INFORMATION TO USERS

The negative microfilm of this dissertation was prepared and inspected by the school granting the degree. We are using this film without further inspection or change. If there are any questions about the content, please write directly to the school. The quality of this reproduction is heavily dependent upon the quality of the original material

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.

2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.

3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps.

A dissertation entitled

Comparative and Internal Reconstruction

of Proto-Tupi-Guarani

submitted to the Graduate School of the
University of Wisconsin-Madison
in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

by

Charles Owen Schleicher

Date of Final Oral Examination: January 9, 1998

Month & Year Degree to be awarded: **December** **May** 1998 **August**

* *

Approval Signatures of Dissertation Readers: **Signature, Dean of Graduate School**

COMPARATIVE AND INTERNAL RECONSTRUCTION

OF THE TUPI-GUARANI LANGUAGE FAMILY

by

Charles Owen Schleicher

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Linguistics)

at the

UNIVERSITY OF WISCONSIN - MADISON

1998

ABSTRACT

This study examines the Tupi-Guarani (TG) language family a language family comprising languages spoken in Brazil, Paraguay, Argentina, Bolivia, and French Guiana. This study applies the comparative method of reconstruction to attested TG languages to arrive at a reconstruction of Proto-Tupi-Guarani (PTG). The protolanguage, in turn, is subjected to internal reconstruction to arrive at a Pre-TG phonology and morphology.

The proposed reconstruction of PTG is significant in that it differs in a number of significant ways from earlier attempts at reconstruction by Jensen (1984). Points of difference are as follows:

- The PTG phonemic inventory reconstructed here is analyzed as having a heretofore unrecognized phonological categorization, called fortis and lenis (which can be identified in other Amerindian languages). Fortes are voiceless stops and lenes are voiced continuants homorganic with the fortes. This captures the fact that, in PTG and many TG languages, voicing and continuance are redundancy features. This leads to an internal reconstruction of PTG.

- PTG is seen here to have three series of fortes: plain, labialized, and palatalized. This also contributes to the internal reconstruction.

• PTG is here found to have had three autosegmental phonemes, two at the morpheme level and one at the level of the syllable.

Additionally, the phonological reconstruction of Pre-TG is shown to contribute positively to the ability to identify cognates between TG and non-TG languages in the Tupian stock.

These discoveries, and the resultant reconstruction of Pre-TG, lead in turn to important conclusions about the morphology of PTG:

• Agentive, circumstantial and abstract nominalizers are found to be related morphemes.

• Form classes of nouns and verbs are found to be originally phonologically conditioned.

• A more accurate reconstruction of the pronominal system, both formally and functionally, is possible.

From these morphological conclusions, the following syntactic conclusions are also made:

- PTG was not ergative, as suggested by Jensen (1990).

- Subordination was expressed by nominalized verb phrases, rather than by separate clauses.

- The distinction between modifiers and nouns was essentially non-existent in PTG and the distinction between nouns and verbs was minimal and may in fact have been more pragmatic than syntactic.

The study concludes with a review of the question of internal classification within PTG and a proposal of a PTG homeland.

ACKNOWLEDGEMENTS

First of all I would like to thank my dissertation commitee: Prof. Andrew Sihler for his patience, thoroughness, and insight in directing this paper, as well as his considerable and positive influence in shaping my appreciation of methodological skill and integrity, especially in the field of historical linguistics. I commend him and thank him for his faithfulness in taking my best interests to heart, and in his tremendous encouragement, both professionally and personally, throughout my career at the UW-Madison. Thanks to Prof. Mary Lou Daniel, for her friendship and her professional and moral support. I greatly appreciate Prof. Ray Harris' professional encouragement and support of this project. Many thanks also to Prof. Manindra Verma and Prof. Tom Cravens for their helpful comments and suggestions. I would also like to thank Prof. Hazel Carter for the generous contribution of her insight in critiquing an earlier draft of this work.

I would like to acknowledge here a special debt of thanks I owe to the professors who helped to instruct me in this most delightful science of historical linguistics: Prof. Andrew L. Sihler, Prof. Robert Howell, Prof. Jeffrey Wills, and Prof. Valdis Zeps, all of the University of Wisconsin-Madison; Prof. Robert L. Rankin and Prof. Karl Rosen of the University of Kansas; and Prof. John Rea of the University of Kentucky.

Great thanks to the many native speakers of the Tupi-Guarani languages here discussed, past and present, whose work and inestimable heritage has informed this work

express my thankfulness. The difficulty that this project has brought you, both when I was with you, and moreso in my absence, has finally borne fruit. Your patience, support, understanding, and encouragement stand as an undeniable testament to the Love Who governs your life. To that Love I owe the greatest debt of all, Who has not only made this work possible, but has made us possible as well.

This dissertation is dedicated to Jesus Christ and to my family, my best friends on this earth.

Table of Contents

PART TWO: MORPHOLOGY

INTRODUCTION

0.1 The Tupi-Guarani language family

The Tupian language stock is one of the most widespread in South America, spanning the entire northern and eastern periphery of Brazil as well as much of the interior; much of Bolivia and Paraguay, especially the latter where Paraguayan Guarani is an official language; and smaller regions of Peru, Argentina, and French Guiana. By far the largest language family in this stock is the Tupi-Guarani (TG) family, which includes all of the area described above; the remaining language families of Tupian are represented in the west-central interior of Brazil. These different groups are listed in the table below.

Non-TG Tupian Families and Languages
(Based on Rodrigues 1986)

Arikem			Ramarama	
Karitiana	109		Arara	92
			Itogapuk	95
Juruna	126			
			Tupari	
Kokama			Makurap	215
Kokama	411		Tupari	56
Omagua	240		Wayoro	?
Monde			Aweti	36
Arua	?			
Cinta-Larga	953		Purubora	?
Gavião	220			
Mekem	40		(Satare-) Mawe	3,000
Monde	?			
Surui	340		Siriono (?)	
Zoro	175			
Munduruku				
Kuruaya	52			
Munduruku	1,460			

According to Rodrigues'(1986) classification, there are twenty-five languages in this family; generally they have defied internal classification, although attempts have been made Rodrigues (1984) and Dietrich (1990). Of the languages in Rodrigues' classification, Kokama has recently been called into question as truly Tupi-Guarani (Dietrich 1990; Cabral 1995). I also exclude Kokama from TG for reasons given in Chapter 5. For reasons given in Chapter 2 I also exclude Sirionô from TG, which has been accepted as TG in all other sources.

These languages can be shown in the following table:

Tupi-Guarani Languages and Dialects
(based on Dietrich 1986, Jensen 1984, Rodrigues 1986)

Akwawa			Lingua Geral		3,000
	Asurini do Tacantins	131	Mbya Guarani		2,248
	Surui do Tocantins	101	Parintintin		
	Parakana	297		Diahoy	13
Amanaye		?		Juma	9
Anambe		61		Parintintin	118
Apiaka		65		Tenharin	256
Arawete		136	Siriono (?)		
Asurini of the Xingu		53	Tapirape		202
Ava		101	Tenetehara		
Chiriguano		65,000		Guajajara	6,776
Guaja		240		Tembê	410
Guarani			Uruewauwau		215
	Kaiwa	7,000	Urubu-Kaapor		494
	Nandeva	4,900	Wayampi		
	P. Guarani	3,000,000		Amapari	279
Guarayo		6,000		Jari	12
Kamayura		207		French Guiana	400
Kayabi		620	Xeta		5

Tupinamba dated from the mid-sixteenth century and is attested until about the mid-seventeeth century. OG is attested from about the turn of the seventeenth century to the middle of the eighteenth. After that there are virtually no records of the language until the 1860s. This is the historic starting point for Modern Guarani, which in the nineteenth century was only slightly different from the language today. By this time Guarani is differentiated into at least three distinct languages: Chiriguano, Guarani, and Mbya (Guarani) formerly called, respectively, Bolivian Guarani, Paraguayan Guarani, and Brazilian Guarani. This study looks chiefly at the dialect of Guarani known as Kaiwa, which is spoken mostly in Brazil, but also in Paraguay and seems to be mutually intelligible with other dialects of Paraguayan Guarani. Only occasionally will stadard Paraguayan Guarani be compared with Kaiwa in this study. Although Old Guarani was spoken in Asuncion, it is almost certainly not the dialect ancestor of modern Paraguayan Guarani. This is even more clearly true of Mbya, which has *í* for OG *s*. Lexical comparison also seems to support a more distant relationship. Nevertheless, Mbya almost certainly developed from a dialect that was mutually intelligible with the Asuncion dialect of OG.

0.2 Previous studies in TG linguistics

Of all the indigenous language families of South America, few have been so extensively studied as the Tupi-Guarani family. No other family is as extensive: Tupi-Guarani languages are spoken in Brazil, Argentina, Paraguay, Bolivia, Peru, and French Guiana. Almost no other language family is as well attested: There are about twenty-four languages in the TG family: Akwáwa, Amanayé, Anambé, Apiaká, Araweté, Asurini of the

Xingu, Avá, Chiriguáno, Guajá, Guaraní, Guarayú, Kamayurá, Kayabí, Kokáma, Lengua Geral (or Modern Tupí), Omágua, Parintintin, Sirionó, Tapirapé, Tenetehára, Uruewauwáu, Urubu-Kaapor, Wayampí, and Xetá.

In addition to these there is one extinct language, Tupinamba (or Old Tupí), which may or may not be directly ancestral to one or more of the modern languages. Given the wide variety and distribution of TG and the extent of their literary record, a brief historical sketch of TG studies is in order.

The earliest grammar of a TG language was that of Tupinamba by Joseph Anchieta, published in 1595. The next major work to appear and arguably the most extensive single work on any TG language was that of Antonio Ruiz de Montoya, a Jesuit missionary to Paraguay, who wrote the Arte de Grammatica de la lingua Guaraní. The Tesoro de la Lingua Guaraní, Diccionario de la Lengua Guaraní, and the Catecismo, a bilingual catechism in Spanish and OG. These were published from 1639 to 1640 in Madrid. The next major work on Guarani was the *Arte de la Lengua Guarani* (1724) by Paulo Restivo. Restivo's grammar reads much less like a language learning textbook than Ruiz de Montoya's, and more like a linguistic description, with close attention to dialectal, individual, and free variation in Guarani. Another published work was by Nicolas Yapaguay, a native Guaraní speaker who published a *Catecismo* (1724) as well as a collection of *Sermones y Exemplos en la Lengua Guaraní* (1727)).

Early MG (eMG) is represented by the *Pequeño Ensayo del Idioma Guarani* (1891), which was the first grammar to be written on MG. It was written by seminary priests in Asunción and edited by Juan G. Granado. Citations from this work are indexed by his initials.

In 1892, the same year that Restivo's *Arte* was republished, Henri Coudreau published the Vocabulaires méthodiques des langues Ouayana, Aparai, Oyampi, Emerillon. It was the first work of a Tupi-Guarani language aside from Guarani, Tupinamba, or Lingua Geral. This was a comparative word list of four languages, two of them Oyampi and Emerillon, being from the same branch of TG (the other two, Aparai and Ouyana, were Cariban).

In 1910, C. Tatarin published La Langue Tapihiya, dite Tupï ou ñeēngatu (belle langue): grammaire, dictionnaire, et textes).

The next language to be documented was Guarayo, recorded in two works by Alfredo Hoeller in 1932: Grammatik der Guarayo-Sprache and Guarayo-Deutsches Wörterbuch. Apart from these, most extensive study in TG has been done in the last half-century. In that short time so much has been published that a full account would be inordinately long for an introduction. It will suffice to mention some of the major figures in TG studies: Aryon Dall'Igna Rodrigues (Tupinamba), Cheryl Jensen (Wayampi), David Bendor-Samuel (Guajajara), Carl Harrison (Guajajara, Asurini), Lorraine Bridgeman (Kaiwa), La Vera D.

Betts (Parintintin), Lucas Espinosa (Kokama), Rose Dobson (Kayabî), Helga Weiss (Kayabî), Robert Dooley (Mbya Guarani), Lucy Seki (Kamayura), James and Kiyoko Kakumasu (Urubu-Kaapor), Carole Schuchard (Chiriguano), Yonne Leite (Tapirape), Gary and Roberta Olson (Wayampi), Audry and John Taylor (Kaiwa), Velda Nicholson (Asurini).

When data from all of the TG languages began to become available for analysis, historical linguistic analysis for TG became possible. The pioneer in this effort was Aryon D'Alligna Rodrigues, who has been unquestionably the dominant figure in TG linguistics beginning with Rodrigues (1953). In this article, Rodrigues was the first scholar to apply modern linguistic analysis to a TG language.

The first linguist to publish a reconstruction of PTG was Miriam Lemle (1971), who reconstructed over 220 protoforms in PTG using data from Asurini of the Tocatins, Guajajara, Guarani, Guarayo, Kamayura, Kokama, Parintintin, Siriono, Tupinamba, and Urubu-Kaapor. In addition to reconstructing the phonemic inventory and a portion of the lexicon, Lemle also attempted to arrive at a conclusion concerning the internal relationships within TG. Later, a different arrangement of internal relationships was proposed by Rodrigues (1984/85).

During this same time, more detailed research was done on the diachronic phonology of TG, and, to some degree, the synchronic phonology of PTG. This work was done chiefly by Yonne Leite and Marilia Facó Soares (Leite 1977, 1982; Soares 1978, 1979; Soares and

Leite 1991). One of the latest comparative/historical works in TG and one of the most impressive to date is Jensen's (1984) work on Wayampi, recently revised (1990). Although, as the title implies, the work focuses on Wayampi, it involves a comparative/historical analysis of the entire family. In addition to revising Lemle's lexical reconstruction, Jensen uses Rodrigues' work on Tupinamba to analyze possible synchronic phonological rules of PTG, and, together with data from other TG languages, reconstruct proto-forms for the bound morphology of PTG.

0.3 The present study

The present work aims to fulfill two goals which have not yet been achieved: to make a comprehensive survey of TG morphology and phonology, referring to and, when appropriate, revising Jensen's (1984) reconstruction, in order to give a diachronic account of TG phonology and morphology from PTG to each of the languages; and to apply the method of internal reconstruction to the phonology and morphology of PTG to account for both regularities and irregularities in the system; and finally, to use grammatical and lexical comparisons to reconsider, one more time, the internal relationships within TG.

0.4 The data and orthography

Due to the scope of this study, field work was not possible as a direct source of information. Therefore the data used here are drawn from the field work of other linguists, chiefly Cheryl Jensen (Wayampi), David Bendor-Samuel (Guajajara), Carl Harrison (Guajajara, Asurini), Lorraine Bridgeman (Kaiwa), La Vera D. Betts (Parintintin), Rose

Dobson (Kayabí), Helga Weiss (Kayabí), Robert Dooley (Mbya Guarani), Lucy Seki (Kamayura), James and Kiyoko Kakumasu (Urubu-Kaapor), Wolf Dietrich (Chiriguano), Ionne Leite (Tapirape), and Velda Nicholson (Asurini). Textual sources for extinct languages were: Joseph Anchieta (1595) (Tupinamba), Antonio Ruiz de Montoya (1640) (Old Guarani), and Paulo Restivo (1724) (Old Guarani).

Although some languages have established orthographies of their own, none of these will be used. All examples and citations will be written in phonetic transcription unless otherwise indicated.

0.5 Layout of Chapters

The study is organized as follows: in two parts, Phonology and Morphology. Syntax is not covered in this study because of the unavailability of sufficient data in this area for many of the languages. Phonology is broken down into two chapters: the first re-examining the reconstruction of PTG phonology by means of the comparative method; a comparative word list shows a reconstructed word list of 224 items; the second chapter applies internal reconstruction to understand PTG phonology better and to arrive at a tentative reconstruction of Pre-TG phonology. Morphology is broken down into four chapters; each of the last three of them has two sections, one dealing with a comparative reconstruction of PTG and one dealing with the internal reconstruction of Pre-TG: the chapter on morphophonology transitions from the Phonology chapters to morphology. Chapter 4 examines the noun system and the root class system in general for PTG and Pre-TG. Chapter 5 examines the

pronominal or cross-referencing system in TG. Chapter 6 examines the structure of verb complexes and by extension, noun complexes to gain insight into PTG word classes. The concluding chapter will summarize the findings of each of the chapters, explain the contribution these findings make to TG linguistics, and makes some proposals about the PTG homeland, possible routes of migration, and internal classification.

The theory of linguistic geography is used to amplify and interpret Dietrich's analysis for internal classification. In proposing the PTG homeland, I will use Dyen's (1956) Migration Theory.

The scope of the present work seems broad, but this is necessary for two reasons: all languages for which there are available data needed to be used to make the reconstructions as meaningful as possible; and because, as will be seen, all the issues dealt with in this work are interconnected in such a way that it is impossible to satisfactorily analyze one problem without taking all the other aspects into account.

Synchronic phonology will be analyzed on the basis of autosegmental phonology. Chapter 1 will examine current theories about the phonological changes from PTG to the modern languages starting with a comparative word list, phonological correspondences will be established. The reconstructed proto-system will be shown to lend itself to internal reconstruction to arrive at a pre-TG phonemic system.

Finally, the analysis of the first two chapters will be used to examine the internal relationships of TG, emphasizing grammatical and lexical resemblances over shared phonological innovations.

0.6 Formal notations and abbreviations

1 - first person
2 - second person
3 - third person
s - singular
p - plural
n - inclusive
x - exclusive
trans - transitive
intr - intransitive
loc - locative
distr - distributive
nom - nominative
acc - accusative
poss - possessive
obtop - oblique topicalization
foc - focus
com - comitative
caus - causative
agnt - agent
circ - circumstantial
pm - pronominal marker
refl - reflexive
pass - passive
recip - reciprocal
int - interrogative
emph - emphatic
fut - future
frust - frustative
MAR - multi-attachment resolution
* - unattested
§ - ungrammatical/unrealized
- - morpheme boundary
= fused boundary (also for multi-word glosses of single morphemes)

Languages

Akwawa   (AT)
Assurini of the Xingû   (AX)
Chiriguano   (Ch)
Mbya Guaraní   (GM)
Kaiwa   (Kw)
Paraguayan Guaraní (G)
Old Guarani (OG)
Guarayo   (Gy)
Kamayura   (Km)
Kayabi   (Kb)
Parintintin (Pt)
Tapirape   (Tp)
Guajajara   (Gj)
Urubu-Kaapor (UK)
Wayapĩ   (W)
      Amapari   (Wa)
    Jari   (Wj)
Tupinamba   (Tb)

Sources

Joseph Anchieta   (A)
David Bendor-Samuel   (B-S)
Cheryl Jensen   (J)
James Kakumasu   (K)
James and Keiko Kakumasu (K&K)
Paulo Restivo (PR)
Antonio Ruíz de Montoya   (RM)

PART ONE: PHONOLOGY
CHAPTER ONE: COMPARATIVE PHONOLOGY

1.0. Introduction

In this chapter we will consider the comparative evidence to reconstruct the phonology of PTG as well as evidence within PTG for earlier stages in pre-TG. As we shall see there has been very little change in the consonant systems of TG languages, probably owing to the symmetry of the PTG consonant system. The same is true of the vowel system. Again, most of the languages have maintained the same vowel system as the parent language. In both consonant and vowel systems, however, destabilizing changes have taken place, almost entirely in the Central Innovation Area. These changes include both conditioned and unconditioned changes. The discussion on the sound change rules of TG will make reference to comparative phonemic data in the cross-linguistic basic word list in Appendix 1.

1.1. Numbers in TG

There are only four numbers which can be reconstructed for PTG: 1-4. Of these only one, *mokõy* 'two', is straightforward; the others are clearly related, but more problematic, and 'four' does not actually exist in one modern language (Pt).

The 'one' word in TG languages has almost as many variations as there are languages; but all the forms have one particle in common: *pe*. In fact, every 'one' word in TG is directly derived from *pe* plus some combination of one or more recognizable

morphemes. The meaning of *pe* itself, ubiquitous in TG but never translated as 'one', is also straightforward: it is a demonstrative, '(proximate) that'.

'Three' *mocapír* is almost as easy as 'two', but in AT and AX the number is *na?iroihi* and *iroma?e* respectively, words that appear to be related to the various 'four' words. Meanwhile, the reflex of *mocapír* in Pt means 'a few, three or more'; Pt is one of the only modern TG languages lacking a word for 'four'.

Finally, we return to the 'four' words, and find that, as with 'one', the forms are many and varied, but have one common element: this time, the morpheme *irũ*. This happens to be a word which commonly means 'companion, close friend' or something even more intimate. In Pt, however, the word simply means 'other' or even 'one more'. If this is the original meaning of the word it would be the logical word to use for four if one's language only had words for 1-3. Furthermore it would seem more than coincidence that PTG had two homophonous roots, one meaning 'another' the other meaning 'companion' (or 'significant other'). The overall picture for PTG is a language which in truth had no number words per se; instead of 'one', 'two', 'three', 'four', one would more correctly gloss these as 'that', 'pair/couple', 'few', and 'another'.

1.2. Correspondence sets

The following phonemes have been reconstructed for PTG (Lemle 1971):

Stops	p	t		k	?
Fricative	β				
Nasals	m	n		ŋ	
Flap			r		
Affricates		c	č		
Glides			y	w	

Apart from *w, c, č,* and *?,* identity sets exist for all these phonemes. The phonemic status of the last two is challenged here, and three phonemes are reconstructed for the proto-language not previously reconstructed: p^y, p^w, and k^w.

From the above comparative list we derive the following cognate sets:

*p/mb > Km p, Pt p, UK p, W p, Kb p, GM p, AT p, AX p, Kw p, Tp m,
    Tb p/mb; 165, 166, 172

*p > Km h, Pt p, UK p, W p, Kb φ, GM p, AT p, AX p, Kw p, Tp p,
    Tb p; 181, 182

This correspondence set is identical to that of *p^w below, and occurs when *p* is immediately followed by *u.* Therefore, we not only reconstruct *p for this set, but also conclude that in Km and Kb (and perhaps originally in PTG as well) *p* was labialized before *u.*

*t > Km c, Pt t, UK š, W s, Kb c, GM č/š, AT č, AX č, Kw t, Tp č, Tb t; 6, 19, 42, 44,
    178, 189, 190

This set only occurs before *i,* therefore it is very straightforward that this set is also from

*t which assibilates before *i*. The very different developments in the daughter languages indicate that this is an independent development in the individual languages (or at the very least, may be; in any case the isogloss for this development contributes nothing to arriving at an internal classification. Those languages which show palatalization, for example, do not come from a single branch, whose ancestor, intermediate between PTG and these languages, exhibited spirantization.

*β > Km w, Pt β, UK w, Wa β, Wj w, Kb w, GM β, AT w, AX w, Kw β, Tp w, Tb β;
    1, 18, 19, 48, 49, 50, (51), (52), 55, 67, 89, 90, 91, 92, 216

This correspondence points to *β, since there is an identity set for *w (immediately below). That correspondence set, showing w for all languages except Pt, GM, and Kw, which have γ^w, clearly points back to an original *w. Obviously this could go either way with PTG *γ^w or *w, but not only does *w fit better into the phonemic inventory of PTG, on the basis of the fact that the strong majority of languages show w, it is better to reconstruct *w for PTG. Given that, it makes the most sense to ascribe the above correspondence set to *β, since there is no environmental conditioning to determine, for example, when γ^w or β appears in Kw.

*w > Km w, Pt γ^w, UK w, W w, Kb w, GM γ^w, AT w, AX w, Kw γ^w, Tp w, Tb w; 199,
    200, 201, 203, 204, 206

*y > Km y, Pt ǰ, UK y, W y, Kb y, GM ǰ, AT č, AX ǰ/ž, Kw ǰ, Tp č, Tb y; 205-224

This points to original *y in PTG, since affricatization is more likely than lenition of an affricate to *y*, a sound change uncommon enough to make it implausible for nearly half

the languages to independently undergo that change, while $y > j$ is common enough as to make it perfectly acceptable that slightly more than half the languages have undergone it.

*β > Km p , Pt β, UK Ø, W Ø, Kb p, GM Ø, AT w, AX w, Kw Ø, Tp p,
　　Tb β; 2, 10, 114, 149, 195

*β > Km p, Pt β, UK Ø, W Ø, Kb p, GM Ø, AT m, AX p, Kw Ø, Tp m,
　　Tb β; 3, 25, 58, 74, 99, 105

These sets, and the ones for *r and *k below, all occur at word-final position, hence the development is different from that of *β above. Indeed, note that when word final consonants are reflected as word final in AT, they are stops; when followed by an unstressed vowel, they have the same reflexes as PTG medial continuants. [1]

*r > Km t, Pt r, UK r, Wa r, Wj Ø, Kb t, GM Ø, AT r, AX r, Kw Ø,
　　Tp n, Tb r; 11, 39, 97, (100), 206

*r > Km t, Pt r, UK r, W Ø, Kb t, GM Ø, AT n, AX t, Kw Ø, Tp n,
　　Tb r; 7

*r > Km t, Pt r, UK r, Wa r, Wj Ø, Kb t, GM Ø, AT r, AX r, Kw Ø,
　　Tp t, Tb r; 43, 78, (140), 161, 179, 188, 223

*r > Km t, Pt r, UK r, W Ø, Kb t, GM Ø, AT n, AX t, Kw Ø, Tp t,
　　Tb r; 112, 198, (216)

---

[1] Because Tb shows a free variation β/p and r/t it seems likely that this variation dates back to PTG rather than that β and r were (re)strengthened word-finally. See (2.1.) below for a detailed analysis of this idea.

Note that in verbs AT has nasals as reflexes for word-final continuants.

Tapirape has sometimes a stop, sometimes a nasal, and there is no conditioning factor, phonological or grammatical, to suggest when either reflex appears. At the same time it is clear that this discrepancy should not lead us to propose two new phonemes for PTG; rather, the most obvious explanation seems to be dialect borrowing, since AT always realizes word-final lenes as nasals. A perhaps less likely possibility is that the Tp data reflect a lexical-diffusion change in progress.

*k > Km k, Pt γ, UK k, W Ø, Kb k, GM Ø, AT ŋ, AX k, Kw Ø, Tp k, Tb k; 5, 62, 67,
  77, 89, (119), 127, 134, 150, 191, (219)

The change *k > γ in Pt clearly is a structurally motivated change, parallel to the lenition of p and t word finally.

*m > Km m, Pt m, UK m, W Ø, Kb m, GM Ø, AT m, AX m, Kw Ø, Tp m,
  Tb m; 4, 24, 53, 60, 65, 106, 158

*n > Km n; Pt n, UK n, Wa n, Wj Ø, Kb n, GM Ø, AT n, AX n, Kw Ø,
  Tp n, Tb n; 26, 141

*n > Km n; Pt n, UK Ø/n, W Ø, Kb n, GM Ø, AT n, AX n, Kw Ø,
  Tp n, Tb n; 46, 64, 81, 96, 160, 203

*ŋ > Km ŋ, Pt ŋ, UK Ø, Wa ŋ, Wj Ø, Kb ŋ, GM Ø, AT ŋ, AX ŋ, Kw Ø,
  Tp ŋ, Tb ŋ; 23, 42, 86, 90, 107, 139, 164, 189, 202, 215

*pʷ > Km hʷ, Pt kʷ, UK kʷ, W kʷ, Kb φ, GM kʷ, AT kʷ, AX kʷ, Kw kʷ,
  Tp kʷ, Tb pʷ; 183, (184)

*kʷ > Km kʷ, Pt kʷ, UK kʷ, W kʷ, Kb kʷ, GM kʷ, AT kʷ, AX kʷ, Kw kʷ,
  Tp kʷ, Tb kʷ; 99, 100

In spite the fact that the majority of languages show k^w we reconstruct $*p^w$, since there is already an identity set for k^w on the one hand and another set containing k^w in most languages and p^w in Tb, one of the oldest attested languages, h^w in Km and ϕ in Kb. Hence we have two labialized consonants in PTG, $*p^w$ and $*k^w$, and only three documented languages (one ancient and extinct, two modern) which have reflexes of $*p^w$ distinct from the reflex for $*k^w$, and neither of the two modern languages has phonetically preserved $*p^w$. The existence of a voiceless fricative ϕ in Kb may eliminate the fortis/lenis distinction in that language, although ϕ does not contrast with a corresponding voiced fricative (there is no β). The absence of β, however, leaves only one originally 'lenis' consonant: r, not enough to justify a claim that Kb had a class of 'lenis' consonants. For this reason the loss of β alone is probably enough to make Kb, or any other TG language, cease being a fortis/lenis language.

$*\check{c}$ > Km Ø, Pt h, UK h, W Ø, Kb Ø, GM Ø, AT h, AX h, Kw h, Tp Ø, Tb s

Finally, there is the development of the affricate or, in terms of a fortis/lenis phonology, the sibilant. There is no single correspondence set for this affricate, as seen below.

	Km	Pt	UK	W	Kb	GM	AT	AX	Kw	Tp	Tb
		h	h	h	h	Ø	h	Ø/h	h	_	
11	Ø̄	h	Ø	Ø	Ø	Ø	h	h	_	Ø̄	
55	h	h	h	Ø	_	č/š	č	Ø	s	_	
56			s/Ø	Ø		_	Ø	Ø	s	_	
57	Ø̄	Ø̄	s	Ø	Ø̄	ʔ	_	_	_	Ø̄	
58		h	h	_	_	_	_	_	_	Ø̄	
59	Ø̄	h	s	s	_	č/š	Ø̄	_	s	Ø	
60	Ø	Ø	Ø/h	Ø	Ø̄	Ø/h	h	Ø	Ø/h	Ø	
61	Ø	Ø	h	Ø	Ø	č/š	h	h	s	Ø	
62	Ø	h	h	_	_	č/š	_	_	s	_	
63	_	_	_	Ø̄	_	č/š	_	_	s	_	
64	_	h	_	_	_	_	h	Ø̄	_	_	
65	_	h	_	_	Ø̄	Ø̄	h	h	_	_	
68	Ø̄	Ø	h̄	Ø̄	Ø	č/š	h	h	s	Ø̄	
		h	h	Ø	_	Ø	h	h	h	_	s
110	Ø̄	h	s	s	_	č/š	Ø	ʔ,h	s	_	
111	Ø	h	Ø	Ø	Ø̄	Ø	Ø	Ø	h	Ø̄	
120	Ø	h	h	Ø/s	Ø	Ø	_	_	h	Ø	s
152	Ø	h	h	Ø	Ø	Ø	h	h	h	Ø	s
163	Ø	h	h	Ø	_	Ø	h	h	_	Ø	s
164	h	h	h	_	_	Ø	h	h	h̄	_	
189	Ø	h	h	Ø̄	Ø̄	č/š	h	h	s	Ø̄	s

Compare some of the more complete correspondence sets:

55	h	h	h	Ø	_	č/š	č	Ø	s	_
59,110	Ø	h	s	s	_	č/š	Ø	Ø	s	Ø̄
61,68	Ø	Ø	h	Ø	Ø̄	č/š	h	h	s	Ø
60	Ø	Ø	Ø/h	Ø	Ø	Ø/h	h	Ø	Ø/h	Ø
111	Ø	h	Ø	Ø	Ø	Ø	Ø	Ø	h	Ø

The existence of these different correspondence sets with no discernable conditioning environments to predict which will occur has led to the positing of two sibilants *č* and *c* by Rodrigues, followed by Lemle (1971), Jensen (1984, 1990), and

Dietrich (1994). In spite of the fact that more than two correspondence sets occur here, no more than two sibilants have ever been reconstructed. It is clear, however, that these correspondences could be interpreted as evidence for three to five different sibilants for the proto-language. There is so much randomness in the correspondences that it is impossible to posit one basic correspondence set for one or even two sibilants; the variations in correspondence sets therefore seem not to justify positing two sibilants any more than positing three, four, or more sibilants. [The reconstruction of two affricates is due to the fact that of the various reflexes of the sibilant(s), two are affricates: c and $\check{c}$. Since it is more likely for an affricate to become a fricative than the other way around, affricates were reconstructed for PTG.] The reason for positing two (and not more) proto-phonemes for these reflexes was not based on correspondence sets, but on the fact that within each of the Guaranian languages, only two variations are predominant ($\check{c}/s$ ̌ vs. $\emptyset$ in Mbya, s vs. h in Kaiwa and Paraguayan Guarani); the 'stronger' reflex supposedly from $*\check{c}$, the 'weaker' from $*c$. However, as has already been seen, the different languages are not always in agreement as to which words reflect which proto-phoneme. The reconstruction of $*\check{c}$ and $*c$ does not adequately explain the chaotic collection of correspondence sets above.

The variations are explained by lexical borrowing across languages and phonological influence among dialects. I would like to show here that such explanations actually make it unnecessary to posit two sibilants, and that only one, $*\check{c}$, is necessary.

Among those lexemes that are reconstructed with *č by Dietrich are the following:

*roʔɨčam 'cold'
*čem 'go out'
*pɨčɨk 'grasp'
*kɨče 'knife'
*waču 'big'
*yačɨ 'moon'
*ačɨ 'illness'
*aču 'left hand'
*apɨča 'hearing'
*tačo 'worm'
*čɨ 'mother'
*čuʔu 'bite'
*čam 'string'
*čɨrɨk 'slip/flow'

Of these, *roʔɨčam is an incorrect reconstruction. The reflex in Pt is roʔitiaŋ, There is

no affricate in this word to begin with, rather ti, which becomes č in many TG languages.

Of the others, *čɨ, *čɨrɨk, are simply examples of a rule that preserved č

immediately after ɨ, in these cases, the third person marker ɨ. *čem and *pɨčɨk have

doublets: even languages which seem to show clearly different reflexes of c and č such

as Mbya Guarani, don't have agreement on pɨčɨk and čem. So, for example, *čem is so

reconstructed because, a language like Paraguayan Guarani, which has s for č and h for

c, has sẽ as the reflex of this proto-form. However, Guayakí, a Guaranian language, has

ẽ, which would point back to *cem; Guayakí does have pɨčɨk for *pɨčɨk, but another

language, Apopucava has pɨɨ (§pici would be expected here, cf. yaci from *yačɨ).

Dietrich explains this by suggesting dialectal variants in the proto-language, čem/cem,

pičik/picik. But if this is so why don't most or all TG languages show such variation? As it happens, only certain languages — Mbya Guarani, Kaiwa (Guarani) Paraguayan Guarani, Chiriguano (also known as Bolivian Guarani), and various Guaranian dialects — show frequent variation of this sort. Languages in the central innovation area as well as relic area languages (such as Parintintin and Guarayo) show no such variation. In other words, the existence of two apparent reflexes for affricates in PTG only shows up in languages descended from OG. It would make more sense, then, to propose dialect variation of this sort as an innovation in OG from a single affricate in PTG. This is supported by OG data which show such variation. Montoya, for example, attests *sē/hē*, 'to go out'. The other languages only point to a single affricate in PTG. The OG variation was inherited by the daughter languages, resulting in the apparent contrast we see today.

Exceptions of the sort mentioned above are not completely limited to Guaranian languages, however. There is a much smaller set of examples of the same kind of variation in other languages. These are so few, however, that they can be (and have been) easily explained as lexical borrowings. Jensen (1990a:23) noted such exceptions in Wayapí, where $*č$ and $*c$ both disappear without a trace with the following exceptions:

> kise 'knife'
> pisaye 'heavy'
> wasu 'big'
> usu 'big'

Rodrigues suggests that these are borrowings from Lingua Geral (Nheengatu) and indeed Jensen reports that the first two of these are considered words 'of the forefathers' ('dos avos'). This strongly suggests that the same is true of *wasu* and *usu* as well. If these words were borrowed from Lingua Geral into Wayapĩ, then they may have been borrowed into other languages as well. Note further that these words are all reconstructed with *č* in PTG. This at least hints that the words which have alleged reflexes of *č* are in fact borrowings from Lingua Geral or even older borrowing from Tupinamba or some other language in which the reflex of the affricate was still *s* (and hadn't yet weakened to *h* or disappeared altogether). In fact, the Wayapĩ words above would no doubt also be touted as evidence for *č* but for the fact that they are so exceptional, most of the words reconstructed with *č* (on the basis of Guaranian evidence) having $\emptyset$ as the reflex for the affricate in Wayapĩ; therefore the borrowed nature of these words is very clear. But there is no reason for not supposing that all the lexical items in all TG languages used as evidence for *č* are not also borrowings. That such examples are more common in Guaranian languages is simply due to the fact that in OG the borrowings were dialectal, not cross-linguistic, and therefore more common. The only other variations in the correspondence sets regarding *č* can be explained as random loss of *h* where it might have otherwise been the expected reflex, a perfectly reasonable hypothesis (cf. Sirionó *ïra* 'fish' < pre-TG *pirá*, vs. *hē* 'you pl.' < pre-TG *pē*).

To sum up, then, the different reflexes for the affricate in PTG are not best

explained as evidence for two proto-phonemes, $*\check{c}$ and $*c$, but by the hypothesis that there was one phoneme $*\check{c}$ and the duality found in different TG languages is due to dialect borrowing (in Guaranian languages) and some cross-linguistic borrowing; that in some languages $*\check{c}$ became c, elsewhere it remained $\check{c}$ and later deaffricated as $\check{s}$; elsewhere $\check{c}$ became c, then s, which was later weakened to h; elsewhere it was lost altogether, and through mutual contact of these dialects more than one reflex for $*\check{c}$ existed in several dialects with no phonological conditioning to predict their appearence.

$*ʔ >$ Km ʔ; Pt ʔ; UK ʔ; W ʔ; Kb ʔ; GM ʔ; AT ʔ; AX ʔ; Kw ʔ; Tp ʔ;
    Tb ʔ

$*ʔ >$ Km Ø, Pt ʔ, UK Ø, W Ø, Kb ʔ, GM Ø, AT Ø, AX Ø, Kw Ø, Tp Ø,
    Tb Ø (?)

The second of these two correspondence sets shows ʔ appearing consistently in Pt and Kb, but in none of the other languages. Tb and OG texts are not always reliable about recording ʔ, but Montoya does seem to be consistent about noting the presence of ʔ in a word for OG. Based on this fact and the fact that all modern Guaranian languages have no preglottalized consonants we conclude that the absence of any notation or mention of ʔ immediately before a consonant signifies that preglottalization was absent in OG. This means that ʔ was already interpreted as a segment which disappeared next to other consonants.

This second correspondence always occurs immediately before a consonant eg. Pt

ja?ɣʷar, Kb *ya?war*, vs. UK *yawar*, Kw *jaɣʷa*, etc. It is much easier to see how *?* could disappear before consonants in most TG languages than that it sprang up randomly in both Pt and Kb, since the appearance of *?* before consonants in Pt is always matched in the Kb cognate; there is no phonological conditioning to trigger the appearence of *?* before consonants in the words where this occurs in Pt and Kb; and Pt and Kb are not close either geographically or in their internal classification in TG. Therefore it is more reasonable to propose that Pt and Kb are conservative in this regard and that *?* occurred before consonants in PTG. This is interesting because it violates a very strict CV syllable structure in Pt, Kb, and PTG (with the exception of word-final consonants). Certainly medial syllables are/were always CV in these languages. This suggests that *?* was not a segment in PTG, nor is one in Pt and Kb, but rather a segmental or even an autosegmental feature. In other words, PTG had pre-glottalized vowels and consonants (specifically continuants, since there is no evidence for pre-glottalized stops in PTG). So for example, PTG **?ú* is not a sequence of two segments, *?* followed by *u*, but rather a single segment, pre-glottalized *u*, perhaps better transcribed **ˀú*.[2] Similarly, the glide in PTG **ya?wár* is a pre-glottalized *w*, not a sequence of *?* followed by *w*, so the syllable division is really *ya-ˀwár*. The loss of *?* in the vast majority of TG languages simply represents the reanalysis of glottal stop as a segment; pre-glottalized consonants were merged with 'simple' stops while pre-glottalized vowels were reanalyzed as sequences of *?* plus vowel. Interestingly, the only kind of segments which are not pre-glottalized in Pt

---

[2]From here on I will use two different symbols to distinguish between glottal stop as a segment (?) and as an autosegment (ˀ).

and Kb, and therefore the only segments which cannot be reconstructed as pre-glottalized in PTG, are stops.

vowels:	*ʔú, 'to eat' *kaʔá, 'forest'
glides:	*yaʔwár, 'jaguar'
lenes:	*iʔβá, 'fruit'
nasals:	*ʔŋā, 'they'

With this better understanding of glottal closure in TG, we can reformulate the above correspondence sets:

*ʔ > Km ʔ; Pt ʔ; UK ʔ; W ʔ; Kb ʔ; GM ʔ; AT ʔ; AX ʔ; Kw ʔ; Tp ʔ;
    Tb ʔ;

*ʔ > Km Ø, Pt ʔ, UK Ø, W Ø, Kb ʔ, GM Ø, AT Ø, AX Ø, Kw Ø, Tp Ø,
    Tb Ø

1.3. Sound change rules

From the above correspondence sets we derive the following sound change rules for each language.

*t > č / _ i GM (č/š), AT, AX, Tp
*t > c / _ i Km, Kb
*t > š / _ i UK
*t > s / _ i W

*β > w / _ Km, UK, Wj, Kb, AT, AX, Tp
*w > γ^w Pt, GM, Kw

*y > ǰ   Pt, GM Kw, AX
*y > č   Tp, AT

*p/β > p   Km, Kb, (Tp)
*t/r > t   Km, Kb, (Tp)

*p/β > β   Pt
*t/r > r   Pt, UK, (Wa)
*p/β > ∅ / _# UK, W, GM, Kw
*t/r > ∅ / _# Wj, (Wa), GM, Kw

*m > ∅ / _#   W, GM, Kw
*n > ∅ / _#   (UK), Wj, (Wa), GM, Kw
*ŋ > ∅ / _#   UK, Wj, (Wa), GM, Kw
*p/β > m   AT, AX, (Tp)
*t/r > n   AT, AX, (Tp)
*k > ∅ / _#   W, GM, Kw
*k > ŋ / _#   (Wa), AT, AX, (Tp)
*k > γ / _#   Pt

*pʷ > kʷ   Pt, UK, W, GM, AT, AX, Kw, Tp
(*p > pʷ / _u) *pʷ > hʷ   Km
(*p > pʷ / _u) *pʷ > φ   Kb

*c > č/š GM
*c > s   Kw, Tb
*c > h   Pt, UK, AT, AX, Kw
*c > ∅   Km, W, Kb, GM, Tp

*ʔC > C   Km, UK, W, GM, AT, AX, Kw, Tp
*ʔV > ʔV   Km, UK, W, GM, AT, AX, Kw, Tp

*e > i / C   Pt
        [-ant]
        [-bk]

*a > ā   Tp
*o > a   Pt, AT, AX, Tp
*u > o   AT, AX, Tp
*ā > ɨ̄   Tp

The role of these sound change rules in the internal classification of the TG family will

be examined in Chapter 6.

1.4. Conclusion

We have used comparative data from basic vocabulary items from modern TG languages to reconstruct a basic lexicon of PTG and have traced the sound changes from PTG to the daughter languages. We have seen that, contrary to previous analyses, the comparative evidence leads to the reconstruction of only one affricate for PTG: *c. We have also seen from comparative evidence that the glottal stop *ʔ, though reanalyzed as a segment in most daughter languages, was originally an autosegmental feature. In the following chapter, we will learn more about these and other aspects of PTG phonology as we apply the method of internal reconstruction to the protolanguage.

PART ONE: PHONOLOGY
CHAPTER TWO: INTERNAL RECONSTRUCTION OF PRE-TG

2.0. Introduction

In this chapter we will consider the comparative evidence for the reconstruction

of the phonology of PTG as well as evidence within PTG for earlier stages in pre-TG.

As we shall see there has been very little change in the consonant systems of TG

languages, owing probably to the symmetry of the PTG consonant system. The same is

true of the vowel system. Again, most of the languages have maintained the same vowel

system as the parent language. In both consonant and vowel systems, however,

destabilizing changes have taken place, almost entirely in the Central Innovation Area.

These changes include both conditioned and unconditioned changes.

2.1. The Fortis/Lenis Distinction

In reconstructing Proto-Tupi-Guaranian (PTG) consonants, Lemle (1971) arrived

at the following system:[1]

Stops	p	t	k	?
Fricative	β			
Nasals	m	n	ŋ	
Flap		r		
Affricate		c		
Glides		y	w	

---

[1]The reconstruction is implicit in Lemle's word list. The consonant chart as shown below appears in Jensen (1987). Jensen slightly modified the system (adding č in particular) based on personal communication with Rodrigues.

This comes from a straightforward analysis of Tupi-Guaranian (TG) languages, but some interesting facts about this system and its reflexes in the daughter languages lend themselves to deeper analysis.

2.1.1. Preglottalization

In addition to the above consonants, Guarani and Parintintin have /ɣ/, but in all other languages this corresponds to /k/. This /ɣ/ patterns with /k/ in the same way that /β/ does with /p/ and /r/ does with /t/. /ɣ/, /β/, and /r/ never appear in initial position.[2] Such a relationship can explicitly be seen in examples such as the following:

t/r as in class II markers (see **4.1** for discussion on root classes)

1.   [tenda] 'place'   [še renda pe] 'at my place'
    [túβa] (OG [tuβ]) 'father'   [še rú] (OG [če ruβ]) 'my father'
    [tesa] 'eye'   [še resa] 'my eye'

p/β as in the enclitic *pe,* 'at, to' when preceded by an atonic pronominal marker:

2.   ók pe 'at the house'   še βe 'to me'
               teté pe 'on the body'   nde βe 'to you (sg.)'
               ǰawá pe 'to the dog'   ore βe 'to us (excl.)'

