>
</tr>
<tr>
<td>match</td>
<td>reflect on, ponder (prefix ma-)</td>
</tr>
<tr>
<td>mantoh</td>
<td>to be virtuous, well-behaved</td>
</tr>
<tr>
<td>satolen</td>
<td>not properly cooked (suffix -en)</td>
</tr>
<tr>
<td>santolen</td>
<td>not properly cooked</td>
</tr>
<tr>
<td>asoh</td>
<td>running off (of water)</td>
</tr>
<tr>
<td>masohan</td>
<td>to float off (prefix ma-, suffix -an)</td>
</tr>
<tr>
<td>mansohan</td>
<td>to float off</td>
</tr>
<tr>
<td>pasohan</td>
<td>something that floats off often (prefix pa-, suffix -an)</td>
</tr>
<tr>
<td>pansohan</td>
<td>something that floats off often</td>
</tr>
<tr>
<td>bako</td>
<td>trickling, dripping</td>
</tr>
<tr>
<td>bangko</td>
<td>trickling, dripping</td>
</tr>
<tr>
<td>takis</td>
<td>drive off, beat off</td>
</tr>
<tr>
<td>tangkis</td>
<td>to parry, ward off</td>
</tr>
<tr>
<td>potong</td>
<td>an abbreviation, shortening</td>
</tr>
<tr>
<td>pontong</td>
<td>part, portion</td>
</tr>
<tr>
<td>gasik</td>
<td>gambling; money game</td>
</tr>
<tr>
<td>gan't'ik</td>
<td>gambling; money game</td>
</tr>
<tr>
<td>kasin</td>
<td>bolt, rail, bar</td>
</tr>
<tr>
<td>kan't'in</td>
<td>bolt, rail, bar</td>
</tr>
<tr>
<td>tikan</td>
<td>bird bone</td>
</tr>
<tr>
<td>tingkan</td>
<td>bird bone</td>
</tr>
</tbody>
</table>
TABLE 8. (continued)

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>halibak</td>
<td>to be smooth</td>
</tr>
<tr>
<td>halimbak</td>
<td>to be smooth</td>
</tr>
<tr>
<td>bua</td>
<td>fruit</td>
</tr>
<tr>
<td>hambua</td>
<td>to look for fruit (prefix ha-)</td>
</tr>
<tr>
<td>habua</td>
<td>to carry fruit (prefix ha-)</td>
</tr>
<tr>
<td>d'akit</td>
<td>encroach upon, spread; infection (illness)</td>
</tr>
<tr>
<td>had'akit</td>
<td>to be contagious (prefix ha-)</td>
</tr>
<tr>
<td>han'd'akit</td>
<td>to climb across, over</td>
</tr>
<tr>
<td>d'ad'egor</td>
<td>to have a long neck</td>
</tr>
<tr>
<td>d'ad'enggor</td>
<td>to have a long neck</td>
</tr>
<tr>
<td>lugai</td>
<td>to be worthless</td>
</tr>
<tr>
<td>lunggai</td>
<td>to be worthless</td>
</tr>
<tr>
<td>gagulon</td>
<td>boiling, rage, foam up</td>
</tr>
<tr>
<td>gangulon</td>
<td>boiling, rage, foam up</td>
</tr>
<tr>
<td>d'abak</td>
<td>a trap for small animals</td>
</tr>
<tr>
<td>d'ambak</td>
<td>a trap for small animals</td>
</tr>
<tr>
<td>kadut</td>
<td>bundle, parcel</td>
</tr>
<tr>
<td>kandut</td>
<td>bundle, parcel</td>
</tr>
<tr>
<td>sagada</td>
<td>weathervane</td>
</tr>
<tr>
<td>sangada</td>
<td>weathervane</td>
</tr>
</tbody>
</table>
In Section 5.7 I argued that both PAN medial nasal and the prefixes associated with the Pränasalierung of initials have their source historically in *-ng-. This being the case, a new strategy for determining the function of *-ng- suggests itself. If the nasal that occurs as the finals of such prefixes, as *paN-, *maN- is the *-ng-infix, it should be possible to isolate its function through a semantic analysis of these prefixes. That is, if one compares the functions of pairs like ma- and maN-, we should be able to 'factor out' ma- and ascertain the function of *-ng-.