This coincides with the fact that, as shown in 1.2. above, *p* and *t* follow the same pattern with *β* and *r* respectively in correspondence sets.

First of all, since /r/ interacts with /t/ the same as /β/ does with /p/, we might put

---

[2]'Words' like *ramo* are no exception since they are clitics. Indeed, when *ramo* is used as a separate word (meaning 'if only...') it becomes *tamo.*

the two together in a new category, 'lenis'.[3] This, with other changes, would revise our reconstruction this way:[4]

PLAIN

Fortis	p	t	k
Lenis	β	r	
Nasals	m	n	ŋ
Affricate		č	
Glides			y	w

PRE-GLOTTALIZED

Lenis	ʔβ	ʔr	
Nasals	ʔm	ʔn	ʔŋ
Glides			ʔy	ʔw

This assumes that the lenis consonants are in complementary distribution with the stops. Nevertheless these pairs are separate phonemes as can be demonstrated with near-minimal pairs like /pɨtã/ 'red', /pirá/ 'fish', /aβá/ 'person', /apó/ 'deed'.

We can propose a rule which states that stops lenited medially in pre-TG. How then to explain all the medial stops? Consider the following possibility:

---

[3] I do not use these terms 'fortis' and 'lenis' in their usual phonetic sense. In this study 'fortis' and 'lenis' refer to a phonological categorization of consonants based on the merger of the features voicing and continuance; consonants are either [−vce] and [−cont] or [+vce] and [+cont].

[4] Glottal stop has been removed because this is now seen to be an auto-segmental phoneme; I reconstruct *č instead of *c.

$$
\left.\begin{array}{l} p > \beta \\ t > r \\ k > k \end{array}\right\} \ / \ V\_V
$$

$$
\left.\begin{array}{l} \text{ʔp} > p \\ \text{ʔt} > t \\ \text{ʔk} > k \end{array}\right\} \ / \ V\_V
$$

In other words, the medial stops are from an earlier pre-glottalized series. Recall that in the previous chapter we reconstructed preglottalization for all segments except stops. This reconstruction would remove that anomaly: the pre-glottalized stops became plain stops (fortes) in the mother language while originally plain stops became lenes.

If this is correct, then pre-glottalized stops could occur medially or at the beginning of words (there is no evidence for word-final pre-glottalized stops). Since β, r, and γ do not appear word initially, we can conclude that pre-glottalized stops fell together with simple stops in initial position.

An apparent exception appears in nouns beginning with t. Consider these examples from OG:

3.
                                             če resá < *tʸe tečá
                                             če tī < *tʸe ʔtī

However, this is not the case because the t/r alternation always occurs with Class 2 nouns

(see (4.1.) below) and should be analyzed as *tʸe-t-ecá.*[5] Such an environment might have allowed lenition. This is in fact what is going on, since /p/ never lenites in the same environment e.g. [pó], 'hand'; [še pó], 'my hand'. If that is the case, then the neutralization rule holds across the board, in which case the only way to discern initial pre-glottalized stops is to see whether the initial stop is vulnerable to nasalization, e.g. OG variants *po ~ mbo.* Unfortunately, such data are not always available.

Apparently, there is a pattern in Guarani that bilabials and dentals change first, then, trailing behind, the velars. This is evident in OG word-final consonants. These also eventually were dropped in a process which began before the historical period, but by that time it had not finished. So a 1640 vocabulary gives us a snapshot of the apocope in progress. We find virtually no words with original final *r* or final *β*, and a plethora with final *γ*. In other words, apocope of *r* and *β* was first and had essentially been completed by the historical period. There is almost no sign, however, that *γ* had even begun to drop out (*-eča* 'see' does appear alongside *-ečaγ*). Basically, *γ* dropped out during late OG. Precisely the same ratios are found for nasal stops of the respective points of articulation. It is this hierarchy — bilabial/dental > velar — which explains the historical manifestation of the lenition rule.

While G *r* and *β* generally correspond to the same phones throughout TG, G *γ*

---

[5] In TG, vowel-initial noun stems were prefixed with a class marker *t-* (see 4.1.).

corresponds to *k* elsewhere in TG. There is no γ in TG except in G and Pt. This conforms to the above hierarchy; dentals and bilabials had both undergone lenition early in TG (probably PTG, or even earlier) while velars did not lenite until after the break-up of TG, and then only in OG and Pt. This also indicates that the hierarchy is common to TG.

2.1.2. Pre-TG Word-Final Consonants

Note that word final consonants are always lenes. This indicates that pre-glottalized stops fall together only with simple stops word initially. The rules are:

1a. P $\Big\}$ > P / #__
    ʔP   __#

b. P   > lenis / V__V
   [−bk]

c. ʔP > P / V__V

2. P lenites finally

In this analysis pre-glottalized stops and simple stops fall together in a natural environment; the absence of fortes word-finally is what we would expect, since this reflects an expected absence of pre-glottalized stops word-finally.

2.2. Labialized and Palatalized Stops

The phonological system of PTG lends to another internal reconstruction. Consider the labialized and palatalized stops (fortes) in PTG.

$$
\begin{array}{lc}
p^w & k^w \\
(p^y) &
\end{array}
$$

The parent language has $*p^w$ and $*k^w$, but not $*t^w$. However, PTG does have $*\check{c}^w$.[6] Note also that $*p^y$ is attested for a very few words in PTG. These two anomalies in the system seem to be related. It is odd that an affricate ($\check{c}$) could be labialized while t could not, but other fortes could (p and k). This is strongly suggestive that $*\check{c}^w$ is in fact from $**t^w$. The reason for marginal existence of attested p^y is suggested by its development in most of the daughter languages as $\check{c}$ or c. It is likely that $**p^y$ and $**k^y$ parallel to $*p^w$ and $*k^w$ also existed and changed to $*\check{c}$ in PTG. The survival of one or two examples of $*p^y$ in PTG would be due to the fact that the change took place by lexical diffusion. If $**t^w$ was possible, then $**t^y$ makes sense also. We then have the following possibilities for Pre-TG:

p^w	t^w	k^w
p^y	t^y	k^y

The boxed area shows sources for PTG $*\check{c}$. All palatalized stops became $\check{c}$ (with the exceptions in p^y) and t + glide always became $\check{c}$: $**t^y > *\check{c}$, $**t^w > *\check{c}^w$.

This internal reconstruction opens up new possibilities for comparison with non-TG

---

[6] The only examples of $*\check{c}^w$ in PTG are suffixes and so do not appear in the list of roots in (1.1.). It did not appear in earlier reconstructions for the same reason that p^w and k^w did not, because these have been construed by Rodrigues and Jensen (for no clear reason) as being sequences in PTG, pw, kw, and $\check{c}w$ respectively and as unit phonemes (p^w, k^w, and $\check{c}^w$) in all attested languages.

languages of the Tupi stock. If *č is originally from palatalized stops, then we should seek comparisons with these stops rather than with fricatives or affricates. For example,

4. a. 'go'   PTG *čó   Surui ka   Pre-TG **kʸo
   b. 'eye' PTG *ečá Tupari epa   Pre-TG **epʸa
   c. 'I'   PTG *če   Kayabi te   Pre-TG **tʸe

The first of these is somewhat questionable; Sataré-Mawé, which is much more closely related to TG, has *to* 'go'. This either means that Surui *ka* is not a cognate, or that the shift of all palatalized stops to *č began before the break between Sataré-Mawé and PTG. It's a close call to make: Sataré-Mawé is indeed so closely related to TG that it was until recently thought to be a TG language; on the other hand, affrication of palatalized stops is also very recent, since isolated lexemes like *epʸák* 'see', show that even by the breakup of PTG, affricatization had not completely eliminated the palatalized stops. In any case, the second example is rock solid, being further confirmed by PTG *epʸák* 'see'.

2.3. The syllable as domain of pre-glottalization

We saw above the following pre-glottalized consonants can be reconstructed for PTG:

Lenes   ʔβ   ʔr
Nasals   ʔm   ʔn   ʔŋ
Glides   ʔy   ʔw

Note that this constitutes the entire PTG consonant inventory with the exception of stops and *č. This strongly supports the hypothesis that fortes in PTG were derived from pre-glottalized stops. But if it is true that PTG *p and *t are from pre-TG **ʔp and **ʔt respectively, and that PTG *β and *r are from pre-TG *p and *t respectively, then where did PTG *ʔβ and *ʔr come from? The answer is that *ʔβ and *ʔr only occur as a result of a morphophonemic rule that still exists in Pt and Kb. This rule has been called 'metathesis' by Jensen (1990) because glottal occlusion has been regarded as a segment in PTG.

$$[+\text{con}] \quad \rightarrow \quad \begin{bmatrix} -\text{syl} \\ -\text{con} \\ -\text{vcd} \end{bmatrix}$$

1   2

12   >   21

Since, as we have seen in the previous chapter, ʔ is an autosegmental feature, metathesis is not an appropriate explanation. To see the real explanation we need to observe the syllable divisions in words that exhibit this morphophonemic rule:[7]

5. ipit 'skin'   ≠ ʔók 'take off'   >   /i pi ʔrók/ 'to skin'
    tiŋ 'white'   + ʔí 'diminutive'   >   /ti ʔŋí/ 'small white thing'
    ʔɨβ 'tree'   ≠ ʔá 'fruit'   >   /ɨ ʔβá/ 'fruit off the tree'
    kam 'breast' ≠ ʔú 'consume'   >   /ka ʔmú/ 'nurse'

---

[7]The symbol כ indicates a morpheme boundary between two compounded roots; the symbol + indicates a morpheme boundary between a root and a bound morpheme.

As may be expected, languages which interpret the glottal stop as a segment do not exhibit this rule, including early languages like OG:

6.   a + i + potáβ + ʔóγ   > aγpotaβóγ   (RM Tes. p. 321v)
     1s-1 3-2 part separate   'break into pieces'

It is clear from these data that ʔ cannot occur anywhere other than at the beginning of the syllable. If ʔ is restricted to syllable-initial position, then we must conclude that what were preglottalized in PTG (and in Pt and Kb) were not segments at all, but syllables. In these languages, then, there are no preglottalized phonemes; rather, there is an autosegmental phoneme called preglottalization, which has the syllable as its domain, in just the same way as nasality and orality are autosegmental phonemes which have the word as their domain (2.4 below). This being the case, the examples above don't exhibit a morphophonemic rule at all, at least not in a transformational sense. Rather, they are the direct consequence of an inherent property of the preglottalization phoneme. They could more accurately be represented this way (O=orality):

```
     O              O              O
     |              |              |
     μ              μ              μ
     |              |             / \
     σ             ʔσ           σ  ʔσ
     |              |           /\  /\
    CVC             V          C VC  V
    | | |           |          | | |  |
    k a m    +      u     >    k a m  u
    [kam]          [ʔu]        [kaʔmbú]
```
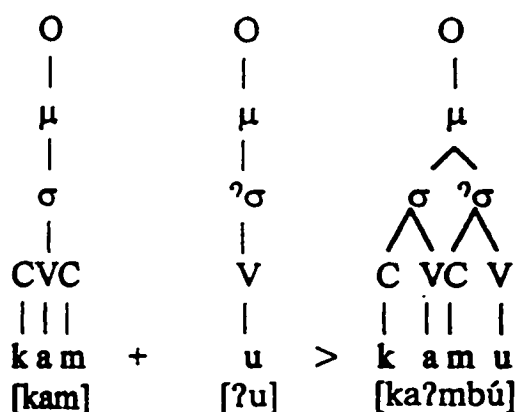
2.4. Nasalization

In addition to the sounds reconstructed for PTG there is an autosegmental phoneme to be reconstructed as well, based on a fact of Tb and OG phonology. In these languages, if the last syllable in a word is a nasal vowel then all vowels in the word are nasal. Likewise, if the last syllable in a word is an oral vowel, then all vowels in the word are oral. Since the stress in most TG languages falls on the last syllable, it seems reasonable not only to reconstruct ultimate stress for PTG, but to draw a connection between stress, nasality and orality. Rodrigues has done this by reconstructing two kinds of stress in PTG, nasal and oral stress. This is further substantiated by the occurrence of nasal and pre-nasalized stops in TG languages.

Jensen (1990) discusses the rule that determines the occurrence of prenasalized stops. The rule is that prenasalized stops are the oral allophones of nasal stop phonemes, i.e. /N/ > $_{\text{[}}$ND] / _ V[−nas]. In addition to this there is a rule of leftward nasal assimilation within roots.

The rule is as follows: a nasal phoneme is phonetically a prenasalized stop if it immediately precedes a vowel with oral stress.

7. /ené/ > [endé]   'you (sg)'   (J 1990a:52)
    /namí/ > [nambí]   'ear'

If a nasal segment is in a syllable followed by a syllable with oral stress, but two or more syllables to the left of the stressed syllable, the nasal segment may be prenasalized.

8. /maraká/ > [maraká] > [mbaraká] 'cat'   (J 1990a:52)
    /noúri/ > [noúri] or [ndoúri]   'he did not come'

Jensen (1990) states that this second rule may have been obligatory originally because of data from GM and Pt which, by phonological and morphological comparison to OG and Tb, show themselves to be two of the most conservative TG languages:

9. a.   /namí/   [nambí]   *[namí]   'ear'   (J 1990a:52)
   b.   /mo'é/   [mbo?é]   *[mo?é]   'teach'
   c.   /mokõy/   [mõkõy]   'two'

In other TG languages, the tendency has been to weaken this rule. Apart from the Tb examples above, W has completely dropped the second half of the rule and applies the first only optionally:

10. /momá?e/ > [momá?e] ~ [mombá?e]   'thing'   (J 1990a:52)
     /aŋé?e/ > [aŋé?e] ~ [aŋgé?e]   'now'

The same holds for UK (Kakumasu 1986:399):

11. /mói/ > [mói] ~ [mbói]   'snake'
     /itaŋwá/ > [itaŋwá] ~ [itaŋgwá] 'his buttocks'

Interestingly, languages that have lost autosegmental nasality, Asurini, Tapirape, and Guajajara (Bendor-Samuel 1972:74) have lost this rule altogether and have only one allophone (nasal) for each of the nasal consonants.

In light of the above, it should not be surprising that OG has the same rule as GM and Pt, supporting Jensen's proposal that the rule as it holds in GM and Pt was the original form of this rule in the proto-language.

The morphophonemic treatment of nasals in Tb seems to point to the original source of prenasalized stops in TG.

The nasalization rule for Tupinamba as stated by Rodrigues (1981) and Jensen (1990) is as follows:

$$
\begin{bmatrix} +cons \\ -vce \end{bmatrix} > [+nas] \;/\; [+nas] \left\{ \begin{bmatrix} -acnt \\ +syl \\ +cons \\ -syl \end{bmatrix} \right\} \begin{Bmatrix} + \\ \neq \end{Bmatrix} \; \underline{\quad} \; \begin{bmatrix} -cons \\ -syl \\ +vce \end{bmatrix} \begin{bmatrix} \begin{bmatrix} +syl \\ -nas \end{bmatrix} \begin{bmatrix} -syl \\ -nas \end{bmatrix} \end{bmatrix}_0^n
$$

Examples of this rule in Tupinamba (J 1990a:59)

12.

nupã 'beat'	katu 'good'	nupãŋatú 'beat well'	[nupãŋgatú]
mo 'causative'	só 'go'	monó 'send'	[mondó]
yũ 'field'	pe 'in'	yũme 'in the field'	[ñũme]
pɨtun 'night'	pe 'in'	pɨtũme 'in the night'	[pɨtũme]
mo 'causative'	pór 'boil'	momór 'make boil'	[mombór]
(e)mi 'passive'	pʷáy 'give orders to'	mimʷáy 'employee'	[mimbʷáy]

This rule, as Jensen explains, no longer holds for Wayampĩ. It survives in frozen forms like the following (J 1990a:60):

13.

t-emi 'pass.'	kwáy 'order'	temiŋ^wáy 'subordinate person'	([§]temĩk^wáy)
mo 'caus.'	có 'go'	monó 'send'	([§]moó)
mo 'caus.'	pór 'jump'	momó 'throw'	([§]mopó)

This contrasts with the productive rule of nasalization exemplified in the forms below:

14. emi   +   pete   emipete   'person or thing beaten with
     pass. 'give a slap'   the hand'

     mo   +   popo   mopopo   'throw away'
     caus.   'jump
             repeatedly'

What Rodrigues and Jensen fail to point out, however, is that even in Tupinamba this rule is contradicted (Anchieta 1595:49):

15. mo 'causative'   +   p^wám 'arise'   mop^wám 'make arise'
     mo 'causative'   +   sém 'to leave'   mosém 'make leave'

This shows that the rule stated by Rodrigues was already non-productive in the historic period. The forms listed in (13) were anomalous in all languages throughout the historic period. The phonological process which generated them therefore dates back to PTG or even earlier. In light of this it may be more productive to seek a diachronic explanation even for the Tupinamba forms, and therefore for PTG (not just, say, W and other modern languages). In fact, data from OG suggest that the rule by Rodrigues did

not exist prehistorically, but rather that they are relics of a sound change within PTG, and the examples that seem to exemplify Rodrigues' rule date from the time of this change.

One could easily explain this sound change by recognizing, in light of the OG data above, that /mo/ was originally /mbō/, /emi/ was originally /embī/ or /ēmbī/, etc.

16.   mbō + karú > mbōŋgaru > moŋgarú
                  'caus' 'eat' 'feed'   'feed'

                  embī + pwáy > embīmbwáy > emimbwáy
                  'pass' order 'someone   'someone
                                  ordered'   ordered'

Implicit in this hypothesis is the idea that all prenasalized stops in PTG were originally not /N/, or even /ND/, but /T/ followed by a vowel with nasal accent.

The reason there is no *nd-* initial variant is because all *t-* initial words were, by the historical period, analyzed as composed of the stem classifier *t-* plus a vowel-initial root. Hence there were few or no root initial t's in OG. The few that do exist seem to have been from pre-glottalized stops, e.g. *tĩ* 'nose', where the presence of a fortis immediately before a nasal vowel points back to **?tĩ* (***tĩ* would have become *tí* or *ní* [ndí]). In fact there are cases of nasalization of *t* in this environment, but the morphemes in question are all enclitics. Consequently, the initial *t-* only appears as an *r* so that the rule formulated by Rodrigues (1981) and repeated in Jensen (1990a:69) is for the nasalization

of *r*.

17. tĩ + rám   >   /tĩnám/ [tĩndám]
     nupã + reme   >   /nupãneme/ [nupãneme]
     irũ + ramo   >   /irũnamo/ [irũnamo]

The best explanation for why word-initial *k* was not nasalized in this way is that it was

not subject to lenition and *ŋ* was restricted to medial and final position.

This also explains Jensen's morphophonemic rule #4. This rule states that, in

compounding, a root-initial *p* lenites to *β* before a root-final lenis.

18. ʔáβ≠pukú   >   /ʔáβukú/   'long hair'   (J 1990a:63)
     okár≠pɨtér   >   /okaβɨtér/   'the center of the clearing'

The same reasons for the restriction of word initial nasalization account for why this

lenition is observed for *p* only, and not for *t* and *p*.

Hence /maniʔók/ [mandiʔók] 'manioc' is originally from /mãtĩʔok/ [mãndĩʔɔ́k].

The loss of pre-tonic nasality led to the phonological situation where [m] only occurred

before nasal vowels or consonants, [mb] only before the absence of both. This led to the

reanalysis of [mb] as the oral allophone of /m/.

OG itself contains some other data that add to the problem of nasalization in TG.

In Ruiz de Montoya, we find doublets like the following:

19. põrãŋ ~ mõrãŋ   'beautiful'
     pĩtãŋ ~ mĩtãŋ   'small'

What could be the cause of the alternation p~m? The most obvious answer is that the nasal consonant produces a nasal accent, which nasalizes the initial consonant. This means there existed in OG a rule of optionally nasalizing an oral stop in the same environment that preserves a nasal consonant.

The picture is further complicated, however, by such examples as the following:

20. po ~ mbo   'hand'
     pɨ ~ mbɨ   'foot'
     pɨtá ~ mbɨtá   'remained'
     pɨtú ~ mbɨtú   'steam','mist'

The alternation p~mb is also attested in Tupinamba. What might explain the nasalization of initial stops when no nasal phoneme is present? One explanation might be the random nasalization of consonants. But it seems strange to explain that some consonants were nasalized by a regular rule while others were nasalized randomly. And there seems to be a pattern in the fact that *p* alternates with *m* in a nasal environment and with *mb* in an oral environment. Might these be two manifestations of a unified rule?

One such rule could propose that at an earlier stage, PTG had an alternation like this:

Stage 1   po ~ põ
      poraŋ ~ põrãŋ
      pitaŋ ~ pĩtãŋ

This resulted in the nasalized form of *p*, i.e. *mb:*

Stage 2   po ~ mbõ
      poraŋ ~ mbõrãŋ
      pitaŋ ~ mbĩtãŋ

Later, there was a general reduction in the nasality of vowels such that vowels no longer

remained nasal after *mb, nd,* etc. Vowels would remain nasal, however, if there

were a nasal consonant at the end of the word.

      po ~ mbo
Stage 3   poraŋ ~ mbõrãŋ
      pitaŋ ~ mbĩtãŋ

In the environment of nasal stress or a final nasal consonant, vowels remained nasal and

further nasalized prenasalized stops, hence:

Stage 4   po ~ mbo
      poraŋ ~ mõrãŋ
      pitaŋ ~ mĩtãŋ

By this time the nasalization of *p* to *mb* would have become non-productive.[9]

---

[9] Tempting as it may be, it is far from certain that there is a connection between OG *mbó* '(someone's) hand' and the causative prefix *mbo-* (/mo/).   See 6.1 about this prefix.

This would also explain certain verbs attested from OG and still in use in modern TG languages:

21.   moŋgarú   'to feed'
                                moŋgaraí   'to baptize'

These verbs are composed of /mo-/ and *k*-initial roots /karú/ and /karaí/, but this should result in *mbokarú* and *mbokaraí* respectively. However, if the real makeup was actually *pō̃ + karú* the regular development would have been:

22.   mbō̃ + karú
       mbō̃ŋgarú
       mõ̃ŋgarú
       moŋgarú

This denasalization of vowels can be seen as part of a long-term process that has been progressive through the history of TG languages.

One possible reason for the loss of nasality on pretonic vowels involves the origin of nasal stress itself. What is the origin of two kinds of stress, oral and nasal?

Most attested Tupian languages have nasal vowels; in other words, nasality has the syllable as its domain. Pre-TG was probably the same. This means that the domain of nasality expanded from the syllable to the word in the same way that the domain of tone expanded from the syllable (as in a tone language) to the word (as in pitch-accent and stress languages). Pre-TG acquired a rule that assimilated the nasality or lack of

same of the stressed syllable to the unaccented syllables. This means that the shift from tone to stress occurred first, while there still existed contrasting series of nasal and oral vowels; this situation still exists in many Tupian languages outside TG, including those closest to TG (Sataré-Mawé, Awetí, and Sirionó).

It is at this stage of oral and nasal autosegments at the syllable level that PTG would have been at the time of (Stage 1) and (Stage 2). When PTG applied stress at the word level instead of the syllable level, the language acquired the rules of leftward nasal assimilation and further nasalization of prenasalized stops, which made the re-analysis of prenasalized stops inevitable (Stage 3 and 4).

The doublets listed above survived in MG just as other dialectal doublets did, with a semantic restriction on one of the variants:

23.
porãŋ	→	porã 'beautiful', 'good'
morãŋ + katú	→	maraŋgatú 'righteous', 'holy'
pitãŋ	→	no reflex in Guarani (but note, for example, Kamayura *pitaŋ* 'child')
mitãŋ	→	mitã 'child'

If this hypothesis is correct, TG languages have continued the reduction in scope of autosegmental nasality that began from preTG to PTG: Tapirape has lost autosegmental nasality while Asurini and Guajajara have lost all phonemic nasality in the vowel system.

The only question remaining is, if this hypothesis is correct, why are there medial

t's in nasal environments? In other words, how were forms like *mūāŋ* possible (as opposed to *mindáŋ* or *mīnāŋ*? The reason relates to the lenition rule discussed above. At this stage in pre-TG fortis consonants were still pre-glottalized stops, or perhaps geminates (at least phonemically)[10] while lenis consonants were still stops (at least phonetically). Simple stops nasalized to prenasalized stops while pre-glottalized stops did not.

Another question is raised by words like **irū*. If this derives from earlier ***itū*, this should have become **indú*. Such cases are easily handled if PreTG ***r* existed as an independent element that merged with ***t* as PTG **n* in a nasal environment while ***ʔr* became **r* in a nasal environment. Hence, ***matī* became **mandí* while ***iʔrū* remained **irū* (this also explains why **porāŋ* above (< ***poʔrāŋ*)was not **ponāŋ*).

2.5. Conclusion

Based on the observations in (2.1.) above, our revised reconstruction of PTG sound system is as follows:

CONSONANTS

Fortes	p	t	k
labialized	p^w		k^w
palatalized	p^y		
Lenes	β	r	

---

[10]In Tembe, Urubu-Kaapor, Kamayurá (Saelzer 1976:135), and perhaps other TG languages, medial stops at the onset of stress are still realized phonetically as long, though they are no longer phonemically geminates.

Nasals	m	n	ŋ
Strident		c	
labialized		cʷ	
Glides		y	w

VOWELS

```
  i       ɨ     u
     e       o
        a
```

AUTOSEGMENTS: ʔσ μN μO

Having examined the sound changes from PTG to the daughter languages, and reconstructed pre-TG phonology from PTG via internal reconstruction, we can frame the following hypotheses about pre-TG:

1. Preglottalization was an autosegmental feature whose domain was the syllable. This continued to be the case in PTG.

2. Preglottalized stops became PTG stops, possibly passing through a stage of gemination and degemination, while Pre-TG stops became lenes medially, merging with preglottalized stops at word boundaries. This resulted in the fortis/lenis distinction of PTG.

3. Pre-TG had three series of stops: plain, labialized, and palatalized.

4. Pre-TG had two series of vowel phonemes: nasal and non-nasal.

5. Nasalization of consonants by vowels led to a greater domain for nasality (and orality) in PTG, from the segment or syllable to the morpheme. Gradually this domain has diminished in the daughter languages.

6. Prenasalized stops arose, at least in part, from the Pre-TG nasalization of stops; the assignment of nasality to the domain of morpheme in PTG led to the reinterpretation of all prenasalized stops as oral allophones of nasal consonants.

The Pre-TG sound system can be summarized as follows:

CONSONANTS

Stops	p	t	k
labialized	p^w	t^w	k^w
palatalized	p^y	t^y	k^y
Nasals	m	n	ŋ
Flap		r	

VOWELS   i   ɨ   u   ĩ   ɨ̃   ũ
          e   o   ẽ   õ
           a   ã

AUTOSEGMENT: ʔσ

This reconstructed system is conspicuous for its lack of fricatives of any kind. Given the general trend of lenition from pre-TG to PTG (as well as lenition from PTG to the daughter languages), it is possible that Pre-TG had fricatives which are unrecoverable because they have disappeared without a trace in attested TG languages, perhaps in PTG as well.

We have seen that the conclusions drawn from this internal reconstruction show promise for aiding in the continued comparative study of TG and non-TG languages of the Tupi stock. We now turn to another type of comparative evidence: comparing synchronic rules of phonology and morphophonology in modern TG languages to see what these can tell us of the phonology of PTG.

PART 1: PHONOLOGY
CHAPTER 3: COMPARISON OF MORPHOPHONOLOGICAL RULES

3.0. Introduction

In Chapter 1 we examined comparative evidence for the PTG sound system and from this comparative evidence, established sound change rules for the daughter languages. From these sound change rules, we have arrived at a lexicon of PTG roots. In this chapter we will look at morphophonological rules in the various daughter languages and see which rules may apply to PTG applying the revised reconstructed sound system of PTG arrived at in Chapter 2.

What we shall see here is that the morphophonological rules that exist are not themselves inherited, but are the interpretive result of forms inherited from the preceding stage of the language. In the diagram below, x = any free morpheme and y = x + some other morpheme. At Stage 1 y_1 is a transparent combination of morphemes without any morphological rules.

$$
\begin{array}{lcc}
\text{Stage 1} & x_1 & > & y_1 \\
 & \downarrow & & \downarrow \\
\text{Stage 2} & x_2 & > & y_2 \\
 & \downarrow & & \downarrow \\
\text{Stage 3} & x_3 & > & y_3 \\
\end{array}
$$
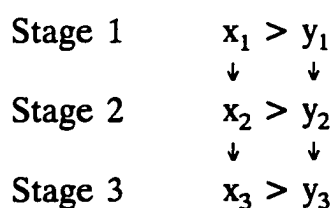
Diachronic rules apply to both forms, but apply differently because of the differences in phonological conditioning. The result is Stage 2. Here y_2 is not totally transparent as a combination of two morphs, but is recognizable as such with some differences. The

synchronic interpretation of these differences is called a morphophonological rule. The process continues until either y is no longer recognizable as a derivative of x and becomes suppletive, and/or an analogy with another form reshapes y. A third possibility is that y is replaced altogether with a truly suppletive form.

3.1. Rodrigues' morphophonological rules

Jensen (1990) describes nineteen morphophonological rules formulated by Rodrigues (1981) for Tb based on texts in that language. These nineteen rules will form the basis of this comparative study. We will summarize the rules and describe how each language manifests a similar rule.

1. A stop (fortis) is always nasalized when compounded to a preceding root ending in a nasal consonant or vowel.

This is the synchronic interpretation of what has been shown in Chapter 2 to be the result of a diachronic process of nasal assimilation before the diachronic change of nasality spreading from the domain of segment to that of word. In fact the nasalization in rule #1 shows that nasality was expanding beyond the morpheme boundary to the whole word, though this is not completely 'achieved'.

(1)   nupã + katú > nupãŋgatú   (J 1990a:59)
        'beat' 'good'   'beat well'

This nasalization of fortes is decidedly rightward, as with the second element in the

above compound and the compounds below, but it also works leftward, as with the causative prefix in the following examples (J 1990a:59).

(2) a.   mbo + có > mbondó > mondó 'send'

    b.   mbo + karú > mboŋgarú > moŋgarú 'feed'

    c.   nupā + katú > nupaŋgatú 'beat soundly'

The doublets discussed in Chapter 2 (e.g. *pitāŋ ~ mitāŋ*) clearly show a leftward nasal assimilation. Whether or not these rules are valid as synchronic rules, they reflect a diachronic reality: that there was at one stage of PTG a rightward nasal assimilation and later a leftward nasal assimilation.

The fact that nasalization has gone in both directions at different stages, means that these are two parts of a greater phenomenon that has taken place, namely, the process of nasal assimilation. The nasal assimilation is itself similar to the metathesis of *ʔ. On the surface at least, the metathesis of *ʔ is something that does actually take place, but it is more accurate to describe it as the operation of preglottalization having the syllable as its domain. Similarly so-called nasal assimilation is really not so much nasal assimilation as it is diachronically the acquisition of the morpheme as the domain of nasality. Nasality in Pre-TG had the segment as its domain and then later acquired the morpheme as its domain. This produced a phonetic effect that has been perceived and described as nasal assimilation. What is actually happening is that the feature nasal is taking over the whole word, or at least the whole morpheme, rather than just being

restricted to a single segment. In PTG a word has nasality or it does not; if it has it at all the whole word is nasal, resulting in what could be described as the nasalization of all the segments in the word. This is the assimilation that has taken place.

Rules of nasal assimilation obscured the underlying nasality of every syllable in a word except for that of the stressed syllable. The status of the stressed syllable, therefore, determined that of the whole word: if it was nasal, the whole word was nasalized; if not, the whole word was oralized.[1]

There is reason to suspect that nasality first acquired the domain of the syllable because there is no word in PTG having a nasal vowel in the stressed syllable, but ending in an oral consonant. There is no word like §*pōr* or §*kāβ*. If the last vowel in a word is nasal, any final consonant will be nasal. And, as we have already seen, if a word ended in a nasal consonant, that same syllable had a nasal vowel as well. So the final syllable is either all nasal or all oral. It could be that some words with word final nasals are cases of nasality spreading rightward from a nasal vowel. In any event, nasal assimilation amounts to the autosegment of nasality or orality applying to the whole word.

---

[1] This does not mean that *m, n, ŋ* ever became *p, t, k* or *b, d, g*. But they came close, becoming prenasalized stops, or perhaps here more accurately postoralized nasals, *mb, nd, ŋg,* and in fact in word initial position, these were phonetically voiced stops.

2. If a root ends in a nasal consonant [having an oral accent] and is compounded with a following root, the accent of the first root will become nasal.

The second rule pertains to the nasalization of a morpheme's accent. The first thing to be said about this is that it is assumed by Rodrigues and also by Jensen that words ending in nasal consonants in Tb and therefore in PTG had oral stress. The probable reason for this is that Anchieta in his grammar does not mark such words with nasal stress but with either an acute accent or a circumflex accent, both used for oral stress. The fact is that Rodrigues never indicates nasal stress orthographically when the last syllable contained a nasal consonant. The only place where Anchieta consistently indicates nasal stress is when it occurs in the absence of a nasal consonant .

(3) a.   yamanô   /yamanō/   'we die'   (A 1595:6)

    b.   açô   /asó/   'I go' (A 1595:6)

    c.   monhàng   /moñáŋ/   'to create' (A 1595:26)

    d.   paranâ   /paranã/   'river' (A 1595:4v)

    e.   nupã   /nupã/   'beat' (A 1595:3)

    f.   irũ   /irũ/   'friend' (A 1595:6)

In all of these examples we know, both from comparative evidence and from internal evidence in Tb, that the PTG word had nasal stress. Anchieta does not indicate nasality, however, because nasality is apparently already clear from the presence of a nasal consonant in the final syllable. And since nasals and prenasalized stops are distinguished

orthographically in Anchieta, he is right about that: the only time a nasal consonant (as opposed to a prenasalized stop) occurs in a final CV syllable is before a nasal vowel. If /manó/ 'die' were obligatorily [mandó] in Tb, then /mén/ 'husband' should also be [mbén]. It never is; in a final syllable of the structure NVN, the first nasal is never a pre-nasalized stop in Tb or in any other TG language. This indicates that the stressed words ending with nasals are nasal as well. So that just as for example, /manõ/ is [mãnõ] so also /mén/ was [mẽn]. One perhaps could argue that the stress was oral and the initial nasal consonant was kept nasal by the final nasal consonant. But there is no other instance in TG of a nasal consonant nasalizing a previous consonant without the intervening vowel(s) also being nasalized. It seems clear, then, that stress was always nasal in words ending in a nasal consonant. In fact the same is true of vowels between nasal consonants in pre-tonic syllables, for example [mõŋgarú] /moŋarú/ 'feed'.

This can be seen by looking at comparative evidence as well. Consider the Wayapí examples in (4).

(4)   akãŋ   'head'   (J 1990a:61)
        amãn   'rain'

All languages that still have word-final nasals and still have autosegmental nasality at all or even traces of it still have nasal vowels next to word-final nasal consonants. I say 'traces of' because of Tapirape, which does not have autosegmental nasality, but does have relics of it. For example the Tp word for 'head' is *ãkɨ̃ŋ* and in AT *akɨŋa,* which

point back to *akāŋ (*akáŋ would have yielded Tp ākāŋ and AT akaŋa). Similarly, Tp amin, AT amina, AX amina, all from *amān, etc.

Anchieta does indicate nasality on vowels which underlyingly precede nasal consonants only when they are later followed by oral vowels. This is why both Rodrigues and Jensen say that the nasalization of a vowel before a nasal consonant only occurs when the word is compounded with another morpheme. An example of this is seen in (5) below.

(5)   akāŋ ≠ méβ > /akāméβ/ [akāmbéβ]   (J 1990a:61)
        head   flat   'flat head'

The only reason nasalization is acknowledged in the literature is because it was explicitly indicated by Anchieta. But the only reason he did so was that the nasal vowel was followed by oral stress and therefore is not predictable.

We can see, then, that vowels were nasal not only before morpheme-final nasals in compounding, but before word-final nasals as well. This means that, contrary to Jensen's and Rodrigues's claim, the nasalization rule in Tb did not become more general in later languages, but was always general.

3. A vowel ɨ is inserted between a non-labial final consonant and the initial consonant of a following suffix.

(6)   ʔár + βo >   ʔáriβo   (J 1990a:62)
        day distributive   'daily'
             loc.

This rule only applies when the first consonant is not labial because otherwise the haplology rule would apply. This rule, not mentioned in Jensen (1990), is that when a labial is followed by a vowel and another labial, the first labial and following vowel are deleted as seen in (7).

(7)   páβ + eme > páme   (A 1595:26)
                  finish when

This strongly supports the argument that the nominal case is simply an epenthetic vowel added to the end of consonant-final words to break up the consonant cluster between words. Basically what we have here is the insertion of ə, which is raised to i at a morpheme boundary and lowered to a at a word boundary. More will be said about 'nominal case' in Chapter 4.

4. Initial labial consonants p and m are spirantized (become β) when compounded to a preceding root ending in a non-nasal continuant (i.e. a lenis).

The fourth rule, the spirantization of labial consonants, is basically the lenition of p. What is really going on there has already been discussed (2.1), i.e. the simple lenition rule, that when a lenis is in contact with a fortis consonant the fortis will assimilate by

becoming lenis.

(8) okár + pɨtér >   okaβitér   (J 1990a:63)
village middle=of 'middle of the village'

ʔáβ + pukú > ʔaβukú   (J 1990a:63)
hair   long   'long hair'

This rule also applies to nasals:

(9) kwáβ + meʔēŋ > kwaβeʔēŋ   (J 1990a:63)
pass   give   'hand over'

Here the continuant feature is transferred to the nasal *m* with the redundancy feature

of [–nas].

5. A root obligatorily loses its final consonant when the following compounding element

is consonant-initial. A word-final consonant is optionally dropped when the following

word or clitic is consonant-initial.

6. When two roots are compounded, the first ending in an obstruent, the second

beginning with a glottal stop, the two consonants metathesize.

As has already been discussed, this is really not a metathesis, but the manifestation of

the fact that preglottalization is an autosegmental phoneme which has the syllable as its

domain.

7. The initial vowel of -*áβo* and -*áβ* assimilates to a preceding mid-vowel.

This is a rule which applies only to the suffixes -*áβ* and -*áβo* (really *-*áβ-βo*) and even then, except in a few instances (**4.2.1**), is an intermediate rule with no overt manifestation in the data. An exception comes from those vowel-final verbs which take -*áβ* and -*áβo* (**4.2.1**).

8. A nasal root nasalizes a labial continuant (lenis) in a suffix.

9. A nasal root will nasalize a suffix-initial *r*.

Rules 8 and 9 are actually two parts of the same rule: rule #9 is for *r* [–lab] and rule #8 applies in a slightly different environment than rule #9. Because Rodrigues does not use the categories of fortis/lenis, he does not recognize that the consonants of the two rules belong to the same class.

The formula for rule #9 in Jensen (1984) differs from that of #8 in the following respects: it is claimed that the spread in #9 can cross theme boundaries as well as morpheme boundaries. The only evidence given for this, however, is *ñ* ≠ *rám* and there is no reason whatsoever for thinking that *rám* was anything more than a stressed suffix. Other suffixes bear stress (*-*áβ* and *-*áβo*). There seems no support therefore for setting up any other boundary than that of the morpheme.

$$8 \; / \quad \begin{bmatrix} \text{+syl} \\ \text{+acnt} \\ \text{+nas} \end{bmatrix} \quad + \quad \begin{bmatrix} \text{+syl} \\ \text{+acnt} \end{bmatrix} \text{---}$$

$$9 \; / \quad \begin{bmatrix} \text{+syl} \\ \text{+acnt} \\ \text{+nas} \end{bmatrix} \quad + \quad \text{---}$$

In other words, #8 'needs' a stressed vowel between the morpheme boundary and the target vowel, while such a vowel 'cannot' occur in Rule #9. In fact there seem to be no data to show that the presence of such a vowel would prevent #9 from applying, or that the absence of such a vowel would prevent #8 from applying, and in any case the role of stress in the application of the rule is nil. We can then conflate these two rules thus:

$$\begin{bmatrix} \text{+cons} \\ \text{+vce} \\ \text{-nas} \end{bmatrix} \quad > \quad [\text{+nas}] \quad / \, C \begin{bmatrix} \text{+syl} \\ \text{+acnt} \\ \text{+nas} \end{bmatrix} + \; Ci \text{ ---}$$

This would capture the fact that r behaves the same as β with respect to nasalization, as we would expect in a f/l language.

Given the nature of nasality in TG (autosegmental, on the level of the morpheme) the rules #8 and #9 are best further reformulated as follows:

$$\begin{array}{cc} \text{N} \quad \text{O} & \text{N} \quad \text{N} \\ | \quad | & | \quad | \\ \mu + \mu & \mu + \mu \end{array} >$$

Rodrigues (1981) and Jensen (1984) show this as nasalization of the sonorant only as in

(10) below:

(10)   kɨtī   + áβo > kɨtī + ámo   (J 1990a:68)
           'cut'   'nom.'

while the surrounding vowels remain completely oral. This is simply not the case. We

have seen that nasality was not marked with a diacritic but was understood in Tb when

a vowel was adjacent to a nasal. Further, synchronic evidence in modern TG languages

bears this out also: Kb *āmū* (Dobson), OG *rāmõ* (Montoya), Pt *āmõ* (Pease). We see

then that in a case of nasal assimilation, an entire morpheme was nasalized, not just one

segment of that morpheme. This is what we would expect, given the nature of nasality

in TG, which Rodrigues and Jensen themselves point out is autosegmental (of the

languages cited above Kb does not have autosegmental nasality but the fact of nasal

vowels in the suffix *āmū* points back to original N).

$$
\begin{array}{c}
\mu \\
\sigma \quad \sigma \\
V \quad C \quad V \\
a \quad m \quad o
\end{array}
$$
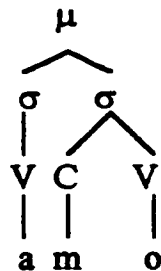
This combines with two other nasalization rules: nasalization of *r* (#9), which is a lenis,

and leftward nasalization of lenes (#13). They seem to have slightly different ways of

operating, but what they demonstrate is that nasal assimilation has spread in two

directions, rightward and leftward. Nasalization of *r* is the same as that of *p*; it's just that

word initial *t* is a class marker, which never occurs in compounding, so the only *t*'s are

those at the beginning of clitics, which, being clitics, never are realized phonetically as

t but as *r*.

(11) a.   tĩ + rám   tĩnám 'nose-to-be'   (J 1990a:69)
                         'nose'   'fut.'

     b.   nupã + reme   nupãneme 'when beating' (J 1990a:69)
                         'beat'   'when'

     c.   irũ + ramo   irũnamo 'as a companion' (J 1990a:69)
                         'friend' 'as'

10. When two stressed morphemes are compounded, the first one loses its stress.

11. When a root loses its stress because of rule #10 and that root ends in a high vowel, that high vowel becomes a glide.

12. When like non-high vowels are conjoined through compounding or affixation, one vowel elides.

13. A nasal root may nasalize a continuant (lenis) in a prefix.

Leftward assimilation of lenes refers to prefixes. So eβo + nã > emonã

(12) a.   ero + ĩn > eroĩn ~ enoĩn (J 1990a:72)
                com   be 'be with'

     b.   ero + sẽm > enosẽm   (J 1990a:72)
                com 'leave' 'leave with'

(12)   c.   ero + puʔãm > enopuʔãm   (J 1990a:72)
             com   stand up 'stand up with'

Note that, traditionally, the verbs in the above examples would be written with oral

stress. But it is hard to see how the word final nasal in each of these words could

nasalize *ero-* without nasalizing the vowel in the verb root itself.

14. A glide (*y*) is inserted between the prefix *i* and a root beginning with *i* or *ɨ*

y-insertion just means that a *y* is inserted between *i* and any other vowel

(13) a.   i + itá > iyitá   (J 1990a:73)
             3   stone 'their stone'

     b.   i + ɨpɨ > iyɨpɨ   (J 1990a:73)
             3   beginning 'his beginning'

     c.   i + u + áβo > iyuβomo   (J 1990a:73)
             3   eat   nom 'the eating of it'

15. When the agentive suffix *-ár* is added to a root ending in *y*, *t* is inserted between the

two morphemes.

Rule #15 is not a morphophonological rule at all, but a strictly morphological rule,

dictating the use of /-tár/, an allomorph of the morpheme /-ár/. This will be discussed

in Chapter 4.

16. An *i* is inserted between a root-final lenis and a suffix-initial *s*.

Epenthesis of *i* before *c* is the same as the other epenthesis rule, but *c* fronts *i* to *i*.

17. A suffix *-i* becomes *-y* when added to a root ending in a vowel.

18. A suffix *-i* disappears when added to a root ending in *y*.

19. When a rounded vowel is followed by another vowel in any environment, *w* optionally intervenes.

Rules #17 and #18 are self-explanatory. Insertion of *w* is the same for back rounded vowels as insertion of y for *i*. Unlike that rule, however, insertion of *w* is optional, perhaps reflecting the status of *w* in PTG as a marginal, or even non-existent phoneme.

For reasons given above, rules #6 and #15 will not be included in the comparison below, and rules #8 and #9 will be combined in the discussion.

3.2. Comparison of morphophonological rules in attested TG languages

Wayampĩ

1. As we have seen, this rule in Tb was not really limited to the boundary of compound elements. We could then reformulate it this way:

$$\begin{bmatrix} +\text{acnt} \\ -\text{nas} \end{bmatrix} > \quad > [+\text{nas}] / \_ \quad \begin{array}{l} +\text{cons} \\ +\text{nas} \end{array} \begin{bmatrix} + \\ \neq \\ \# \end{bmatrix}$$

This is exactly the sound change rule proposed by Jensen (1990) for Wayampĩ. We see then that the diachronic rule isn't diachronic at all: it describes the state of affairs in Tb, and probably in PTG, subsequently continued in W.

(14)　　　akā　　　　amā　Wj　　(J 1990a:61)
　　　　　akāŋ　　　amãn　Wa
　　　　　'head'　　　'rain'

Assurini of Trocará

1. This and other rules of nasalization have been negated in AT since autosegmental nasalization has been lost without a trace.

Parintintin

1. This rule applies in Pt (Betts 1981):

(15)　iβaɣa 'sky　　　　　　　　　iβaŋapiraŋuhũ 'sunset'
　　　ja?ɣʷar 'jaguar'　　　　　　ñaŋʷapinim 'jaguar'
　　　ñandu 'spider'　　　　　　　ñanupoamiraŋuhũ 'spider monkey'
　　　-nambi 'ear'　　　　　　　　-namihɨmuhũ 'to have keen hearing'
　　　-tiŋ 'white'　　　　　　　　-ti?ŋgi 'a small white thing'
　　　-kaβur 'black and smooth'　　-moŋgaβur 'to make black and smooth'
　　　-katu 'good'　　　　　　　　-ñimohiŋatu 'to be improved'
　　　ɣaretɨmakaŋ 'his shin'　　　ŋanetɨmakaŋ 'their shin'
　　　ɣara?ɨr 'his son'　　　　　　ŋanda?ɨr 'their shin
　　　-paɣ 'wake up'　　　　　　　-mombaɣ 'melt'
　　　-ho 'go'　　　　　　　　　　-mondo 'send'
　　　-pepo 'wing'　　　　　　　　-nambiβepo 'ears that really stick out'
　　　-mokõy 'two'　　　　　　　　-mokoñate 'few, more than two'

Kayabi

1. This is inherent in autosegmental nasality and vice versa. Kayabí no longer exhibits autosegmental nasality. Hence the rule no longer applies.

Wayampi

3. The epenthesis rule is unpredictable in Wj because the loss of final consonants eliminates the environment of the rule. In Wa, however, the rule is still productive:

(16)   iar pe > iaripe   (J 1990a:63)
           canoe in 'in the canoe'

           wir pe > wiripe   (J 1990a:63)
           low in   'below'

           *ʔaŋ βo > ʔaŋiβo   (J 1990a:63)
           this dist. 'throughout this'
                  loc.

AT

3. The epenthesis law has changed only to the extent that the epenthetic vowel is now *i* instead of *ɨ*:

(17)   ar -mo > arimo   'during the day'
       ɨpɨton -mo > ɨpɨtonimo   'during the night'
       kʷarahɨpɨter -pe > kʷarahɨpɨteripe   'in the middle of the sun'

Pt

3. Parintintin preserves this rule as exemplified in (18) below.

(18) -ep^yaγ   -βo   > -epiaγaβo
      'see'   nom   'seeing'

Kayabi

3. Like AT, Kayabi has preserved this rule while changing the epenthetic vowel to *i*.

<div style="text-align:center">

ar + mu   > arimu
day dist=loc   'daily'

mɨter + pe   > mɨteripe
middle loc   'in the middle'

</div>

OG

3. This rule did not exist in OG. In OG morph-final consonants were always last unless

immediately followed by a vowel-initial morpheme. So in the environment

___ + [+cons]

all morpheme-final consonants are lost

(19) ɨβáγ + pe > ɨβápe   (RM 1639:166)
      heaven in   'in heaven'

      a+h+endúβ ≠ katú > ahendúkatú   (RM 1639:152v)
      1s-3-hear   good   'I'm listening attentively'

      ar ≠ katú + pɨrɨ >   á katúpɨrɨ   (RM 1639:5)
      day good little=by=little 'good day'

      a + ár + βɨβí >   aáβɨβí   (RM 1639:4)
      1s fall on=the=point=of 'I'm about to fall'

(19)   o + ár + βa'e > oáβa'e 'he who was born'   (RM 1639:4)
        3 fall rel

Anterior lenes and nasals drop in word-final position as well.

(20)   a + h + endúβ > ahendú (RM 1639:151)
        1s 3   hear   'I heard'

       n + a + h + endúβ + i > ndahendúβi   (RM 1639:151)
       neg 1s   3   hear   neg 'I don't hear'

       o + úr > oú (RM Arte, p. 63v)
       3   come 'someone comes'

       n + o + úr + i > ndoúri   (RM 1640:63v)
       neg 3   come neg 'she doesn't come'

       a + h + ečaγ > ahečáγ   (RM 1639:375v)
       1   3   see   'I saw it'

       h + ečaγ + eĩ > hečaγeĩ (RM 1639:375v)
       3   see   neg   'you didn't see it'

       a + hēm > ahē   (RM 1639:147)
       1s   leave   'I left'

       i + hēm + áβ > ihembáβa   (RM 1639:147)
       3   leave nom '(the occasion of) his leaving'

       če + tĩm > četĩ   (RM 1639:387)
       1s   bury   'my burying'

       a + ye + tĩm + i > añetĩmĩ   (RM 1639:387)
       1s   pass + bury + obtop

       a + mo + porãŋ > amoporãŋ 'I make it beautiful'   (RM 1639:317)
       1s   caus beautiful

What this points to is an earlier stage in which there was no epenthetic vowel between

morpheme-final consonants and suffix-initial.

(21)   iβáɣ pe > iβáɣpe
        ↓   ↓   ↓
       iβáɣ pe > iβápe

Urubu-Kaapor

3. This rule applies as an optional alternative to #5 below, with two differences: first there are three allophones to the epenthetic vowel: [ɛ] following *i* or *e*; [ə] following *ɨ*, *a*, and *o;* and [u] following *u.*

(22) a.   u-hem tĩ [uhemɛtĩ]   (K&K 1990:135 l. 190)
               3-leave also
               'he also left'

     b.   mair r-aʔɨr-ta [raʔɨrəta] (K&K 1990:135 l. 4)
               Mair cm-son-fut
               'Mair's future son'

     c.   h-ok pe   [hokəpe]   (K&K 1990:137 l. 24)
               3-house to
               'to his house'

     d.   tatimbuk pe har ke [tatimbukupeharəke] (K&K 1990:160 l. 23)
               ashes   in agnt foc
               'the one in the ashes'

The second difference is that the epenthetic vowel is not only inserted between two bound morphemes, but between words, even when those words belong to different phrases.

(23)   mɨkur pandu [mɨkurupandu]   (K 1990:400)
               opossom speak
               'the opossom spoke'

Conclusion

The status of #3 in PTG and the data examined above point to a more complex situation than that of Tb. The OG evidence suggests that no epenthetic vowel was inserted in PTG, while the evidence from most other languages suggests that in PTG one was inserted. UK seems to hold the key to this puzzle. The three epenthetic vowels are environmentally conditioned variants of ə. A PTG epenthetic vowel ə would explain *i* in Tb and Wj, *i* in AT and Kb, and *a* in Pt. The fact that any epenthesis at all is optional in UK is supportive of the idea that either epenthesis was optional in PTG, or it was dialect-specific. If it was dialect-specific, then OG is derived from the non-epenthesizing dialect while Tb and others were descended from an epenthesizing dialect or dialects while the state of affairs in UK would have been the result of some earlier cross dialectal/cross-linguistic influence. If epenthesis was generally optional, then most languages have simplified the situation one way or the other, while UK has conserved the optionality of epenthesis.

Wayapí

4. The lenition rule does not occur because loss of final *p* eliminates the environment.

Assurini

4. unknown

Parintintin

4. The lenition rule still occurs across morpheme boundaries in Pt.

(24) a. -eβir + -kʷar > -eβiγʷar (Betts 1981:62)
  'backside' 'hole' 'anus'

 b. *ʔa + -paγ > -ʔaβaγ (Betts 1981:23,49,158)
   'head'   'awake'   'turn with a start'

 c. iriβi + -piraŋ > iriβiβiraŋ (Betts 1981:165,211)
   'neck' 'red'   'red neck'

 d. nambi + -pepo > -nambiβepo (Betts 1981:136,161)
   'ear(s)' 'wing'   'ears that stick out'

Kayabí

4. This rule also seems to apply in Kayabí (data from Dobson 1988).

(25) a. iat + pitet > iawitet (Dobson 1988:131)
   'canoe' 'middle' 'middle of a canoe'

 b. yetik + piaramũ > yetiwiaramũ (Dobson 1988:131)
   'potato' 'look for' 'look for potatoes'

 c. way + φuku > waywuku (Dobson 1988:131)
   'tail' 'long'   'long tail'

The first of these would correspond to the following in PTG:

(26) *iar + *piter > *iaβiter

a classic example of this lenition rule. In Kb however, a change has taken place. In the

first example above, there is not a conjunction of a lenis and a fortis, but of two fortes.

One could perhaps argue that the former really is underlyingly a lenis and only appears as fortis in word final position. In the second example, *-yetik* may or may not end in a lenis underlyingly.[2] It certainly did not in PTG. The third example in any case shows what is happening here; no final obstruent of any kind appears in *way* while *ɸuku* (< *pukɨ*) has changed to *-wuku* (< *-βuku*). At least in this third case what seems to have happened is lenition of a fortis between vowels or at least non-consonantals. That means that the first two examples can be interpreted in one of two ways: either they too show lenition of a fortis between non-consonantals (in which case this rule must be preceded by a rule simplifying consonant clusters) or that final obstruents in Kb really are underlyingly lenis and the first two examples above are to be written like (27) below.

(27) a.   ɨar   +   pɨter
               'canoe' 'middle'

     b.   yetɨɣ + piaramū
               'potato' 'look for'

These therefore exemplify the rule (4) of Tb (which would have to be followed by a consonant cluster reduction rule as in Tb). The former of these two propositions turns out to be the better for this reason: other data show independently that word final

---

[2]In combination with another morpheme *-k* final words lenite to either -ɣ or ŋ ex.

     piɨk + ukat > piɨɣukat or piɨŋukat
     'catch' caus.   'have someone catch something'

Because of the absence of any nasal consonants (or vocalic nasality, past or present) Dobson (1988) favors -ɣ as the usual morphophonemic outcome, considering [ŋ]here to be an unusual free variation of /k/. This need not be the case, however, as indeed it is not in AT where *k#* consistently becomes -ŋ# regardless of nasality in the word historically or synchronically.

consonants in Kb are really fortes.

Dobson (1988:130) shows that there is a rule of lenition between vocoids.

(28)   ɨat + a > ɨara
  'canoe'   nom. 'canoe'

      yaamep + uu > yaamewuu
      'animal' 'big'   'big animal'

      pɨɨk + -ukat > pɨɨɣukat
      'catch' caus.   'have someone catch something'

While it is natural for *p* to lenite to *w* it is harder to imagine that *w* would strengthen to

p. We can conclude then that the word-final obstruents in Kb really are fortes.

The 'lenition' or spirantization of stops also applies to nasals as with the following

example:

(29)   kʷáβ meʔē̃ŋ > kʷaβeʔē̃ŋ
       ↓   ↓   ↓
       kʷap meʔeŋ > kʷaweʔeŋ
       'pass' 'give'   'hand over'

has led to a reinterpretation of such examples as the following:

(30)   kʷap meʔeŋ > kʷa meʔeŋ > kwaweʔeŋ

This has spread by analogy to other examples not affected in (and therefore not

inherited from) PTG.

(31) -ata -manɨrɨk > -atawanɨrɨk   (Dobson 1988:131)
      'fire' 'draw'   'light a fire'

      -ɨwɨ -nupā   > -ɨwɨrupā (Dobson 1988:131)
   'earth' 'beat'   'beat into the dust'

      -nupiʔā -ka   > -nupiʔāŋa   (Dobson 1988:131)
      'knee' 'hit'   'hit the knee'

These examples demonstrate the phenomenon described in (3.0.) where a complex form

is replaced by an innovation through analogy.

(32) ɨβɨ nupā > ɨβɨnupā
       ↓   ↓
      ɨwɨ nupā > ɨwɨrupā
   'earth' 'beat'   'beat into the dust'

This rule of lenition in Kb also applies to nasal consonants, which have the same

lenis allophones as the fortes:

(33) -ata -manɨrɨk > -atawanɨrɨk   (Dobson 1988:131)
      'fire' 'draw'   'light a fire'

      ɨwɨ + -nupā > -ɨwɨrupā   (Dobson 1988:131)
   'earth' 'beat'   'beat into the dust'

      -sī + -mukup > -sīwukup (Dobson 1988:131)
   'tip' 'heat up' 'heat up the tip'

      -nupiʔā + -ka > -nupiʔāŋa   (Dobson 1988:131)
      'knee'   'hit' 'hit the knee'

Lenition does not occur in the following example:

(34)   moʔɨt + pɨtaŋ > moʔɨ pɨtaŋ   (Dobson 1988:132)
      'necklace'   'brown' 'brown necklace'

The easiest explanation is that the *p* of *pitaŋ* is from an original pre-glottalized stop.

OG

4. Because of final consonant loss in OG, the application of this rule was sporadic in OG.

(35)   okar + piter > okaβiter ~ okapiter   (RM 1639:255v)
       'village'   'middle' 'middle of the village'

       okar + pa?ū > okaβa?ū ~ okapa?ū   (RM 1639:255v)
       'village' 'the midst' 'in the midst of the village'

UK

4. The rule of spirantization, which has been shown to be a lenition rule, does not occur

in UK. No consonant obligatorily lenites in any environment; *p* never becomes *w* or *β*

(*β* has lenited to *w*). However, optional lenition does occur for *k* only, at any word

boundary.

(36)   usak oho [usaγ oho] 'he went to see'   (K 1986:399)
       kutuk katu [kutu γatu] 'she washes well' (K 1986:400)

As the second example illustrates, *k* lenites in word initial position when it is the second

of two consecutive consonants, the first of which has deleted by rule #5 below.

This development might suggest that the contrast of fortis/lenis is still alive in UK.

But this appears less likely when one notes that in more careful speech, *k* only lenites

to *g,* and that unstressed, morpheme internal *p* and *t* lenite not to *β* and *r* respectively,

but to *b* and *d.*

(37)   arapuha [arapuha] or [arabuha] 'deer'   (K 1986:399)
        heta tipe [heta tipe] ~ [heta dipe] ~ [heta tibe] ~ [heta dibe] 'there are many,
                                                        but in vain'   (K 1986:399)

It seems then, that while lenition in UK may hearken back to the fortis/lenis

categorization of PTG, it equally looks ahead to possible dawning of a distinction

between voiced and voiceless stops.

Conclusion

The fortitioning of word-final lenes has resulted in a different lenition rule in Kb

and UK than that of PreTG.

(38)   yacameβ + ucu > yacameβucu
             ↓   ↓   ↓
        yaamep + uu > yaamewuu   (Kb)
        'animal' 'big' 'big animal'

        ocepʸák ocó > ocepʸák ocó
             ↓   ↓   ↓
          usak oho > usaɣ oho   (UK)
          'see' 'much' 'see greatly'

Here the lenition rule has spread to velars not by word-final strengthening, but by medial

lenition of velars, a later development in Kb and UK.

Consequently what was originally a lenition/spirantization rule in PTG, has merged with

a lenition rule in Kb and UK so that the examples of the spirantization rule of PTG have

become examples of the lenition rule in Kb and UK.

Wayapí

5. Simplification of clusters C1C2 > C2 /_ {≠} [−syl]

<center><#></center>

only applies in Wa:

(39)   tareʔɨr ≠ siŋ   >   tareʔɨsiŋ   (J 1990a:65)
        traira white   'white traira'

        tukan ≠ siŋ ≠ pirã >   tukãsĩpirã   (J 1990a:65)
        toucan white red   'species of toucan with a red beak tip'

        akãŋ ≠ so   >   akãso   (J 1990a:65)
        head   'to bang one's head'

Assurini

5. unknown

Parintintin

5. Simplification of consonant clusters can be seen in Pt examples as in (40).

(40)   -apɨter + kutuɣ   >   apɨtekutuɣ (Betts 1981:44)
      'crown of head' 'make hole' 'punch a hole in the head'

          -apɨakʷar + -βɨɣ   >   -apɨakʷaβɨɣ (Betts 1981:42)
          'inner ear' 'to close'   'shut one's ears'

Kayabí

5. The simplification of consonant clusters has already been demonstrated in Kb above.

It is ordered before the rule of lenition.

(41)   ɨat + pɨtet   (Dobson 1988:130)
        ɨa + pɨtet
        ɨawɨtet

        yetɨk + piaramū   (Dobson 1988:131)
        yetɨ + piaramū
        'potato' 'look for'
        yetɨwiaramū
        'look for a potato'

However, this rule is not limited to the compounding of noun and verb roots. The conjunction of verb complexes and clitics also displays this consonant reduction, but without the lenition rule:

(42)   a + kutuk ye > akutu ye   (Dobson 1988:132)
        1s-1 - bore 1s   'I bore'

        a-ʔu-pap + ko > aʔupako   (Dobson 1988:132)
        1s-eat-finish pst 'I ate everything'

        pē-esak + te > pēesa te   (Dobson 1988:132)
        2p-1 - see int 'Do you see him?'

The one apparent exception to this rule is the clitic *ramū* 'when'.

(43)   ʔŋa kutuk + ramū > ʔŋa kutuɣ amū   (Dobson 1988:132)
        he bore   when   'when he bores a hole'

OG

5. The rule on consonant sequences has been amplified to the general rule of final consonant loss discussed in (3) above.

$$[+con] > \emptyset \; / \; \underline{\quad} \begin{bmatrix} <+> [\text{-syl}] \\ <\neq> \end{bmatrix}$$

$$\begin{bmatrix} +con \\ \text{-bk} \end{bmatrix} > \emptyset \; / \; \underline{\quad} \quad \#$$

$$[+con] > \emptyset \; / \; \underline{\quad} \begin{bmatrix} +[\text{-syl}] \\ \neq[\text{-syl}] \\ <\#> \end{bmatrix}$$

Urubu-Kaapor

5. The simplification of consonantal sequences occurs across both word and morpheme boundaries.

(44) mani?ok yande yapirok [mani?o yande yapirok] 'we peeled manioc'
      manioc 1p   peel   (K 1986:400)

      aman pe [amã pe]   'in the rain'   (K 1986:400)
      rain in

This is certainly a progression from PTG when the simplification of consonant sequences applied only within compounds. Also, to the extent that one may talk of ordering rules, the reinterpretation of #4 means that #5 applies first, then applies the new #4, unlike Tb where the rules applied in reverse order.

Conclusion

The general progression of apocope in TG as evidenced by comparing Tb with modern languages, raises a doubt as to its application in PTG. Since apocope was exhibited less and less as one moves backward toward PTG perhaps it did not occur at all in PTG. It is a productive rule in all the languages where its effects are observed, so

the rule is not a relic of an earlier diachronic change. Apocope amounts to regularizing an otherwise aberrant syllable structure in which every syllable is V or CV unless it is morpheme-final, where it may be V or CV or CVC. An analogous change of this sort (changing the syllable structure of root-final syllables by analogy with all other syllables) is possible as an independent change in various daughter languages, and need not be a reflection of PTG apocope. Nevertheless, evidence for TG apocope is very early, and it is possible that PTG had at least an optional rule of apocope parallel to that of Tb.

Wayapĩ

7. Does not apply: *-áβo no longer exists, and *-áβ, reduced to -a, no longer blends with vowel-final stems (all verb stems are vowel-final in both dialects).

Jensen's statement that this rule does not apply in W because the environment is not the same begs the question. It does not explain why the following shift did not take place.

(45)   co-aβ
            go-nom

            coβo
            coβ
            §co 'going'

The point here is that -a is no less liable to assimilate to a preceding vowel simply

because it is not followed by a fricative.[3]

Assurini

7. This law does not seem to apply in AT because the environment no longer exists. PTG *-áβo has been reduced to -w in all cases. The nominalizing suffixes -cáβ and -cár have been generalized to -hawa and -hara respectively (PTG *c > AT h) for all verbs ending in a vowel.

Parintintin

7. unknown

Kayabí

7. Vocalic assimilation holds for Kb. The nominalizing 'gerund' suffix is -áw or -ámū depending on the nasal environment. The rules of truncation are different, however, with respect to their conditioning environments. Generally, if the vowel preceding -aw is a mid vowel, the vowel in -aw is deleted.

---

[3] The occurrence of -a as the serial verb suffix is a reflection of the fact that *-áβo was not assimilated to the preceding vowel in at least some dialects of PTG.

There are, of course, alternative explanations. One is that -a is really 'nominal case' applied to verb stems. But 'nominal case' in W is still largely phonosyntactic and never is applied to vowel-final roots. Another possibility is that the pre-TG suffix **-ca, used only on consonant final verbs in PTG, was still used also on vowel-final verbs; or at the least that *-áβo having disappeared, the verbal nominalizing suffix for consonant-final verbs was applied to vowel-final verbs as well.

(46)   yuká + aw > yukáw   (Dobson 1988:42)
       kill   nom 'killing'

       apo + aw > apow   (Dobson 1988:42)
       make nom 'making'

       moŋe + aw > moŋew   (Dobson 1988:43)
  make enter   nom 'putting to sleep'

If a non-high vowel was preceded by ? there was no assimilation, but an synezesis

instead. But *aw* does not assimilate to *e* when *e* is preceded by *s*, *m*, or *w:*

(47)   se + aw > (siaw >) saw   (Dobson 1988:43)
       sleep nom   'sleeping'

       wewe + aw > wewiaw   (Dobson 1988:43)
       fly   nom   'flying'

       mu?e + aw > mu?yaw   (Dobson 1988:42)
       teach nom   'teaching'

       apẽ?ẽ + aw > apẽ?yãw   (Dobson 1988:43)
       arrange nom   'arranging'

High vowels *i* + *u* are deleted before *a*:

(48)   po?ru + aw > po?raw   (Dobson 1988:42)
       use   nom 'using'

       apisi + aw > apisaw   (Dobson 1988:43)
   kill=many nom 'killing many'

This is always true of *u*, but when *i* is preceded by *m* or *w* it remains:

(49)   tɨamī + aw > tɨamīãw   (Dobson 1988:43)
       twist   nom 'twisting'

When *u* is nasal *a* does assimilate and is replaced by *mū*.

(50) manū + aw > manūmū   (Dobson 1988:42)
      die   nom 'dying'

This last example is significant because it shows that either the underlying form of the suffix is never realized on the surface (something like *-awu*), or the allomorph after *ū* is not derivable form the underlying form.

7. Vocalic assimilation not only holds in OG, it expands to all vowels resulting in the reinterpretation of -*áβo* as -*βo*.

(51) i + mboʔé + βo > mboʔéβo   (RM 1640:16)
      3   teach nom 'teaching him'

      i + kiβú + βo > ikiβúβo   (RM 1639:332)
      3   delouse nom 'his delousing'

However, forms ending in -*áβo* also appear.

(52) mombeʔúβo ~ mombeʔwáβo   (RM 1640:28)
      'telling'

Words like *mombeʔwáβo* are the source, by metanalysis, of the -*wáβo* that appears on all vowel-final verbs:

(53) wi-yɛpeé-βo ~ wi-yepeé-wáβo[4]   (RM 1640:28)
1s get=warm-nom
'my getting warm'

i-yɨkɨ-βo ~ i-yɨkɨ-áβo   (RM 1640:28)
3-husk-nom
'the husking of it'

i-suʔu-βo ~ i-suʔu-wáβo   (RM 1640:28)
3-bite-nom
'the biting of it'

soʔó   >   isoʔowáβo   (RM 1640:28)

potí   >   wipotyáβo   (RM 1640:28)
clean   'my cleaning'

porú   >   iporwáβo   (RM 1640:28)
use   'the using of it'

tɨpɨ   >   itɨpɯáβo   (RM 1640:28)
take stew   'her taking stew'

ʔe   >   wiyáβo   (RM 1640:28)
say   'my saying'

A couple of examples seem to show Kb-type morphophonemics:

(54) aɨtɨpeʔi   >   ɨtɨpeʔyáβo ~ ɨtɨpeʔíβo   (RM 1640:29)
     'I sweep'   'sweeping'

amoiŋgie   >   imoiŋgiáβo ~ imoŋgiéβo   (RM 1640:28)
'I make   'making him enter'
him enter'

---

[4] This derives from an earlier alternation: yepeʔé-áβ-βo ~ yepeʔé-wáβ-βo. Evidence from Tb confirms that in PTG serial verbs ending in a vowel were marked by áβo when a verb ended in two identical vowels separated by glottal closure (V₁ʔV₁). In OG such serial verbs began to be marked with -βo by analogy with other vowel-final verb roots.

This situation was further complicated by the loss of word-final consonants, which led consonant-final verbs to be reinterpreted as vowel-final.

(55)   o-eyɨβ >   o-eyɨpa (RM 1640:26)
         ↓   ↓
       w-eyɨ >   w-eyɨ-βo   ~   w-eyɨpa
   3-go=down   3-going=down-nom 3-going=down=nom

         o-potár   i-potáta   (RM 1640:26)
           ↓
         o-potá >   i-potá-βo
         3-want   3-want-nom

Urubu-Kaapor

7. This rule doesn't apply. The only environment is -áβ, which was lost by regular sound change rules.

Conclusion

The data for #7 are problematic. Every language where there is the conjunction of like vowels across a morpheme boundary exhibits this rule to at least some extent. Evidence from Pt and Kb however, show that $á$ did not assimilate to preglottalized mid-vowels. And certain verbs in Kb point to non-assimilated forms in PTG even where preglottalization was not a factor:

(56)   βeβe > βeβiáβo
         ↓   ↓
       wewe > wewiaw
       fly   flying

See #11 below for the Kb example of *se* 'enter' > *saw* 'entering'

In Kb, W, and perhaps other languages, the rule that assimilates two non-high vowels and then deletes one is not productive.

(57) yaamep 'animal' (Kb)   (Dobson 1988:130)
      eraáa 'bringing' (W)   (J 1990a:72)

Therefore the assimilated forms in TG languages are inherited from PTG rather than the result of a productive rule.

The rest of the data in Kb and Pt indicate that the rule of assimilation in Tb existed in PTG, but was more restricted than in Tb.

$$\acute{a} > \begin{bmatrix} \alpha\,bk \\ -\alpha\,rnd \\ -hi \end{bmatrix} / [+int] \begin{bmatrix} +syl \\ \alpha\,bk \\ -\alpha\,rnd \\ -hi \\ +stress \end{bmatrix} + \underline{\quad}$$

When *á* is preceded by a morpheme boundary and a non-glottalized mid-vowel preceded by a fortis, it assimilates to that vowel.

Wayapï

8&9. Do not apply for same reason as given in (7) above.

Assurini

8. As noted above, autosegmental phonology has been lost in AT. However, the following nasalization rule appears to apply: $T > N / \_\#$. This is exemplified by data

such as the following.

(58)　　　we-čɨm-a > we-čɨp-a
　　　　　1s-descend-nom　　1s-descend-nom

　　　　　we + hem + a > we-hem-a
　　　　　1s-leave-nom　　　1s-leave-nom

Synchronically, this would be explained by saying that the above roots are underlyingly

/čɨp/ and /hem/ respectively and adding the nasalization rule in AT, even though there

is no direct evidence that *čɨm* and roots like it ever ended in a stop phonetically. In

PTG, they ended with lenes as the diachronic derivation of the above forms shows:

(59)　　　*yɨβ　　*yɨp-a　　　　*cem　*cem-a
　　　　　　↓　　　　↓　　　　　　　↓　　　　↓
　　　　　čɨm　　čɨp-a　　　　　hem　hem-a
　　　　　'descend' 'descending' 'leave' 'leaving'

But consider the following:

(60)　　　*ye-r-erúr　　*nde-r-erúr　　*pe~n-erúr
　　　　　　↓　　　　　　　↓　　　　　　　↓
　　　　　če-r-eron　　ne-r-eron　　pe-n-eron
　　　　　1s-cm-bring　2s-cm-bring　2p-cm-bring

In PTG 2p accusative marker *pe*~[5] nasalized a following class marker because

it was underlyingly nasal (though unstressed). In AT all autosegmental nasality is gone;

2p accusative marker is wholly oral: *pe-*. Yet in combination with vowel-inital roots

---

[5]The form *pe˜* is meant to express that while this prefix is not nasal on the surface (because it does not bear stress) it transfers the autosegmental feature of nasality to the immediately following stem.

which take the class marker -r-, pe- still has a nasalizing effect on the following class marker. Apparently what was a morphophonological regularity in PTG has become a strictly morphological anomaly in AT.   With the 2p accusative/possessive one uses -n- for the class marker instead of -r-; that's the way the paradigm works. Since the choice of class marker is inexorably bound up with the choice of person/number cross-reference marker and with nothing else, it may be more accurate to analyze class markers not as class markers at all, but as part of the cross reference system. We would say, then, that there are two series of accusative/possessive cross-referencers; one used with consonant-initial roots, the other with vowel-initial roots. Indeed, root classes under this analysis would be redefined in terms of which series of cross-referencers they took.

(61)   *ye-r-ecá   *ne-r-ecá   *pe-n-ecá
                ↓   ↓   ↓
          čer-eha   ner-eha   pen-eha
          1s-eye   2s-eye   2p-eye

Parintintin

8&9. In the examples below (61a) is an example of #8 while (61b) is an example of #9.

(62) a.   nupã + -βo >   nupãmo   (Pease 1968:27)
           beat + nom   'beating'

     b.   ŋã retɨmakaŋ   >   ŋã netɨmakaŋ   (Betts 1981:8)
          3p shin   'their shins'

Kayabí

8&9. Nasalization of lenes does occur although nasality is no longer an autosegmental feature. In fact we may say that it is precisely here that segmental assimilation does take

place while in languages having autosegmental nasality it is the entire morpheme that is nasalized.

(63) -ru + -ʔẽ > -nũʔẽ   (Dobson 1988:134)
      bring

      i-nupã + ramũ > inupãnamũ   (Dobson 1988:134)
      3-beat   when   'when beating him'

OG

8&9. This applies chiefly to the *-áβo suffix. It still does in OG.

(64) h-eroĩrõ-mõ (RM 1640:27)
      3-scorn-nom

      wi-βãhẽ-mõ (RM 1640:27)
      1s-arrive-nom

      i-mõpĩ-mõ   (RM 1640:27)
      3-swing-nom

      h-epeñã-mõ (RM 1640:27)
      3-attack-nom

What is true about word-final consonants is also true about word-final nasals:

(65) wi-wẽʔẽnã   ~   wi-wẽʔẽ-mõ (RM 1640:27)
      1s-vomit=nom   1s-vomit-nom

      i-kotɨrũ-mõ ~   i-kotɨrũŋgã   (RM 1640:28)
      3-lie=in=wait-nom 3-lie=in=wait=nom

Urubu-Kaapór

8&9 These and other nasalization rules are non-productive in UK due to the complete

loss of autosegmental nasality. For example, in *ta?ir namō* 'as a son', the *namō* is from **ramō*. Diachronically, the *r* has nasalized because of the nasal vowel. Synchronically the *n* of *namō* is underlyingly *n*, not the nasal allophone of anything.

Conclusion

We have seen that #8 and #9 are really one rule nasalizing lenes. Since PTG was a f/l language, it is possible that this was a productive rule in PTG. This is especially likely for two reasons: nasal assimilation, a common phenomenon, was a diachronic change that actually happened from Pre-TG to PTG; and daughter languages maintain forms that together point back to PTG forms that would be interpreted by a nasalization rule. TG languages preserve relics of such nasalization even when those languages have lost autosegmental and even segmental nasality.

Wayapī

10. In Wayapī the stress has shifted from the ultimate to the penult: the rule therefore cannot apply in its original form. There is in fact no conflict between two stressed morphemes, since the W nominalizing suffix *-a* is always unstressed. There is simply the general rule of penult stress on the word level. Hence
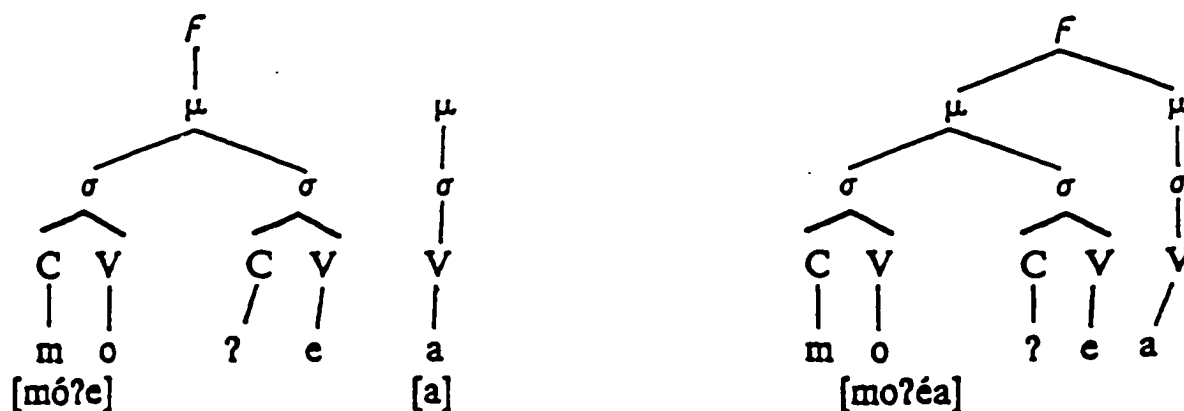
Assurini

10. As noted in (7) above, the suffix *-áβo* has been reduced to atonic *-w* so the conditioning environment for this rule is non-existent in AT.

Parintintin

10. This is another rule in the formation of abstract verbal nominalizations and is formulated accordingly. There is another case of verbal nominalization, however. This is the set of suffixes: *-áβ*, *-ár*. In those cases, the environment is different, but the result is the same: suppression of the first of two accents.

Kayabí

10. Suppression of accent also takes place in Kb. Unlike most TG languages, Kb manifests suppression of accent on the surface level, but only with verbs ending in *i*.

(66)   -porowɨk - áw > porowɨkɨáw   (Dobson 1988:42)
       work   nom   'working'

OG

10. Suppression of the accent applies where it is appropriate (since *-áβo* is largely replaced by *-βo*, this is rare).

Urubu-Kaapór

10. Suppression of the accent can be observed with the nominalizer *-há.*

(67)   pukú-há > puku-há (K 1986:345)
        long-nom long-nom

Conclusion

The suppression of stress rule, like the 'metathesis' of ?, is actually only the manifestation of an autosegmental feature. Since stress dominates the morphosyntactic unit of 'complex' (discussed in Chapters 4 and 5) There can be only one stress per complex (roughly the equivalent of 'word'). There is then no suppressing going on; except that when two stressed elements are bound in one complex, that complex will have only one stress.

Wayampĩ

11. Wayampĩ does not have sequences of stop and glide, hence rule (11) does not apply. However, in all other environments, *i* does change to a *y* when adjacent to a vowel.

(68)   i + ɨ > iyɨ 'his mother' (J 1990a:72)
            3 mother

            i + upe > iyupe   (J 1990a:72)
            3   to   'to him'

Here the environments were created by loss of c/č: *iyɨ* < **ičɨ*; *iyupe* < **ičupe.*

Assurini

11. As with (7) and (10) above, the reduction of *-áβo* to *-w* eliminates the possibility of this rule taking place.

Parintintin

11. This rule has been reformulated in Pt for the following reason: Cy > Ci ex. *ep^{y}ak* > *epiaγ*. When *y* or *w* begins a glottalized syllable, however, they are preserved as glides and this rule applies to them. In fact the rule applies to all non-central vowels in final glottalized syllables (data based on Pease 1968:27).

(69)   ʔu + áβo > ʔg^{w}óβo
         eat nom   'eating'

         kaʔ$\bar{e}$ + áβo > kaʔñámo
         be=cured nom 'being cured'

         auʔí + áβo > auʔjáβo
         cover nom 'covering'

         jeheʔo + áβo > jeheʔg^{w}óβo
         cry   nom   'crying'

         pɨkaneʔõ + áβo > pɨkaneʔŋ^{w}óβo
         be=tired   nom   'being tired'

This means we have to reformulate the rule for Pt:

$$\begin{bmatrix} +\text{syl} \\ \alpha\,\text{bk} \\ \alpha\,\text{md} \end{bmatrix} > [-\text{syl}] / ʔ\_+ \begin{bmatrix} +\text{syl} \\ +\text{acnt} \end{bmatrix}$$

Kayabí

11. There is synezesis in Kb only when the final syllable of a verb root is glottalized.

(70)   muʔe + aw > muʔyaw   (Dobson 1988:42)
       teach nom   'teaching'

Synezesis holds for both root-final *e* and root-final *i*.

When preceded by *m* or *w*, *i* and *e* neither undergo synezesis nor drop out:

(71)   wewe + aw > wewiáw   (Dobson 1988:43)
       fly   nom   'flying'

       tɨamĩ + aw > tɨamĩãw   (Dobson 1988:43)
       twist nom   'flying'

i and *e* drop after *s*.

(72)   apisi + aw > apisaw   (Dobson 1988:43)
       kill=many nom 'killing many'

       -se + -aw > -saw   (Dobson 1988:43)
       enter nom   'entering'

The reason for this rule is also historical. The sibilant *s* derives from *t* and *k* before front

vowels. The synezesis that would have taken place 'blended' with the preceding affricate:

	kill=many nom killing=many
stage 1	apiti + aw > apityaw
stage 2	apici + aw > apicyaw
stage 3	apici + aw > apicaw
stage 4	apisi + aw > apisaw

```
                        enter nom 'entering'
         stage 1        ke + aw > keaw⁶
         stage 2        ke + aw > kiaw
         stage 3        ce + aw > ciaw
         stage 4        ce + aw > cyaw
         stage 5        ce + aw > caw
         stage 6        se + aw > saw
```

OG

11. It is hard to tell if synezesis took place in OG. Based on the orthography of Montoya, there seems to be evidence that it did. Glides in OG are designated in two ways. One is by the corresponding high vowel with a circumflex accent.

(73) w or γ^w = <gû> gûara = γ^wara ~ wara 'benefactive'
       y = <î> <îu>   = yu 'needle, thorn'
                        quaî = k^{w}ay 'command, rule'

Another convention, used as often, is the circumflex over the vowel following the glide:

(74)   guâg   'ornament, to adorn'
            guâra 'benefactive'
            quây   'command, rule'

It is this orthographic convention we find in the environment of the synezesis rule.

(75)   wi-poti - áβo > wipotyáβo <guipotiâbo>   (RM 1640:28)
            1s-excrete-nom 'my excreting'

            wi-poru-áβo > iporwáβo <iporuâbo>   (RM 1640:28)
            3-use-nom   'the using of it'

---

⁶This indicates that the synezesis rule did exist at an earlier stage of Kb.

(75)   i-ʔu - áβo > iɣʷáβo   <iguâbo>   (RM 1640:28)
              3-eat-nom   'the eating of it'

Urubu-Kaapór

11. unknown

Conclusion

In the comparison of rules and forms the similarity of forms in Kb and Pt suggest that if a rule of synezesis existed in PTG, it most likely followed the pattern of Pt and Kb. They both have the identical morphophonological rule, which goes hand in hand with the fact that they both preserve ʔ as an autosegmental feature. Likewise the change in the rule in other daughter languages is bound up with the reinterpretation of ʔ as a segment.

Kayabí

(76)   ʔú + áβo > ʔwáβo   karu > karuáβo
       ↓   ↓   ↓   ↓
       ʔú   > wáβo   karú > karuáβo
       eat   'eating'   eat   'eating'

Since ʔ is reinterpreted as a segment, ʔ, we have, on the one hand, Cu > Cw, but on the other hand, we have Cu > Cu in the same morphological environment. The irregularity was smoothed out by analogy of verbs like *karu* with verbs like *ʔú.*

(77)   ʔú > ʔwáβo   karú > karuáβo
       ↓   ↓   ↓   ↓
       ʔú   wáβo   karú   karwáβo
       eat   'eating'   eat   'eating'

This is what happened in Tb and other daughter languages that have a similar morphophonological rule.

Wayapĩ

12. The elision of non-high vowels does not exist in W. This is due primarily to the fact that in PTG the nominalizing suffix for vowel-final verbs was *-cár* or *-cáβ*; the conditioning environment for the elision rule was not present. When *c>∅* in W, this left adjacent unelided non-high vowels.

(78)   eracá aβo
      ↓ ↓
     eráa + a > eraáa (J 1990a:72)
     take nom taking

     yúka + a > yukáa (J 1990a:72)
     kill nom killing

     ména + a > menáa   (J 1990a:72)
     marry nom wedding

Assurini

12. As with (7) and (10) above, the reduction of *-áβo* to *-w* eliminates the possibility of this rule taking place.

Parintintin

12. Because of the change in rule (11) above, (12) needs necessarily to be reformulated as well.

$$\begin{bmatrix} +syl \\ -hi \\ \alpha\ bk \\ \beta\ md \\ -acnt \end{bmatrix} > \emptyset \quad -^? \underline{\quad} \begin{bmatrix} +syl \\ \alpha\ bk \\ \beta\ md \\ +acnt \end{bmatrix}$$

where -ʔ denotes the absence of an immediately preceding glottal stop.

Kayabí

12. The elision of non-high holds in Kb with the above exceptions.

(79) -yuka + aw > yukaw   (Dobson 1988:42)
      kill   nom 'killing'

      -apo + aw > -apo + ow > -apow   (Dobson 1988:42)
      do   nom do   nom   'doing'

      -moŋe + aw > moŋe + ew > moŋew   (Dobson 1988:43)
      make=enter nom make=enter nom 'making enter'

The occurrence of this rule in Kb as well as in other TG languages seems to point to this rule occuring in PTG. But the development of *keaw (or *keáβo) above suggests otherwise.

OG

12. The rule of non-high vowel elision is followed in OG.

This rule in OG sometimes extends to high vowels as noted above.

The assimilation of non-high vowels has a notable exception recorded in Anchieta (1595), but not in Jensen (1990).

(80) ayepeʔe >   yepeʔeáβo ~ yepeʔewáβo (A 1595:28)
      I heat up >   'heating up'

      aypoʔo   poʔoáβo ~ poʔowáβo   (A 1595:28)
      I seize   'seizing'

      aయšoʔo   šoʔoáβo ~ šoʔowáβo   (A 1595:28)
      I bite   'biting'

This lack of assimilation after pre-glottalized syllables is also recorded by Montoya in

OG.

(81) wi-yepeʔe-βo ~ wi-yepeʔe-áβo   (RM 1640:28)
      1s-heat-nom

      wi-yaβo (< wi-ʔe-áβo)   (RM 1640:28)
      1s-say-nom

      i-soʔo-áβo   (RM 1640:28)
      3-bite-nom

Urubu-Kaapór

12. There is no environment to which this rule would occur.

Wayapī

13. At least one frozen example of this rule's application exists in W (Jensen 1990:72):

enoẽ 'leave with (someone)', but the rule is unproductive today.

Assurini

13. unknown

Parintintin

13. This rule still applies in Pt (Betts 1981:181).

$$\begin{array}{l} \text{ero} + \text{nupã} > \text{enonupã} \\ \text{con} \quad \text{beat} \qquad \text{'beat together with'} \end{array}$$

$$\begin{bmatrix} \text{cons} \\ \text{+voc} \\ \text{-nas} \end{bmatrix} > + \text{nas} / \_ V + C \begin{bmatrix} \text{+syl} \\ \text{+nas} \end{bmatrix} \text{+nas}$$

$$\mu > \mu / \_ \overset{N}{+} \overset{N}{\mu}$$

Kayabí

13. See (8) and (9) above.

OG

13. Leftward nasalization across morpheme boundaries appears to have been optional.

(82)   a-ro-ñã > aroñã 'I ran with (him)'   (RM 1639:241v)
       1s-con-run

       a-ro-ĩ > anoĩ 'I've been standing with him.'   (RM 1639:163v)
       1s-con-be

Urubu-Kaapór

13. UK has lost autosegmental nasality and therefore this rule does not apply in UK.

Conclusion

The occurrence of this rule in disparate TG languages and the naturalness of such

a rule leave little reason for supposing that this rule did not apply in PTG, especially

since, as seen in Chapter 2, leftward nasal assimilation was a diachronic change from Pre-TG to PTG.