As might be expected there are no studies that address themselves to this task. By way of a preliminary study, I have looked into Tagalog which exhibits full triplets, for example, pa-, ma- and na- beside paN-, maN-, and naN-. Of these six prefixes, all are associated with verbal derivation or inflection except paN-. According to Schachter and Otanes (1972) paN- is used to form reservational and instrumental adjectives. Bloomfield (1917) glosses it: 'the thing used for doing so and so or such and such.' Consider the following examples: simba 'church', pansimba 'for church (e.g. clothing)'; dalawa 'two', pandalawa 'for two (e.g. a table for two)'. paN- does occur in the verbal paradigm in connection with the suffixes -in 'direct passive (object focus)' and -an 'local
passive (locative focus). In both cases the \textit{paN-}...-	extit{-in/-an} corresponds to the \textit{maN-} actives\textsuperscript{13} (see below).

The prefix \textit{ma-} is used to form direct passives. According to Bloomfield (1917:283) \textit{ma-} in this sense denotes 'the object directly affected by an action which an actor is able to perform.' \textit{na-} represents the completed aspect form of \textit{ma-}. It is formed by infixing \textit{-in-} 'completed aspect' into \textit{ma-}; \textit{m-in-a-} + \textit{mn-}a- + \textit{na-}. \textit{ma-} is also a stative prefix denoting (Bloomfield; 1917:288) 'that which by nature or circumstance possesses such and such, usually quality.'

The prefix \textit{maN-} expresses (Bloomfield, 1917:239):

...action more deliberate, selective, or in larger quantity (professional, habitual) than \textit{mag-}; like the latter, but more clearly, it is used for the making or using of such and such an object.

Bloomfield also notes, however, that there are a few forms that express a static meaning. Considering the following examples, all with reduplication: \textit{isda} 'fish' \textit{mangingisda} 'fisherman', \textit{sulat} 'write' \textit{manunulat} 'writer', \textit{putol} 'cut', \textit{mamumutol} 'cutter (e.g. wood-cutter)'.\textsuperscript{14}

\textit{pa-} is a causative. Bloomfield notes (1917:299):

'The root with prefix \textit{pa-} denotes an object or person ordered, caused, or allowed to undergo such and such an action.' This function seems to be totally unrelated to \textit{paN-}. However, Bloomfield also notes that \textit{pa-} (most
certainly a distinct prefix) (1917:299) '...as conjunctive attribute...may express manner....' And Schachter and Otanes note (1972:217) that pa- plus base denotes the 'equivalent to adverb of manner expressions in English.' Consider, for example, bulong 'whisper' bumulong 'actor focus, to whisper', pabulong 'in a whisper' and ngiti 'smile', ngumiti 'actor focus, to smile', pangiti 'smilingly'.

Table 9 (p. 219) summarizes the functions of ma-, maN-, pa- and paN-. This table is very difficult to interpret. On the one hand, it is possible to see ma-stative recur in maN- reduplications. This is encouraging. On the other hand, however, I am unable to isolate a constant function for *-ng-, even with the benefit of the results of the preceding section. At the present, I do not see any way to resolve this unexpected and rather puzzling state of affairs. Perhaps when more is known about PAN morphology, a solution will be discovered. (See also my remarks in 5.10.)
TABLE 9. Tagalog pa-, ma-, paN-, and maN-

pa- causative
pa- conjunctive attribute expresses manner adverb of manner
paN- reservational and instrumental adjectives
pa-...-an (local passive) that caused to undergo an action
paN-...-an (local passive) corresponds to actives with maN-
pa-...-in (direct passive) person ordered to do so and so
paN-...-in (local passive) corresponds to actives with maN-