Wayapí

14. See #11 above.

Assurini

14. This rule applies in AT with the slight modification that it is fed by the rule deleting *c from PTG, and č (< PTG *y) is inserted rather than y. So for example, i 'third person' + ope 'to' > ičope 'to him' (<*iyupe < *iupe < *ičupe). However, since *c almost always remains in AT as h, the application of this rule is very rare.

(83)　　*ičɨ > ihɨ 'his mother' (not §iɨ > §iyɨ > §ičɨ)

Parintintin

14. It is unclear whether the rule applies in Pt or not. Its application seems to have been very rare, primarily to cases of possessed nouns not usually possessed.

(84)　　　　i + ita　　　> ijita
　　　　　　3　hard　　　'is hard'

Ita ordinarily would be seen as a class 3 noun (non-possessed); since it is not Class 2 it takes the 3 poss i- even though it is vowel-initial. In modern languages this rule still

applies, but principally where an originally intervening *č has disappeared due to the regular sound change *č > ∅. Since Pt has an overt reflex of *č (h), there is little or no occasion of the conditioning environment for this rule.

Kayabí

14. Insertion of y

> No record of this rule in Kb.

OG

14. Epenthetic y is recorded for OG:

(85)   (i+ipɨ) > iyɨpɨ oúβaʔe če 'I am the first who came'   (RM 1639:176)

Urubu-Kaapór

14. No record of this rule in UK.

Conclusion

At this point no sure conclusion can be drawn as to whether this rule applied in PTG. Forms like *iyitá* 'his stone' appear in various daughter languages, but the rule is such a natural one, both from a synchronic and a diachronic point of view, that one cannot rule out the possibility of independent innovation.

Wayapí

16. The conditioning environment for this rule was a suffix-initial sibilant. The original PTG sibilant *č was lost in W and the new sibilant *s* is from PTG *t* before *i.* Such an *s* does not occur at the beginning of a suffix. Hence this rule can never apply in W.

Assurini

16. This rule applies in AT, less frequently.

Parintintin

16. Does not occur, since there is only one silibant in Pt,č, which does not appear in the conditioning environment for this rule.

$$\emptyset > \begin{bmatrix} +\text{syl} \\ +\text{hi} \\ -\text{bk} \end{bmatrix} / [+\text{con}] + \underline{\qquad} \begin{bmatrix} +\text{con} \\ +\text{str} \end{bmatrix}$$

Kayabí

16. As noted above there is a rule inserting an epenthetic *i* but this *i* turns out to be from the PTG epenthetic vowel *i,* and consequently the rule of epenthetic *i* in Kb has the same evironment as that of PTG epenthetic *i.*[7] The C which was part of the environment in the original rule of epenthetic *i* in PTG however, is reduced to ∅ in Kb. Consequently, the evironment for this rule is eliminated and either there is no need for

---

[7] This is probably originally the same epenthesis in *i* which was fronted by the following *c* or *s.*

epenthesis at all, or the environment is the same as for the regular epenthesis rule.

OG

16. The epenthesis of *i* is recorded in OG.

(86) če-maʔe-kʷaáβ-sé > čembaʔekʷaaβisé   (RM 1639:114v)
      1s-thing-know-want
      'my wanting to know something'

Urubu-Kaapór

16. The environment that conditions this rule is not found in UK.

Conclusion

This rule is a rule of PTG and the epenthetic vowel is certainly a variant of the

*ə in rule #3 above which is clearly fronted and raised to *i* by *č.

Wayapī

17. This rule still applies in W although its occurrence is much less frequent, owing to morphological change in the language: the fact that the object prefix -*i*- is no longer incorporated into the verb; and the loss of the partitive locative suffix -*i.

Assurini

17. No evidence of this rule in AT.

Parintintin

17. $\begin{bmatrix} +syl \\ +hi \\ -bk \\ -acnt \end{bmatrix} > \begin{bmatrix} -con \\ -syl \\ +vce \\ -rnd \end{bmatrix} \quad / \quad +syl \underline{\quad} \begin{bmatrix} \{\#\} \\ +syl \end{bmatrix}$

This rule can in point of fact be combined with (11) above as the following rule: i> y when it is adjacent to any vowel other than high unrounded vowels.

$\emptyset > \begin{bmatrix} -con \\ -syl \\ +vce \\ -rnd \end{bmatrix} / \underline{\quad} \begin{bmatrix} +syl \\ +hi \\ -rnd \end{bmatrix}$

11 and 17 i > y elsewhere

Kayabí

17. Diphthongization (i > y / V__) still holds for Kb, but it is rarely attested, due to the fact that the object prefix *i*- and locative suffix -*i* have been lost.

(87) na   ko   pe arŭy ye oy (oy < o-i)   (Dobson 1988:62)
            neg clearing to neg 1s go   go-OBTOP
            'it wasn't to the clearing that I went'

OG

17.

(88) pɨtá-i > pɨtáy ⟨pitây⟩   (RM 1639:162v)

            apɨ + i > apɨ̃y ⟨apỹî⟩   (RM 1639:162v)
            point on   'on the point'

n-a-mbo'e-i   > nambo'ey ⟨namboeî⟩ (RM 1640:13)
neg-1s-teach-neg   'I didn't teach them'

a-iko > ayko   (RM 1639:171)
1s-live   'I live'

Urubu-Kaapór

17. This rule applies, but not just across morpheme boundaries; it applies within morphemes as well.

Conclusion

As with #16 above, #17 and #18 are so natural that no conclusion can be drawn from the appearence of this rule in modern TG languages.

Wayapí

18. The rule also applies in W (not clear whether still after *-w)*.

(89)   n + o + poray + i > noporay   (J 1990a:77)
          neg-3-

          n + e + kasi + i > nekasi   (J 1990a:77)
          neg-2s-

Assurini

18. No evidence of this rule in AT

Parintintin

18. $i > \emptyset\ /\ Vy\ \_\_\ \#$

Kayabí

18. Absorpton of *i* by *y* holds for Kb.

(90)   i-poey-i   > ipoey
           3-wash-ОВТОР

OG

18. The absorption of *i* after a glide is also recorded in OG:

(91)   če-aséy   -i > čeaséy   (RM 1639:162)
           1s-shoulders-on 'on my shoulders'

           nda-če-awɨyeteí-i > ndačeawɨyeteí
           neg-1s-2-good-neg 'I am not good'

Urubu-Kaapór

18. No evidence of this rule in UK.

Wayapí

19. Still applies but optionally. The rule may additionally merge the back vowel with the

epenthetic *w:*

(92)   o-eraa > oeraa ~ oweraa ~ weraa   (J 1990a:78)
           3-bring

        o-erek^{w}a > oerek^{w}a ~ oerek^{w}a ~ owerek^{w}a ~ werek^{w}a   (J 1990a:78)
           3-wife

Assurini

19. No evidence of this rule in AT.

Parintintin

19. In Pt *w became g^w. Since g^w patterns the same as *w*, it is hard to know whether this is still phonemically a simple glide /w/. The insertion rule has modified somewhat.

$$\emptyset > g^w / C \begin{bmatrix} +syl \\ +bk \\ +rnd \end{bmatrix} \_\_ +syl$$

$$\begin{bmatrix} +syl \\ +bk \\ +rnd \end{bmatrix} > g^w / \# \_\_ [+syl]$$

This makes the change of {o,u} to g^w more analogous to the change of *i, e* to *j̃*.

Kayabí

19. No evidence of this rule in Kb.

OG

19. This rule is attested in OG.

(93) o-ata > oata ~ owata   (RM 1639:70)
      3-walk   'he walks'

Urubu-Kaapór

19. As with #17 above, this rule in UK applies within morphemes as well as across morpheme boundaries. Furthermore, it appears after all back vowels not just after rounded vowels, even when there is an intervening *š*.

(94) koĩ [kowĩ] 'tomorrow'   (K 1986:400)
      yašer [yašwer] 'it is old/useless'   (K 1986:400)

Conclusion

This rule (of inserting *w* between a rounded vowel and a following vowel) is so natural that it is very likely indeed that such an optional rule applied in PTG. Its manifestation in UK is more mysterious and must, at least tentatively, be regarded as an innovation.

3.3. Conclusion of morphophonological comparison

We have seen that examining the forms that reflect past as well as present synchronic morphophonological rules is more helpful in reconstructing the phonology of PTG. We have also seen how the internal reconstruction of PTG lends important insight into morphophonological rules, especially in recognizing that certain morphophonological rules are not morphophonological at all (Jensen's rule #6 and Rodrigues' rule #10) as well as the re-evaluation of rules such as #8 and #9 in light of the f/l distinction. This should give us a more complete and accurate analysis of PTG phonology and morphophonology, which in turn will be useful in the reconstruction of Proto-Tupí.

PART TWO: MORPHOLOGY
CHAPTER FOUR: NOUNS

4.0. Introduction

Having examined the phonological development of PTG into the daughter languages and having further applied the method of internal reconstruction to PTG thus arriving at pre-TG, we will turn our attention to the development of affixes in TG and the impact of the above phonological hypothesis on the development/reconstruction of these affixes.

4.1. Form classes

Any discussion of morphology in TG must begin with a look at the form classes in TG, since all other morphological phenomena turn on this issue. First we will look at examples of different form classes of nouns in the various languages under examination in this study and then draw our conclusions about the nature of form classes in PTG.

The system of labelling these was devised by Rodrigues (1981) specifically based on data from Tb. However, as we shall see, form classes are organized along very similar patterns in TG languages, and this class system can be used to compare data in all TG languages. Rodrigues' system, echoed in Jensen (1984, 1990), divides roots into two classes, I and II, based on which third person possessive crossreferencing marker is used. Each of these classes, in turn, is divided into sub-classes based on variation in possessive marking.

Classification of roots in PTG

I   combine with third possessive *i-*

         Ia roots which don't begin with *p*
                 combine with Ø-

         Ib roots which begin with *p*
                 combine with *m-*

II   combine with third possessive *c-* and *t-*

         IIa combine with *t-* and with third possessive *c-*
         IIb   "   with *t-* and with *t-*
         IIc   "   with Ø- and with *c-*
         IId   "   with (V>Ø) and with *c-*

III   do not combine with possessive prefixes (nominals only)

These could be further generalized as follows: those roots that take possessives (I & II)

and those that do not (III).

.

Tupinamba

Ia	káŋ 'bone'	še-káŋ 'my bone'	i-káŋ 'his/her/its bone'
	sĩ 'mother'	še-sĩ 'my mother'	i-šĩ 'his/her/their mother'
Ib	mbó 'hand'	še-pó 'my hand'	i-pó 'his hand'
	mbĩ 'foot'	še-pĩ	i-pĩ 'his foot'
IIa	t-embé 'lip'	še-r-embé 'my lip'	s-embé 'his lip'
	t-endáβ 'place'	še-r-endáβ 'my place'	s-endáβ 'his place'
	t-áβ 'hair'	še-r-áβ 'my hair'	s-áβ 'his hair'
	t-úβ 'father'	še-r-úβ 'my father'	s-úβ 'his father'
IIb	t-endĩr 'older sister'	še-r-endĩr 'my older sister	t-endĩr 'his older sister'
	t-aʔĩr 'son'	še-r-aʔĩr 'my son'	t-aʔĩr 'his son'
IIc	ók 'house'	še-r-ók 'my house'	s-ók 'his house'
IId	peé 'road'	še-r-apé 'my road'	s-apé 'his road'
	panakū 'basket'	še-r-epanakū 'my basket'	s-epanakū 'his basket'
III	amãn 'rain'		
	kʷár 'sun'		
	ár 'day'		
	yasĩ 'moon'		

Assuriní of Trocará

Ia	kɨŋ(a) 'head'		
	hɨke 'mother'		
	mena 'husband'	i-mena 'her husband'	Note: Ib is gone by merging
	ʔawa 'hair'	i-ʔawa 'his hair'	with Ia

Ib	pɨʔa 'liver'	i-pɨʔa 'his liver'
	paa 'hand'	i-paa 'his hand'
	pɨa 'foot'	i-pɨa 'his foot'

IIa	t-era 'name'	če r-era 'my name' h-era 'his name'	Note: IIb is gone
	t-ɨroa 'canoe'	če r-ɨroa 'my canoe' h-ɨroa 'his canoe'	(merged with IIa)
	t-áwa 'person'	če r-awa 'my person' h-awa '	IIc is gone
b	t-aʔɨra 'son'	če r-aʔɨra 'my son' h-aʔɨra 'his son'	
	t-ope 'father'	če r-ope 'my father' h-ope 'his father'	
c	aŋa 'house'	če r-aŋa 'my house' h-aŋa 'his house'	

III	amɨna 'rain'
	kʷarahɨa 'sun'
	ara 'day'
	čahɨa 'moon'

IV	ɨhara 'canoe'	če ɨhara 'my canoe' i-čɨhara 'his canoe'	New class from PTG Class Ia nouns with root initial ɨ.

Kamayurá

Ia	i-kaŋ 'his bone'		
	i-ɨ 'his mother'		
Ib	i-po 'his hand'		
	i-pɨʔa 'his foot'		
IIa	ye r-et 'my name'	h-et 'his name'	Note: apparently Km has
			only possessed
b	ye r-aʔɨt 'my son'	t-aʔɨt 'his son'	forms of possessible
	ye r-up 'my father'	t-up 'his father'	nouns.
c	ye r-ok 'my house'	h-ok 'his house'	

III   aman 'rain'
      etc.

Tapirape

Ia ākɨŋ 'head'
  ɨ 'mother'
  itā 'stone'
  ham 'hair'
  men 'husband'

Ib ma 'hand'   i-pa 'his hand'
  mɨ 'foot'   i-pɨ 'his foot'
  mɨhā 'liver' i-pɨhā 'his liver'
  mahɨŋ 'medicine' i-pahɨŋ 'his medicine'

IIa eɨmam 'pet' če r-eɨmam 'my pet' eɨmam 'his pet'
  ekʷam 'net' če r-ekʷam 'my net' ekʷam 'his net'

IIb t-ā?ɨt 'son' če r-ā?ɨt 'my son' t-ā?ɨt 'his son'
  t-op 'father' če r-op 'my father' t-op 'his father'
  t-ɨket 'older sister' če r-ɨket 'my older sister' t-ɨket 'her older sister'

IIc t-akan 'house' če r-akan 'my house' akan 'his house'
  t-et 'name' če r-et 'my name' h-et 'his name'

IId[1] pe 'road' če r-āpe 'my road' āpe 'his road'

IIe wɨ 'blood' če r-owɨ 'my blood' wɨ 'his blood'
  wāy 'tail' če r-owāy 'my tail' wāy 'his tail'

III amɨn 'rain'
  ?an 'day'
  āwɨt 'house'
  konomī 'boy'
  kotātāi 'girl'
  wɨrā?i 'bird'
  patɨt 'flower'
  tam 'village'

[1]This is (IIe) in Almeida et al. (1983). My (IIe) is (IId) in Almeida.

In Tp, curiously, there are a number of roots in class IIc that in other TG languages are in other classes, chiefly IIa. At the same time, the only word attested in most languages for IIc, *akan* (PTG **okár*) is in class IIa in Tp. The readiest explanation for this is that some vocabulary in IIa has stopped being prefixed with *t-*. The determining factor for which roots are not prefixed with *t-* in their non-possessed form seems to be semantic. Hence members of class IIc are almost all roots either referring to the body (*aʔa* 'flesh', *ĕā* 'eye' *eni* 'saliva' *eme* 'lower lip' *awā* 'forehead'), or are compounds based on such a root (*emeʔip* 'river bank', literally 'earth-lip'). The only roots which appear to be exceptions to this class are *ātā* 'fire', *am* 'feather', *ap* 'leaf', and *āti* 'wife', and *et* 'name' and *oʔip* 'arrow' (these included in an open list of thirty-three roots). Of these, *āti* and possibly even *et* could be construed as belonging semantically to this class, if these were seen as being inalienably possessed, depending on cultural beliefs about wives and names.

OG

Ia	káŋ 'bone'	če-káŋ 'my bone'	i-káŋ 'his/her/its bone'
	sɨ 'mother'	če-sɨ 'my mother'	i-čɨ 'his/her/their mother'
Ib	mbó 'hand'	če-pó 'my hand'	i-pó 'his hand'
	mbɨ 'foot'	če-pɨ 'my foot'	i-pɨ 'his foot'
	mbɨʔá 'liver'	če-pɨʔá 'my liver'	i-pɨʔá 'his liver'
	moʔaŋ 'medicine'	če-poʔáŋ 'my medicine'	i-poʔaŋ 'his medicine'
IIa	t-embé 'lip'	če-r-embé 'my lip'	h-embé 'his lip'
	t-endáβ 'place'	če-r-endáβ 'my place'	h-endáβ 'his place'
	t-áβ 'hair'	če-r-áβ 'my hair'	h-áβ 'his hair'
	t-úβ 'father'	če-r-úβ 'my father'	h-úβ 'his father'
IIb	t-endɨ 'sister'	če-r-endɨ	t-endɨ 'his sister'
	t-aʔɨ 'son'	če-r-aʔɨ 'my son'	t-aʔɨ 'his son'
IIc	óγ 'house'	če-r-óγ 'my house'	h-óγ 'his house'
	ɨrū 'container'	če-r-ɨrū 'my container'	h-ɨrū 'his container'
IId	peé 'road'	če-r-apé 'my road'	h-apé 'his road'
	panakū 'basket'	če-r-epanakū 'my basket'	h-epanakū 'his basket'
III	amā 'rain'		
	kʷá 'sun'		
	á 'day, time'		
	yasɨ 'moon'		

Kayabí

Ia		i-kaŋ 'his bone'
		i-ɨ 'his mother'
		i-men 'her husband'
	ɨat 'canoe'	i-ɨat 'his canoe'
		i-sĩ 'his nose'
	kanawa	i-kanawa

Ib		i-po 'his hand'
		i-pɨ 'his foot'
		i-pɨ'a 'his liver'

IIa	t-aitɨ	yer-aitɨ	'ŋar-aitɨ
IIb	t-up 'father'	yer-up 'my father'	'ŋar-up 'his father'

IIc	'ok 'house'	yer-ok 'my house'	'ŋar-ok 'his house'
	'et 'name'	yer-et 'my name'	'ŋar-et 'his name'
	'ɨrũ	yer-ɨrũ	'ŋar-ɨrũ
	'o'o 'flesh'	yer-o'o 'my flesh'	'ŋar-o'o 'his flesh'

IId

III	ɨwɨtɨt 'wind'
	amãn 'rain'
	kʷat 'sun'
	yaɨ 'moon'

The system of form classes in Kayabi has been greatly simplified. Classes I and II no longer have any distinction between their subclasses with one exception: class IIc not only remains distinct, but has grown by the addition of other roots in other subclasses of class II. Hence

-et (< PTG *ér*) is *'et* in its unpossessed form (cf. Tb *t-ér*). In class II, the original third person possessive *c-* in IIa and IIc (which would have been ∅ in Kb) and *t-* in IIb has been replaced by the third person independent pronouns (to be discussed in Chapter 5).

Also worthy of notice is that roots of both classes have further divided into always possessed and sometimes possessed nouns.

Urubu-Kaapor

Ia   i-ākā 'his head'
i-membɨr 'her child'
i-ham 'his hair'
i-nambi 'his ear'

Ib   i-po 'his hand'
i-pɨ 'his foot'
i-pɨʔa 'his liver'

IIa   ihē r-er 'my name'   h-er 'his name'
ihē ramūi 'my grandfather'   h-amūi 'his grandfather'
ne r-āi 'your tooth'   h-āi 'his tooth'
ihē r-endɨ 'my sister'   h-endɨ 'his sister'
ihē r-akehar 'my wife'   h-akehar 'his wife'
ihē r-u 'my father'   h-u 'his father'

IIc   ok 'house'   ihē r-ok 'my house'   h-ok 'his house'

IId   pe 'path'   h-ape 'his path'

III   aman 'rain'
kʷarahɨ 'sun'
ar 'day'
yahɨ 'moon'
kaʔa 'forest'
awaši 'corn'
kupiša 'garden'

IV   ihē kupe 'my back'   šupe 'his back'

elements of I and IV: ihē kɨwɨr   i-šiwɨr
ihē kʷar 'my hole'   i-šʷar 'his hole'

NB: The distinction between Ia and Ib is lost because nouns in UK are not only classifiable as never possessed and sometimes possessed, but also as always possessed. Some of the sometimes possessed nouns became always possessed forms. An unusual exception to this is *ok* 'house', which thus preserves the 'class' of IIc. The 'class IV' derives from a morphophonological change PTG *ik* > *š̆*.

Parintintin

Ia	akaŋ 'head'	i-akaŋ 'his head'
	i 'mother'	
Ib	mbo 'hand'	i-po 'his hand'
	mbi 'foot'	i-pi 'his foot'
	mbiʔa 'liver'	i-piʔa 'his liver'
IIa	t-uβ father	t-uβ 'his father'
	t-er 'name'	h-er 'his name'
IIb		
IIc	okan 'house'	h-okan 'his house'
	oŋga 'house'	h-oŋga 'his house'
IId	pehe 'path'	
III	aman 'rain'	
	kʷara 'sun'	
	ar 'day'	
	jahi 'moon'	

Wayãpi

Ia	iakã 'head'			i-akã 'his head'

Ib	mo 'hand'			i-po 'his hand'
	moã 'medicine'			i-poã 'his medicine'

IIa	t-ea 'eye'	e-rea 'my eye'	h-ea 'his eye'
	t-e(r) 'name'	e-re 'my name'	h-e 'his name'

IIb	t-aʔɨr 'son'	e-raʔɨr 'my son'	t-aʔɨr 'his son'
	t-u 'father'	e-ru 'my father'	t-u 'his father'

IIc	ok 'house'	e-rok 'my house'	h-ok 'his house'

IId	panakũ 'basket'	e r-epanakũ 'basket'	h-epanakũ 'basket'

III aman 'rain'
awasi 'corn'
yaɨ 'moon'
kwaraɨ 'sun'
arɨβ/wo 'day'

From the data we have, it appears that W has remodeled class I nouns after third person possessed forms. We cannot say that possessible nouns are only possessed: note IIa nouns such as *t-ea r-ea h-ea* 'eye', which have a unique prefix for unpossessed forms.

Guajajara

Ia   pɨ 'foot'

Ib   pɨ 'foot'

IIa   he-r-u 'my father'

IIb

IIc

IId

III

Chiriguano
Ia   kā 'bone'
      sɨ 'mother'

Ib   po 'hand'
      pɨ 'foot'

IIa   t-esa 'eye'
      t-ape 'path'   če-r-ape 'path'   h-ape 'path'

IIb

IIc   o 'house'   če-r-o 'my house'   h-o 'his house'

IId

III   āma 'rain'
     kʷaráɨ 'sun'
     a 'day'
     jásɨ 'moon'

Kaiwá

Ia kãɨ 'bone'
  sɨ 'mother'

Ib po 'hand'
  pɨ 'foot'
  pɨʔa 'liver'

IIa t-erɨ 'name' čererɨ 'my name' h-erɨ 'his name'
  túβɨ 'father'

IIb t-uβɨ

IIc óγa 'house' če-r-oγa 'my name' h-oγa 'his name'

IId tape 'path' če-rape 'my path' h-ape 'his path'

III ama/amɨn 'rain'
  kʷarahɨ 'sun'
  árɨ 'day'

Mbyá

Ia kã 'bone'
  čɨ 'mother'

Ib po 'hand'
  pɨ 'foot'
  pɨʔa 'liver'

IIa terɨ 'name' čerɨ 'my name' herɨ 'his name'

IIb

IIc o 'house' če-r-o 'my house' h-o 'his house'

IId tape 'path'

III amã 'rain'
  kʷaraɨ 'sun'
  ára 'day'

Guarayo

Ia   kaŋg^{w}er 'bone'   i-kaŋg^{w}er 'his bone'
       cɨ 'mother'   i-cɨ 'his mother'

Ib   po 'hand'
       pɨ 'foot'
       pɨ^{ʔ}a 'liver'

IIa   t-uɣ^{w}ɨ 'blood'   -r-uɣ^{w}ɨ 'blood' c-uɣ^{w}ɨ 'blood'
      t-er 'name'   če-r-er 'my name'   c-er 'his name'

IIb   u 'father'   če-r-u 'my father'   c-u 'his father'

IIc   o 'house'   če-r-o 'my house'   c-o 'his house'

IId   ape 'path'

III   aman 'rain'
      ar(ɨ) 'day'
      yacɨ 'moon'

4.1.1. Form classes in PTG

From the above data, we can reconstruct the following root classification system: the system of roots in PTG was essentially the same as that of Tupinamba and OG above. There was certainly in the parent language a dichotomy between what here are referred to as class I and class II roots. Further the evidence from W and Tp (and Gy?) suggests that IIb was inherited, while that of IIc is even more widely attested (i.e. every language except Tp). Evidence for Class IId is more scarce, and we may have to be more agnostic for the time being concerning its status in PTG.

PTG

Ia   akã‖ 'head'   i-akaŋ
      cɨ̃ 'mother'   i-čɨ̃
      itá 'stone'   i-y-ita
      ʔáβ 'hair'   iʔáβ
      mẽn 'husband'   i-mẽn

Ib   mbó 'hand'   i-pó 'his hand'
      mbɨ̃ 'foot'   i-pɨ̃ 'his foot'
      mbɨ̃ʔá 'liver'   i-pɨ̃ʔá 'his liver'
      moʔãŋ 'medicine'   i-poʔãŋ 'his medicine'

IIa   t-eɨmbaβ 'pet'   če r-eɨmbaβ   c-eɨmbaβ
      t-ekʷam 'net'   če r-ekʷam   c-ekʷam
      t-uwɨ̃ 'blood'   če r-uwɨ̃   c-uwɨ̃
      t-uwáy 'tail'   če r-uwáy   c-uwáy

IIb   t-ãʔɨ̃r 'son'   če r-ãʔɨ̃r   t-ãʔɨ̃r
      t-úβ   'father'   če r-úβ   t-úβ
      t-ɨkér 'older sister' če r-ɨkér   t-ɨkér

IIc   ʔók 'house'   če r-ók   c-ók
      ʔokár 'house'   če r-okár   c-okár
      ʔér 'name'   če r-ér   c-ér

IId   pé 'road'   če r-apé   c-apé
      panakũ   če r-epanakũ   c-epanakũ

III   amãn 'rain'
      ʔár 'day'
      kunumĩ 'boy'
      kuyãtãi 'girl'
      wɨraʔí 'bird'
      potɨ̃r 'flower'
      táβ 'village'

Having arrived at a comparative reconstruction of PTG root classification, we now apply to this reconstruction the internal method to develop some theories as to the origin of this classifying system.

4.1.2. Form classes in pre-TG

First it should be noted that there is no clear semantic rationale for the system. Certainly there is no grammatical rationale (e.g. nouns predominating in one class, verbs in another) so functional motivation could be ruled out. Phonological motivation seems, at least at first glance, to be the most satisfying, but it does not completely explain the root class system.

The most likely conclusion is that PTG started out with form classes that were determined by phonological characteristics alone and that this system was later complicated and obscured by reanalysis and semantic considerations. The first phonological motive for form classes was the difference between consonant-initial and vowel-initial roots: class I nouns were consonant-initial, class II nouns were vowel-initial. The use of *t-* with generic or non-possessed forms was then restricted to roots which were phonetically vowel initial.

The second phonological phenomenon further distinguishing form classes involved the phenomenon of nasalization discussed in (2.4) and (3.1) above. As mentioned in (2.4), nasalization of stops word-initially was restricted to stops also subject to lenition,

i.e. *p* and *t*. But the use of *t* as a marker of class II roots appears to have led to the reanalysis of all *t*-initial roots as class II forms. Hence, by default, the only root-initial consonant subject to nasalization is *p*. In PTG the nasalized *m*- occurs in a context in which the root is not the head of a possessor phrase, while the unnasalized *p*- forms occur exclusively when preceded by a possessing pronominal. This is because a preceding oral proclitic may block nasalization, while the lack of such preceding morphology allows nasalization to take place.

po ~ mbo   če-po

The preserved root-initial *p* was extended to occurrences of these roots possessed by noun phrases. Only the nasal variation was used to express general possession.

As seen in the above comparison, many TG languages do not have this *m/p* distinction. In some languages, such as Guaranian languages, we know that the original distinction was lost. In some languages, however, the nasalized variant may never have been generalized to all *p*-initial roots and/or to all *p*-initial roots not preceded by a clitic.

The third phonological phenomenon further distinguishing form classes involved the phenomenon of pre-glottalization. Because of preglottalization, roots which began with a glottal stop were phonologically vowel-initial, possessing the autosegmental phoneme of preglottalization:

O
μ
|
ʔσ
|
/ok/   [ʔok]

On the one hand, such roots were marked as other class II (vowel-initial) roots, while at the same time prefixing a stop to such a root was, in PTG phonology, ungrammatical.

O
μ
|
ʔσ
|
/tok/   §[ʔtok]

This led to the first split in class II, namely between so-called class IIa and class IIc. Class IIa originally consisted of roots which were both phonologically and phonetically vowel initial while IIc consisted of roots which were phonologically vowel-initial but phonetically consonant-initial, i.e. beginning with a glottal stop (which, in PTG phonology would have made the root vowel initial underlyingly). Preglottalization in roots like ʔók, and ʔirū prevented the prefixation of t-, but strangely this was not so when preceded by a possessive pronominal or noun, since in PTG, the class marker r- used with possessives was considered a separate phoneme, one which could be preglottalized:

```
        O
        μ
        |
       ʔσ
        |
     /rok/    √[ʔrok]
```
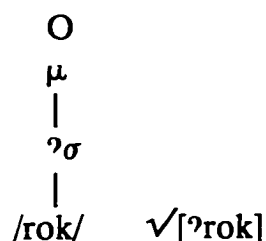
In the earliest TG manuscripts glottal closure was not transcribed consistently; in any given word there is no consistent way of indicating whether there is glottal closure in that word. This makes it impossible to say for sure whether forms like *če-ʔrók* existed in the historical period, and hence can be reconstructed with confidence for PTG. It may be that in PTG, just as in Kb, preglottalization was reinterpreted as the non-possessive morpheme for class IIc roots, just as nasalization was reinterpreted as the non-possessive morpheme for class I(b) roots. If so, this meant that preglottalization could be dropped when a noun was the head of a possessive construction. The difference between class Ib and IIc roots is that Ib added an autosegmental feature not present before in the non-possessed form and did not add it when the head of a possessive phrase; while class IIc nouns kept an autosegmental feature which was originally part of the root, and deleted that feature when the noun was the head of a possessive phrase. At any rate, almost all modern TG languages changed *ʔr into r by means of natural sound change independent of morphological considerations. Naturally, as TG languages developed the phonological rule that automatically adds a glottal stop before a word-initial vowel, the glottal stop at the beginning of IIc nouns expressing general possession was not significant and more and more IIc nouns were reinterpreted as members of one of the other subclasses of class II.

The fourth phonological phenomenon to differentiate form classes in PTG was the optional loss of initial vowels before p. this meant that roots beginning with Vp could lose the initial vowel in 'generally possessive' form while retaining it when prefixed by a possessing noun or pronominal proclitic. This gave rise to the very small class IId.

After these phonologically factors, the different form classes came to be reinterpreted as semantically motivated. Hence the m- of class Ib was reinterpreted as a non-possessed prefix. This reinterpretation was extended even further as entire classes were reinterpreted as having a particular semantic field. Hence certain class II roots which were considered inalienable were set aside as a separate subclass (distinguished in PTG by using t- as the third person possessive as well as the generally possessive form).

Another, later, result of this reanalysis was that class I forms came to be seen as less marked than class II forms and so some roots shifted from class II to class I or from one subclass to another in a way similar to the reanalysis of Old English strong verbs as weak verbs. So for example, *$?okár$ 'house', a class IIc root, is a class IIa root in Tp (*akan*) while in the same language, almost all p initial roots have moved from class Ib to class Ia, except for roots based on *pa* 'hand' and *pɨ* 'foot' and a few others. To a lesser extent this change has also taken place in Pt. This contrasts with Tb where, according to Anchieta, all p initial roots without exception belonged to class Ib.

Many roots which are vowel initial exist in class Ia. This perhaps is the result of a reanalysis of class IIc roots. In OG, for example, and perhaps in PTG as well, there was a rule automatically adding *?* to a vowel-initial word. This means that word-initial preglottalization would no longer be discernable and many roots beginning with *?* were taken into class I (as consonant initial) and later reanalyzed as vowel-initial. An example of this is **akáŋ,* clearly not preglottalized (Tb *iyakáŋ,* OG *iñakāŋ,* 'his head', not §i$^{?}$akáŋ). But this word is equally clearly from ***?á* 'head' + ***káŋ* 'bone' (***káŋ* 'bone' is the same in PTG and is attested in many TG languages (see **1.1.** #107); for ***?á* 'head', cf. non-TG Tupian examples: Karitiana *?e* 'head', Munduruku *-?a$_2$* 'head').[2]

We can now summarize the development of the root case system in the following way:

Phonologically motivated
       consonant vs. vowel (I vs. II)
       p-initial vs. other consonant (Ia vs. Ib)
       ?-initial vs. vowel initial (IIa vs. IIc)
       Vp vs. V+other C (IId vs. rest of II)

Semantic reinterpretation
       kinship vs non-kinship (IIb vs. rest of II)
       possessible vs. obligatorily possessed

Marking the third person possessive as *t-* or *c-* may have been due to the root-initial

---

[2] In non-TG Tupian languages, *?* seems to be a segment rather than an autosegment. Based on this evidence I suppose here that *?* was at an even earlier stage before PTG, a segment, then became an autosegment in PTG. Munduruku is a tone language; the supscript numeral in *-?a$_2$* denotes a mid-level tone.

vowel, assuming that *c-* is from ***t-i-* (see below). A root-initial front vowel might result in *-i-* assimilating or dropping out on the one hand (leaving *t-* as the possessive marker), or a root-initial vowel might cause synezesis: *t-i-> tʸ- > c-*. This would result in two alternative forms of the possessive marker, again phonologically conditioned. The semantic reinterpretation of these prefixes, however, led to the use of *t-* with kinship terms while *c-* was used with all other roots. Hence all roots originally using *t-* for the possessive marker shifted to class IIa while all kinship terms shifted to class IIb, regardless of whether they originally used *c-* or *t-* for the possessive marker.[3]

The comparative evidence points to a root class system like that of Tb and OG. We have used nouns to exemplify this system; however, as we shall see, TG languages have very fine lines between nouns and other parts of speech. In some languages, there is no morphological distinction at all.

4.2. Nominalizers in PTG

The phonological analysis given in Part I can also shed light on the development and origin of nominalizers in TG. In this section we will make a comparative survey of nominalizing affixes in TG languages to ascertain the PTG nominalizing system. Again, we will first take a comparative overview of TG nominalizing suffixes, leading to a reconstruction of PTG nominalizers.

---

[3] The fact that the word 'mother' is not a class II root is due to the fact that it is a consonant-initial root, although it is possible that it was originally class II *t-yɨ > *cɨ.

Tupinamba

The nine nominalizers in Tb all have cognates in at least some TG languages. No other TG languages have any different nominalizers.

The prefix *mi-* derives from a transitive verb a noun which is the object or goal of the action of the predicate. It makes roots into Class IId stems. All other nominalizers are suffixes.

The clitic *-βa?e* was attached to a verb phrase rather than a root or stem. It indicated the subject of an intransitive predicate.

The suffix *-pĩr* indicates the patient of a transitive verb. No functional difference seems to exist between *mi-* and *-pĩr.*

The clitic *-sár* marks the agent of transitive verbs. It has three allomorphs: *-sár, -ár,* and *-tár.* These will be discussed more detail in **4.2.1** below.

A clitic called 'circumstantial' in function is *-sáβ* . It derives nouns which are neither agents nor patients of the root predicate. It has the same morphophonemics as *-sár.*

Another agentive suffix is *βór;* this suffix marks habitual agent.

The suffix -*a* simply makes a verb an abstract noun. It is only used with verb roots ending in consonants, with the morphonological rule that a root final lenis becomes fortis when this suffix is added.

The abstract nominalization used for verbs ending in a vowel is -*aβ-βo*. It is functionally identical to -*a* above. Heretofore this has been written -*áβo*, but the problem with doing so is that no other morpheme in Tb, free or bound, bears penult stress. It must therefore be analyzed as -*áβ-βo*, that is, the *s*-less allomorph of -*sáβ* plus the distributive locative suffix -*βo*. Because of the rule of haplology in Tb (see **3.1**), this would have been -*áβo* on the surface.

Guajajara

mi-   goal or object of action
            he-mi-reko 'his wife' (B-S 1972:117)

-har   agentive
            karaiⱱ wa-mume?u-har 'one who told the white man' (B-S 1972:118)

ma?e   agentive
            uhu-katu ma?e 'one that really is big' (B-S 1972:119)
            i-zupe u-ze?eŋ ma?e 'one who talks to him' (B-S 1972:119)

-haw   instrumental, abstract, resultative
            i-zuka-haw 'the killing of him' (B-S 1972:120)
            herezepin-haw 'scissors' (B-S 1972:121)

pɨr   patient
            acapo tɨm pɨr 'rice, which is planted' (B-S 1972:122)

-a   abstract noun

AT

emi-

-awa instrumental, locative; -awa, -tawa,
    oʔam > oapawa 'standing up'; aha > aatawa 'going'; očoka >
    iče ičokatawa 'the place where I killed him'; oken > čenekehawa
    'the place where we sleep'; opɨhɨ(ŋ) > ipira-pɨɨkawa 'the fishing net
    (device for catching fish)'; esak > mo-r-esak-awa 'camera'

-ara agentive
    oeraha > heraatara 'someone who takes it'
    opɨhɨŋ > ipira-pɨɨkara 'someone who fishes'
    omoʔe > konomia-moena 'someone who teaches children'

-waʔe subject (stative verbs only)
    hahywaʔe 'someone who is sick'
    iroʔɨwaʔe 'someone who has a fever'

-o abstract noun   o-ata-o

-a abstract noun γ + -a = ka r + -a = ta; w + -a = pa; y + -a = -ta

	/C_	/V_	/y_
circ	á(βa)	há(βa)	tá(βa)
agent	ára	hára	tára
	a	o	ta

Tp

-āwa (-am-a) circumstance, instrumental, locative, etc.
    āičāk nenopɨāwa 'I saw the circumstance of hitting you'
        (Aldeia 1983:35)

    amapen nekāwipɨwāwa 'he broke your neck from drinking liquor'
        (Aldeia 1983:35)

emi- patient
    čeremiʔo hēʔē 'my food is tasty' (Aldeia 1983:35)

-a	abstract
-ana	abs. agent (tr. verbs)
-maʔe	relative agent (intr. verbs)
-pɨra	patient
emi-	object
-āwa	circumstance
-maʔe	habitual agent (intr. verbs only)

GM

embi-	patient
-βɨ	abstract noun
aŋʷa	circumstance, instrument
-βaʔe	agent

Kw

embi-	patient
-βa	relativizer
haŋwā	circumstance, purpose
-há	subordinator, instrument
-hára	agent
-βɨ	subordinator

Ch

mbáe   abstract noun

      -kʷa   mbaékʷa 'knowledge, knower'

      -u   mbaeporóu '

      -nũpa   mbaenũpa 'azote'

      -mõña   mbaemõña

      -ápo   mbaápo 'work'

      -káwi   mbaekáwi

      kɨra   mbaekɨra

      agʷɨ̈ye   mbagʷɨ̈ye 'perfection, maturity'

      potɨ 'flower' mbaepotɨ 'any flower'

embi-   patient

      -u   tembíu 'food'

      -monde temimonde 'clothes'

      -ápo   tembiápo 'creation' 'work'

      -póta temimbóta 'desire'

      -eča   tembieča 'sight,panorama, spectacular'

-wa   relativizer

mbïa kõe oyápo-wa 'the man who makes the >>

-βae

-a   kõe iyapóa the maker of >>

-a   abstract noun

      če-āyñ-a

      óu če-āyñ-a-pe

      õe háe-õyñ-a-gʷi

Km

emi-   patient

-ap/-tap   abstract noun

-at/-tat   abstract noun

| -pɨt | patient |
| -maʔe | agent |

Gy

embi-	patient
-βaʔe	agent
-car	agent
-caβ	circumstance, instrument
-pɨr	patient

Pt

embi-	patient
-βeʔe	agent/relativizer
-har	habitual agent (-γ-har > -kar)
-haβ	instrument, circumstance
	(-g-haβ > -kaβ; ŋ-haβ > ŋgaβ; ʔu-haβ > ʔgʷaβ; m-haβ > mbaβ; n-haβ > ndaβ; β-haβ > paβ; r-haβ > haβ; y-haβ > ytaβ)
-ndɨβ	place
-pɨr	patient

UK

| -ha | instrument, agent, abstract |
| | maniʔo sosok-a 'manioc pounder'; u-sak-iha 'the one who was seeing (him)'; o-ho-ha 'his going' |

·

-me?ẽ	subject
	i-ho-hon-me?ẽ 'the one who went'; katu-me?ẽ 'the one who is good'; pitɨm ra?ɨr pukek-me?ẽ ke 'cigarettes which were wrapped' (which
he	wrapped?)

-har	patient
	yete-har 'the thing that is true'; rake-har 'one who lives beside'; ka?a rupi-har 'thing of the forest'; ɨman-te-me?ẽ rehe-har 'something that pertains to ancient times';

(i)mi-	patient (non-productive)
	h-imi-?u 'his food'; arar mi-?u 'macaw's food'; h-imi-yar 'his hunt'

-kʷer	nominalizer (nonproductive);
	katu-kʷer 'good things'

Kb

-at	agent
-ap	instrument
-a	action
-pɨt	patient of a temporal action
-mi-	" " " habitual action
-ma?e	patient
-?wat	determined objective agent
-?wap	" " instrument
-?ap	" " abstract action
-wat	determined place

W

emi-   object
                emi-yuka 'someone who was killed'
                er-emi-esa 'someone who I met'

-a(r)   agent
                mo?e-a 'teacher'

-a   locative
                mo?e-a 'place of teaching'; pira rekɨy-ta 'place or instrument of
                fishing'; ɨwɨ pɨkũy-na 'instrument for digging the earth'

ma?e   relative
                o-ata ma?e 'someone who walks'

-ka/-pa   instrumental

-wa(r)   agent of postpositional phrases
                ɨtuwasu pe-wár 'resident of Cachoeira'
                yawa rewa   'someone who always works with jaguars'

-we   agent
                iyā-we 'runner'

OG

embi-   object
                h-embi-reko 'his wife'

-há(ra)   agent   -hara/-ara/-tara

-βa?e   agent

-há(βa)   instrumental, abstract, resultative -saβ/-aβ/-taβ

-pɨr   patient

-βo   abstract noun
                a-ha i-yuká-βo 'I am going to kill him'

-βo   habitual agent
                  mboraɨhú-βo če 'I am someone who loves'

-sé   habitual agent
                  nda-če-ho-sér-i 'I am not someone who wants to go'

	/C_	/V_	/y_
circ	á(βa)	há(βa)	tá(βa)
agent	ára	hára	tára
	a	áβo	ta

4.2.1 Comparative and internal reconstructions of the nominalizing suffixes

From the above comparisons, we reconstruct the following nominalizers for PTG:
PTG

embi-	patient
-cár	agent
-βaʔe	agent
-cáβ	'circumstantial' i.e. place, manner, instrument
-pír	patient
-áβ-βo	serial verb suffix (i.e. circumstantial + distributive locative)
-βór	habitual agent
-ca	serial verb suffix

Three of these nominalizers form a set in most TG languages, in which all three nominalizers have the same morphophonemics. According to Jensen (1990a) these nominalizers also shared the same morphophonemics in PTG:

	/C_	/V_	/y_
circ	áβ	cáβ	táβ
agent	ár	cár	tár
serial	a	áβo	ta

This chart can be simplified by recognizing that -áβ and -cáβ are historically the same morpheme, likewise -ár and -cár. As Jensen points out, 'quando os suffixos nominalizadores *+ár e *+áβ seguem um tema que termina em fricativa bilabial esta última se torna oclusiva (β>p).' (1990a:116) She uses the example from Tupinamba kʷáβ > kʷapár. Many other examples could be given:

    a-h-ečáγ   h-ečáka   (RM 1640a:26)
    1s-3-see   3-see=sub.

    a-h-endúβ   h-endúpa   (RM 1640a:26)
    1s-3-hear   3-hear=sub

This occlusion is seen even in languages where the final consonant has been lost.

The morphophonemic rule /lenis/ > /fortis/ / _ + áβ , ár

remains a mystery only if we regard this rule as strictly synchronic. It is not a mystery, however, if we apply the diachronic phonology of pre-TG to see the origin of this rule. I propose that *-áβ and *-ár were originally **-cáβ and **-cár, respectively, which originally applied to all verb stems, vowel final and consonant final. Vowel finally there would be no change in the suffix:

    *karú + cáβ > *karucáβ

Consonant-final verb stems, however, would result in an assimilation:

    *endúp + cáp > *enduppáp

After the lenition rule given in the previous chapter, this would give us *endupaβ*. This also explains the behavior of nominalizers after nasal-final consonants:

ye?éŋ + ca > ye?éŋga
speak nom speaking

According to the oralization rule the above could be analyzed synchronically as

ye?éŋ + a > /ye?éŋa/ = [ye?éŋga]

mo?é mo?eár
peté peteár

In some cases in Wayampi the consonant form is preserved:

apík apik-áβ
↓ ↓
apɨ apɨka (reanalyzed as *apɨ-ka*)   (J 1990a:115)

kʷáβ kʷápaβ
↓ ↓
kuaa kuapa (reanalyzed as *kua-pa*)   (J 1990a:116)

in Wayampi:
　　ex. kusiwa-ka　　'instrument for drawing'　(J 1990a:116)
　　　　kusiwa-pa

This reduces the list of morphemes to two in TG:

circ. cáβ táβ
agent cár tár

That there should be a separate circumstantial and agentive suffixes just for verbs ending

in /y/ is highly unlikely; it is more likely that these, too, are allomorphs of *-cáβ* and *-cár*.