ma- direct passive
ma- actor-focus; person able to do so and so
ma- stative
maN- deliberate action (professional, habitual)
ma-...-an (local passive)
ma-...-in person given to doing so and so
maN-...-in deliberate action
5.10 Some Problems. Up to this point in the present chapter I have considered the morphophonological evidence for the Nasal Infix Hypothesis, various versions of this hypothesis, and broached the difficult problem of determining the function of *-ng-. I have argued that the morphophonological evidence indicates that there was a pre-AN relational pattern of infixation. The relational pattern has been characterized as the systematic pairing of wordbases of the shape CVCVC with CV-ng-CVC (CVCVC::CVngCVC) and CVCVCVC with CV-ng-CVCVC (CVCVCVC::CVngCVCVC). I have argued that the function of this infix was to derive verbs from nouns. The infixed forms result in PAN CVNCVC and CVNCVCVC. I have also argued that PAN Pränasalierung is triggered by the same infix, *-ng-, which has come to be reanalyzed as the final segment of certain prefixes, for example, PAN *maN-. In short, it has been possible to establish most of the details of the Nasal Infix Hypothesis. Nonetheless, there are still certain problems that must be cleared up. In this section I will consider the loss of productivity of *-ng-, the different reflexes of *-ng- in medial position and in connection with Pränasalierung, and finally, some general problems associated with apparently excrescent nasals in Austronesian languages.
Throughout this chapter I have suggested that *-ng-* lost its function very early in the development of Austronesian languages. I have three reasons for making this suggestion. First, no Austronesian language is known that preserves *-ng-* infixation intact. However, to the extent that the histories of individual languages are known in any detail, the remnants of infixation have been inherited from PAN. Second, at least in medial position in PAN, there is no evidence that infixation with *-ng-* was productive. Indeed, there is evidence that it was not in the large number of PAN forms—that is, -(N)C- forms—whose OG and NG reflexes are identical in meaning (see 5.8). I do not think that this widespread synonymy of OG and NG forms is consistent with the claim that infixation of *-ng-* was productive in PAN times. Rather, if *-ng-* lost its function before PAN, one could expect to find many synonymous forms in addition to those few archaic forms which point to the original function. In short, I would argue that the results presented in Figures 6-9 are consistent only with a hypothesis that *-ng-* had long ceased to mark any clear grammatical distinction. Finally, I have been able to characterize the function of *-ng-* in only the vaguest of terms. If infixation was productive in PAN times, I think a clearer picture of the function of *-ng-* would emerge from the data.
If in fact *-ng- had ceased to be productive by PAN times as I suggest, one would do well to ask why this might have happened. The key to the answer to this question lies in a problem I noted earlier when discussing the directionality of the OG-NG relationship (5.8). Recall, that in many cases it was difficult if not impossible to determine whether the OG or the NG was the basic form and the other derived. Most Austronesian languages defy description of lexical items in terms of strict category; that is noun, verb, adjective, and so on. If the syntactic reconstructions that have been proposed to date for PAN are correct (see for example Wolff 1973, 1975; Reid and Pawley 1975), then this categorization problem penetrates deeply into the history of Austronesian languages.

I suggest that this lack of strict categorization, in conjunction with the proposed function of *-ng-, a verbalizer of some sort, would lead to a rather difficult situation. That is to say, the function of *-ng- became opaque as a consequence of the ability of wordbases to function as nouns, verbs, adjectives and the like quite freely. *-ng- became fossilized in most forms—hence the predominance of PAN CVNCVC and CV(N)CVC. It may very well be that this set of changes is the consequence of a more pervasive syntactic change.\(^{15}\) However, this issue is beyond the scope of the present paper and will have to await future research.
There is a second problem with the proposed Nasal Infix Hypothesis in the reflexes of Pränasalierung and the medial nasal clusters.

Recall that Pränasalierung take two distinct forms—nasaler Ersatz and nasaler Zuwachs—depending on whether a wordbase begins with a voiced or voiceless initial stop. A summary of the Tagalog, Toba Batak and Javenese examples cited earlier (Table 5) is presented below in Table 10.

<table>
<thead>
<tr>
<th>Initial</th>
<th>Tagalog</th>
<th>Toba Batak</th>
<th>Javanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>p *p</td>
<td>m (mp)</td>
<td>m (mp)</td>
<td>m (mp)</td>
</tr>
<tr>
<td>t *t</td>
<td>n (nt)</td>
<td>n (nt)</td>
<td>n (nt)</td>
</tr>
<tr>
<td>T *T</td>
<td>n</td>
<td>T (NT)</td>
<td></td>
</tr>
<tr>
<td>t' *k'</td>
<td>n</td>
<td>n</td>
<td>n' (n't')</td>
</tr>
<tr>
<td>s *t'</td>
<td>n</td>
<td>n</td>
<td>n' (n't')</td>
</tr>
<tr>
<td>k *k</td>
<td>ng (ngk)</td>
<td>ng (ngk)</td>
<td>ng (ngk)</td>
</tr>
<tr>
<td>b *b</td>
<td>mb</td>
<td>mb</td>
<td>mb</td>
</tr>
<tr>
<td>d *d</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>D *D</td>
<td>nd</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>d' *d'</td>
<td>nd</td>
<td>n'd'</td>
<td>n'd'</td>
</tr>
<tr>
<td>g *g</td>
<td>ngg</td>
<td>ngg</td>
<td>ngg</td>
</tr>
</tbody>
</table>

Recall that I argued that nasal substitution (nasaler Ersatz) was not really a substitution at all, but a perceptually conditioned deletion. The clusters in parentheses represent Pränasalierung of voiceless initials before the deletion rule applied. Compare the reflexes
of Table 10 with the reflexes of medial clusters presented below in Table 11.