The derivation of *c* from ***p^y*, ***t^y*, and ***k^y* discussed above may explain the

morphophonemic rule:

$$-ca > -ta \ / \ y\_$$

This would have originated from a dissimilation.

<div align="center">

k^{w}áy-t^{y}a

k^{w}áy-ta

</div>

This also indicates that the dissimilation took place after the lenition rule. Consider that,

if the dissimilation had taken place before lenition, we would have *kway-ra.*

endúp-t^{y}a	kwáy-t^{y}a	kwáy-t^{y}a
endúppa	kwáyt^{y}a	kwáyta
endúpa	kwáyt^{y}a	§kwáyra
____	kwáyta	____

The **t^{y} like other palatalized stops does not lenite; there is no evidence of β^{y} or r^{y} in

PTG.

Hence historically we can see one suffix each for circumstantial and agent, *cáβ*

and *cár* respectively. The same analysis applies to the serial i.e.

-a	-áβo	-ta
C_	V_	y_

where *-a* and *-ta* come from earlier *-ca* (the puzzle of *áβo* will be dealt with below).

This in turn points to three nominalizing suffixes:

-ca   gerund (general nominalizing suffix)
-cáβ   circumstantial
-cár   agent

Because all of these contain the element *ca we can propose that the latter two

nominalizers are composed of *-ca and *-β (**-p) was circumstantial and *-r (**-t) was

agentive. The fact that there is a second morpheme added to *-ca in these forms may

correlate with the fact that they bear stress, while *-ca 'gerundive' is unstressed.

The idea that *ca was originally a nominalizing suffix is supported by data from

Anchieta.

> ycatûcepiâcagoera   'it was pleasant to see'   (A 1595:28v)
> (i-katu s-ep^yáka wéra)
>
> ycatûcepiâcao ~ áma 'it will be pleasant to see' (A 1595:28v)
> (i-katu s-ep^yáka-o-áma)

Here Tb uses nominal tense on a verb with -a (< *-ca).

The idea that *-p by itself was the actual 'circumstantial' morpheme and *-t was

the actual agentive morpheme has some further support in *-βór 'habitual agent' attested

only in Tupinamba and OG. This looks very much like *-βo + r where -βo is the PTG

suffix meaning 'distributive locative' or habitual adverbial suffix added to nouns (e.g.

*ariβo 'daily'). The combination of the two suffixes would mean 'doer of x over a length

of time', i.e. 'habitual agent'.

This same -βo may be linked to -áβo gerundive listed above. Evidence for the

shape *-áβo* comes from several languages.

OG: Restivo 1724:1,47
Kayabi: Weiss 1972:vi
Parintintín: Betts 1981:23-4
Tupinamba: Rodrigues 1981:4-5,8

We can imagine that there was a stage in which the nominalizers were formally the same

in all environments:

/C_	/V_	/y_
cát	cát	cát
cáβ	cáβ	cáβ
cár	cár	cár

We have noted the semantic similarity between *-βo,* 'distributive locative' and *-áβo.*

Since many verbal stems could also function as nouns, *-βo* may have been applied to

mean habitual or simultaneous action in relation to another verb.

agent	circ.	general
cat	cap	ca

After lenition these would have the following allomorphs:

/C_	/y_	/V_
-ár	-tár	-cár
-áβ	-táβ	-cáβ
-a	-ta	-ca
-áβo	-táβo	-cáβo

Other suffixes which contain the *-´r element are *-cʷár 'nominalizer of circumstantial complement' and -cʷér 'having the tendency to do x'.

The remaining nominalizers with -´r, -iwár 'person originating from' and -pír 'patient', could hardly have anything to do with an agentive suffix.[4]

4.2.2. Development of nominalizers in the daughter languages

With this comparative reconstruction as background, we will here consider the development certain nominalizers whose function has changed significantly over time.

4.2.2.1. Relative clauses

Keenan and Comrie (1977) undertake to frame some generalizations about relativization based on a survey of fifty languages. Their conclusions were that a) every language can relativize on subjects, and b) any relative clause strategy must cover a continuous segment of the accessability hierarchy: subject > direct object > non-direct

---

[4] But consider Greek -εύς which is both agentive and ethnonymic, and German ethnonymic suffix -er related to the agentive. As for -pír, note that -er in English does refer to a patient in a closed set of examples (eg. fryer). Such examples, however, do not lend strong support to a hypothetical relationship between -´r and -iwár, and between *-´r and *-pír: if there were a relationship between these morphemes, it presumably would not be because of a functional shift from the agentive function of *-´r to ethnonymic and patient, as in Greek, German or English; but because of a combination of the agentive marker *-´r and some other morpheme. To date no independent evidence has been found for PTG *-iwa 'origin' or *-pi 'patient'.

object > possessor. That is, there are languages which relativize only subjects, those which relativize only subjects and direct objects, and so on. However, there are exceptions to this pattern. Malay, for example, relativizes subjects and possessors, but not other kinds of relations (Yeoh 1979). Different languages in TG all use a reflex of the PTG nominalizer *-βaʔé to relativize noun phrases, but in each language the relative clause strategy covers a different segment of the accessibility hierarchy.

4.2.2.1.1. Relativization in OG

OG was a language which relativized only subjects of verbs and possessors. Ruiz de Montoya stated that -βaʔé could only be used with the third person subject prefix, the logical result if the antecedent must be the subject of the verb in the relative clause, ex. *če o-mboʔé-βaʔé* 'I am he who teaches'. A sentence with a relativized non-subject could have a first or second person subject, which would require first or second person prefixes and this construction did not occur in OG. A sentence like the following, with relativized direct object, was ungrammatical.

(37)   §yagua a-hečá-**βaʔé**
                dog 1s-see-**rel**
                'the dog which I see'

The functional equivalent of such clauses had to be expressed by some other construction.

(38)   'the sin which we contracted from our first father'

ñandé ruβ-ïpï angaipá ñande-ri o-yá-βaʔé   (RM C. p. 121)
1pn father-origin sin 1pn-to   3-join-**rel**
'our first father's sin which is joined to us'

A sentence with a relativized possessor in English or Spanish such as in (39)

would also be translated into Guarani by a paraphrase. For example, the following

sentence:

(39)   El Cacique a cuyos vassallos yo enseñé la música, me lo dió
       (The chief, whose vassals I taught music, gave it to me)

is paraphrased in Old Guarani as seen in (40) below.

(40)   o-meʔē če-βe aβa-ruβiča, y-βoya-reta   če **remi**-mboʔé-**cué** mburahéy rehe
       3-give me-to man-chief   3-vassal-many 1s **ptnt**-teach-**past** song for (PR 1724:188)
       'the chief gave it to me, his vassals were my former students for music'

In this example, no relative marker was used. A different construction is seen in (41)

below:

(41)   Tupa-cï y-yahoyá-βaʔé kʷarahï (PR 1724:188)
       God-mother 3-robe-**rel** sun
       'the mother of God, that which is her robe is the sun'
       la madre de Dios, cuyo manto es el sol
       (the mother of God, whose robe is the sun)

Here what was expressed by a relatived possessor in Spanish is expressed by placing the

relativizing suffix at the end of a genitive construction. As for non-direct objects, these

were expressed either by paraphrases or by the abstract nominalizing suffix *-háβa* (see

below). This could express, among other things, the place or time of an action.

Restivo, however, in the eighteenth century recorded that many Indians used first and second person prefixes with relativized verbs, primarily with intransitive verbs.

(42)   če-rako ka^ʔá-tí pe **a-ha-βaʔerã**
       1s-emph maté-group loc **1s-go-rel=fut**
       'I am the one who will go to the place where maté is grown.'
       yo soy él que ha de yr al hierbal

Ordinarily, the relative clause in (42) would be expressed as *o-ho-βaʔerã*. This could also occasionally happen with transitive verbs.

(43)   **y-yayé-βaʔerã a-i-kʷaá-βaʔe**   niche (PR 1724:155)
       3-finish-rel=fut 1s-3-know-**rel** emph
       'I am the one who knows the things which will be finished'
       yo soy él que sabe las cosas que se han de cumplir

In modern Paraguayan Guarani, *-βaʔe* has been reduced to *-βa*.

(44)   kuimbaʔe **a-heša-βa** ou še róγa pe.
       man **1s-see-rel** 3=come 1s house loc.
       'The man whom I see is coming to my house.'

4.2.2.1.1.1. Expansion to include direct and indirect objects

On very rare occasion, objects were relativized in OG.

(45)   **a-mboʔé-βaʔe**   (PR 1724:155)
       1s-reach-rel
       'the one whom I teach'
       á quien yo enseño

The more common way to say this, according to Restivo, would not be to relativize it,

but rather to use the nominalizing prefix *tembi-*, i.e. *če remi-mbo'é*, 'my student'.

The use of *-βa* to relativize direct objects is now firmly established in MG.

(46) Péva kuña re-heša-**βa**.
that woman 2s-see-**rel**
'That is the woman whom you saw.'

One must therefore depend on context to disambiguate phrases like *kuña o-mbo'é-βa*,

'the woman who is teaching', or, 'the woman who is taught.'

OG *-βa'e*, like nouns, could take the nominal tense suffixes *-kʷe(ra)*, *-ráma*, and

-rangue(ra). In the case of *-βa'e*, these yielded *-βa'ekue(ra)*, *-βa'eráma*, and

-βa'erangue(ra).

(47) o-mbo'e-**βa'e-kʷéra**   (RM, p.17)
3-teach-**rel-past**
'he who taught'
el que enseñó

The first of these, *-βa'ekue*, is now used in Paraguayan Guarani to indicate past tense

for relativized as well as nonrelativized clauses.

(48) ja-japo-va'ekue
1pn-make-past
'we made it (non-recent)'

Unlike *-βaʔekue, -βaʔerã* is not used as a relativizer in Paraguayan Guarani. A relative clause in the future tense is formed by simply using the future tense form of the verb and adding *-βa*.

(47) a-heša pe kʷimbaʔe o-ñeʔẽ-**ta-va**
     1s-see that man 3-speak-**fut-rel**
     'I see the man who will speak.'

4.2.2.1.2. Relativization in Wayãpi

The nominalizer **-βaʔé* has undergone a similar extension in Wayampi. The following examples are from Jensen (1990a:118).

e-nupã maʔẽ   'someone who beat me'
2s-beat rel

o-ata maʔẽ   'someone who walks'
3-walk rel

i-kasi maʔẽ   'someone who is strong'
3-strong rel

iru-sĩ maʔẽ   'someone whose clothes are white'
clothes-white rel

a-nupã maʔẽ   'someone whom I beat'
1s-beat rel

a-meʔẽ i-yupe maʔẽ   'someone to whom I gave it'
1s-give 3-to rel

a-a ikoti maʔẽ   'the place which I went to'
1s-go place rel

As can be seen from the above example, Wayampi has extended the functional domain

of *-βaʔé much further then has Guarani. Note also that Wayampi maʔé is no longer a clitic, but a separate word with its own stress. This is an innovation shared by Guajajara, Tapirape, Kayabi, and Kamayurá, languages within the innovation area.

4.2.2.2. The nominalizer -haβ

The suffix -βo was used in OG to indicate simultaneous time with another action, often implying either purpose or cause and therefore is equated by Ruiz de Montoya and Restivo with the gerundive or supine in Latin.

(103)a. ex.: a-há i-mboʔé-βo (RM 1640a:16)
1s-go 3-teach-siml
'I go to teach him.'

(103)b. če-kaneʔõ i-mboʔé-βo (RM 1640a:16)
1s-tired   3-teach-siml
'I am tired of teaching him.'

The form -βo was used for vowel-final verbs. Verbs which were consonant-final underwent devoicing of the final consonant and added -a.

4.2.2.2.1. Direct and indirect complements

While OG had no specific word or affix for making a verb the direct or indirect complement of another verb, this type of subordination could be expressed through the use of the suffix -haβ. The suffix -haβ was an abstract nominalizer which could indicate a wide variety of meanings, depending on the context. Examples of these various uses

follow.

time: karu-háβa (PR 1724:101)
      eat-nom
      'time of eating'

place: če-reko-ha-rupi o-kʷá   (PR 1724:102)
      1s-stand-nom-about 3-pass
      'he passed by where I was'

instrument: iβira-kɨti-há (PR 1724:102)
         stick-cut-nom
         'instrument with which the stick is cut'

         karu-háβa (RM 1639:135)
         eat-nom
         'eating utensils, table cloths, napkins, etc.'

abstract instrument: na Tupã grasia mo-cañɨ uka háβa ruŋʷáy rako aŋgaypá miri
                this God grace tr-lose cause nom neg this sin little
      'venial sin is not a thing, is not a sin which makes one lose the grace of
                                  God' (PR 1724:102)

content: ñandeyara remi-mborara-kʷé niko če-ne-mo-ñẽʔēē haŋʷā.   (PR 1724:102)
        Our Lord pass-suffer-past emph 1s-pass-caus-say nom=fut
      'the subject of my sermon will be the passion of our Lord'

manner: če-mboʔé epe   če i-yapó haŋʷā (PR 1724:102)
        1s-teach   2s→1 1s 3-do   nom=fut
        'teach me the manner in which I must do it'

        nd-a-ikʷaá-i   če i-porú haŋʷā (PR 1724:102)
        neg-1s-know-neg 1s 3-use nom=fut
        'I don't know how I should use it'

effect: ñande- rúβ -ɨpɨ   angaipá-haɣʷe niko tekó-así tetirō   (PR 1724:102)
        2p=exc-father-first sin-nom=past emph life-sick any
        'sicknesses are the effect of our first father's sin'

ability: ñandeyára yesus kristo ñõ-te o-mẽʔẽ paʔí abare upe angaipa mo-caəỹ-haɣʷã
       Our Lord Jesus Christ only-emph 3-give priest to sin   caus-lose-nom=fut
                                            (PR 1724:103)
       'our Lord Jesus Christ only gives priests the power to absolve sins'

 kaʔá-tɨ   -pe če-hó haɣʷã rehe a-yerure paʔi upe (PR 1724:103)
mate-group-loc 1s-go nom=fut for 1s-ask priest to
'I asked the priest for permission to go to the place where mate is grown'

cause/motive: misa-rendú haɣʷã ño i-tú (PR 1724:102)
            mass-hear nom=fut only 3-come
            'the purpose/cause/motive of his coming will only be to hear masses'

intent: akóy-ramo-βe če-ne- ŋʷahẽ haɣʷãma a-rekó (PR 1724:103)
       then-when-βe 1s-pass-arrive nom=fut 1s-have
       'from that time forward, I have the intention of escaping'

Like the other nominalizers, -haβ uses nominal tense markers: -haβ + -kuér(a) >
-haɣʷér(a)

                        hoʔá-haɣʷéra (RM 1640a:18)
                        3=be=born-nom=past
                        'his birth'

        -haβ + -ŋʷãm > -haŋʷãm

                  mondó-haŋʷãma (RM 1640b:305)
                  send -nom=fut

        -haβ + -raŋgʷér > -haβaŋgʷér

                  i-mboʔé-haβaŋgʷéra (RM 1640a:18)
                  3-teach-nom=fut.perf
                  'place, etc., where one was to be taught'

The future of -haβ was expressed by the suffix -ŋʷãm, which both Ruiz de Montoya and

Restivo record as equivalent to -rãm.

(48) ex.: če-rembiapó ŋʷāma (RM 1639:129)
          1s-work   fut
          'that which I have to do'

It has already been noted above that even in OG, -haŋʷācould express purpose as it does

in MG.

(49) ex.: Ey- nupã oy-kʷaá-haŋʷā   (JG 1891:94)
          sg.imp.-hit 3-know-pur
          'hit him so that he will know'

          o-pɨra o-pɨtuʔú haŋʷā (JG 1891:94)
          3-stay 3-rest pur
          'he's staying in order to rest'

Verbs nominalized by -haβ also served as the objects of verbs like roβʸá, 'to

believe', which may take an entire clause as its object. Compare the OG of (50) with the

corresponding Spanish sentence in (51).

(50) ex.: i-maraŋgatú iβa -pe herahá haŋʷāma ere-roβʸá raʔé (RM 1640b:276)
                3-good   -rel heaven-to take nom=fut   2s-believe int=emph
                'do you believe in the taking to heaven of the holy?'

(51) ex.: crees,   que a los buenos ha de dar la gloria eterna
          believe-2s that to the good 3=fut give the glory eternal

Here, the nominalizer -háβ corresponds to the Spanish conjunction que (-ŋʷām

corresponds to ha de).

In Paraguayan Guarani -haŋʷā is no longer the future tense of -haβ, but is used

exclusively for purpose expressions. So, for example, the future tense of the following

indirect statement (52a), is not (52b), but (52c).

(52) a. ex.: ere   re-ju -ha
          2s=say 2s-come-comp
          'you say that you are coming'

    b.   ex.: ?ere   re-ju-haŋʷa
          2s=say 2s-come-pur
          'you say it in order for you to come'

    c.   ex.: ere   re-ju-ta-ha
          2s=say 2s-come-fut-comp
          'you say that you will come'[5]

By adding -iša, 'similar to' to -haβ, MG has -háiša to express similarity between two actions.

(53) a. ex.: a-japo nde re-japo-háiša   (JG p.94)
          1s-do   2s 2s-do -like
          'I do as you do'

    b.   e-japo re-japo-se-háiša (KC p.126)
          sg.imp.-do 2s-do-want-like
          'do as you like'

Tense applies to -háiša as well: past -hawéiša, future -haŋʷáiša.

---

[5] This is true in the Asunción dialect. Pederson (1977) recorded that in Eastern Paraguay, this shift is still in transition. The sentence 'He says that he will come' can be expressed in any of the following ways:

            he?i oú-haŋʷã
            3=say 3=come-comp=fut

            he?i oú-ta-ha
            3=say 3=come-fut-comp

            he?i oú-ta-haɲʷã
            3=say 3=come-fut-comp=fut

4.2.2.2.2. The nominalizer -haβ as subordinating conjunction

(104) ex.: a-hečáɣ 'I see'; hečáka 'in order to see him' (RM 1640a:26)

a-hendúβ 'I hear'; hendúpa 'in order to hear' (RM 1640a:26)

The ending -a was also used for nasal vowel-final verbs with apparently underlying final nasal stops.

(105) ex.: a-meʔē 'I give'; i-meʔēŋga 'to give to him' (RM 1640a:27)

As can be seen in the examples of -hár (2.1.1 above), h-initial suffixes placed immediately after a consonant resulted in the loss of the h and the devoicing of the stem-final consonant (this also occured with -háβ, ex. sóɣ, 'to pound'; íɣa sokáβa 'instrument for removing nails', from -sóɣ + háβ) (RM 1639:116).

Even in Ruiz de Montoya's time many verbs were losing these final consonants as is evidenced by such pairs as:

(106) ex.: a-weyí 'I go down'; weyípa or weyí-βo 'going down' (RM 1640a:26)

This was also the case with nasal vowels.

(107) ex.: a-weʔē 'I vomit'; weʔēna or wi-weʔē-mo 'vomiting'
a-ñe-mo-saʔenā; wi-ñe-mo-saʔenāna or wi-ñe-mo-saʔenānga
or wi-ñe-mo-saʔenā-mo 'being perceived' (RM 1640a:27)

Here -mo is -βo assimilated to the nasal vowel. In the last example, the nasal

purpose suffixes vary completely at random because the final nasal consonant is no longer predictable from the base form.

Throughout the OG period, the occurrence of *-haβ* was unchanged, but as final consonants disappeared the distribution became lexical rather than phonological. In Yapaguay's *Sermones,* there is no appearance of final consonants, even when followed by vowels. Nevertheless, Yapaguay consistantly used forms such as *o-h-ečáka.* This temporarily created unpredictable classes of verbs, which occurred with *-ka, -pa -ta -na, -nga,* and *-βo.* In Paraguayan Guarani, Kw, and Ch, *-βo* is used with all verbs and has come to be interpreted as a subordinating conjunction, rather than as a nominalizer (see **5.4.2** below).

4.3. Nominal Tense

One of the typologically unusual characteristics of TG languages is that they have a separate tense system specifically for nouns; and that furthermore this tense system is more complex that that of the verb system. Here we will compare the nominal tense of TG languages.

All TG languages have at least three tenses for nouns: past, present, and future. Present tense is unmarked. In addition to these three, the most conservative of TG languages combine the future and past tense markers to form a fourth tense, 'that which was to be and is not'.

Tb

rám	'future'	
kʷér	'past'	
raŋgʷér	'what was to be and is not'	

OG   rã (rám), wã(ma)

       če-r-embi-apo-wãmã   (RM 1639:127v)
       1s-cm-pat-make-fut
       'that which I will make'

       če-r-oɣʷãma   (RM 1639:127v)
       1s-cm-house=fut
       'may house that will be'

   kʷé (kʷér)
   raŋgʷé (raŋgwér)

Tp   rɨm   Tãipa kʷe a-ãpa če-hãɨpɨrɨro-rɨma   (Aldeia 1983:28)
            Tãipa   3-make 1s-shirt-fut
            'Tãipa is going to make that which will be my shirt'

            konomĩ pãčē-rɨma   (Aldeia 1983:28)
            boy   shaman-fut
            'this boy is going to be a shaman'

   kʷer   ʔipira   ʔipira-kʷera   (Aldeia 1983:28)
          'fish'   'that which was a fish'

          kɨče   kɨče-kʷera   (Aldeia 1983:28)
          'knife'   'that which was a knife'

UK   wam   kapitã-wam   (K 1986:374)
            captain-fut
            'someone who will be a captain'

   kʷer   kã-ŋgʷer   (K 1986:373)
            bone-past
            'extracted bone'

AT   kʷer   hatɨ-kʷera
            wife-past
            'former wife'

        o-kɨkano-waʔe-kʷera
        3-maimed-rel-past
        'someone who was maimed'

rom   če-kɨhe-roma
        1s-knife-fut
        'the knife which will be mine'

nom.verbs
        oro atat-awera-hi   oron
        1px-hunt-nom=past-from 1px=come

        i-čekɨy-tawera-hi   i-čehow-i
        3-die-nom=past-from 3-get=up-овтор

        ne-ro-awoma
        2s-arrive-nom=fut
        'your arrival (in the future)'

AX   -om   a-pɨwɨn topaw-om-amo   (Nicholson 1982:17)
        1s-spin hammock-fut-nom
        'I am spinning in order to make a hammock'

Gy   rã   ko meca-rã   (Hoeller 1931:40)
        this table-fut
        'this will be a table'

        a-maʔe-ra
        1s-look-fut
        'I will look'

        aere mini-rã
        later something-fut
        'it will be something later'

        nde katu-pɨrɨ-rã
        2s good--fut
        'you will or should be good'

kʷer   ko če mbaʔe kʷer
        that 1s thing past
        'that was one of my things'

     raŋgʷer ko če reko-raŋgʷer
          that 1s life-fut=past
          'that would be my life'

Gj   rəm
        kʷer

Note that the future suffix is used for not only nouns, but verbs, adjectives, and even postpositional phrases.

Ch   rã, ta, kʷe

Here, inverse to the case of Gy, Ch not only uses nominal tense suffix *rã*, but also *ta*, a suffix used only for verbs in other TG languages.

GM   rã, kʷe
Kw   rã, kʷe, raŋgʷe

The morphology for noun tense is virtually identical in all the TG languages. The comparative data do prompt two questions, however: was the combination of future and past markers original; and, why is the UK future marker *wam* instead of the expected [§]*ram*?

    The portmanteau form *raŋgʷer* appears in both the oldest (Tb, OG) and most conservative (Pt, Gy) TG languages, indicating that this was a feature of PTG. The UK nominal future marker *wam* actually has cognates in Tb and OG *wám* or *γʷám* and *wā* or *γʷā* respectively. The Tb data indicates that *-wám* is really *-o-ám*. This suggests something about the 'other' future suffix **-rám;* namely, that it was not

actually a different suffix at all, but the same suffix *-am, which was treated as a class II root: taking the class marker *-r- in some cases, taking the marker *-o- in others. The conditioning factors for using one marker or the other has been long obscured even by the early historical period, but will be discussed in **5.5.1.1.**

4.4. Conclusion

In this chapter we have examined the noun system in TG and the reconstructed system of PTG. We conclude that the present PTG class system for nouns first evolved from phonological distinctions in roots, which later was interpreted as having semantic significance. Nouns in Pre-TG may not have been organizable into any sort of class system; or they may have had a different system, which was replaced by the current systems. Comparisons with other Tupian families are not yet conclusive, but such as have been made suggest a very early date for the distinction between class I and class II roots (consonant initial and vowel initial) by the use of *t-* with vowel initial roots.

We have also examined the inventory of nominalizing suffixes in various TG languages and from them made a reconstructed list of such affixes. Internal 10-reconstruction applied to this comparative reconstruction yields further insights into the morphology of Pre-TG which will be extremely useful in the reconstruction of the PTG and Pre-TG verb systems.

Finally, we have seen how the function and scope of these affixes varies from language to language and what this tells us about the development of TG morphology from the time of the parent language.

PART TWO: MORPHOLOGY
CHAPTER FIVE: PRONOUNS

5.0. Introduction

Our look at the noun system in the previous chapter gave some hint as to the importance of the pronominal cross-referencing system in understanding root classes in TG. A more complete study of pronominal cross-referencing is even more revealing about the development of pronouns and the development of syntax and case in TG.

To begin with, we will examine the cross-referencing systems in the different TG languages to arrive at a reconstructed proto-system. Following are the systems in the different languages.

5.1. Comparison of cross-referencing systems

In comparing the cross-referencing systems of TG languages, we will use the pronominal marker (pm) categories used in Pease's (1968) descriptive analysis of Parintintin. Pease's pm1 and pm2 are for intransitive and transitive verbs, respectively; hence they can be conflated to a single column, which comprises nominative cross-reference markers with the exception of those prefixes which cross-reference both subject and object. These we will put in a separate category. Also note that in Pt the 1st plural inclusive marker is different for transitive and intransitive verbs. We will still conflate the two columns, noting the distinction in Pt and other languages having the same distinction for the 1pn. Accusative pronominals are found in Pease's pm4 (third person accusative *i-* is left out because Pease analyzes it as a class marker and therefore part

of the verb stem). Pease's pm3 corresponds to Jensen's (1990b) Set 3 while pm4 corresponds to Jensen's Set 2. This differs from pm5 only in the 3rd person because Pt, like Kb, has separate pronouns for possessives and for the subjects of stative verbs (or the possessors of descriptive nouns). The pronouns in the column marked pm6 are emphatic topicalized forms. They also occur minus the -hV element and this shorter, more basic form Pease identifies with pm5, which she therefore treats as free pronouns; however, as possessives, pm5 markers are bound morphemes, except those of the third person.

Parintintin

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro-[1] opo-	i-	ji-	ji	jihi
2nd sg imp	ere- e-	ere- e-	e-	nde-	nde	ndehe
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-		hẽ ga ŋã ahe	hehẽ ŋãhã ŋãhã ahe
1st pl excl	oro-	oro- oro- opo-	oro-	ore-	ore	ore
1st pl incl	ča-	ti-	ñande-	ñande-	ñande	ñande
2nd pl command	pe- pe-	pe- pe-	pe- or peji-	pe-	pe	pehe

---

[1] Here, and in the other tables as well, the three forms of pm2 in the first person are, respectively, first person subject/third person object; first person subject, 2s object; first person subject, 2p object.

Old Guarani

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro- opo-	wi-	če-	če	če
2nd sg imp	ere- e-	ere- e-	e-	nde-	nde	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-	i-/h-	i-/h-	ha'e ase
1st pl excl	oro-	oro- oro- opo-	oro-	ore-	ore	ore
1st pl incl	ya-	ya-	ya-	ñande-	ñande	ñande
2nd pl command	pe- pe-	pe- pe-	pe-	pe-˘	pe-˘	peẽ

Kayabí

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro-	te-	ye-	ye-	ye
2nd sg imp	ere- e-	ere- e-	e-	ene-	ene-	ene
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-		ēē ʔŋa ʔŋã ae	ēē ŋã ŋã ae
1st pl excl	oro-	oro-	oro-	ore-	ore-	ore
1st pl incl	sa-	si-	yare-	yane-	yane-	yande
2nd pl command	pe- pe-	pe- pe-	peye-	pẽ-	pẽ-	pẽ

Assurini

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro- opo-	we-	če-	če-	iče
2nd sg imp	ere- e-	ere- e-	e-	ne-	ne-	ene
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-	i-/h-	i-/h-	aʔe ahe
1st pl excl	oro-	oro- oro- opo-	oro-	ore-	ore-	ore
1st pl incl	ča-	ča-	čere-	čene-	čene-	čane
2nd pl command	pe- pe-	pe- pe-	peče-	pe-	pe-	pehe

Wayampi

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro- poro-		e-	e-	ye
2nd sg imp	ere- e-	ere- e-		ne-	ne-	ne
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-		i-/∅-	i-/∅-	aʔe ahe
1st pl excl	oro-	oro- oro- poro-		ore-	ore-	ore
1st pl incl	ya-	ti-		yane-	yane-	yane
2nd pl command	pe- pe-	pe- pe-		pe-ˉ	pe-ˉ	pehe

Kamayura

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- aro- apo-	we-	ye-	ye-	iye
2nd sg imp	ere- e-	ere- e-	e-	ne-	ne-	ene
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-	i-/h- i-/h- i-/h- Ø-/t-	i-/h- i-/h- i-/h- Ø-/t-	ae awa
1st pl excl	oro-	oro- oro- opo-	oro-	ore-	ore-	ore
1st pl incl	ya-	ya-	yere-	yene-	yene-	yane
2nd pl command	pe- pe-	pe- pe-	peye-	pe-	pe-	pee

e

Tupinamba

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro- opo-	wi-	(i)še-	(i)še-	iše
2nd sg imp	ere- e-	ere- e-	e-	(e)nde-	(e)nde-	ende
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-	i-/s- i-/s- i-/s- m-/t-	i-/s- i-/s- i-/s- m-/t-	aʔe ahe
1st pl excl	oro-	oro- oro- opo-	oro-	ore-	ore-	ore
1st pl incl	(t)ya- /(t)i-	(t)ya-/ (t)i-	yere-	yande-	yande-	yande
2nd pl command	pe- pe-	pe- pe-	peye-	pe~-	pe~-	peyẽ

Urubu-Kaapor

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a-		hẽ-	hẽ-	ihẽ
2nd sg imp	ere- e-	ere- e-		nde-	nde-	nde
3rd sg fem sg masc plural general	u- u- u- u-	u- u- u- u-		i-/h- i-/h- i-/h- Ø-/t-	ŋa ŋa ŋa ahe	ŋa ŋã ahe
1st pl excl	ya-	ya-		ore-	ore-	ore
1st pl incl	ya-	ya-		yande-	yande-	yande
2nd pl command	pe- pe-	pe- pe-		pehẽ-	pehẽ-	pehẽ

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Assurini of the Xingú

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro-		ǰe-	ǰe-	
2nd sg imp	ere- e-	ere- e-		ende-	nde	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-		i-/h- i-/h- i-/h- Ø-/h-	ga ga ŋā ahe	ahe
1st pl excl	oro-	oro- oro-		ore-	ore-	ore
1st pl incl	ča-	ča-		žande-	žande-	žande
2nd pl command	pe- pe-	pe- pe-		pende-	pende-	pehe

Guajajara

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- apu-		he-	he-	ihē
2nd sg imp	ere- e-	ere- e-		ne-	ne-	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-		i-/h- i-/h- wa-/wan- Ø-/t-	i-/h- i-/h- wa-/wan- Ø-/t-	ahe
1st pl excl	uru-	uru- apu- (urupu-)		ore-	ore-	ore
1st pl incl	ža-	si-		žane-	žane-	žane
2nd pl command	pe- pe-	pe- pe-		pe-	pe-	pee

Tapirape

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	ã-	ã- ãpa-	we-	če-	če-	ie
2nd sg imp	ere- e-	ere- e-	e-	ne-	ne-	ane
3rd sg fem sg masc plural general	a- a- a- a-	a- a- a- a-	a- a- a- a-	i-/Ø- i-/Ø- i-/Ø-	i-/Ø- i-/Ø- i-/Ø-	aheŋã aheŋã aheŋɨ ahe
1st pl excl	ara-	ara- ara-	ara-	are-	are-	are
1st pl incl	ča-	či-	čere-	čane-	čane-	čane
2nd pl command	pe- pe-	pe- pe-	peče-	pe-	pe-	peẽ

Paraguayan Guarani

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- ro- po-		še-	še-	še
2nd sg imp	re- e-	re- e-		nde-	nde-	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-		i-/h-		haʔe
1st pl excl	ro-	ro- ro- po-		ore-	ore-	ore
1st pl incl	ǰa-	ǰa-		ñande-	ñande-	ñande
2nd pl command	pe- pe-	pe- pe-		pende-	pende-	peẽ

Chiriguano

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- ro- po-		če-	če-	če
2nd sg imp	re- e-	re- e-		nde-	nde-	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-		i-/Ø-	i-/Ø-	haʔe ahe
1st pl excl	ro-	ro- ro- po-		ore-	ore-	ore
1st pl incl	ža-	ža-		žande-	žande-	žande
2nd pl command	pe- pe-	pe- pe-		pende-	pende-	pẽẽ

Mbya

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a-		če-	če-	če
2nd sg imp	ere- e-	ere- e-		nde-	nde-	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-		i-/Ø- i-/Ø- i-/Ø- i-/Ø-	i-/Ø- i-/Ø- i-/Ø- i-/Ø-	haʔe ae
1st pl excl	oro-	oro-		ore-	ore	ore
1st pl incl	ǰa-	ǰa-		ǰande-	ǰande	ǰande
2nd pl command	pe- pe-	pe- pe-		pe-	pe	peẽ

Kaiwa

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro- apo-		še-	še-	še
2nd sg imp	ere- e-	ere- e-		nde-	nde-	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-		i-/h- i-/h- i-/h-	i-/h- i-/h- i-/h-	ha?e
1st pl excl	oro-	oro- oro- oropo-		ore-	ore-	ore
1st pl incl	ja-	ja-		ñande-	ñande-	ñande
2nd pl command	pe- pe-	pe- pe-		pende-	pende-	pee

Guarayo

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro- opo-	wi-	če-	če-	če-
2nd sg imp	ere- e-	ere- e-	e-	nde-	nde-	nde
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-	i-/c- i-/c- i-/c-	i-/c- i-/c- i-/c-	ahe
1st pl excl	oro-	oro- oro- opo-	oro-	ore-	ore-	ore
1st pl incl	ya-	ya-	yere-	yande-	yande-	yande
2nd pl command	pe- pe-	pe- pe-	peye-	pe˜-	pe˜-	peẽ

5.2. Reconstruction of PTG cross-reference markers

The pronouns as they are presented here are in four basic cases: emphatic/topic (pm6), nominative (pm1 and pm2), accusative/non-refl possessive (pm4 and pm5), and reflexive possessive (pm3). Of these four, it is clear that emphatic/topic and accusative/non-reflexive possessive were originally the same case, specifically that the function of topic was originally handled by the accusative case, although formally, the emphatic forms are the more conservative. With this understood, three of the pronouns can be readily reconstructed for the accusative case.

5.2.1. Third person

The third person cross-referencers are identical in all languages for pm1, pm2, pm3 and pm4 with two exceptions: 1) Tp has *a-* where other languages have *o-*, but this is what one would expect in light of the sound change rule given in 1.3. changing *$^*o$ to *a* in all environments; 2) the pm4 marker for class two nouns has considerable variation, but all these point back to *$^*c$. There is less uniformity among languages on pm5 markers. In OG, AT, W, Km, Tb, Tp, Gj, Ch, GM, Kw, and Gy pm5 markers are identical to pm4 markers. In Pt, Kb, UK, and AX, however, pm5 markers are free pronouns. In UK and AX these are very simple; in Pt the pronominal system is more complex, marking gender as well as number. In Kb, the pm5 markers are more complex still. Because of their apparent absence from other languages, the origin of these pronouns is not immediately obvious, but the diversity of TG languages in which they occur suggests that they were present in PTG. The greatest likelihood is that these pronouns have their origin as noun classifiers. In Pt, the pronouns *hē*, *ʔγa*, *ʔŋā*, and the relative marker *-βeʔe* are also noun classifiers (see 6.2. below). In Kb, the pronominal/classifier system is so complex, involving the age, gender, number and other factors of both the speaker and hearer, that it has not as yet been completely analyzed. Apparently, these classifiers belong to the vague class of bound and free morphemes that are referred to in the literature as 'particles'. Such morphemes, generally serving various pragmatic functions, are common in Amazonian languages (see Dooley 1990). Some languages apparently began using these particles as pronouns in the absence of cross-refence markers for the pm5 functions. We can conclude then, that while these

'particles' existed in PTG, they were not used as pronouns, certainly not as pm5 markers, since they to not serve this function in most TG languages, including the earliest attested languages.

5.2.2. First person inclusive

5.2.2.1. The 1pn cross-referencers *ya-* and *ti-*

Jensen mentions a distinction in cross-referencing which occurs in Wayampi and several other TG languages, but not in Tupinamba or Guaranian languages: 1pn markers *ya-* and *ti-*. In Wayampi, Guajajara, Tapirape, Kayabi, and Parintintin, *ya-* marks only subjects of intransitive verbs (except verbs of the so-called 'stative' class) while *ti-* marks subjects of transitive verbs. Jensen (1990a:96) interprets this as a conservative feature which 'é uma distinção em línguas ergativas.'

But it is odd, to say the least, that if a truly ergative distinction is conservative, it is only marked in the 1pn. Further, *ya-* in these languages cross-references the absolutive case — but in Jensen (1990b) Set 2 markers are equated (*yande* for 1pn) with the original markers of absolutive case. There is also the problem concerning *ace* above, that cross-referencers are normally derivable from a common morpheme:

Set I	Set II	Set III	Set IV	Other
a	iye	wi		
ere	ende	e	oro	
ya	yande	yere		ti
oro	ore	oro		
pẽ	pe	peye	opo	

The cross-reference marker *ti-* is conspicuously out of place here with respect to its shape.

Finally, the languages which distinguish *ti-* and *ya-* in this way lie almost entirely within the innovation area, not in relic areas.

Tupinamba, which does not use *ti-* in this way, nevertheless attests its use for the 1pn. Its origin is bound up with the morpheme *ta-*, which expressed in PTG (and still does in many TG languages) a hortatory mood.

1.   Pedindo licença.
        t-açô   va eu
        T-oro-çô   vamos-nos
        t-o-çô va elle ou vão elles

        Concedendo permittindo, mandando ut.
        T-ere-çô   vas tu ou vay tu ou iras tu
        Ta-pe-çô   vos.

        Exortando, inuitando, imperando, ut.
        Tia-çô vamo nos
        (A 1595:23v)

Anchieta further made this observation about *t-ia*:

2.

        Na primeira plural que tem, *tiâ*, se soe tirar o *a*, e ainda o *t*
        eleganter nos verbos activos, porque tambem com *iâ*, se usa o
        Indicativo, pro Imperativo, *yarû* pro *tiarû*, *irû*, pro *tirû*.
        ut, *Tiarû*, tragamus, protemus, *tirû*. 1. *irû*.
          (A 1595:23)

Here then we see that *tya-* alternated with *ti-* in 'active' (i.e. transitive) verbs.

Furthermore *ya-* alternated with *tya-* (and *ti-*) freely to express both indicative and hortatory moods.

The reason for the alternation *ya-/i-* and *tya-/ti-* may be explained by comparative evidence outside of TG.

The 1pn marker in Tuparí is *ki*. Other cognates show a correspondence between Tuparí *k* and PTG *ʔ*.

3.

Tup.	PTG	
ku	ʔú	'eat'
ke	ʔé	'say'
kɨp	ʔɨβ	'tree'

Based on this correspondence the cognate for *ki* in PTG could have been *ʔi*. PTG may have remodeled the 1pn marker form **ʔi* to **ʔi-a*, compounding it with the 1st sg marker, later becoming **ya*. Other Tupian languages outside TG show similar remodeling of the 1pn marker, usually with the 1s marker before *i-*.

4.

	1sg	1pn
Munduruku	o-	wɨy- (< *(w)oy-)
Sataré-Mawé	a-	ai-
Karitiána	i-	iy-
Surui-Mondé	o-	oy-

The forms in *i-* and *ti-* may then preserve the original form of the 1pn marker.

In conclusion, we see that *ya-* alternating with *ti-* in transitive verbs led to

permanently assigning exclusive *ti-* to the function of 1pn marker for transitive verbs, leaving *ya-* as the marker for intransitive verbs. That this was an innovation in TG is further confirmed by the fact that almost all the languages exhibiting this distinction lie within the innovation area. The fact that one language exhibiting this innovation (Parintintin) lies in a relic area, together with the fact that the originally free variation spawning this innovation is attested in Tupinamba, suggests that the free variation in question dates back to PTG.

The loss of *ʔ* in this environment is attested elsewhere in TG ex. *-ʔu* eat + *-áβo* > *-wáβo*.

5.2.2.2. First person inclusive vs. all-inclusive

The cross-referencing picture of TG is made even more complicated by data in Tupinamba analyzed by Rodrigues (1990) in examining grammars of Anchieta (1595) and Figueira (1687), showing the existence of a pronoun *ase* in addition to the pronouns already discussed.

There are two grammatical features connected with this pronoun. One is that it marks a distinction between 1pn including only 1 and 2 persons, and 1pn including 1st, 2nd, and 3rd persons.

yande = you and I
ase = you and I and him/her/it/them

The other significant feature is that it distinguishes between focused and non-focused 3rd person transitive subjects. The pronouns *yande* and *ase* correspond to the Set 1 agreement markers *ya-* and *o-* respectively. These, together with other Set 1 markers, interact in a fashion represented by Rodrigues in the following diagram where 3^f refers to pronouns with third person in focus and 3^{-f} to pronouns with the third person out of focus (Rodrigues 1990:402):

		Contrast 1/2		
		+		−
		1	2	
3^f	−	iše	ene	yande
	+	ore	peẽ	ase

So *ya-* can refer to either 12 subject or to a 3^{-f} subject.

5. a. ya-i-kutuk
        12-3-pierce
        'we pierced it'

  b. ya-rasó temō sapiʔá iβák-ipe tupána še-r-úβa mã.
      3^{-f}-carry opt soon sky-to God 1s-cm-father oh
      'Oh that God would soon take my father to heaven!'

Likewise *o-* can refer to a 123 subject or to a 3^{+f} subject.

6. a. o-yeroki pe asé Jesus ʔe-reme
        123-bow int 123 Jesus say-when
        'Do we bow when we say "Jesus"'?

6. b.   ase yuká 'somebody kills one, i.e. me, us, etc'
             'A homem matão, i. a mĩ, a nos, &c.'

c.   Tupã o-manõ, memẽ-tipó ase o-manõ-mo
             God 3-die so=much-more 123 123-die-ger
             'God died, so much more we shall die'

Further there seems to be an element of ergativity in Tupinamba in addition to those stated by Jensen (1990b): namely, that focus is grammaticalized in the 3rd person only in the ergative case for subjects of transitive verbs. For intransitive subjects ('active' or 'stative') and patients, focus is not grammaticalized. This also suggests the historically original value for these markers, since *ya-* when marking intransitive verbs is only 1pn, and *o-* is only 3rd person. Moreover, *yande* is only 12, never 3^{-f} while *ase* is only 123 never 3^{+f}, since these are used for patients and oblique cases. However, this is not really evidence that Tb was an ergative language; it is, rather, a logical result of how focused and non-focused nouns were contrasted. Grammatical marking of focus is not needed with patients or intransitive subjects because focus on 3rd person is irrelevent without the presence of another 3rd argument in the same verb phrase. If the verb is intransitive there is only one argument; if the verb is transitive, the subject is already marked for focus and the object by default has the opposite focus value as the subject.

7. a. moruβišáβa moná ya-y-namí-ʔók-ukár
   judge   thief 3^{-f}-3-ear-remove-caus
   'The judge (-f) ordered the ear of the thief (+f) to be taken off.'

b. moruβišáβa moná o-y-namí-ʔók-ukár
   judge   thief 3^{+f}-3-ear-remove-caus
   'The judge (+f) ordered the ear of the thief (−f) to be taken off.'
   (A 1595:36v)

In the first example the fate of the thief is focused on, while in the second example the severity of the judge is the focus.

The fact that intransitive verbs *ya-* means only 1pn (12 or 123) while *o-* means only 3rd person (±focus) indicates that *ya-* and *o-* were originally 1pn and 3 respectively, the distinctions between 12 and 123, and between 3^{+f} and 3^{-f} being later innovations. This might help to explain the anomaly of *ase*. Although this pronoun makes the pronominal paradigm very symmetric in Tupinamba, there is a morphological anomaly: consider the Set 1 and independent pronouns:

iye	a
ende	ere
yande	ya
ore	oro
peē	pe

Each of these pairs as we have seen in **5.2** can be traced to a single common morpheme. The exception to this is *ase*, which, in PTG would have been *ače*:

ače   o

There is no way to trace these two morphemes to a common proto-form. Further, the Set 2 markers in Tb are all identical to the pronouns, allowing for the drop of initial front vowels, again with the exception of *ase*:

išé	še-
endé	nde-
yandé	yande-
oré	ore-
peẽ	pe($\sim$)-
asé	**i-**

Rodrigues (1990) remarks that the 3rd person had no pronoun, but in the place of a pronoun demonstratives were used (*a?é*, *aipó*, etc.). The situation of *ačé* in relation to *o-* and *i-* looks similar to this. And if *o-* and *i-* originally were only 3rd person markers and *ačé* was the corresponding pronoun, this leads to the possible conclusion that **ačé*, if reconstructable for PTG as a pronoun, originally was either a demonstrative, or a noun that acquired a pronominal function.[2] In Parintintin, *ahe* has a meaning '(dead) person'. Based on our internal reconstruction of palatalized stops in **2.2** above, *ačé* can be traced backwards to ***ap^{y}e*, a noun related to pre-TG ***apa* and PTG **aβá*, 'person'. Note, however, that unlike demonstratives and like pronouns, *ase* took *-βe* rather than *-pe* as a marker for indirect object in Tb.

This supports the idea that the distinctions 12 vs. 123 and 3^{+f} vs. 3^{-f} were

---

[2] Such a demonstrative may in fact be related to c(e)- of class II stems. If this is the case then **c(e)-* derives from **ace* by dropping the initial vowel as with *iše* > *še* and *ende* > *nde*.

innovations in Tupinamba, although they may have existed in PTG as well. The fact that *yande* and *ače* did not each originally refer to both 1st and 3rd person, is reflected in patient and intransitive subject cross-referencing. When pragmatics encouraged the development of a distinction between 3^{+f} and 3^{-f}, this was only needed where two 3rd arguments occurred in a clause, or whenever 1pn occurred (ergative or otherwise), since 1pn originally included both 3^{+f} and 3^{-f}. This inevitably resulted in a grammatical distinction marked only for subjects of transitive verbs (but note that the accompanying distinction between 12 and 123 was **not** restricted to subjects of transitive verbs). Here then we have a grammatical distinction which was not truly ergative/absolutive in nature, but whose grammaticalization inevitably took on the appearance (in part) of an ergative/absolutive distinction. The same can be said for the verb *paβ* in conjuction with other verbs and possibly even for monosyllabic reduplication. Other evidence presented by Jensen for absolutive structures is as easily or more easily explained as examples of nominalizations with oblique cross-referencing. The conclusion that we reach then is that PTG was originally a nominative/accusative language which early in its development acquired certain constructions reminiscent of an ergative/absolutive distinction; these were maintained in most TG while others (W, K, MG) lost these features shortly after the beginning of the historical period, replacing them with nominative/ accusative structures. The focus distinction was apparently the first quasi-ergative structure to disappear, since it remains in none of the modern daughter languages. At the same time, lexical, phonological and syntactic changes have contributed to the emergence of a quasi-active/stative distinction in cross-referencing

which may develop into a full-blown active/stative distinction replacing the older nom/acc distinction. This is true to a greater extent in languages outside the Guarani group, but has also begun in varying slight degrees in languages inside the Guarani group. In Guarani, the non-adjectival stative verbs are all still analyzable as cases of subject incorporation, but two are not: *esaráy* 'forget' and *mandu?á* 'remember'; 'adjectives' are no longer nouns, however, but a class of verbs (possessive and adjectival constructions differ: *túβa po*, the father's hand, but *túβa i-porā* 'the father is good'; *túβa r-āi* 'the father's tooth', but *túβa hasy* 'the father is sick'). By contrast, in Chiriguano, noun incorporation seems to be unproductive as a syntactic process, certainly according to Dietrich (1990:306); 'adjectives', however, are still analyzable as nouns. The same is true in Guarayo, although there noun incorporation still seems to be a very productive syntactic process (see **6.3**).

5.2.3. First person exclusive and second person

Because of the history of prenasalization outlined in **2.4**, *ende* is seen as more conservative than *ene* and similarly, *yande* is more conservative than *yane*. The reconstruction of 1px *ore* is virtually tautologous for pm4, pm5, and pm6.

It is generally agreed that Set 2 pronouns are derived from independent pronouns and that they are nearly identical to one another. But in fact there is evidence in Tupinamba that there was no morphological difference between the pronouns and Set 2 prefixes.

The difference between these two sets of morphemes is mainly in the 1ˢᵗ and 2ⁿᵈ persons singular:

Pron.	Set 2
išé	še-
endé	nde-

Other persons are virtually identical:

Pron.	Set 2
yandé	yande-
oré	ore-
peẽ	pe⁻-

Tupinamba gives evidence that at that early date, the differences between pronouns and Set 2 markers is not a grammatical, but a phonosyntactic one.

The difference between the pronouns and Set 2 markers is that when the pronouns are bound they lose their initial vowel. However, the initial vowel is not always lost, nor is it always kept when the pronoun was not bound. Hence dative case is given as *iše-βe, iše-βo, še-βe,* and *še-βo.* Likewise for the second person: *ende-βe, ende-βo, nde-βe, nde-βo.* The free pronoun as subject is also attested without an initial vowel. The following pair of sentences shows the contrast between pronouns and pm3 markers.

8. a.   išé še-suy
           I   1s-from
           'from me'

8. b.   endé nde-suy
               you 2s-from
               'from you (sg.)'

               (A 1595:12v)

Compare this with the following, noting the loss of initial *e* from *ende*.

9. a.   išé še-só-reme
               I   1s-go-if/when
               'if I go'

    b.   ndé nde-só-reme
               you 2s-go-if/when
               'if you go'

    c.   išé še-yuká
               I   1s-kill
               'he killed me'

    d.   ndé nde-yuká
               you 2s-kill
               'he killed you (sg.)'

               (A 1595:12v)

In Gj the 1s pm4 *he* (<*če) always appears as *ihe* (<*iče) with the locative postposition *pe*: *ihe-ße* < *iče-ße.

The 2s pm4 and pm5 in Kb is *ene* from *ende. The reason for the variation *peẽ* ~ *pe(˜)* is simply that when *peẽ* (or *pẽ*) is bound to a verb or noun, it does not bear stress. Since nasality on a vowel is an autosegmental feature in PTG associated with stress, as discussed in 2.4. above, no unstressed element does bear nasality on a vowel

(underlyingly). So when *pē(ē)* is stressed, its representation is as follows:
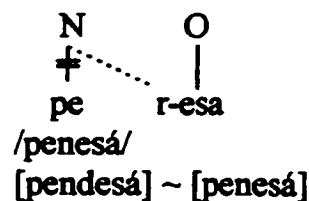
```
      N
      |
      pe
```
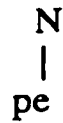
When it is bound, the autosegental feature is still there, but is unattached to the segment and is reattached to the following segment:

```
      N        O
      ╪ ·······  |
      pe       t-eča
```
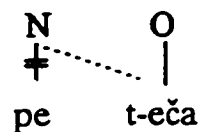
This results in **pe-ndečá*. After the re-interpretation of prenasalized stops as only oral allophones of nasal stops, this may have changed to the following (in Tupinamba, for example):

```
      N        O
      ╪ ·····  |
      pe       r-esa
      /penesá/
      [pendesá] ~ [penesá]
```

While some variation between *iše* ~ *še* and *ende* ~ *nde* is recorded, *peē* seems to be limited to the status of free pronoun, while *pe ̄*- is only attested as a Set 2 marker.

The point of all this is to show that while it may be necessary to reconstruct pronouns and Set 2 markers as different sets of morphemes, the difference between them is not as readily perceptible in the reconstructed forms as in the attested forms.

This may well reflect a closer affinity between pronouns and Set 2 markers in the parent language.

Finally, 2pl points to original *pee*, since intervocalic *h* in Guajajara and Asurini would point back to an original glottal stop, which would not be missing in OG. In this case, however, the pm5 form seems to be actually more conservative than the pm6. This conclusion is reached based on the paradigm as a whole.

pee   nominative
pe   accusative/non-reflexive possessive
pe   reflexive possessive
opo   accusative (1st person subject)

OG *pee* could be an alternation of *pehe*, but this would point back to *pece*, which, in turn, would become [§]*pese* in Tb and [§]*pece* in Gy. One way to avoid this complication is to suggest that the emphatic of the 2pl is simply a lengthened form of original *pe*. This will be further discussed below.

Most TG languages attest the pm3 category. OG, however, though older, was already on its way toward losing the reflexive possessive case by merging it with nominative case forms in the plural. This is by analogy with 3 *o-* and 1px *oro-,* which were already identical in pm1, pm2, and pm3. Taking phonological rules into account, this would result in the following forms for pm3.

	Sg	Pl
1pn		yere
1	we	oro
2	e	peye

Of the 1s forms OG, Tb and Pt point to *wi-, while Kb, AT, Km, and Tp all point to *we-. The form we- is probably original, it being a simple matter of assimilation (of height) to change we- to wi- in OG, Tb, and Pt.

There is some variation in the form of 1s Set 2 (or pm4/pm5) markers in TG languages. The attested forms in Tb, OG, W, UK, Gj, Ch, MG, Kw, and Gy point to PTG *če-; Pt, AT, AX, and Tp point to *ye; Kb ye could point to either *iče or *iye. The 1s pm3 in Kb, te, is unique in TG and explaining it leads to the explanation of the double reconstruction *če-/*ye-. The closest thing to a clue is the Sirionó 1sg prefix a- ~ ta- and Kokama ta-. Reference to these Macro-TG languages indicates that the explanation for Kb te- lies in pre-TG. The Sirionó and Kokama forms are and could be explained by the fusion of optative t- with a regular verb prefix.

The best explanation for *t-we- is that it is the sole example of an old dialectal variation of the first singular Set 3 marker in PTG where the optative prefix t- was prefixed to the pronominal marker. The idea here is that in PMTG, the *i-/*t-i-/*ya-/ *t-ya- confusion, specifically the free variation between *i- and *t-i- on the one hand (eventually causing *t-i- and *t-ya- to be reinterpreted as single morphemes *ti- and *tʸa)

led by analogy to an *a-* ~ *ta-* variation in the first singular, which then spread to all first singular markers. Set 2 **ye-* ~ *ťʸe-*, Set 3 *we-* ~ *ťʷe-*. Only Kokama and Sirionó maintained the *ta-* form in Set 1. The *ťʸe-* form in Set 2 is the source of **če-* and the slightly remodeled pronoun form *iče* (also by analogy: *ťʸe* and *ťʸe-* > *ce* and *ce-* respectively; *nde : ende :: pe-¯ : epē :: ye-/ce- : iye/ice*. Phonetically, *ice* was realized as *iče,* which influenced Set 2 *ce-* → *če*).

We see under closer examination, then, evidence for a variation in the first singular parallel to that of the first plural inclusive: *a-/ta-*, *ye-/ťʸe-*, and *we-/ťʷe-*. This variation, not at all preserved in Sateré-Mawé, is preserved in Sirionó and Kokama, and is partly preserved in PTG: no dialects preserve the variation in Set 1, but some do in Set 2 and Set 3, that is, with oblique case cross-referencers. Some dialects, ancestral to Pt, AT, AX, Tp, and Km did not preserve the variation even here. Only one language, Kb, is descended from a dialect preserving the variant forms in both Set 2 and Set 3; *ye-* < *iče-*, *te-* < **twe*. Tb, OG, Gj, UK, W, and possibly Tp, and Km are descended from dialects that preserved *iče*; but lost the alternate Set 3 marker **twe-*. Possibly the reason why **twe-* did not fare well was that it was clearly a variant of **we-*; PTG held a preference for the more conservative form. In Set 2, however, *ye-/če-* were so suppletive that neither appeared to be a derivative of the other, and each variant had an equal chance of being the preferred one in any given dialect.

5.2.4. Portmanteau forms

Finally, we come to the accusative forms for the 2nd person. In most TG languages, person hierarchy is manifested morphologically:

če-	1s obj	ore-	1px obj	a-	1→3
oro-	1→2s	opo-	1→2p	ere-	2s→3
nde-	3→2s	pe⁻-	3→2p	pe-	2p→3
		i-	3 obj		

In this chart, 1→2s signifies that the marker is used to reference a first person subject and a second person singular object, and so on. That is, objects are marked according to the relative position in the hierarchy of the subject, since 1 and 3 are at the top and bottom of this scale only one object is needed. For second person, two forms are needed and the one marking an object with a 'higher' subject is used exclusively for this purpose. In Pease's model, these markers are included with first person markers, but they are better understood as second person object markers. This means that the pm5 markers for second person are object markers with third person subject only. Only UK has lost this system.

There has been considerable remodeling of these markers, but their reconstruction is not difficult. Apart from the loss of initial vowels in Ch and Kw, all the irregularities in the comparison can be attributed to remodeling after the first and second person pronominals. Seki's (1990) assertion that these are originally portmanteau forms

fails to account for how *oropo-* contracted to *opo-*, or why the initial vowels in these prefixes changed the way they did.

Several languages also distinguish second person subjects with third person object from second person subject with first person object. Markers indicating second person subject with a first person object are not prefixes, but enclitics on the verb complex. This suggests that this part of the system is a later development than the rest of the cross-referencing system, but they are attested in enough languages that we can confidently reconstruct them for PTG.

Based on the above comparison we can reconstruct the following paradigm for PTG.

		yande
		(t)y(a)
		yere
(t)ye		ore
a		oro
(t)we		oro
(e)nde		pe̜
ere		pe
e		peye
oro		opo

Based on the above analysis, we arrive at the following reconstruction of PTG cross-reference markers:

PTG

	pm1	pm2	pm3	pm4	pm5	pm6
1st sg	a-	a- oro- opo-	we-	iye-/iče-	iye-/iče-	iye/iče
2nd sg imp	ere- e-	ere- e-	e-	ende-	ende-	ende
3rd sg fem sg masc plural general	o- o- o- o-	o- o- o- o-	o- o- o- o-	i-/c-	i-/c-	aʔe, aipo, etc. ace
1st pl excl	oro-	oro- oro- opo-	oro-	ore-	ore-	ore
1st pl incl	ya-	ya-	yere-	yande-	yande-	yande
2nd pl command	pe- pe-	pe- pe-	pe(y)e-	pe˜-	pe˜-	pe(y)ẽ

As a result of this comparison we see that there is no distinction between transitive and intransitive 1pn. There is therefore no difference between pm1 and pm2 except for the synthetic morphemes *oro-, *opo-, and the pronouns *epe and *epe(y)epe. Likewise, there is no difference between pm4 and pm5 and almost no difference between these two categories and pm6, and no gender or number distinction in any categories. For these reasons, the PTG system of cross-referencing lends itself to Jensen's (1990b) 4 categories or "sets". Set 1 corresponds to the markers in pm1 and pm2 (minus the synthetic morphemes). Set 2 corresponds to pm4 and pm5; Set 3 corresponds to pm3; Set 4 comprises the synthetic cross-referencers in pm2. Since the emphatic/topic pronouns in pm6 are not cross-reference markers, they are not included in these

categories.

Using this means of charting the cross-reference markers, PTG cross-referencing system looks like this:

PTG

	Set 1	Set 2	Set 3	Set 4
1st sg	a-	(i)ye-/(i)če-	we-	
2nd sg	ere-	(e)nde-	e-	oro-epe
3	o-	i-/c-	o-	
1st pl incl	ya-	yande-	yere-	
1st pl excl	oro-	ore-	oro-	
2nd pl	pe-	pe ~ -	peye-	opo-epeyepe

5.3. Changes for Cross-Referencing Systems

Tupinamba: Phonology: $č > š$, changing *če- to še- in Set 2; and raising of *we- to wi-.

Morphology: None.

Parintintin: Phonology: $y > ǰ$, $e > i/y\_$ $t^y > č$, hence *ye > ǰi '1s-2', *peye- > peǰi-, '2p-3'; *t^ya- > ča- '1p-1 intrans'.

Morphology: Pt has made the innovation of resolving *i-/*ya-/*ti-/*t^ya- by creating a transitive/intransitive distinction in Set 1 members. In Pt, both of these are from t- (optative) forms: *t-ya-, *t-i-. Also, the 2/1 Set 4 forms *epe '2p/1' and *epeyepe '2p/1' have been lost.

Urubu-Kaapor: Phonology: o > u so *o- has become u- except in roots having an o. č > ∅/h so *če- has become hẽ- with nasal ẽ that may be due to influence form pehẽ-.

Morphology: a radical innovation in UK is that all cross-referencing for objects has ceased; consequently Set 4 markers have been lost and Set 2 markers are only used for possessives and other oblique relations.

Verbs in subordinate or serial constructions are cross-referenced with Set 1 markers. Consequently Set 3 markers have been lost as well.

The inclusive/exclusive distinction has been lost; ya- and yande- cross-reference any 1pl referent; *oro-and *ore- have been lost. The 3-1 marker is only used with monosyllabic roots. 2pl-2 *pe- ~ has been replaced with pehẽ- from the pronoun pehẽ itself from the PTG free (topic) pronoun peẽ.

Assuriní of the Xingú: Phonology: y > ž, ǰ; ty > č

1pl incl-1 *ča-* points back to **tya-*; if this is correct, either AX made the transitive/intransitive distinction and later lost it, or it is the only TG language to resolve the *i-/*ya-/*ti-/*tya-* confusion by adopting a marked form (**t-ya-*) **without** adopting the transitive/intransitive distinction. Like Pt *ǰi-*, AX *ǰe-* seems to point back to **ye-*, since **če-* would have yielded §*he-* (exx. **-čó* > *-ho*, **puračéy* 'dance' > *porahey*; **kʷaračí* > *kʷarahi* 'sun'). This might support the hypothesis that PTG **če-* and **iče* were really **iye* (see under Kayabí below).

Morphology: As in Paraguayan Guaraní **pe-˜* has been remodeled as *pende-* due to a reanalysis of **pe-˜* with Class 2 roots e.g. **pe-˜ t-ečá* [pendecá] → *pende-(e)čá*. This may or may not explain 2sg-2 *ende-* (cf. PTG **nde-*). Either **nde-* was remodeled as *ende-* because of *pende-* or it represents, like **ǰe-*, a preservation of an earlier form of the cross-reference markers than is otherwise reconstructed for earlier or more conservative dialect than PTG may also be supported by the fact that AX has phonemic preasalized stops (it may seem to result from the loss of autosegmental nasality, which is decidedly innovative) but usually such loss results in the loss of prenasalized stops (though not in UK Again dialect borrowing may be a factor here).

A definite innovation is the generalization of the singular Set 4 markers to the plural.

Old Guaraní: Phonology: no phonological change with the possible exception of y/ñ free variation. The dominant form seems to have been with *ñ*.

Morphology: The Set 3 markers have undergone some remodeling. Original Set 3 markers **o-* 3-3 and **oro-* 1pl excl-3 are identical to Set 1 markers leading to the replacement of **yere-* and **peye-* with *ya-* and *pa-* respectively. Interestingly, these are the same two Set 3 markers replaced in Pt but by Set 2 markers in that langauge. This change in OG had the result that all but 1 sg and 2 sg markers of subordinate verbs are identical to Set 1 markers, leading to the reinterpretation of verb nominalizations subordinate to verbs as subordinate verbs and the subsequent complete disappearence of Set 3 markers in modern Guaranian languages.

Kayabí: Phonology: *ty > č > s *t-ya- > *ča- > sa-* ; **t-i- > *či- > si-*; loss of autosegmental nasality resulted in the loss of prenasalized stops. At the same time, nasal vowels became phonemic, hence **pe-~* became *pẽ-*.

Morphology: Kb kept the 2→1 Set 4 markers but lost the 1→2 Set 4 markers. Kb adapted the transitive/intransitive distinction for 1pl

incl-1 markers, both from the marked *t-* forms. Kb has reshaped

1pl incl-2 *yane-* to *yene-* by analogy with 1pl incl-3 *yere-*. Strangely,

2p-2 *pẽ-* has influenced the 2p-1 marker to also have a nasal vowel:

pẽ-. Stranger still is the 1sg-3 marker, *te-*. The most likely

derivation for this form is from **t-we-*, as discussed above.

Assuriní of Trocará: Phonology: *y* > *č* hence 1sg-2 **ye-* > *če-*, 1pl incl-1 **ya-* > *ča-*,

1pl incl-2 **yande-* > *čene-*, **yere-* > *čere-*. Loss of autosegmental

nasality and orality result in loss of prenasalized stops. Unaccented

e in **epe* raises to *i* perhaps as a result of contact with verb-final

vowels (i.e. it becomes a front off-glide).

Morphology: Unlike Kayabí, which remodeled the 1 pl incl-3 after the

corresponding Set 2 marker, AT remodeled **yande-* to *čene-* after

the Set 3 marker, although the free pronoun is still *čane*. Also the

2/1-4 marker *epe* has been generalized for the plural, replacing

*epeyepe.

Wayapí: Phonology: *č* > *∅* so **če-* > *e-*. Greater restriction of assimilation to

autosegmental phonemes has meant loss of prenasal stops in cross-

reference markers **yande-* > *yane-*, **nde-* > *ne-*.

Morphology: Set 3 markers are gone. 2/1 Set 4 markers are gone.

Kamayurá: Phonology: restriction of nasal and oral assimilation. Like AT and unlike
            Kb, Km has changed the 1pl incl-2 marker by analogy with the
            corresponding Set 3 marker: *yane- → yene.

            Morphology: 1>2 Set 4 markers are reanalyzed as portmanteau forms: a-
            (1s Set 1) + -ro- (2s) → aro- from *oro- reanalyzing the last two
            segments as a 2s object marker, by analogy with 2s Set 1 marker
            ere- and a- (1s Set 1) + -po- (2pl) → apo from *opo- reanalyzing
            the last two segments as a 2pl object marker, by analogy with 2pl
            Set 2 marker pe-. 2>1 Set 4 pronouns were lost.

Guajajara: Phonology: o > u results in *o- 3-1 > u-, 1px-1 *oro- > uru-, 1px-2 ore-
            > ure-, 1/2 pl-4 opo- > apu- 1/2sg-4. The remodeling of this prefix
            is explained below. y > z changed 1plincl-1 *ya- to za- and 1pn-2
            *yande- to zane-. Loss of autosegmental nasality resulted in the loss
            of prenasalized stops. Lenition of č to h resulted in 1s-2 *če- >
            he-; ty > s changed *ti- to si- 1pn-1trans.

Morphology: The Set 4 markers are here also remodeled as portmaneau forms: *opo- → a-po reanalyzing the last two segments as a 2p object marker, by analogy with 2p Set 2 marker pe¯-. Later, however, this was generalized as the a person marker for both numbers and Set 4 markers are no longer distinguished by number. In one dialect of Gj, a 1px form corresponding to the 1s form: a-pu, uru-pu (Harrison 1986:432). 2→1 Set 4 markers are lost.

Tapirape: Phonology: a > ã, o > a. 3-1 *o- > a-, 1px-1 *oro- > ara-, 1px-2 *ore- > are-, 1s-1 *a- > ã-, 2s/1-4 *oro- > ara- 2p/1-4 *opo (> *apo, see below) > ãpa-. y, ty > č hence 1pn-1 *(t)ya- > ča-, 1pn-1 *ti- > či-, 1pn-2 *yande- > čane-, 2p/1-4 *epeyepe > pečepe, 2p-3 *peye- > peče-. Loss of autosegmental nasality and orality led to the loss of prenasalized stops.

Morphology: Set 4 markers have been seriously altered *opo- was reanalyzed as in Gj to *apo-, which after sound change became ãpa-. Strangely this reanalysis did not take place in the singular. pečepe was reanalyzed as pe- 2p + čepe 1sg obj because of 1s-2 če-. Hence *epe was remodeled to čepe, which was analyed as če-pe. Consequently a corresponding form was made for 1px: are-

1px-2 *pe*. However, while *čepe* is 2s/1s-4, *arepe* is not 2s/1p excl, but rather 2/1px-4.

Chiriguano: Phonology: Assibilation has changed *y* to *ž* in 1pn pm1 and pm2 *ža-* and pm4, pm5, and pm6 *žande-*.

Morphology: The initial vowel was dropped from all (disyllabic) 2nd person markers, apparently by analogy with 2p *pe-*. Set 3 (pm3) has been replaced by Set 1 (pm1). 2>1 Set 4 markers have been lost. The combination of *pe⁻-* with class II roots, resulting in forms like *pe-ndesa* 'your eyes' has been reanalyzed as *pende-sa*, then as a result of levelling, *pende-r-esa*.

Mbya: Phonology: Affrication has changed **y* to *ǰ*. **c* has disappeared.

Morphology: Set 3 has been replaced by Set 1. Set 4 2>1 markers are gone.

Kaiwa: Phonology: Affrication has changed **y* to *ǰ*. **c* has disappeared and **č* has deoccluded to *š*.

Morphology: Set 3 has been replaced by Set 1. Set 4 2>1 markers are

gone.

Guarayo: Phonology: The 1s pm3 marker *we- has changed to wi-.

Morphology: None.

5.4 Case in PTG

There has already been considerable discussion of the case system of PTG

pronouns. The hypothesis has been advanced that PTG pronouns were ergative. Seki

regarded PTG as active/stative. In this section we shall see that the pronominal system

of PTG is best reconstructed as a nominative/accusative system, but which in modern TG

languages is beginning to evolve into an active/stative system. As our background to this

study we will look at the most comprehensive, and the most lucid, presentation of current

thinking on PTG pronominal case, by Jensen (1990b).

5.4.1. Jensen's theory of ergativity

In her paper, *Cross-Referencing Changes in Some Tupi-Guarani Languages,* Jensen

makes a thorough cross-linguistic study of cross-referencing pronominal elements in TG

languages and draws some conclusions about the development of cross-referencing from

PTG to modern languages.

Jensen begins by pointing out that the vast majority of TG languages, including its oldest attested member, Tupinamba, have the same system of cross-referencing on verbs. Using Tupinamba for her examples she demonstrates the system for independent verbs:

10. a. a-só   še-nupã³ ⁴
            1s-go 'I go'   1s-hit 'he/she/they/it/you hit me'
            S(ag)   P

    b.   a-i-nupã   še-katu
             1s-3-hit 'I hit it'   1s-good 'I am good'
             A-P   S(nom-ag)

Jensen notes that active subjects of intransitive verbs as well as agents use one set of pronominals while non-active subjects and patients use another set of pronominals.

These are designated Set 1 and Set 2 respectively:

    1s   a-   če- (r-)
    1px   oro-   ore- (r-)
    1pn   ja-   jane- (r-)
    2s   ere-   ne- (r-)
    2p   pe-   pe- (n-)
    3   o-   i-, c-

The determiners which follow Set 2 markers (before class II roots) are included in

---

³ Jensen (1990) supposes [še] to be /syé/; Anchieta records it as <xe>. I am not aware of any independent evidence for claiming that Tupinamba [š] was /sy/.

⁴ Jensen, follwoing Rodrigues (p.c.) writes set 2 markers as separate words; for reasons mentioned in Part I, I regard them as bound.

parentheses. Note that *pe-* is followed by *n-*. This is because *pe-* is from *pĕ*. The fact that it is not nasal as a Set 2 marker while still retaining nasal morphophonemics indicates that it has not denasalized, but at the same time does not bear nasal stress itself — strong support for the claim that Set 2 markers, like Set 1 markers, were bound morphemes. In addition to Sets 1 and 2, Jensen also refers to Set 3, which are the oblique markers referred to above. Finally she mentions Set 4, which are the portmanteau forms *oro-* and *opo-*. (Jensen does not mention *epe* or *epeyepe*).

In the following section, Jensen exemplifies the cross-referencing system; that is, Set 1 markers used when there is a third person patient, or where in intransitive verbs where the subject is active; Set 2 markers where there is a non-third person patient hierarchically superior to its agent for transitive verbs and for non-active subjects of intransitive verbs.

11.

a-só	'I went'
oro-só etc.	'we (excl.) went'
a-i-kutúk	'I pierced it,' etc.
oro-i-kutúk	'we (excl.) pierced it'
a-s-ekár	'I looked for it'
oro-s-ekár	'we (excl.) looked for it'

These examples also demonstrate the difference between class I and class II verb stems (*epʸák* 'see' is class II).

12.

še-r-epʸák	'he etc. saw me'
ore-r-epʸák	'he etc. saw us (excl.)'
oro-epʸák	'I/we saw you sg.'
opo-epʸák	'I/we saw you pl.'

Jensen goes on to demonstrate that in all other types of verbal constructions a different system of cross-referencing is employed, one using only Set 2 markers.

13. a. i-paw-ramo   Assurini
         3-finish-when

   b.   i-nopo-ramo   Assurini
         3-hit-when

   c.   pajé só-reme   Tupinamba
         shaman go-when

Tupinamba did not require a marker with a nominal subject. The same was true of OG:

14.   čuã ho-ramõ
         João go-when

This holds true for transitive verbs as well:

15. a. ma'é-asɨ-βór-a   pajé i-šubán-eme
         thing-pain-nom   shaman 3-suck-if
         'If the shaman sucks [treats] the patient'
                                           Tupinamba

   b.   pajé ma'é-asɨ-βór-a   šubán-eme
         shaman thing-pain-nom   suck-if
         'if the shaman treats the patient'

   c.   čuã perú juká-ramõ
         John Peter kill-when   Old Guarani
         'when John killed Peter'

Serial verbs (a similar rule holds for serial verbs):

16. a. o-úr s-ep^yɛ̌k-a   'He came to see him'
          3-come 3-see-SER

    b.   o-úr   i-kwáp-a   'He came to meet him'
               3-come 3-meet-SER

    c.   o-úr   kunumí kwáp-a   'He came to meet the boy'
               3-come boy   meet-SER

Intransitive serial verbs employ Set 3 person markers; these indicate the subject of

the serial is the same as that of the main verb.

Oblique-Topicalized Verbs

Guajajara:
17.a. a?e-pe h-eko-n   'He is there'
     that-at 3-be-OBTOP

    b. ka?a-pe ure-r-eraha-n   'He took us to the jungle'
      jungle-at 1px-det-take-OBTOP

Tupinamba:
18.a. k^wesé   i-só-w
      yesterday 3-go-OBTOP
      'yesterday he went'

    b. k^wesé jané-só-w
      yesterday 1pn-go-OBTOP
      'yesterday we went'

    c. k^wesé paJé se-šubán-i
      yesterday shaman 1s-suck-OBTOP
      'yesterday the shaman cured (lit. sucked) him'

Nominalizations
19.a. s-ep^yák-(a)
      3-see-(nom)
      'the seeing of him'

19.b.   ore-r-ekár-(a)
        1px-cm-seek-(nom)
        'the seeking of us'

   c.   ne-só
        2s-go
        'the going of you'

Agent

20.   i-mojàn-ár [imojandár]
        3-create-agnt
        'the creator of it'

Circumstance

21.a.   i-juká-sáβ
        3-kill-circ
        'the time/place/manner/instrument of killing them'

   b.   i-moján-áβ [imojandáβ]
        3-create-circ
        'the time/place/manner/instrument of creating them'

Patient Nominalizations

22.   -pɨr
        i-juká-pɨr(-a)
        3-kill-pass
        'the one who was killed by him'

23.a.   emi-
        še-r-emi-'u
        1s-det-pass-eat
        'the one eaten by me'

   b.   s-emi-moján(-a)
        3-pass-creat
        'the one created by him'

The prefix *emi-* presents a problem for Jensen's analysis because it demonstrates Set 2

prefixes marking an **agent**. But even in an ergative language one does not expect absolute markers (as Jensen interprets these) to mark agent. Jensen explains this by comparing *-emi-* with an incorporated object:

24.a. kunumí a-i-nupā   'I hit the boy'
        boy   1s-3-hit

    b. a-kunumí-nupā   'I hit a boy'
       1s-boy-hit

    c. se-r-emi-nupā   'the one whom I hit'
       1s-det-pass-hit

The idea is that just as incorporating an object detransitivizes a verb, so incorporating *emi-* detransitivizes a verb also. Hence 'I' is not an A but an S.

The problem with this argument is that it overlooks the crucial fact that *emi-* is a **nominalizing** prefix. I.e. (3) above is not a verb phrase, as (1) and (2) are, but a noun. This is also indicated by the presence of *-r-* marking *emi-nupā* as a class II noun (*nupā* as a verb stem is class I). This therefore can't be parallel to noun incorporation, which results in a verb phrase, since this is a noun phrase. Further confirmation of this can be seen in the fact that forms with *-emi-* take nominal tense.

25.a. če-r-emi-mboʔé kʷéra 'my past taught one', i.e. 'the one I taught'
       1s-ii-pass-teach past

    b. če-r-emi-mboʔé ráma 'my future taught one', i.e. 'the one I will teach'

    c. če-r-emi-mboʔé raŋgʷéra 'the one I was to teach, but did not'

The fact that stems prefixed by *emi-* are nouns or noun phrases not only weakens Jensen's explanation for this apparent kink in the data; it in fact suggests an entirely different analysis for all the verbal constructions Jensen describes.

In all TG languages, and reconstructable for PTG, Set II markers are used not only for patients and adjectival verbs, but also for possessives and postpositional objects; in short, for relations 2, 3, and OBL. Knowing this opens the possibility for interpreting Set II markers in nominalization as possessives. Jensen has already done this for nominalized actions.

26.   s-ep^{y}ák(a)   'his being seen'
                 3-see

         i-juká-sáβ 'his death place'
         i-moján-áβ 'its making place'
         i-moján-ár   'its maker'

This could apply to nominalizations involving *-pir* and *emi-*:

27.   se-r-emi-ʔu   'my food'
      s-emi-moján(a) 'his handiwork'

We have seen above (**4.2.1**) how the serial verb suffix was historically a nominalizing suffix and related to *-čár* and *-čáβ*, and how *-áβo* was derived from the circumstantial nominalizer *-áβ* plus the diffusive locative marker *-βo*, implying that the verb was functioning as a noun. Therefore we can interpret serial verbs as nominalizations as well.

28.   o-úr   s-ep^{y}áka   'He$_i$ came for his$_j$ seeing'

        o-úr   i-kwápa   'He$_i$ came for his$_j$ meeting'

        o-úr   kunumí kwápa   'He came for the boy's meeting'

                                       (i.e. he came to meet the boy)

Two facts about these examples, one morphological, one syntactic, strengthens this interpretation. Note that the noun *kunumí*, above, is not only in the syntactic position of an object, relative to a verb; it is also in the position of possessor, relative to a noun. Also, the person marker for *i-* 'his$_j$' is used with patients of transitive serial verbs because the object is not co-indexed with the subject. The possessor *i-* is used as a non-reflexive possessor. With intransitive serial verbs the reflexive possessor *o-* is used:

29.   o-só o-puraséy-ta   'He$_i$ went for his$_j$ singing.'

        3-go 3-sing-SER

Subordinating conjunctions (if/when) have been shown to behave similarly to class II stems. These may derive from either nouns or postpositions. The fact that forms such as *č-eme* are not attested is easily understood, since the conjunction is only used with verb stems; *č-eme* would mean something like 'if it', which is unacceptable as a clause. Even here, it is by no means certain that verbs in if/when clauses were interpreted as nouns in PTG. But their origin as such would give an alternate explanation to saying that PTG was a mostly (ergative) absolutive language. What seems to have happened is that pre-TG used unmarked verbs for independent predicates and expressed all other verbal concepts by different types of nominalizations. In PTG, changes in the pronominal markers (i.e. Set II markers ceased to be used for nouns) led to these constructions being reinterpreted as clauses. This created the cross-referencing system for PTG (adapted

from Jensen (1990b)).

	Set I	Set II	
	S	S(adj)	
	A	P	independent verbs (main)
	A	S	if/when, serial, ob-top verbs
		P	monosyllabic reduplication
			(suffix *-paβ)

We have seen how *páβ in PTG could mean simply completion of action (for either transitive or intransitive verbs) or it could mean a kind of completion of a plural subject (for intransitive verbs; i.e. all x did this). This alternate interpretation of *páβ may have been influenced by a reinterpretation of transitive verbs with *-páβ, understood as referring to the totality of the object, not the action: o-ʔu-páβ 'I completely ate it', i.e. I ate all of it.

It is a small semantic step from one to the other and may have been encouraged by the (newly) absolutive construction of subordinate clauses. The same may (or may not) be true of monosyllabic reduplication. Again, note that while such reduplication always has reference to the object when the verb is transitive, it can refer to the subject of an intransitive verb only when that subject is plural:

30.   a-i-mokó-kón   'I swallowed one after another'
            oro-pó-pór   'we jumped, one after the other'

Again, as with *-páβ, there is little difference between the frequentative reduplication referring to the action only and reduplication referring to the object 'I swallowed them

over and over' ~ 'I swallowed one after another'. This shift to an absolutive function of monosyllabic reduplication may have been due to the arising of quasi-absolutive subordinate clauses. If this is so, then disyllabic reduplication may be a relatively more recent development to express the purely frequentative notion originally expressed by monosyllabic reduplication.

Jensen cites two other verbal constructions which seem to demonstrate ergativity in PTG. One is the use of the completive suffix *-*páβ* 'all', which as Jensen points out, is derived from the verb *páβ*, 'to finish'. When used with intransitive verbs, *-*páβ* refers to the subject, when it is used with transitive verbs, it refers to the object (exx. from Jensen):

31.   o-có-páβ   'he already went, they all went'
           o-ʔu-páβ   'he ate it all'

Thus the thing whose totality is indicated by *-*páβ* corresponds to absolutive case.

Note, however, that with a singular subject of an intransitive verb, *-*páβ* refers only to the action itself, not to the subject, and in fact the notion of completing the action noted is little different from the totality of the object. Since *páβ* as a verb means to 'finish', it seems likely that this principally referred to the completion of the action without respect to subject or object.

The other construction is monosyllabic reduplication:

32.   a-i-mokón   'I swallowed them'

                  a-i-mokó-kón   'I swallowed one after another'

Here the reduplication refers to more than frequentative action, but specifically action on one patient after another. With intransitive verbs this refers to one subject after another:

33.   oro-po-pór   'we all jumped one after the other'

Hence reduplication refers to a repetition of the argument which corresponds to absolutive case.

5.4.2. Subordinate clauses as noun phrases

There is another way to interpret the facts that are shown in Jensen's chart. In the Guaranian languages (Kaiwa, Chiriguano, Mbya, Guarani), there is a division in cross-referencing between nominalizations of intransitive verbs and nominalizations of transitive verbs. But this is because these are nominalized with different suffixes: intransitive verbs are nominalized with *-βa* and transitive verbs are nominalized with *-hára*. In these languages, *-βa* is no longer interpreted as a nominalizer,[5] but as a

---

[5] If it ever was, in the strict sense; OG *-βaʔe* was an enclitic attached to a verb phrase, not suffixed to a stem like *-háβ* and *-hár*.

relative conjunction joining a relative **clause** to a main clause. On the other hand *-hára* still seems to be a nominalizing suffix which turns verbs into nouns. It is not accurate to say that such nouns have cross-referencing, since they need not bear any prefix at all: *a-heša mboʔehára*, 'I saw the teacher'; cross-referencing is just a possessive prefix on a noun. *Pedro mboʔehára*, 'Peter is a teacher'; *Pedro mboʔeharakʷé*, 'Peter is a former teacher,' i.e. 'was a teacher', showing the use of nominal tense with *-hára*. Note that nominal tense is no longer used with *-βa* as such: *-βaʔekue* and *-βaʔerā* have been reinterpreted for different functions.

The circumstantial suffix *-ha* (Guarani) or *-a* (Chiriguano, Kaiwa) is no longer interpreted merely as a nominalizer, but also as a subordinating conjunction, linking a subordinating clause with a main clause, even though *-(h)a* is still a suffix. This is confirmed by the fact that unlike OG, *-(h)a* cannot be used with a bare verb stem.

34. še ha-ʔe o-ñeʔē-ha
     1s 1s-say 3-speak-that 'I said that he spoke to me'

By contrast, nominalizations in Wayampi are still interpreted as nominalizations, not as subordinate verbs.

35. moʔe-a 'place of teaching'

     pira-r-ekɨy-ta 'instrument for fishing'

This explains why they only bear possessive prefixes, never nominative (or, in truth,

accusative) cross-referencing. Verbs with *reme*, and serial verbs, however, have been reinterpreted as verb + subordinating conjunction and so are cross-referenced like other verbs. The Ob-Top suffix no longer exists as a productive suffix, according to Jensen (1990b). So this has been lost, not reinterpreted, and clauses with formerly Ob-Top verbs are now treated as any normal main clause without special marking on the verbs.

What looks like absolutive cross-referencing in Jensen's chart, then, is simply possessive marking on verbal nouns. There was only one cross-referencing system for PTG: that for verbs. The spread of verbal cross-referencing to other constructions simply followed the reinterpretation of verbal nouns as (conjugated) verbs. As verbal nouns were reinterpreted as verbs heading subordinate clauses, verbal cross-referencing was naturally used for them.

Other examples of *have* as a linking verb:

36. a. yandê yombaê 'we have each other's things' (A 1595:16)

    b.   to be: Tûba, oyeûbamo ceçôu or ogûbamo ceçóu
                   oyecïramo, ceçou or ocîramo ceçóu
                                (A 1595:17)

5.4.2.1. Serial verb suffixes as nominalizers

In Tupinamba, serial verb suffixes have undergone a reanalysis:

37.   s-ep^{y}ák-a   cepiáca   'seeing him'
                              3-see=nom

                              s-ausúpa   çauçúpa   'loving him'
                              3-love=nom

37.(con't)   káy-a   cáya   'falling'
                   fall-nom

         (potáa >)   i-potá   potá   'wanting it'
                      3-want=nom

That this is a reanalysis is evident form the fact that verbs with final *r* lost their final *r*;

if these were from verb + *a*, we should expect *potára*. Note the corresponding forms for

OG:

38.   a.   h-ečáka
                               3-see=nom

       b.   h-aihúpa
                               3-love=nom

       c.   káy-ta
                               fall-nom

       d.   potáta
                               want=nom

We have already seen (**4.2.1** and **5.4.1**) how these derive form verb + *ca* in pre-TG. The

data in Tupinamba, however, suggest that there was an alternate development of this

suffix with coronal-final verbs. The above examples of serial verb forms are shown below

as reconstructed for Pre-TG:

39.   a.   ep^{y}ák-t^{y}a
                               see-nom

       b.   ait^{y}úp-t^{y}a
                               love-nom

       c.   káy-t^{y}a
                               fall-nom

39.   d.   pottát-tʸa
                                    want-nom

The coronal-final verbs are unique in that they end in a consonant homorganic with that of the consonant in the suffix. There are two ways that these could develop. One is for them to develop like other stop-final consonants.

40.   a.   epʸákka
                                      see=nom

       b.   aɨtʸúppa
                                      love=nom

       c.   káy-tʸa
                                      fall-nom

       d.   pottátta
                                      want=nom

The lenition rule then resulted in the PTG forms

41.   a.   epʸáka
                                      see=nom

       b.   aɨcúpa
                                      love=nom

       c.   káy-ca
                                      fall-nom

       d.   potáta
                                    want=nom

These would give us the forms we find attested in OG. Pre-TG ****pottattʸa* could develop

another way, however. That is for *t* to assimilate to *t^ya* or remain unchanged:

42. a.   ep^yákka
                     see=nom

    b.   aɨt^yúppa
                     love=nom

    c.   káyt^ya
                     fall=nom

    d.   pottá(t)t^ya
                     want=nom

After lenition, this would give us:

43. a.   ep^yáka
                     see=nom

    b.   aɨt^yúpa
                     love=nom

    c.   káyt^ya
                     fall-nom

    d.   potát^ya
                     want=nom

After dissimilation and affricatization, we would have the following PTG forms:

44. a.   ep^yáka
                     see=nom

    b.   aɨcúpa
                     love=nom

    c.   káyta
                     fall=nom

44. d.   potáca
                                 want=nom

So at this stage we have two alternative forms for the serial verb from the t-final stems:

potáta and *potáca*.

Recall that velar stops did not lenite in PTG and anterior stops had not

completely undergone lenition word-finally so that a verb like *aɨcúβ* freely alternated with

aɨcúp. This would make the above forms analyzable as follows:

45. a.   epʸák-a
                                 see-nom

     b.   aɨcúp-a
                                 love-nom

     c.   káy-ta
                                 fall-nom

     d.   potá-ca
                                 want-nom

Without any morphophonemics one could analyze the serial verb suffix as -*a* attached

to the end of a verb stem. Since epenthetic *a* was often added to nouns and serial verbs

were, in fact, nouns in PTG, this would seem a logical reanalysis. Apparently anomalous

forms like -*ta* and -*ca* were then levelled out in favor of this suffix as we see in

Tupinamba:

46. a.   ep^{y}ák-a
             see-nom

   b.   aɨcúp-a
             love-nom

   c. *káy-ta →   káy-a
             fall-nom

   d. *potá-sa →   potá-a (> potá)
             want-nom

In the TG dialect(s) where *pottátt^{y}a > potáta, it fit in with the rest of the fortis-final verbs and so no levelling took place.