TABLE 11. Tagalog, Toba Batak, and Javanese Reflexes of PAN Medial Nasal Clusters

<table>
<thead>
<tr>
<th></th>
<th>Tagalog</th>
<th>Toba Batak</th>
<th>Javanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>*mp</td>
<td>mp</td>
<td>pp</td>
<td>mp</td>
</tr>
<tr>
<td>*nt</td>
<td>nt</td>
<td>tt</td>
<td>tt</td>
</tr>
<tr>
<td>*NT</td>
<td>nt</td>
<td>tt</td>
<td>NT</td>
</tr>
<tr>
<td>*n't'</td>
<td>ns</td>
<td>ts</td>
<td>ngs</td>
</tr>
<tr>
<td>*ng'k'</td>
<td>ns</td>
<td>ts</td>
<td>n't'</td>
</tr>
<tr>
<td>*ngk</td>
<td>ngk</td>
<td>kk</td>
<td>ngk</td>
</tr>
<tr>
<td>*mb</td>
<td>mb</td>
<td>mb</td>
<td>mb</td>
</tr>
<tr>
<td>*nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>*ND</td>
<td>nd</td>
<td>nd</td>
<td>ND</td>
</tr>
<tr>
<td>*n'd'</td>
<td>nd</td>
<td>n'd'</td>
<td>n'd'</td>
</tr>
<tr>
<td>*ngg</td>
<td>ngg</td>
<td>ngg</td>
<td>ngg</td>
</tr>
</tbody>
</table>

Tagalog and Toba Batak have merged *T and *t to t and *k' and *t' to s. In Javanese, *T is retained and *k' > t'. In all three languages *t' > s, but see below. Tagalog merges *D, *d' and *d to d. Toba Batak merges *D and *d to d. Javanese retains *D, *d' and *d in tact.

It is clear at a glance that Javanese n' (Table 10) for initial s is problematic. Recall that in order to explain the ngs reflex (Table 11) Blust and Dahl argued that *t' > s before ng assimilated to the point of articulation of the stop, *t'. Thus, -ngt'- > -ngs- and -n't'- that Dempwolff established never actually existed. In
order to explain Javanese n' under Pränasalierung, it will be necessary to assume that nasal assimilation applied before *t' > s, that is, while *t' retained a palatal articulation point. We are thus presented with an apparent relative chronology paradox.

I am unaware of any obvious solution to this problem. The only explanation that I can offer is quite speculative. At the earliest stages in the development of Pränasalierung and the medial clusters, both n' and ngs would have had the shape -ngt'. If we consider where these clusters would have occurred within a PAN word, and assume that the Austronesian tendency for penultimate stress was an aspect of its earliest history, then (1) and (2) below would obtain.

1.) Medial Cluster: CVngt'VC
2.) Pränasalierung: CVngt'VCVC

Perhaps, the assimilation change occurred first in unstressed syllables and only later came to affect the nasals of stressed syllables. Thus, (2) became CVn't'VCVC while (1) remained CVngt'VC, the stress preserving the velar articulation of -ng-. From CVn't'VCVC, (2) either became CVn'sVCVC (and then CVn'VCVC) or CVn'VCVC directly. The latter alternative is probably closer to being correct because the potential for perceptual confusion would be greater for the n't' cluster than for the n's cluster.
If \( n'a \) becomes \( n'z \) (following Herbert's principle), there is no corresponding voiced initial pair for the sequence to be confused with.

While the explanation presented above is not phonetically implausible, it rests crucially on the assumption that pre-AN had penultimate stress. As I mentioned in 5.1, it is not possible to say with any certainty where stress occurred.