5.4.2.2 Active/stative case

Perhaps the most striking problem with Jensen's analysis is the split subject division in independent verbs. This split according to Jensen is between active subjects and inactive subjects. Consider however the following :

47.   o-manō
       3-die
       Set 1

While one may argue that dying in TG culture is considered an active verb, this may not have any bearing on the grammar. Furthermore consider:

48.   o-ikó 'he lives (exists)'
       o-īn 'he is (in some place)'
       o-ʔãm 'he is standing'
       o-ʔár 'he fell'

One would be hard put to it to explain how these verbs are more active than the

following:

49.   če-rorɨ   'I am happy'
               c-acɨ   'they are sick'

In fact all independent verbs using Set 2 markers fall into two categories. One are verbs

exhibiting apparent incorporation of a subject:

50. a. če- pɨtu-có   'I breathe'
             1s-breath-go

    b.   nde-r- eca-ráy   'you forget'
              2s-det-eye-shrink

These will be dealt with in Chapter 6. The other category, and the one constituting the

vast majority of verbs taking Set 2 markers, are adjectival verbs like those above.

Seki (1990) also notes this distinction for Kamayurá showing that verbs which

mark subjects with Set 2 markers are made up (primarily) of adjectival verbs, but also

include verbs like the following:

51.   earay   'to forget'
             eakwap   'to remember'
             ʔangehɨy   'to sigh'
             ʔangeraha 'to mourn'

Because these verbs all seem to have a common semantic feature, which Seki refers to

as 'lack of control', Seki, like Jensen, regards the distinction between verbs using Set 1

markers and verbs using Set 2 markers as an active-stative distinction. Seki recognizes,

however, that in Kamayurá as with other TG languages, not all 'active' verbs are active:

52.   -mano   'to die'
             -in   'to be sitting'
             -kuy   'to fall down'
             -kɨye   'to fear'
             -ko   'to be (in a place)'

The real explanation for this distinction in marking seems to be a historical, not a synchronic, semantic one.

Most TG languages have stative verbs like those above. In Paraguayan Guarani nearly all of these are analyzable as instances of what is referred to by some as 'subject incorporation'.

53.   še-pɨtu-hó
             1s-breath-go
             'I breathe'

In these verbs, a possessed subject is incorporated into the root of the verb, hence *še-pɨtu o-hó* becomes *še-pɨtuhó*. This process will be analyzed in **6.2.1.1**.

In Guarani, only two verbs in this category are not readily analyzable as instances of subject incorporation: *esaráy*, 'forget' (cognate with Kamayurá *earay*) and *manduʔá*, 'remember'. These derive from OG *esaráy* and *maʔenduʔár* respectively. In OG these too are analyzable as instances of subject incorporation:

54. a. če-r-esa=ráy Peru-wi
        1s-det-eye=shrivel Peter-from
        'my eyes shrink/shrivel from Peter' = 'I forgot Peter.'

   b.   če-maʔē̃=tuʔár   Peru-rehe
         1s-vision=remains   Peter-concerning
         'my vision concerning Peter remains' = 'I remember Peter'

Lexical and phonological changes have made it impossible to interpret these as incorporations in the modern languages and so they have become verb roots. The same is true for some/most examples of 'subject incorporation' in Kamayurá. But in PTG, these constructions were not verb roots at all, but examples of a syntactic process in PTG. Hence the only verbs which took Set 2 prefixes were descriptive verbs.

In PTG the distinction was between descriptive verbs and all other verbs: all non-descriptive stative verbs used Set I markers and all descriptive verbs used Set 2 markers. There is no semantic rationale for this distinction, and one cannot suggest synchronically that Set 2 markers were 'just used' with adjectival verbs; that is a description, not an explanation. We can give an adequate account of these forms in terms of Relational Grammar theory.[6] We have to assume that Set 1 verbs marked only subjects (agents

---

[6] Relational Grammar (RG) is a syntactic theory that analyzes clauses purely in terms of grammatical relations within clauses, with no regard for the ordering of constituents in the clause (such ordering is considered by the theory to be language specific). The primary relations possible between constituents are arranged in the following hierarchy: Subject > Direct Object > Indirect Object > Oblique. The positing of 'oblique' means that all relations other than the first three (instrumental, benefactive, possessive, and so on) have no hierarchical relationship to one another. This hierarchy is based on strong independent syntactic tendencies (wrongly called 'universals') attested in dozens of widely varying languages (the relativization hierarchy mentioned in 4.2.2.1 is an example). Schematically these four relations are labeled 1, 2, 3, and OBL respectively. The matrix of relationships of constituents within a clause is

of transitive verbs, and subjects of all intransitive verbs) while Set 2 markers denoted the grammatical relations 2, 3, and OBL. Why then, in a clause translatable as 'I am good' is 'I' in 2, 3, or OBL relation? More disturbingly, because of the Final 1 Law, which states that for every language each clause must have exactly one constituent bearing the 1 relation at the final stratum, the question arises: what bears the 1 relation at the final stratum in the descriptive clause of PTG?

To answer these questions, we must first identify the grammatical relation borne by the Set 2 marker that is attached to the adjectival verb. In the terms of Relational Grammar, it cannot be the 3 relation because it does not function as an indirect object, and further because when a Set 2 marker denotes the 3 relation it is always immediately followed by the suffix $-\beta e$. If the Set 2 marker were initially a 2 that would make sense semantically, but not grammatically because there is absolutely nothing in the language to keep an initial 2 from raising to a final 1, as in stative verbs:

---

represented by curved arrows (called 'arcs') connected to a single point of origin (the clause node).
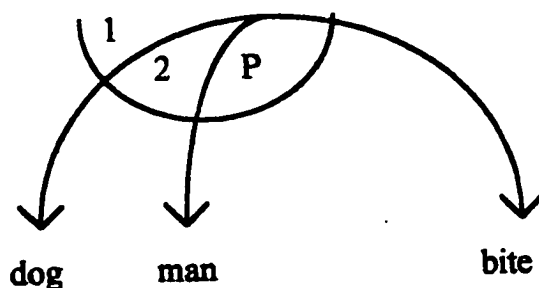
In some clauses, such as passive sentences, for example, a constituent may have more than one relation (direct object with regard to theta role, subject with regard to case assignment). Usually, instead of representing this by multiple arcs, RG represents this by multiple strata on a single arrow.

o   ʔa   o-ʔa 'they fell'
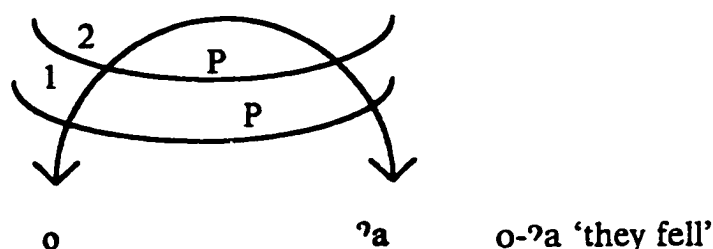
That leaves OBL, and there is one OBL relation which is expressed with a Set 2 or independent pronoun and no postposition or other affix: the relation of Possessor (POSS).

That Set 2 markers in PTG marked POSS is supported by evidence from Kamayurá. In Kamayurá, according to Seki (1990), descriptive (adjectival) verbs cannot take the same imperative construction as other verbs:

55. a.   e-yan   'Run!'   (Seki p. 370)
                 2s imper-run

(55) b.   i-katu e-ko   'Be good.'   (Seki p. 371)
                          3-good 2s=imper-be

Note that the subject of *katu* is 3rd person, while the imperative of 'be' is 2nd person. One reason for the necessity of this type of construction, is that the real subject of the verb is a null expletive for which there is no imperative form. Imperatives, therefore, must be expressed by using a verb taking Set 1 prefixes (*ko*) and making the adjectival

verb subordinate to this verb.

In OG the verb 'to have' was phonologically null so that a construction like *če-róγ* could mean 'my house' (a noun phrase) or 'I have a house' (a clause) depending on the context. Adjectives in PTG may not have been treated as verbs at all, but rather as nouns.

While all of this sheds light on the nature of cross-referencing in PTG, it does not address the fact that in Kamayurá there are non-adjectival verbs which take Set 2 markers to indicate subject; and even adjectival verbs taking Set 2 pronouns cannot be explained the same way they are for PTG, Tb, or OG, since the verb 'to have' and appositive nouns are not handled that way in Kamayurá. We therefore have a distinct class of verbs made up largely, but by no means entirely, of adjectival verbs. The significance of the foregoing is to show that the existence of this class has not originated from the grammaticalization of a semantic distinction; rather the emergence of these two classes is the purely accidental result of historical changes from PTG to Kamayurá. Nevertheless it is worth noting that even though the so-called active verbs in Kamayurá include stative verbs, none of the so-called stative verbs is semantically active.

So we have two classes of verbs in Kamayurá: one class consisting entirely of stative verbs and another consisting of active and stative verbs. In Kamayurá, unlike PTG, the Set 2 prefixes that cross-reference the stative verbs can only be interpreted as

either subjects of the stative class verbs or patients of active class verbs. In other words, Kamayurá is not a language with a straightforward active-stative distinction in verb morphology; but it is a formerly nominative-accusative language which, through a combination of otherwise unrelated historical changes, is seeing the emergence of a genuine active-stative distinction. It would be interesting to see how this system develops in the future: whether some of the stative verbs of the active class become semantically active,[7] and/or whether some of the stative verbs of the active class remains stative sematically and adopt Set 2 prefixes.

5.5 Internal reconstruction

We have arrived at a comparative reconstruction of PTG pronominals; we will now apply internal reconstruction to these proto-forms to arrive at earlier stages of the pronominal system. As with the phonological and stem-class systems, our interest in reconstructing pre-TG stages is in expediting comparisons with more distantly related languages to arrive at a proto-Tupian reconstruction.

---

[7] This has already begun to happen in Wayampi where *eko/iko 'to be/to live' has become *ko, 'to be in movement', making it more active semantically. Wayampi has spread this morphological distinction to other verb constructions (see above). Guarani has spread it even further, but except for 'remember' and 'forget' all verbs of the would-be stative class are adjectival or can be analyzed as examples of subject incorporation (object incorporation is still a productive process in this language).

Dietrich (1990) claims that such verbs in Chiriguano and Guarayo are no longer results of a syntactic process, but are interpreted as (simple) verb stems. Adjectives, however, he interprets as nouns so that a Set 2 marker and adjectival noun is simply a possessive expression, either as a (noun) phrase, ex. *i-puku* 'his longness' or as a clause 'it has length'. Interestingly, while Guarani and Guarayo have a genitive construction involving the juxtaposition of two nouns (Dietrich *gwíra pepo* 'a bird's feather') Chiriguano prefixes using a 3rd person possessive, *gwíra i-pepo*.

5.5.1. Third person cross-reference markers

In 4.1.2. we saw the anomaly in 3 person possessive cross-reference markers in PTG: the third person possessive is marked by *i-* on class I stems, *c-* on class II stems. We have also seen from 2.2. above how *c may be traced to **p^y, **t^y, and **k^y. This phonological change helps to explain the morphology of possessives in PTG. Possessives have two forms: one is noun/pronoun + (determiner +) noun:

56.   a.   iye r-eca   (class II)
                             1s cm-eye
                             'my eye'

      b.   tubica r-uβ (class II)
                             chief cm-father
                             'the chief's father'

      c.   iye po   (class I)
                             1s hand
                             'my hand'

the other is 3 poss + noun:

57.   a.   c-uβ
                             3-father
                             'his father'

      b.   i-po
                             3-hand
                             'his hand'

It is strange that the 3 person possessive has no class marker for class II verbs when they are always used with possessive nouns and pronouns. At the same time, class II nouns take a different 3 possessor than class I. I.e. *c-* is only used with nouns that otherwise

use the determiner *t-*, while *i-* is only used with nouns not using the determiner *t-*. The phonological origin of *c-* may explain the origin of this morphology. If *c can derive from **t^y- this 3 poss. may in fact be **t-i-, i.e. det. + poss. + noun. The only wrinkle here is the ordering of morphemes in a phrase:

58. a. iye-t- up
     1s- det-father
     'my father'

    b. t-i-up
        det-3-father
        'his father'

The strongest argument in favor of ***t-i- is that there is only one reflexive 3rd person possessor for class I and II; one would expect therefore one 3 non-reflexive possessor form for both classes also. And is it a coincidence that *i-* is used with that class which doesn't use *t-* while *c-* is used for those nouns which do use *t-*?

Note, however, that the 3rd reflexive possessor *o-* definitely did not need a class marker.

59. a. ow-ok(a)   'his own house'
     3refl-house

    b. ow-eca   'his own eyes'
        3refl-eye

The main argument against *t-i- is that there seems to be no real grammatical evidence for claiming that *o-* needed no class marker just because it was reflexive. Hence, we must assume for now that, if *o-* did not need a determiner, neither did *i-*.

5.5.1.1. Deixis in third person pronominals

Upon closer examination, there is a grammatical reason for *i-* being prefixed by *t-* while *o-* was not. According to Rodrigues (personal communication) *t-* in Tupinamba signified more than class; it signified possession by an unknown and/or unreferenced possessor. This function also holds true in other languages, such as Kamayurá.

Realizing that possessives do not exist in isolation, we see that there is a relationship between the subject of the verb and the possessor of a noun complex. Consider the following as an example: 'he beat his animal'. The referencer used for 'his' depends on its relationship to 'he'. The possessives could reflect a deictic structure like the following:

```
        'here'          'there'
         o-               t-
                         / \
                        /   \
                'has been    'has not been
                mentioned    mentioned in
                in discourse' discourse'

                    i-           Ø-
```
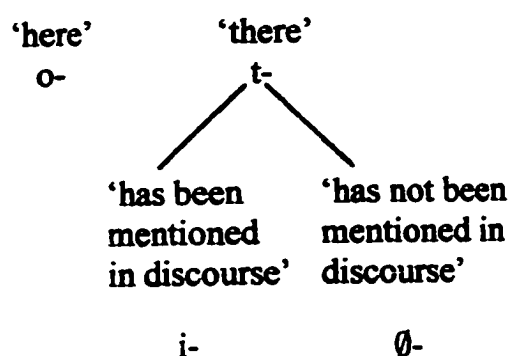
If the antecedent of 'his' was 'he' then the possessor would be 'here' with respect to the subject.

60.   o-nupã o-eimba 'he_i beat his_i animal'   (oweimbá)

If the antecedent of 'his' was someone other than 'he', then the possessor would be

'there' with respect to the subject, hence *t-* would be used; followed by *i-* if reference had been made to it earlier in the discourse, without *i-* if not.

61.   o-nupā t-i-eimba 'he$_i$ beat his$_j$ animal' (ceimbá)
        o-nupā t- eimba 'he beat someone's animal' (teimbá)

I refer to *o-* and *t-* as 'here' and 'there' respectively because they did not at this early stage have a purely, or even primarily, pronominal function, but rather a deictic one. Hence the contrast between Pre-TG **-*o-ām* and **-*r-ām* (**-*t-ām*) 'future' is understood in light of this analysis as a contrast between 'immediate future' or 'near future' and 'unforeseen future'. When the functions of **o- and **t-/r- (which in PTG were two distinct morphemes *t-* and *r-*) took on primarily cross-referencing and/or possessive functions and the original deictic functions were lost, *-wām* and *-rām* became increasingly synonymous. Consequently one suffix supplanted the other, or were separated according to function (so in OG, both *-wām* and *-rām* were used with noun stems, but only *-wām* was used with nominalized words (ending in *-hára*, *-háβa*, etc.; later noun stems are made future exclusively with *-rām*).

5.5.2. First and second person cross-reference markers

  A certain amount of internal reconstruction can be done based on the form *epe*. As mentioned above, *epe* is clearly derived form the 2p paradigm. Note the parallel between *ende-*, *ere-*, and *oro-*. Contrast this with *pē-*, *pe-*, *opo-*. One possibility is that

epe was created by proportional analogy to *ere-*, i.e. *oro* : *ere* :: *opo* : *x*, where *x* = *epe*. If that were the case, though, it would be strange that the ergative did not become *epe*, but rather that *epe* was immediately used for an accusative function.

Another possibility is that the earlier use of *epe* preserves the original form of the 2p pronoun which lost its initial vowel just as *ende* did (and as did *ere*, *oro*, and *oro* 1px in MG). This would suggest then an earlier paradigm:

Nom.		ya
Acc.		yande
Poss.		yere
Nom.	a	oro
Acc.	iye	ore
Poss.	we	oro
Nom.	ere	epe
Acc.	ende	epẽ
Poss.	e	epe
Acc.	oro	opo

On this note, it may be mentioned that *epe* is attested in Munduruku as a 2p accusative marker. That the 2p nominative *epe* was originally vowel-initial would explain the long vowel in the forms *pe(y)ẽ* and *pe(y)e*. As we saw in Part 1, the loss of an initial vowel before *p* could lead to compensatory lengthening of the remaining vowel (PTG **apé* ~ **peé* 'path').

The pronominal forms shown were originally free pronouns. The process of

binding pronouns to verbs is an important aspect of TG historical linguistics and will be developed more fully in Chapter 6.

The only point remaining with regard to pronouns is the question of phonology and case. A look at the PTG paradigm shows a clear pattern in the nominative and accusative cases, and also the accusative case in the 2nd person pronouns:

Nom.	a	oro
Acc.	iye	ore
Poss.	we	oro
Nom.	ere	epe
Acc.	ende	epẽ
Poss.	e	epe
Acc.	oro	opo

One sees an e/o ablaut in these pronouns as well as nasalization in the 2nd person pronouns. This is especially true if we realize that 1s *iye* and *we* could have been *e(y)e* and *oe* respectively (on independent evidence for #wV from #oV, see **3.2** above). Taking these forms before lenition (see **2.1.1**) would give us the following:

Nom.	a	oto
Acc.	ee	ote
Poss.	oe	oto
Nom.	ete	eʔpe
Acc.	ende	eʔpẽ
Poss.	e	eʔpe
Acc.	oto	oʔpo

One possible explanation for the *e/o* ablaut is that the *e/o* forms are from pre-TG *ə*

which split into [o], [e], and [ə] which was grammaticalized in pre-TG. We can propose an earlier stage where *ə* = nominative, *e* = accusative, and *o* = (reflexive) possessive.

Nom.	ə	otə
Acc.	e-e	ote
Poss.	o-e	oto
Nom.	etə	eʔpə
Acc.	etẽ	eʔpẽ
Poss.	e	eʔpe
Acc.	eto	eʔpo

A change occurred assimilating *ə* to the preceding vowel. Later *eto* and *eʔpo* would have changed to *oto* and *oʔpo* by analogy to *eCe* forms in the nominative and accusative. The nasal morpheme, also indicating accusative case is attested elsewhere in Tupian: Karitiana ˉ-*n*, Tupari ˉ-*n*, and Munduruku ˉ-*n*; in the last of these, it marks ergative case. Its application is a bit obscure, however; in Munduruku it is used for both 1st and 2nd person pronouns, but only in the singular. In the case of PTG, the nasal morpheme appears on accusative/non-reflexive possessive case in the second person.

It will be noted that this analysis divides 1s oblique cases into a stem (*e-* or *o-*) plus a suffix *-e*. This suffix is also attested in Munduruku and could have been suffixed to the second person forms, assimilating to the final vowel of each oblique form. That suffix may also help the internal reconstruction of 1pn *ʔi-/ya-*.

Nom.	ʔi/ya
Acc.	yande
Poss.	yere

We could reconstruct parallel 1pn forms based on either *ʔi* or *ya*.

Nom.	ʔi	ya
Acc.	ʔinde	yande
Poss.	yere	yare

A cognate (or reflex) of **ʔinde*, is actually attested as the 1pn Set 2 marker *ini* in Kokama, a language classified as TG, but probably a non-TG Tupian language. From the internal pattern of 1pn markers and the fact that *e* in the second person is the subject of imperatives and the only subject form in non-TG Tupian languages, we can see that **-nde* and **-re* are from the suffix **-tə*.

Nom.		ʔi ~ ya
Acc.		yẽ-tə ~ yã-tə
Poss.		ye-tə ~ ya-tə
Nom.	ə	otə
Acc.	e-e	ote
Poss.	o-e	oto
Nom.	e-tə	eʔpə
Acc.	e-tẽ	eʔpẽ
Poss.	e	eʔpe
Acc.	e-to	eʔpo

This is one way of explaining 1pn Set 2 *yene* in some languages and *yane* in others; likewise 1pn Set 3 *yere* in some languages and *yare* in others. But the main support for this reconstruction is from both internal evidence in the patterns of the system reconstructed for PTG, and comparison with cross-referencing/pronominal systems in non-TG Tupian languages. Indeed, while the above hypothesis already goes a long way

toward a Proto-Tupian pronominal system, an even earlier stage of this system can be reconstructed using comparative Tupian evidence, but such a reconstruction exceeds the scope of this work.

5.6. Conclusion

In this chapter, we have compared the person/number cross-reference markers for verbs and nouns in TG to arrive at a reconstruction of PTG. In the process of so doing, we have seen that reanalysis of the Set 3 third person cross-reference markers as class markers has led to their replacement in the accusative and non-reflexive possessive cases with pronouns derived most likely from discourse particles in PTG. We have also reached the following conclusions about PTG grammar:

1. PTG was a nominative/accusative language.

2. Emphatic pronouns were derived from accusative/non-reflexive possessive cross-reference markers.

3. These markers employed a $e \sim o$ ablaut to distinguish case.

4. The optative prefix *t- optionally blended with the 1pn marker, and by extension, with the 1s marker as well, but apparently not with the 1px markers.

5. The original form of the 1pn nominative pronoun was *$ʔi$; as in other Tupian languages, it was replaced by an extended form, but both $ʔi$- and its extended form ya-

were still in use in the early historical period.

Internal reconstruction has led to the following conclusions about Pre-TG grammar:

1. Cross-reference markers were possibly free morphemes in Pre-TG.

2. Both singular and plural 2→1 Set 4 pronouns were derived from the 2pl nominative prefix/pronoun, indicating the use of 2pl pronoun for 2s in pre-TG.

3. *o-* and *t-* were originally deictics indicating 'here' and 'there' and were applied to cross-referencing nouns and verbs to indicate 'same as subject of the verb' and 'different than subject of the verb'; *i-* indicated previous occurrence in the discourse.

4. Third person non-reflexive possessive was marked by *i-* for all classes of nouns; PTG **c-* for Class II resulted from the fusion of *-i-* with the deictic *t-*.

We have reconstructed Pre-TG pronouns to a considerable time depth, and it has been noted that further reconstruction is feasible using comparative data from non-TG Tupian languages. Such comparisons could be very helpful in understanding the development of non-TG Tupian languages and in relating Tupian to other language stock of South America. Reconstruction of such a depth, however, passes beyond the scope of the present work. Instead, the reference made earlier to the original free status of PTG

cross-reference markers leads us within the scope of this study to examine the history of the verb complex in the next chapter.

PART TWO: MORPHOLOGY
CHAPTER SIX: VERB COMPLEXES

6.0. Introduction

At the beginning of Part Two, we examined the system of noun classifications and saw how the classification of noun roots also applies to verb roots and, to a limited degree, to postpositions as well. The characteristics of different stem classes led to a more in-depth study of person/number cross-reference markers and pronouns in TG. In so doing, we discovered interesting facts concerning the formation of verbs in TG grammar; in particular, that constructs expressed by subordinate clauses in most languages, certainly Indo-European languages, are expressed not by verb phrases at all, but by noun phrases centering on a nominalized verb. We see here two things showing the thin line between nouns and verbs in TG: the common structure of verb stems and noun stems; and the use of nominalized verbs to express verbal concepts.

In this chapter we will further examine the relationship between verb and noun roots. We will also examine the construction of the highest level of morphological complexity in TG, namely the complex. Because we have seen, in the previous two chapters, the basic construction of noun complexes, we will concentrate on the formation of verb complexes cross-linguistically and try to ascertain the construction of verb complexes in PTG.

6.1. Derived Verb Stems

When we compare TG verb complexes we find three basic types of verbs: intransitive, transitive, and stative. All three can be used as the basis for verb complexes. Verbs can be either simple or derived. The two derivative prefixes are the causative and comitative prefixes.

Pt

1. h-enduβ-i   (Pease 1968:1)
    3-hear-foc
    'hears it'

comitative (*-ero-*)
2. h-**ero**-kʷap-a (Pease 1968:4)
    3-**com**- nom
    'passing by with'

causative (*mbo-*)
3. o-ñaka-**mo**-hē   (Pease 1968:6)
    3-head-**caus**-leave
    'stuck its head out'

4. o-**mbo**-hovyovy   (Pease 1968:6)
    3-**caus**-green
    'he painted it green'

Gj

5. he-azutar   (B-S 1972:95)
    1s-help
    'he helped me'

comitative (*-(e)ru-*)
6. ue-**ru**-zewɨr (B-S 1972:98
    3-**com**-return
    'returned with him'

causative (*mo-*)
7.   ere-**mo**-pok   (B-S 1972:98)
            2s -**caus**-burst
            'you made it burst'

In Gj -*mu*- can also be used with a pre-nominal such as *nən*, 'not':
8.   a-mu-nən   (B-S 1972:97)
            1-caus-not
            'I erased'

Tp

9.   ã-ãrõ   (Aldeia 1983:33
            1s-wait
            'I'm waiting for him'

comitative (*era-*)
10.   če-r-**era**-čɨp   (Aldeia 1983:40)
            1s-cm-**con**-go=down
            'they lowered me with them'

causative (*ma-*)
11.   ã-**ma**-petɨwam   (Aldeia 1983:39)
            1s-**caus**-pipe
            'I make a pipe'

12.   ã-**ma**-kane?õ   (Aldeia 1983:40)
            1s-**caus**-tired
            'I make them tired'

13.   ã-**ma**-te?omat   (Aldeia 1983:40)
            1s-**caus**-tired
            'I make them work'

Kb

14.   ?ŋa i-apekũ me ?ɨɨpo munep-a   (Dobson 1988:14)
            3ms indf-gills in cipó put-tn
            'he put the *cipó* in the gills.'

15.   a?erauwe ye tey-u?ɨw-a r-eka-a   (Dobson 1988:13)
            conj   1s 1sr-arrow-mn obj-seek-tn
            'after that happened, I looked for my arrow'

comitative (*eru-*)

16.   yapepoa ye a-**eru**-ʔat   (Dobson 1988:30)
             jug   1s 1s-**com**-fall
             'I fell with the jug'

causative (*mo-*)

17.   a-ekoʔwoɣ ete noko ye i-**mo**- no-w   iɨ̃ pɨpe yepi ʔya
             1s-pour=out emph hab 1s 3-**caus**-go-tn container in hab voc
             'I always pour them out in the container.'   (Dobson 1988:13)

OG

18.   Peru soʔó o-ʔú   (RM 1640a:35)
             Peter meat 3-eat
             'Peter ate the meat'

comitative (*ero-*)

19.   a-**ro**-paɣ če-kaneʔõ (RM 1639:343v)
             1s-**con**-wake 1s-tired
             'I woke up with my tiredness'

causative (*mbo-*)

        a-mo-ŋgarú   (RM 1639:223v)
        1s-caus-eat
        'I made them eat'

Tb

comitative (*ero-*)

20.   a-ro-kér še-r-air-a   (A 1595:49v)
             1s-com-sleep 1s-cm-son
             'I sleep with my son'

causative (*mbo-*)

21.   še-mo-motár Tupã   (A 1595:49)
             1s-caus-want God
             'God makes me want Him.'

22.   a-mo-pʷám   (A 1595:49)
          1s-caus-arise
          'I make it rise.'

W

comitative-causative (*ero-*)
23.   a-**ro**-ʔa   (J 1990a:109)
          1s-**com**-fall
          'I made him fall with me'

          o-**ero**-ʔa   (J 1990a:109)
          3-**com**-fall
          'he made them fall with him'

causative (*mbo-*)
24.   o-mo-ḗ   (J 1990a:108)
          caus-leave
          'he makes them leave'

25.   o-mo-yaw   (J 1990a:108)
          caus-bathe
          'she bathes them'

AT

comitative (*ero-*)
26.   o-ero-čerem   (Nicholson 1982:19)
          3-com-let=go
          'he let go with it'

causative (*mo-*)
27.   o-**mo**-kačɨm   (Nicholson 1982:66)
          3-caus-see
          'she made it lost'

28.   o-w-ahɨ   (Nicholson 1978:36)
          3-caus-sick
          'it made him sick'

AX

comitative (*ero-*)

29.   o-romoi-žerep   (Nicholson 1982:19)
          3-com-let=go
          'he let go with it'

causative (*mo-*)

30.   a-mo-tɨwɨn   (Nicholson 1982:19)
          1s-caus-dry
          'I dried it'

In comparing derived verbs in TG, we find that there is almost complete consistency between them, and we can confidently reconstruct for PTG: **ero-* 'comitative' and **mbo-/* **o-* 'causative'. Class I roots took **mbo-* as the causative while Class II roots used **o-*. The only digression from this is that W has modified the comitative prefix to function as a causative comitative prefix.

6.2. Reflexives, Passives, and Reciprocals

W

31.   o-yi-yuka kupa   'they killed themselves/one another'   (J 1990a:83)
      3-pass-kill pl

AT

32.   o-če-nopo   'he hit himself/was hit'   (Harrison 1975:102)
      3-pass-hit

33.   o-čo-nopo 'they hit each other'   (Harrison 1975:104)

UK

34.   a-yu-pukʷar 'I tied myself up'   (K 1986:340)
      1s-refl-tie=up.

35. yu-tuka-tuka 'he hit himself'   (K 1986:340)
    3=refl-hit-hit

Pt
36. ji-koγ   (Betts 1968:14)
    refl-put=down
    'put himself down'

    ñi-mbo'e   (Betts 1968:14)
    refl-teach
    'is taught'

    jo-pɨhɨγ   (Betts 1968:14)
    recip-catch
    'catch each other'

Tb
37. a-ye-yuká   (A 1595:35)
    1s-refl-yuka
    'I kill myself/am killed'

38. a-ye-'ú   (A 1595:35)
    1-refl-eat
    'I am eaten'

AX
39. o-žo-nopɨŋ   (Nicholson 1982:17)
    3-recip-beat
    'they beat each other'

40. o-žo-ye'eŋ-atɨ   (Nicholson 1982:17)
    3-recip-speak-intens
    'they speak loudly to each other'

Kb
41. ka'ia kĩã o-ye-ɨwu-ukat kĩã upe   (Dobson 1988:27)
    monkey 3ms 3-refl-shoot-caus 3ms by
    'the monkey let himself be shot by him'

42. o-yo-mo-φerap wã   (Dobson 1988:32)
    3-recip-caus-recover 3p
    'they cured each other'

43. o-ye-ɨwu kĩã (Dobson 1988:28)
    3-refl-shoot 3ms
    'he shot himself'

Gj
44. a-ze-cak   (B-S 1968:99)
    1s-refl/pass-see
    'I was born'

45. u-ze-maʔemupu   (B-S 1968:99)
    3-refl/pass-thingbeat
    'he played the drum'

Tp

46. a-če-čokã   (Aldeia 1983:40)
    3-refl-kill
    'he kills himself'

47. či-ča-čokã   (Aldeia 1983:41)
    1pn-recip-kill
    'we kill each other'

Gy

48. a-ñe-meʔẽ Tupã upe   (Hoeller 1932:89)
    1s-refl-give God to
    'I give myself to God'

49. a-ye-epʸa   (Hoeller 1932:89)
    1s-refl-see
    'I see myself'

50. a-ye-roya   (Hoeller 1932:89)
    1s-refl-trust
    'I trust myself/I hope'

51. oro-yo-aɨcu   (Hoeller 1932:89)
    1pe-recip-love
    'we (excl) love each other'

52.   o-ño-pɨtĩβoi   (Hoeller 1932:90)
3-recip-support
'they hold each other up'

53.   o-ño-nupã   (Hoeller 1932:90)
3-recip-beat
'they beat each other'

54.   o-yo-aɨcu or o-ye-aɨcu   (Hoeller 1932:90)
'they love each other'

55.   oro-ye-popɨci or oro-yo-popɨci   (Hoeller 1932:90)
'we offer our hand to one other'

OG

56.   a-ñe-nupã   (RM 1639:90)
1s-refl-beat
'I beat myself'

57.   a-ye-pok^{w}aá   (RM 1639:194)
1s-refl-make=used=to
'I get used to it'

Ch
58.   re-ye-yuka   (Dietrich 1992:102)
2s-refl-kill
'you kill yourself'

59.   ña-ñe-ñ ɨwo   (Dietrich 1992:102)
we-refl-shoot
'we shoot ourselves (with arrows)'

60.   o-ye-péa   (Dietrich 1992:102)
o-refl/pass-open
'it opened itself'/it was opened

61.   ya-yo-γ^{w}ɨráha   (Dietrich 1992:102)
1pi-recip-go=away
'we went away together'

62. yo-γ"éru   (Dietrich 1992:102)
    3=recip-carry
    'they carry each other'

63. yo-γ"ɨráha   (Dietrich 1992:102)
    3=recip-go=away
    'they went away together'

6.2.1. The dual function of -ye- in TG

Many TG languages use the same morpheme to mark both passive voice and reflexive action. We can arrive at one possible solution to the two types of clauses through an RG approach.[1]

64.   a- ye- yuka   'I kill myself'
                       1s-refl-teach
                       ERG

This could be explained as having the following RN:

---

[1] See note #1 in 5.4.2.2 above for brief discription of Relational Grammar. As mentioned there, in some clauses, such as passive sentences, for example, a constituent may have more than one relation (direct object with regard to theta role, subject with regard to case assignment). Usually, instead of representing this by multiple arcs, RG represents this by multiple strata on a single arrow.
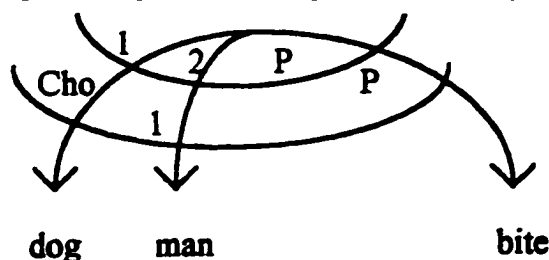
As the above figure shows, when a consituent bears the same relation on one stratum that is borne by another constituent on the immediately higher stratum, the first instantiation of that relation is cancelled or 'put on chomage', abbreviated Cho.

65.
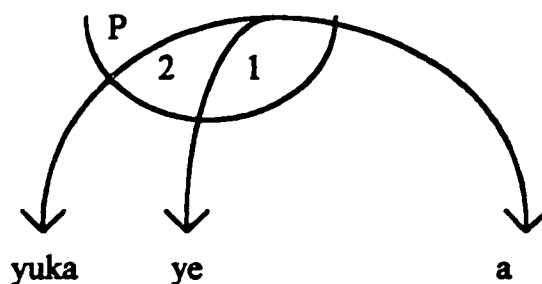
This analysis, however, fails to account for sentences like (66) below:

66.

o- ye-pe?a   čuā   rehe   'It is opened by John'
3- pass-open John by
ERG

There *ye-* is not used as a direct object and indeed *o-* cannot be nominative with the following RN:

67.
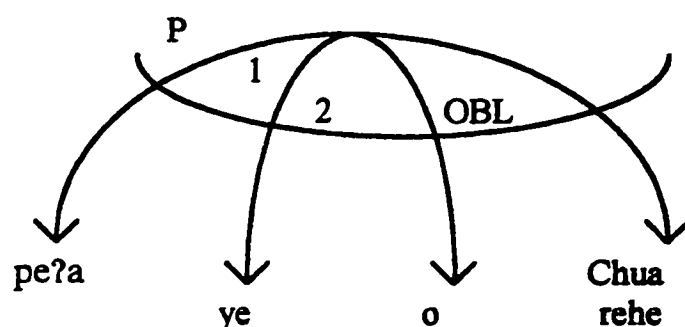
In order to be consistent with (65) above, *ye-* would have to head its own arc. Even if we somehow argue that *ye-* marks the agent of a passive verb while the nominal agent itself is oblique, we are left with two problems. First, why does *ye-* head a 1-arc in

some sentences (such as (67)) and a 2-arc in others (such as (65))? Second, and more troublesome, why is a final 2, which could only be marked as absolutive case, marked with ergative case? The problems with this analysis are so great we are compelled to reject it.

The central problem with the above analysis lies with the assumption that *ye-* bears its own relation. Another possible analysis sees *ye-* as marking a relational phenomenon and not heading an arc. In such an analysis the subject of (65) above would head not only the 1-arc, but also the 2-arc.

68.

yuka   a
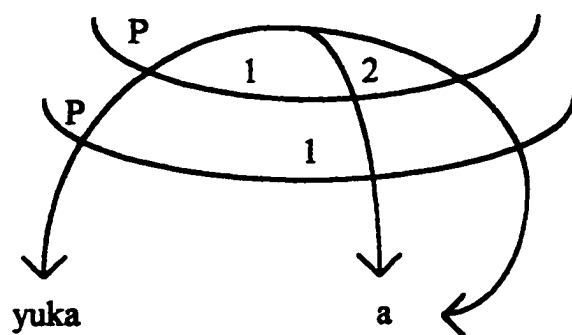
Here *a-* heads a 1-arc which has a 2 in the same stratum. *Ye-* then, marks the resolution of the multi-attachment. This would also explain the use of *ye-* with passive verbs as well as the use of nominative pronominals with passive verbs. A passive verb is not as obvious a case of multi-attachment as reflexivization, but it is possible if we propose retroherent advancement for the subject of a passive verb.

69.

pe?a   o
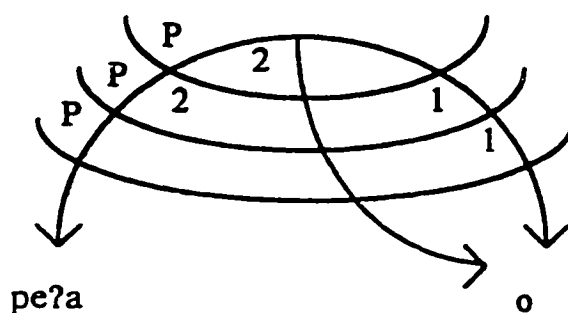
This analysis of *ye-* also is corroborated by the documented use of *ye-* when it cannot possibly show reflexivization or passivization. In early seventeenth century Guarani there was a verb *(r)esareko*, 'to meditate, ponder'. Because this was an intransitive verb in Guarani, one would expect the Set 2 marker to be used as a subject as in (70) below.

70.   če resareko nde rehe
             1s ponder   2s about
             ABS

The RN for such a sentence would be (71) below.

71.

resareko   nde   če
                        rehe
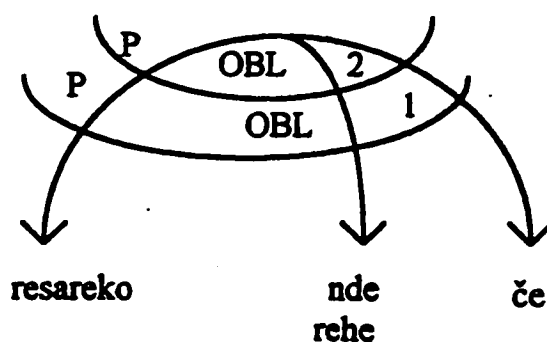
Use of the Set 2 marker is indeed attested. But there are also instances of using the nominative form of the pronominal marker. This might be a consequence of retroherent 2→1 advancement as in (72).

72.
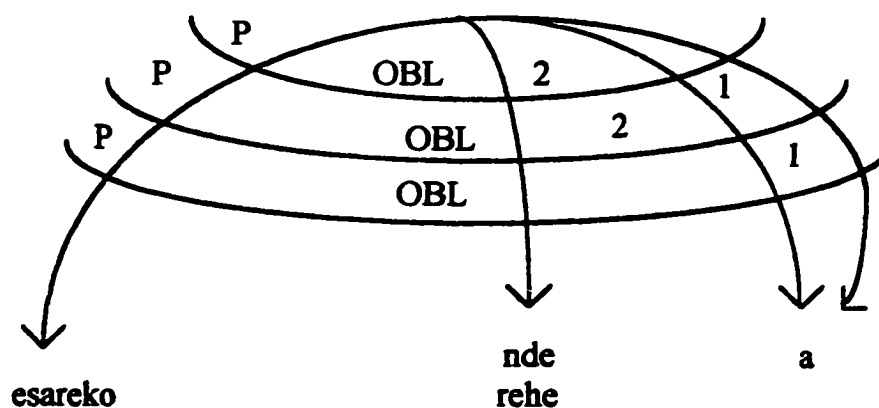
Such an advancement results in multi-attachment, which is then resolved in favor of the 1. But multi-attachment resolution is marked by *ye-* and that is precisely what is attested. When ergative pronouns are used, *ye-* is also used.

73.   a-ye-esareko nde rehe   (RM 1639:374v)
         1s-X-ponder 2s about
         'I ponder over you'

Note here that *ye-* does not, indeed cannot function as a passive or reflexive marker. This corroborates the analysis that *ye-* is the marker of a relational phenomenon: multi-attachment resolution. Based on the above cross-linguistic data, we can reconstruct *-ye-*

for PTG as a prefix denoting multi-attachment resolution, and *-yo- as a reciprocal prefix. In W, UK, and Gj, these have simplified into a single suffix, but each language has done so in a unique way: in W -y(i)- (< *-ye-) is either reflexive or reciprocal, Gj -z(e)- is either reflexive, passive, or middle voice, and UK yu- (< *-yo-) indicates passive generally, reciprocal voice with reduplicated verbs, never passive voice.

6.3. Noun Incorporation

Aside from deriving verb stems by means of verbalizing prefixes, Some TG languages also form complex verb stems by means of noun incorporation (NI). As a lexical process, NI is manifest in a number of TG languages.

Pt
74.   haβ-ekɨ-y   (Pease 1968:5)
      fur-remove
      'removes fur'

75.   pɨtupir-aho-paβ   (Pease 1968:5)
      heel=skin-cover-nom

76.   a-ji-po-er-u   (Pease 1968:5)
      1/-rfl-hand-com-come
      'I beckon'

Gj
77.   ciriŋ-ʔok ire   (B-S 1972:143)
      rubber-tap after
      'when they have finished tapping the rubber'

78.   u-mu-zuru-peka   (B-S 1972:98)
      3-caus-mouth-open
      'he opened it'

79. u-zuru-peka   (B-S 1972:100)
     3-mouth-open
     'he yawned'

Kb
80. poye   ye te-φuaka-paw   -amu   y-ui (Dobson 1988:20)
     suddenly 1s 1s-force-complete-tn indf-from
     'Suddenly I was tired of it.'

OG
81. a-yuru-peká   (RM 1639:267)
     1s-mouth-open
     'I opened his mouth'

Tp
82. ã-tɨro-patokã   (Aldeia 1983:67)
     1s-clothes-wash
     'I'm washing clothes'

Ch
83. añ-aka-asóy   (Dietrich 1990:304)
     1s-head-cover
     'I cover (my) head'

Gy
84. a-c-éko-kwáa   (Dietrich 1990:305)
     1s-3-nature-know
     'I know his nature'

NI also occurs to some extent in UK where it is attested in two frozen forms:

85. maniʔok pirok o-hɔ   (pirok < pir 'skin' + -ʔok 'take out') (K 1986:394)
     manioc 3=peel 3-go
     'he went to peel manioc'

86. i-ãkã rupi pokok   (pokok < po 'fist' + -kok 'touch') (K 1986:394)
     3-head along 3=hit=with=fist
     'he hit (him) around the head with his fist

While NI appears to be a lexical process in most TG languages, there is evidence that in Gy it may be more of a syntactic process:

87. a-rója Tũpa-ñẽe   (Dietrich 1990:306)
    1s-believe God-word
    'I believe in God's word'

88. a-tũpa-ñẽe-re-rója   (Dietrich 1990:306)
    1s-God-word-about-believe
    'I believe in God's word'

The incorporation here of, not just an object, but a postpositional phrase, is counter to the conventional definition of NI. Because of the otherwise conservative grammar and phonology of Gy, this suggests not only that NI is a syntactic process in Gy, at least in some instances, but also that it originates from a syntactic process PTG. To examine this issue further, we turn to evidence in OG.

6.3.1. Development of Noun Incorporation in OG

It has already been proposed that NI in Guarani is a relic of SOV verb phrases at an earlier stage of the language (see Pederson (1977) and Velazquez (1986)). While NI does clearly reflect SOV word order in OG verbs, however, it cannot be analysed as pronoun + object + verb. Rather, they are single words just as in modern Guarian languages, although NI did have more flexibility in OG.

There is one example in Ruiz de Montoya that may provide evidence for a proclitic status of person markers.

89.   a-mba'é-ri yeruré (RM 1639:53)
      1s-thing-to beg
      'I beg for things.'