The final problem I wish to discuss does not present any real problem for the Nasal Infix Hypothesis, but rather represents a problem for Austronesian linguistics in general. The development of nasals and nasal clusters in Austronesian languages is in no way straightforward and contradictory developments are not at all uncommon. I have argued that CVN\(^2\) forms retain the original articulation place of \( \mathbb{N} \) in most cases. At the same time, languages widely assimilate clusters of nasal plus stop. There is no obvious reason why CVN\(^2\) clusters should be preserved.\(^{16}\)

Note also that in many languages affixation with a prefix that triggers Pränasalierung and subsequent reduplication yields rather unexpected results. Consider Tagalog, putol 'cut', pamutol 'for cutting' (pa\( \mathbb{N}-\)) but pamumutol 'a cutting' and not expected pamuputol. Perhaps even more interesting is the fact that Tagalog
'copies the nasal' just in a case a 'substitution' occurs or the baseword begins with a vowel, for example, mang-ingisda 'fisherman', composed of maN- and isda. That is, bases that begin with voiced initial consonants that do not substitute, do not copy the nasal. Ngadju Dayak, however, always copies the nasal (see Dempwolff 1922-23). Satisfactory solutions to these peculiar interactions of nasal assimilation and reduplication have not yet been proposed. Perhaps in connection with this problem, Dempwolff also reports that in Ngadju a medial ⟨⟩ will appear as ⟨⟩ if a nasal consonant occurs previously in the word. Consider for example, ujah 'salt', but mun'ah 'to salt'.

As a final unexplained change involving nasals, notice that Malay unexpectedly shows bimbit for *bitbit. The expected reflex is bibit, with Malay regularly simplifying these obstruent-obstruent clusters. There is no obvious explanation for the appearance of this nasal.

Other examples from all subgroups of Austronesia could be added at will. The linguist interested in processes affecting nasals has at least one life-time's work in just the Austronesian area. The point I wish to make here is that Pränasalierung and medial clusters constitute only a small aspect of the much larger problem of the history and development of nasals in the Austronesian language family. It is hoped that the exposition of the
Nasal Infix Hypothesis presented above represents at least a small step in solving one of these long-standing problems, the origin of PAN medial nasal clusters.

In order to reconstruct the pre-AN relational pattern of nasal infixation, it has been necessary to consider several diverse phenomena in detail. This is by no means unusual (consider the complexity of Saussure's theory of IE ablaut, for example). The depth of the discussion here was necessary to adequately illustrate the method for reconstructing relational patterns proposed in the last chapter. It should be clear from the analysis presented in this chapter that a general discussion of the kinds of evidence one can adduce in support of a reconstruction, the potential problems the language historian encounters, and indeed, the fruits of the method would not have been sufficient to illustrate the richness of the method.
NOTES

1Lincoln, 1977 arrives at the figure 951. I suspect this figure is inflated since Lincoln was not careful in distinguishing languages from dialects of the same language.

2For *t* Dempwolff (1934:39) says '...der stimmlose, vordere palatale Verschlusslaut...'. However, for *d* he says (1934:56) '...stimmhafte palatale Verschlusslaut mit vorwiegender Beteiligung der Vorderzunge.'

3See Dahl, 1973. I adopt his usage instead of root in order to avoid any confusion the use of the latter term might entail with respect to Brandstetter's roots.

4There is a very strong tendency towards disyllables in the Austronesian family. See Blust 1977 for a stimulating discussion of the effects of this tendency on the history of Austronesian languages.

5See below 5.5. There is evidence for medial -RC-clusters.

6The square brackets around the *t* of a form like liman[t]ek or *L of ka[L]an'dang indicate that the testimony of the correspondence sets is indeterminate. For example, the correspondence set for *T differs from the correspondence set for *t only in the Jv reflex, t instead of t. In the absence of a Jv cognate it is not possible to determine whether *T or *t is to be reconstructed. In such cases Dempwolff used brackets and wrote the symbol for the statistically most frequent segment. The superscript (i) on *galing* indicates that more than one item of this shape has been reconstructed.

7The sequence -al- occurs widely in Philippine languages in the names of trees and plants. See below. R. Blust has suggested to me that limantek should be reconstructed as tetra-syllabic alimantek with the collective prefix ali-.

8I have omitted forms which Dempwolff considered to be morphologically complex, for example, *b-in-anga.

9-um- and -in- are prefixed when the root begins with a vowel. -um- is also prefixed if the root begins with a bilabial consonant. See below.
See Pinnow and Stampe, ms.

If -VN- had been the shape of the infix one would expect either a long vowel in the paN- prefixes or a vowel of a different color as the result of coalescence.

Chretien counted nasal clusters with his count of medial oral stops. Thus, ND counts as D. No figures are available on just the nasal clusters.

'Active' and 'passive' do not have their usual references in discussions of Philippine languages. 'Active' generally corresponds to the actor focus and 'passive' to any one of several oblique foci. See Schachter and Otanes (1972) for a discussion of focus in Tagalog.

See Carrier (1977) and Latta (1976, 1977) for analyses of the copied nasal in the reduplication.

For example, the development of the focus system.

See Blust, 1978 and footnote 4, above.
CHAPTER VI

CONCLUSIONS

In the preceding chapters I have discussed in detail methods for internal reconstruction; that is, various principles and procedures that may be employed in reconstructing aspects of the history of a language on the basis of purely internal, structural evidence. I will briefly review my major conclusions concerning the nature of internal reconstruction in the paragraphs that follow.

There is no unitary method of internal reconstruction. Rather, there are a number of principles and associated procedures for working with various kinds of structural evidence. These principles are appealed to when making inferences about the history of a language. They reflect the beliefs and assumptions of linguists about the structures of natural languages and the ways in which those structures may develop.

In the course of this dissertation I have discussed six methodological principles for internal reconstruction. For ease of reference, these are repeated below along with the section number in which each was introduced.
1.) The Principle of Automaticity (2.2): automatic alternations in the forms of morpheme synchronically are assumed to be the result of regular sound changes.

2.) The Principle of Economy (2.3): when there are two alternative hypotheses which attempt to account for the same set of data, the hypothesis which permits the phonetically best motivated statement of the sound changes is the more highly valued hypothesis.

3.) The Principle of Assumed Parallel Histories (2.3): when two linguistic elements, x and y, at one stage in the history of a language are reflected as x' and y', respectively, at a later stage, and if y' bears the relation to y that x' does to x, the histories of x' and y' are assumed to be parallel.

4.) The Principle of Optimal Use (2.3): If it is possible to provide a plausible account for some linguistic phenomenon on the basis of what it has already been necessary to reconstruct to account for some other distinct phenomenon, one should do so.

5.) The Principle of Similarity of Function (3.3): Linguistic elements that function similarly under similar circumstances are to be analyzed similarly.
6.) The Principle of Systematic Pairings (4.1): the existence of the systematic pairing of morphemes of different canonical forms is indicative of the occurrence of a relational pattern diachronically.

Contrary to what some scholars have suggested, the comparison of cognate morphs is not the only basis for internal reconstruction. Of the six principles formulated in the preceding chapters, only one, the Principle of Automaticity, is directly concerned with morphological alternations.

Finally, none of the proposed principles is specific to internal reconstruction, or even diachronic analysis. The principles of Economy and Optimal Use are concerned with formal simplicity and therefore reflect general principles of the philosophy of science concerning descriptions and hypothesis construction. Automaticity and Assumed Parallel Histories are the diachronic correlates of basic principles of synchronic analysis. Likewise the principles of Similarity of Function and Systematic Pairings reflect assumptions linguists have made about the significance of certain kinds of structural data.

It might be possible to reduce the number of principles in some formalistic way. For example, one might propose a general principle of simplicity which would subsume my Economy and Optimal Use and perhaps also
Assumed Parallel Histories and Similarity of Function
which are also concerned with simplicity in a broad sense.
I have not considered such possibilities here since my
goal was to illustrate clearly the different kinds of
arguments and types of evidence one may use in internal
reconstruction. And it seems to me the reduction would
obscure the significant differences involved in the appli-
cation of the different methods for internal reconstruc-
tion.
BIBLIOGRAPHY


Blust, Robert 1972 Proto-Austronesian addenda (revised version of 1971, same title). Oceanic Linguistics. XI.


Chafe, Wallace 1959 Internal reconstruction in Seneca. Language 35.


235


Hoenigswald, Henry 1946 Sound change and linguistic structure. Language 22.


Jeffers, Robert 1977 Morphological reanalysis and analogy: two case histories from Latin and Greek. Lingua 41.


Marchand, J. W. 1956 Internal reconstruction of phonemic split. Language 32.


Topping, Donald 1968 Chamorro vowel harmony. Oceanic Linguistics VII.


von Humboldt, W. 1836-7 Über die Kawi-Sprache auf der Insel Java. 3 vols. Berlin.

Wang, S.W.Y. 1969 Competing changes as a cause of residue. Language 45.