Here a postpositional phrase intervened between the subject marker *a-* and the verb *jeruré*. Even here, however, the postpositional phrase was the complement of the verb and not an adverbial phrase such as *tetã-me* 'in the town' as in the Gy example above. No such phrase could ever intervene in OG. Moreover, the above sentence marked habitual action similar to (2).

90.   a-mba'é-apó (RM 1639:53)
      1s-thing-do
      'I work.' i.e. 'I do things.'

A-mba'é-ri jeruré is listed by Ruíz de Montoya as parallel to sentences like (3).

91.   a-poró-yuká (RM 1639:53)
      1s-people-kill
      'I have in me the practice of killing people'
      yo contengo en mi el exercicio de matarlos

Hence incorporating *mba'é-ri* was a device to make the verb habitual. No other noun + *ri* is attested as intervening between the subject marker and the verb in OG. Finally, constructions of the type in (1) are not mentioned at all in Restivo's grammar. So while the possible clitic status of subject markers may be attested in the earliest OG, the transition to NI was complete early in the OG period.

Nevertheless, the OG and Gy evidence point to a period when pronominal

markers were proclitics rather than prefixes, suggesting that verb complexes were actually more like verb phrases in PTG and that NI developed from lexicalization of what was the original SOV word order of such phrases.

Other evidence suggests that NI developed from the lexicalization of full verb phrases. According to Mithun (1984), there are certain characteristics of NI that hold across languages. In particular, all incorporated nouns have the following restriction: they must be unmarked for definiteness, although they may have a specific reference. So while a verb may incorporate nouns it may not incorporate articles, demonstratives, possessive pronouns, quantifiers, or adjectives. This restriction does not strictly apply to OG.

92. a. ex.: ko   a-ɨbɨ -eya (PR 1724:51)
          this 1s-land leave
          'I am leaving this land.'
          dexo esta tierra

     b. ex.: o-ara-sá irundɨ (PR 1724:51)
          3-day-pass four
          'He spent four days.'
          passó quatro dias

Note, however, that while the incorporated nouns are definite, the demonstrative in (a) and the number in (b) are not incorporated (and according to Restivo (1724), could not be).

Mithun and Velazquez both make reference to the fact that incorporated nouns are mostly limited to body parts, but Velazquez itemizes other classes of nouns which are incorporable as well: domestic animals, common man-made utensils, and things 'usually associated with human beings, e.g. sleep, mind, soul' (Velazqez, 1986:40). Virtually all these are attested in OG, but other nouns appear as well, for example, 'God', 'land', and 'word' as in (91).

93. a. ex.: a-Tupã-r-aɨhú   (PR 1724:51)
            1s-God-love
            'I love God.'
            amo á Diós

    b. ex.: a-ɨβɨ-ečá   (PR 1724:51)
            1s-land-see
            'I look at the land.'
            miro la tierra

The other restriction on incorporated nouns in Paraguayan Guarani is that no adjectives may be incorporated. This restriction does not apply in OG.

94. a. ex.: a-ɨβɨ pukú-ečá   (PR 1724:51)
             1s-land large see
             'I saw a large piece of land.'
             vi una tierra larga

In (93) below, note the parallel between the structure of a. and b. and that of c. and d.

95. a. ex.: čuã perú-yuká ramõ   (PR 1724:75)
             John Peter kill when
             'when John kills Peter'
             matando Juan á Pedro

95. b. ex.: perú čuã i-yuká ramõ (PR 1724:75)
         Peter John 3-kill when
         'when John kills Peter'
         matando Juan á Pedro

     c. ex.: a-Tupã-pisí   (PR 1724:50)
              1s-God-lay=hold=on
              'I lay hold on God.'
              comulgar
              [to receive Communion]

     d. ex.: a- i-pisí   Tupã   (PR 1724:50)
              1s-3-lay=hold=on God
              'I lay hold on God.'
              comulgar
              [to receive Communion]

Note also with Class II verbs:

96. a. ex.:   Perú Tupã-r-aihú-ramõ   (PR 1724:75)
              Peter God love when
              'when Peter loves God'
              amando Pedro á Diós

94 b. ex.: Tupã Pedro-h-aihú ramõ   (PR 1724:75)
            God Peter-3-love when
            'when Peter loves God'
            amando Pedro á Diós

    c. ex.: a-Tupã-r-aihú   (PR 1724:50)
            1s-God-love
            'I love God.'
            amo á Diós

    d. ex.: Tupã a-h-aihú   (PR 1724:50)
            God 1s-3-love
            'I love God.'
            amo á Diós

When the overt object is moved from the verb in a subordinate clause, it must be replaced with the third person marker *i-*, or, in the case of a Class II verb (94), *h-*.

Apparently the same pronoun substitution rule that applies in (93b) also applies in (93d), indicating that the underlying form for (93d) is (93c) and the underlying form for (94d) is (94c), just as the underlying form for (94b) is (94a). This lends support to the idea that pronominal markers such as *a-*, *ere-*, and so on, filled the same syntactic slot as nominal subjects. But clearly, these markers were not words; the obvious NI of the sentences in (91) gives equally clear evidence that pronominal markers were bound. Further, there is strong evidence that verbs with incorporated nouns were considered as a single unit. For example, the comitative prefix *ro-* could only be attached to intransitive verbs. Yet verbs like the one in (95) were possible.

97.　ex.:　a-ro- Tupā- pisí　　če-angaipá (PR 1724:62)
　　　　　1s-com-God-lay=hold=on 1s- sin
　　　　　'I receive communion with my sin.'
　　　　　comulgo en mi pecado

This sentence could have been acceptable only if *Tupāpisí* was not considered a verb phrase, but an intransitive verb.

It is clear then, that OG did use NI. It is equally clear that NI in OG used more syntactic categories (i.e. noun phrases as well as nouns) than NI in MG will allow. NI has become more restricted and more of a morphological process and less of a syntactic

process. This has been shown (Schleicher 1989) to be part of a broader change making clitics in general more closely bound to stems as affixes. We can reasonably, conclude, therefore, that cross-reference markers and other affixes in TG languages were clitics in PTG and that verb and noun complexes were phrases.

6.3.1.1 Subject Incorporation

We now have one more problem to consider: there is a class of verbs which continues to incorporate possessed nouns. For example, -pytuho, 'suffocate' is actually a compound of pytu, 'breath' and ho, 'go'. Since the pm3 pronominals are also used as possessives, še- pytu-ho may be considered at some stratum as 'my breath goes'.

98.   še-pɨtu-hó
     1s-breath-go
     'I suffocate'

99.   še-pɨ-r-ɨrɨ̄y
     1s-foot-cm-fast
     'I am speeding'

100.   še-pɨʔa-wapɨ
     1s-liver-sit
     'I am at peace'

Suppose we consider such an analysis. The RN of (98) would look like (101).

101.
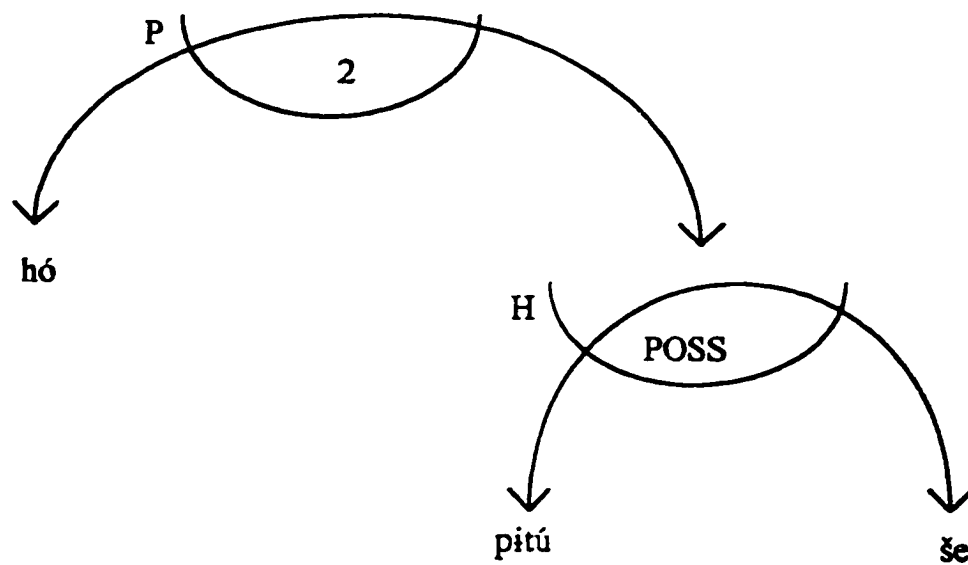
The possessive then advances to a 2 in (102) below (Object Raising):

102.
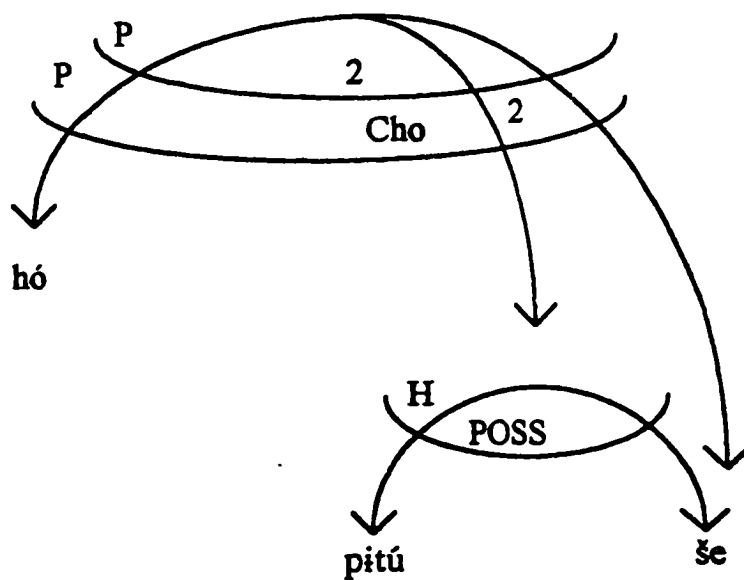
The object raising of the possessor puts the head on chômage resulting in the

incorporation of the head into the verb.[2] Since, in Guarani, only initial 2s advance,[3] an

---

[2] Since the head is incorporated as a 2-Cho, this is really not subject incorporation, but a kind of object incorporation. The other kind of object incorporation found in Guarani is where the object is not the head of a possessor in any stratum:

a   yuka   kure
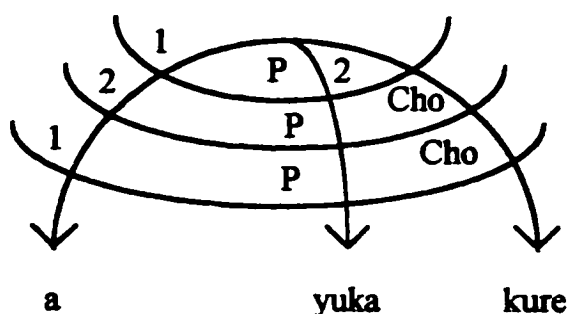
= a-kure-yuka   "I kill pigs."

Here *kure* is put on chômage by an antipassive construction. The phenomenon of a 2-chômeur being incorporated into the verb has been found before, in Hebrew (Dubinsky, 1989:9). There are differences between the two cases also. In Hebrew only indefinite 2 trigger antipassive and are then put on chômage and incorporated. In Guarani, however, indefinite 2s *optionally* trigger antipassive and must then be incorporated. A further difference is that certain 2s which are the head of a possessor phrase *may* also be incorporated by object raising of the possessor which puts the head on chômage.

[3] Independent evidence exists from subordinate clauses that can't function as subjects.
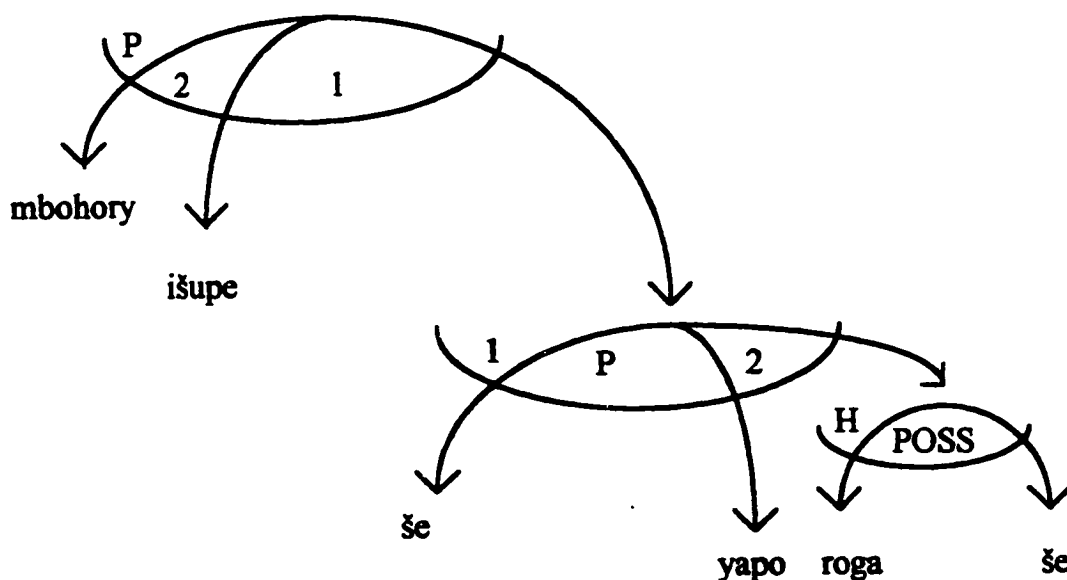
= *o-mbohory išupe še yapo-ha še roga
"That I build my house makes them happy."

expletive must be generated to provide a final 1.

103.

If non-initial 2s could advance, sentences like the above should be acceptable; they are not because of the initial 2 raising rule in Guarani. Generating an expletive won't make the sentence acceptable because clauses are not morphologically marked as 2s; pronouns are. In other words, there is no overt morphological distinction
between a subordinate clause raised to a 1 (unacceptable), and a clause with a 2 relation accompanied by a phonologically null expletive. In the case of (103), however, the accusative marking of *še* makes the presence of the expletive apparent.

The final 2 generates the accusative pronominal *še-*.[4]

A.

Likewise (73) would not have the RN shown in (72) but rather the RN in (B) below:

B.
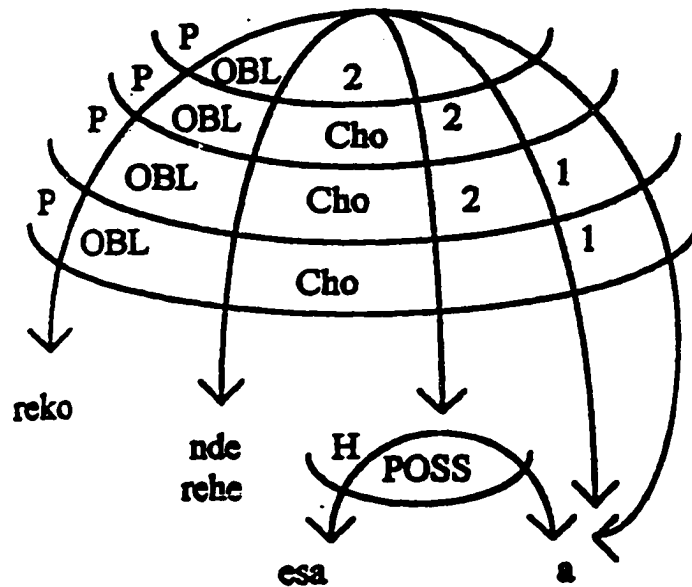
A problem remains in this hypothesis, however. The expletive is a final 1. For first and second person objects, third person subjects are not realized phonetically.

104.   še-nupā
             1s-beat
             'he beats me'

A third person subject is phonetically realized, however, when there is a third person object.

105.   o-i-kuaa
             3-3-know
             'he knows it'

We would therefore expect verbs like (93-97) to be like (106) in the third person, but they are not. Instead they are like (107).

106.   §o-i-pytuho
             3-3-breathe

107.   i-pytuho
             3-breathe

We can still explain the absence of a pronominal marker for the expletive if we suppose some very reasonable characteristics about expletives in Guarani. I propose that expletives in Guarani have no features of person or number. Since it has no feature for person, it can never be third person (or any other person). Therefore it cannot generate a third person pronominal (or any other pronominal). This gives us *če-pituhó* and for

third person *i-pituhó*.

6.4. Negation

In many attested TG languages we find that there is one morpheme for negating verbs (the ambifix *na- -i*) and one for negating nouns (*eʔim*). In Tb, however, we find that there was no absolute restriction of the application of these two negatives applied.

6.4.1. OG Negatives

For indicative verbs in independent VP's negation was formed by *nd(a)* + subject marker + verb + *-i*. Below the verb *atá* 'walk' is used to illustrate the paradigm.

	SG.	PL.
1	nd-a-ata-i	ndi-ja-ata-i
		nd-oro-ata-i
2	nd-ere-ata-i	nda-pe-ata-i
3		nd-o-ata-i

Only *ndi-ja-* and *nda-pe-* show *nd-* + a vowel. But *nda-* is used with pronouns (e.g. *na-če-porāŋ-i*) so *nda-* seems to be the underlying form. What is more, *j-* was apparently originally a glide connecting *i-* and an oral vowel (e.g. *i-* 'his' + *aó* 'clothing' = *ijaó*); so either *ja-* was originally *ija-*, or *nda-* assimilated to *ja-* by becoming *ndi-*.

Restivo attests that in some towns the negative suffix was *-iri* instead of *-i*, ex. *n-a-mboe-iri*, 'I'm not teaching'.

For verbs with a marked tense or mood (e.g. imperatives, future, permissives), OG had special negative markers. For the future tense the form was: *nd* + subject marker + verb + *i-če* (or *i-se*) + *ne*.

108. ex.:   nd-a-haɨhú-iče-ne
                neg-1s-love-neg-fut.
                'I won't love him' (PR 1724:45)

According to Restivo, *-iče* can also be used with other particles such as:

109. tamō raʔé: ex. nd-a-haɨhú-iče tamō ra'é   (PR 1724:45)
                    neg-1s-3=love-neg want emph
                    'I wish I didn't love him'

110. amo: ex: nd-a-haɨhu-iče amō   (PR 1724:45)
                    neg-1s-3=love-neg then
                    'I would not love, or would not have loved him'

111. -βeē amō: ex.: nd-a-haɨhu-iče-bē-ē amō   (PR 1724:45)
                    neg-1s-3=love-neg-since-after then
                    (translation same as (110))

These seem to indicate that -iče could be used with a variety of particles. But only with

the future tense marker ne was it obligatory; in all other cases, -iče is optional.

For imperatives and permissives the form was: verb + ime or eme (both

apparently from eʔime, 'without').

112. a. ex.: e-mboʔé ɨmē   (RM 1640a:14)
             sg=imp-teach neg
             'don't teach'

     b.   t-o-mboʔé ɨmē   (RM 1640a:14)
             hort-3-teach neg
             'don't let him teach'

Unmarked verbs in subordinate clauses generally took eĩ(m) after the verb.

113. a. ex.: a-haɨhú eʔĩ ramō   (PR 1724:45)
             1s-love neg if/when
             'if/when I do not love him'

113. b. ex.: a-mbo?éỹ-ramō (or) a-mbo?é eĩmãmō   (RM 1640a:14)
        1s-teach neg-if 1s-teach neg=if
        'as I don't teach', or 'not teaching' or 'I wouldn't have taught'

In the latter example, the initial *r*- of *ramō* has contracted with *e?ím*.

Restivo also attests the use of *nd*- + verb + -*i* with a subordinator (i.e. *ramō*).

114. ex.: nd-a-i-potar-i ramō   (PR 1724:46)
        neg-1s-3-want-neg if
        'if I don't want it'

When *ramō* was used to indicate cause, it could take *ruŋ^{w}ã*.

115. ex.: na-če-raɨhú-ramō-ruŋ^{w}ã-y   (PR 1724:46)
        neg-1s-love cause affirm-neg
        'because of my not loving him'

The word *ruwã* or *ruŋ^{w}ã* is attested independently of negative constructions as meaning 'perhaps'.

116. ere-yú ruŋ^{w}ã paŋã?   (RM 1639:346)
      2s-come perhaps int
      'did you come, by chance?'

117. če-ruŋ^{w}ã če-poraβɨkɨ́, nde aβé e-poraβɨkɨ́   (RM 1639:346)
      1s-perhaps 1s-work, 2s also 2s=imp-work
      'since I work, you work also'

It could also be used with nominalized verbs

118. na-o-ho-βa?e ruŋ^{w}ã-y če   (RM 1639:346)
      neg-3-go-rel neg-neg I
      'I am not the one who went'

It could be used on regular verb complexes as well.

119. nd-a-h-aihú-ruŋʷã-y   (RM 1639:346v)
      neg-1s-3-love-neg-neg
      'I don't love him'

120. nd-a-karú-ruŋʷã-y   (RM 1639:346v)
      neg-1s-eat-neg-neg
      'I am not eating'

6.4.2. Negation in (modern) Paraguayan Guarani

The negative markers for Paraguayan Guarani verbs are as follows. For non-future verbs the markers are *nd(V)-* + verb + *-i*, but because of the change in the personal prefixes, the system seems somewhat irregular (again we use the verb *watá*, 'walk', for illustration):

	SG.	PL.
1	nd-a-wata-i	nda-ja-wata-i(incl.) ndo-ro-wata-i (excl.)
2	nde-re-wata-i	nda-pe-wata-i
3	nd-o-wata-i	

Although the initial vowels of the personal pronouns have dropped elsewhere, they remain in the negative form of the verb, and so there now seem to be four variants of the negative prefix: *nd-*, *nda-*, *nde-*, and *ndo-*. Note also, however, that OG *ndi-ja-* has changed to *nda-ja-*, preventing a greater degree of irregularity.

For subordinate verbs, OG *eĩ* has become *-ʔĩ in Paraguayan Guarani*, but it is now more common to use the system given for main verbs. For example (121a) is more common than (121b):

121.   a. ex.: pe-je **nda**-pe-karu-i **-ha**
               2p-say **neg**-2p-eat-**neg** -comp
               'you say that you're not eating'

        b. ex.: pe-je **nda**-pe-karu-**ʔĩ-ha**
               2p-say **neg**-2p-eat-**neg-comp**
               'you say that you're not eating'

This is also true with relative clause/tense markers. Hence (122a) is prefered to (122b).

122.   a. ex.: **nd**-a-hecha-i *βaʔekue*
               neg-1s-see-**neg past**
               'I didn't see'

        b. ex.:**nd**-a-hecha-**ʔĩ** *βa'ekue*
               neg-1s-see-**neg past**
               'I didn't see'

The use of *-ʔĩ* with verbs is considered more elegant and poetic, and is very common in songs and poetry. It can also be used for a stronger negative effect than the *-i* suffix. All this confirms that *-ʔĩ* is considered more archaic than *-i* and is being replaced by *-i*.

Hortatory verbs still must use *-ʔĩ* (from OG *-ĩme*).

123.   ex.: t-o-guata-**ʔĩ**
             hor.-3-walk-**neg**
             'may he not walk'

6.4.3. Tupinamba

Tb had the same negativization as OG:

	ni-ya
n-a	n-oro
n-ere	na-pe
n-o-	

Tupinamba has the same basic verb complex structure as other languages.

neg-pron-H-post-neg

Both imperatives and optatives were negated by *ume*:

124. a. t-a-yuka ume   (A 1595:18)
       opt-1s-kill-neg
       'may I not kill it'

      b. t-ere-yuka ume   (A 1595:18)
       opt-2s-kill-neg
       'may you not kill it'

125. a. e-yuka ume   (A 1595:23v)
       2s=imp-kill-neg
       'don't kill it'

      b. nde-ño-ume e-yuka   (A 1595:23v)
       2s-only-neg 2s=imp-kill
       'don't kill it alone (i.e. by yourself)'

Both *e?im* and *ume* were free morphemes in Tb:

126.   yuka-awér-e?ima ~ yuka-e?im-awéra   'not a past killing' (A 1595:20v)
           kill-past-neg   kill-neg-past

       yuka-awām-e?ima ~ yuka-e?im-awāma 'not a future killing' (A 1595:20v)
           kill-fut-neg   kill-neg-fut

Tb could use both *eim* and *n- -i* for 'real' verbs:

127.   a-yuka-eʔim   (A 1595:35v)
                1s-kill-neg
                'I didn't kill him'

Anchieta does not record any semantic difference between a verb negated by *-eʔim*, on the one hand, and the same verb negated by *n-i* on the other.

Like *-če-* in OG, the form *-šo(e)-* appeared between a negated verb and various particles outside the verb complex.

128.   a-yuka-mo   (A 1595:19v)
         1s-kill-yet
         'I am/was yet to kill it'

129.   n-a-yuka-i-šoe-mo   (A 1595:19v)
         neg-1s-kill-neg-yet
         'I am/was not yet to kill it'

or   n-a-yuka-i-šo-mo

130.   a-yuka-meemo   (A 1595:19v)
         1s-kill-imperf
         'I was killing it'

131.   n-a-yuka-i-šo(e)-meemo   (A 1595:19v)
         neg-1s-kill-neg-imperf
         'I was not killing it'

132.   a-yuka-ne   (A 1595:18v)
         1s-kill-fut
         'I will kill it'

133. n-a-yuka-i-šo(e)-ne   (A 1595:18v)
     neg-1s-kill-neg- -fut
     'I won't kill it'

134. a-yuka-te-mo-mā   (A 1595:18)
     1s-kill-emph-yet-opt
     'oh, if only I (had) killed it'

135. n-a-yuka-i-šo(e)-te-mo-mā (A 1595:18)
     neg-1s-kill-neg-emph-yet-opt
     'oh, if only I had not/didn't kill it'

136. a-yuka-meī-mo-mā   (A 1595:18)
     1s-kill-opt-yet-opt
     'oh, if only I should kill/had killed it'

137. n-a-yuka-i-šo(e)-meī-mo-mā   (A 1595:18)
     neg-1s-kill-neg-opt-yet-opt
     'oh if only I should not kill/had not killed it'

Just as verbs can be negated with either *n-i* or *e?im*, nouns can be negated with *n-i* as well as with *e?im*. When they are negated with *n-i*, the meaning 'to be x' is implicit. When negated with yet another negative suffix, *ruŋʷāy* 'to have' is implied.

In Tb negating a noun complex with *n-i* can mean either 'to not be x' or 'to not have x'. In many cases, the meaning of the root distinguished which was implied.

138. še-aβaré   (A 1595:47)
     1s-priest
     'I am a priest' (§'I have a priest')

139. na-še-aβaré-i   (A 1595:47)
     neg-1s-priest-neg
     'I am not a priest'

140. še-pindá   (A 1595:48v)
      1s-fishhook
      'I have a fishhook' (§'I am a fishhook')

141. na-še-pindá-i   (A 1595:48)
      neg-1s-fishhook-neg
      'I don't have a fishhook'

Other roots would need context to disambiguate them.

142. še-rúβ
      1s-father
      'I am a father'~ 'I have a father'

143. na-še-rúβ-i
      neg-1s-father-neg
      'I am not a father' ~ 'I don't have a father'

Note that the null linking verb can be explicitly 'to be', both in the positive and in the

negative; in the positive by placing the pronoun after the noun, in the negative by

negating the complex with *n-(r)uā(-i)*.

144. kawára išé   (A 1595:47)   še-kawár   (A 1595:47)
      drinker I   1s-drinker
      'I am a drinker'

145. na-kawára-uā išé   (A 1595:47)   na-še-kawár-i   (A 1595:47)
      neg-drinker-neg I   neg-1s-drinker-neg
      'I am not a drinker'

146. aβaré išé   (A 1595:47v)   še-aβaré   (A 1595:47)
      priest I   1s-priest
      'I am a priest'

147. n-aβaré-ruā išé   (A 1595:47v)   na-še-aβaré-i   (A 1595:47)
      neg-priest-neg I   neg-1s-priest-neg
      'I am not a priest'

148.   moro-mboʔe-sára išé   (A 1595:47)   še-moro-mboʔe-sár (A 1595:47)
       people-teach-agnt I   1s-people-teach-agnt
       'I am a teacher'

149.   na-moromboʔesára-ruã išé (A 1595:47)   na-še-moromboʔesar-i (A 1595:47)
       neg-people-teach-agnt-neg I   neg-1s-people-teach-agnt-neg
       'I am not a teacher'

In Tb, a noun or pronoun negativized with *ruã* had the effect of relativizing a verb

complex.

150.   na-iše-ruã a-só   (A 1595:48v)
       neg-1s-neg 1s-go
       'I'm not the one who is going'

151.   na-Pedro-ruã a-yuká   (A 1595:48v)
       neg Peter-neg 1s-kill
       'Peter is not the one I killed'

152.   na-mbaʔé-ʔú-potá-ruã a-yúr   (A 1595:48v)
       neg-thing-eat-want-neg 1s-come
       'I come, not because I want to eat'

This form of negation could be used with postpositional phrases

153.   na še-r-úβa-supe ruã a-i-meʔéŋ   (A 1595:48v)
       neg 1s-cm-father-to neg 1s-3-give
       'it's not my father that I gave it to'

154.   na-iše-só-reme-ruã t-úr-i   (A 1595:48v)
       neg-1s-go-rel-neg cm-come-foc
       'he came (but) not because I'm the one who went'
       (more literally, 'not because I'm the one who went did he come')

Note that *iše* is never used as a subject or possessive, but an independent pronoun in

equational clauses. That means that *só-reme* is used here as a noun. Most likely *reme*

is not functioning as a relative marker; that would make complexes like *še-só-reme* to mean 'my one who went'. It must be a locative of some sort; so *iše só-reme* 'I am in the act of going', *še-só-reme* 'in the act of my going'. Understanding the above example confirms that a verb root + *reme* was treated as a noun or pospositional complex.

Finally, *ruā* could be used in conjunction with *eʔim*.

155. na-še-só-eʔime-ruā   (A 1595:48v)
      neg-1s-go-neg-neg
      'not because I didn't go'

Anchieta's grammar does not discuss *ruā* outside of its use as a negator.

6.4.4. Kayabi

	na-ya-
n-a-	n-oro-
n-ere-	ne-pe
n-o-	

It is hard to know if the second plural form is *ne-pe-* or *n-epe-*; in any case, it preserves the original form of the 2p nominative marker **epe-*.

The negation of narrative verbs in Kb is the same as that historically used for nouns: namely, *-eʔem*. The syntax of this suffix is unusual, because its position depends on the shape of the verb root. If the root is vowel final, the negative suffix comes at the end of the complex.

156. ka'a pe te-w-aw-e'em   (Dobson 1988:47)
     forest to 1s-go-nar-neg
     'I didn't go to the forest'

Otherwise, the negative suffix is followed by the narrative verb marker:

157. w-ipɨwɨɣ-e'em-amu (Dobson 1988:48)
     3-afundar-neg-nar
     'He didn't sink'

158. i-kutuɣ-e'em-a   (Dobson 1988:48)
     3-bore-neg-nar
     'It was not bored'

The consistent phonologically conditioned difference points to the idea that the narrative suffix (originally a nominalizing suffix) and the negative suffix were freer morphemes than they are now. Originally both *e'ɨm-a* and *e'ɨm-áβ-βo* must have been possibly as also *-á-eɨm* and *-áβ-βo-eɨm*. Sometime in the development of Kb from PTG, *-áβ-βo* became *-aw*. At that point, Kb seemed to prefer avoiding the occurrence of two adjacent vowels, so between *e'em-a* and *-á-e'em* speakers preferred **e'em-a*. Explaining cases like *wɨpɨwɨye'emamu* is easier. Technically *-ɨpɨwɨɣ* is a noun root and doesn't need to be nominalized. This also explains why the 'narrative' form uses a different suffix than other verb roots do: the narrative suffix for active verbs is historically a nominalizer, while the narrative suffix for stative/descriptive verbs is the same even in Kb as the postposition added to noun roots: *-(r)amū* < PTG *-ramō*. Originally, then, *-(r)amū* was (and perhaps still is) not a narrative suffix filling the 'post' slot on the verb (or noun) complex, but a postposition and the head of a larger complex, the postpositional complex, within which could be a negated noun:

159.　w-ipiwiɣ-eʔem amũ (Dobson 1988:48)
　　　3-sinking-neg　as
　　　Gen-Head-Neg　Postp

　　　Noun complex　Head

　　　　　Postpositional Phrase

6.4.5. Guajajara

Unlike most TG languages, Gj doesn't have a negative ambifix: *na-* can occur with *-i* or

kʷaw, 'neg.', but it sometimes does not:

160.　na-pe-ker　'you didn't sleep'　(B-S 1972:86)
　　　neg-2p-sleep

Imperatives are negated by *-(z)o*, a postverbal suffix appearing in the same slot as *-i* of

the negative ambifix, and as in AX, used with imperatives.

161.　pe-muew-kar-o　　(B-S 1972:95)
　　　2p-go=out-caus-neg
　　　'don't make it go out'

This suffix seems to be cognate with Tb *-i-šo(e)* and OG *-i-če*.　Formally, they go

together: all would point back to PTG *-i-co(e)* ([ičo]~[ič(o)we]).　But while in Tb and

OG these are the last part of the negative ambifix plus a suffix of uncertain function that

occurs between a negated verb and an adverbial particle, in Gj these have fused into

-ico > *-io* > *-yo* > *zo* and now this functions specifically to negate imperatives.

6.4.6. Assurini of the Xingu

The structure of verb complexes differs from those in other TG languages in that two forms of negation *n(e)- -i* and *i?im* are used interchangeably.

162. n-a-koaw-i   (Nicholson 1982:10)
     neg-1s-know-neg
     'I don't know it'

163. a-kohaw-i?im[5]   (Nicholson 1982:10)
     1s-know-neg
     'I don't know it'

As we have seen, *i?im* or its cognate is used in other languages to negate noun complexes, including 'serial verbs' or 'narrative verbs'.

The negating morphemes *n(e)- -i* and *i?im* are used more or less interchangeably, but the ambifix *n(e)- -i* is more common.

164. a-kohap 'I know it'   n-a-koaw-i 'I don't know it'   (Nicholson 1982:10)
     1s-know   neg-1s-know-neg

                         a-kohaw-i?im 'I don't know it'
                         '1s-know-neg'

165. pende-ti?ara 'your are hungry'   ne-pende-ti?ar-i 'you are not hungry'
     2p-hunger   neg-2p-hunger-neg (Nicholson 1982:10)

Note that here, as in Kb, the original initial vowel of *epe* has been preserved in the negative.

---

[5] The difference between *koaw* and *kohaw* is merely one of free variation.

In some cases *d(e)- -i* has been attested for *n(e)- -i.*

166. d-o-paw-i   (Nicholson 1982:10)
     neg-o-finish-neg
     he didn't finish

167. d-ur-i   (Nicholson 1982:10)
     he didn't come

168. d-u-potar-i   (Nicholson 1982:10)
     neg-3-want-neg
     he doesn't want it

169. d-akow-i   (Nicholson 1982:10)
     neg-hot-neg
     it's not hot

The imperative is negated by the suffix *-o/-(i)ñō*:

-o/ C_
-(i)ñō / V_

170. e-mamar-o   (Nicholson 1982:12)
     2s-throw-neg
     'don't throw it'

171. e-džoka-iñō (Nicholson 1982:12)
     2s-kill-neg
     don't kill it!

172. e-k^{w}a-ñō   (Nicholson 1982:12)
     2s-pass-neg
     don't pass!

173. e-ʔo-ñō   (Nicholson 1982:12)
     2s-eat-neg
     don't eat it!

174. e-žar-o   (Nicholson 1982:12)
     2s-come-neg
     don't come!

175.　e-kir-o　　　(Nicholson 1982:12)
　　　2s-sleep-neg
　　　don't sleep!

Like other languages, AX can interpret a noun phrase as a verb phrase in two ways as

'to have x' or 'to be x' and negates the one with *-i?ima* and the other with *n(e)-i* or *d(e)-i*

176.　dže-rerak^war-i?ima　　　de-dže-rerak^war-i　　(Nicholson 1982:11)
　　　1s-husband-neg　　　　　　neg-1s-husband-neg
　　　'I don't have a husband'　　'I'm not a husband'

6.4.7. Urubu-Kaapor

In UK all negation is expressed by *im* (< PTG *e?im*), whether for verb or noun

complexes.

177.　u-hɨk　　we　im　rī　　　(K 1986:358)
　　　3-arrive little neg still
　　　'he hasn't arrived yet'

178.　sawa?e　im　　(K 1986:358)
　　　man　　neg
　　　'he is not a man'

179.　e-raho　im　　(K 1986:358)
　　　2s=imp-take neg
　　　'don't take him'

180.　pe riki　a?i　　　ākā　im wapɨk u-ī　(K 1986:358)
　　　and so　old=woman head neg 3=sit 3-be=sitting
　　　'and so the old woman without a head was sitting'

181.　paite　im　　(K 1986:358)
　　　far　　neg
　　　'it is not far'

This use of *im* for all negation seems to be further evidence that **e?im* was used for

both verbs and nouns in PTG. In UK, the alternation between the *nd-i* and *e?ím* was simply resolved in favor of *e?ím*. The ambifix *nd-i* survives only in the expression *nišoy* (< *n-i* 'neg' + išo 'be' (< PTG *iko*)) 'there is none'.

182.   ihẽ rakehar ɨm   (K 1986:334)
       I   wife   without
       I don't have a wife

183.   eha ɨm riki awa   (K 1986:334)
       eye neg emph people
       People were without sight.

6.4.8. Akwawa

In AT, as in Tb, OG, and AX, reflexes of both *nd-i* and *e?ím* are used to negate verbs.

184.   a-ha-ɨ?ɨm   (Nicholson 1978:13)
       3-go-neg
       'he didn't go/isn't going'

185.   n-a-ha-ihi   (Nicholson 1978:13)
       neg-3-go-neg
       'he didn't go/isn't going'

Note that the second element in the ambifix, -*ihi* is lengthened or expanded in comparison to -*i* in other languages. OG had both *nd-i* and *nd-iri*, but AT -*ihi* and OG -*iri* are clearly unrelated. In AT, *n-ihi* is also used for stronger emphasis; (185) above could be translated, 'he never went'. It is also the only form of negation used with the adverbial partical *we*.

186. n-a-ha-ihi we   (Nicholson 1978:13)
    neg-3-go-neg more
    'he didn't go yet'

Since no special connotations are associated with *i?im*, this seems to be the unmarked

form of negation for verbs.

Imperatives are negated with *-eme*.

187. e-čar-eme   (Nicholson 1978:15)
    2s=imp-go-neg
    'don't go!'

6.4.9. Tapirape

In Tp, verbs are negated exclusively with *nā-i*, but nouns can be negated with

either *nā-i* or *-e?ima* below is the verb *piik* 'to catch' negated by *nā-i*.

	Singular	Plural
1n		nã-čã-piik-i
1x	n-ã-piik-i	n-ara-piik-i
2	n-ere-piik-i	nã-pe-piik-i
3	n-o-piik-i	

As in other TG languages, both *n-i* and *e?im* can be used to negate nouns, but which

form of negation is used determines whether the noun complex means 'to be x' or 'to

have x'. The former is referred to by Almeida as a 'verbalized noun' (all examples below

from Almeida (1983:29)).

	'have'	'be'
188.	če-r-ek^w-āwa 1s-cm-be-nom 'I have a life'	če-r-ek^w-am (ie-) 1s-cm-be-nom 'I am a living being'
189.	če-r-ek^w-āw-eʔɨma 1s-cm-be-nom-neg '	nā-če-r-ek^w-āw-i neg-1s-cm-be-nom-neg
190.	če-r-ek^w-āw-era 1s-cm-be-nom-past	če-r-ek^w-āw-era 1s-cm-be-nom-past
191.	če-r-ek^w-āw-eʔɨ-wera 1s-cm-be-nom-neg-past	nā-če-r-ek^w-āw-er-i neg-1s-cm-be-nom-past-neg
192.	če-r-ek^w-ā-rɨma 1s-cm-be-nom-fut	če-r-ek^w-ā-rɨma 1s-cm-be-nom-fut
193.	če-r-ek^w-āw-eʔɨ-rɨma 1s-cm-be-nom-neg-fut	nā-če-r-ek^w-ā-rɨm-i neg-1s-cm-be-nom-fut-neg

6.4.10. Wayampi

The primary form of negation in **W** is *n-i*.

	na-ya-yuka-y
n-a-yuka-y	n-oro-yuka-y
n-ere-yuka-y	na-pe-yuka-y
	n-o-yuka-y

In Wayampi *eʔɨm* is no longer a productive negator of nouns, but survives in the suffix

-*eʔɨma* (< *eʔɨm + *-áβ), 'place where there is no x'.

194.	ɨ tɨpɨ-eʔɨma	(J 1990a:112)
	water spring-place=without	
	'place without a source of water'	

The suffix *eʔɨ* is productive as the negation of optative verbs.

195. t-o-ʔu-eʔɨ   (J 1990a:112)
     opt-3-eat-neg
     'so that he may not eat'

196. t-o-manũ-eʔɨ   (J 1990a:112)
     opt-3-die-neg
     'so that he may not die'

6.4.11 Parintintin

In the case of 'real' verbs the 'pre-' and 'post' slots can be equated with the negative ambifix (shown below with the verb *ho* 'to go').

```
                    nda-ča-ho-i
   nd-a-ha-i        nd-oro-ho-i
   nd-ere-ho-i      nde-pe-ho-i
            nd-o-ho-i
```

Notice that the 2p form, *ndepehoi*, preserves the original form of the 2p pronominal, *epe*.

Nouns are negated by *-eʔim*.

197. ɨ-eʔim   (Pease 1968:43)
     water-neg
     'without water'

198. kaʔaruɣ-eʔim   (Pease 1968:43)
     late=afternoon-neg
     'not late in the afternoon'

Like other TG languages, Pt has an altered form of *eʔim* for imperatives and optatives.

199.   t-o-ho-ɨme   (Pease 1968:16)
        perm-3-go-neg
        'let them not go'/'they don't want to go'

The negative suffix here is *-ɨme* as it would be if the subject were first person. But if the subject is 2nd person, *-i* is used:

200.   t-ere-ho-i   (Pease 1968:19)
        perm-2s-go-neg
        'may you not go'

This use of *-i* for the negative also holds when the object of a transitive verb is 2s:

201.   ta-nde-juka-i   (Pease 1968:18)
        perm-2s-kill-neg
        'don't let him kill you'

Since *-i* can occur independently of *nd-* in Pt, while *nd-* and *-i* always co-occur in most TG languages, the ambifix *nd--i* has to be analyzed as two affixes: a prefix *nd-* and a suffix.

6.4.12 Guarayo

$$nda\text{-}ya\text{-}c\text{-}ep^ya\text{-}i$$
$$nd\text{-}a\text{-}c\text{-}ep^ya\text{-}i \qquad nd\text{-}oro\text{-}c\text{-}ep^ya\text{-}i$$
$$nd\text{-}ere\text{-}c\text{-}ep^ya\text{-}i \qquad nda\text{-}pe\text{-}c\text{-}ep^ya\text{-}i$$
$$nd\text{-}o\text{-}c\text{-}ep^ya\text{-}í$$

Unlike most TG languages, Gy uses *nd-i* and never *eɨ* to negate nouns.

202.   nda-če-maʔenduʔa-i (Hoeller 1932:95)
        neg-1s-remember-neg
        'I don't remember'

203. nda-če-repanaku-i   (Hoeller 1932:95)
     neg-1s-basket-neg
     'I don't have a basket, it is not my basket'

The negative *eɨ* is only used for nominalized verbs.

204. če i-mboʔe eɨ moce
     1s 3-teach neg when
     'when I do not instruct him'

205. a-mboʔe eɨ moce   (Hoeller 1932:96)
     1s-teach neg
     'when I didn't teach'

206. nd-a-mboʔe-i moce (Hoeller 1932:96)
     neg-1s-teach-neg
     'when I didn't teach'

The alternate forms of negating a when/if nominalization show how this construction is well on its way to being reinterpreted as a clause. The same can be seen with the relativizer, now usually interpreted as a relative clause and with passive nominalization in *-emi-*.

207.a. nd-o-c-ep^{y}a-i-βaʔe   (Hoeller 1932:96)
     neg-3-3-see-neg-rel
     'he wasn't the one who saw it'

   b. o-c-ep^{y}a-eɨ-βaʔe   (Hoeller 1932:96)
     3-3-see-neg-rel

   c. o-c-ep^{y}a-βaʔe-eɨ   (Hoeller 1932:96)
     3-3-see-rel-neg

208.a. če-r-emi-ep^{y}a-eɨ   (Hoeller 1932:96)
     1s-cm-pass-see-neg
     'the one not seen by me'

208 b. nda-če-r-emi-ep^ya-i   (Hoeller 1932:96)
    neg-1s-cm-pass-see-neg
    'the one not seen by me'

Hoeller (1932) points out that (207a) is the most common way to negate a relativized verb complex; (207b) is frequent; (207c), rare. However, the above examples do show a conservative feature: *eĩ* is free with respect to its ordering with nominalizing (or subordinating) suffixes.

209.a. c-ep^ya-car-eĩ   (Hoeller 1932:96)
    3-see-agnt-neg
    'one who didn't see him'

   b. c-ep^ya-eĩ-car   (Hoeller 1932:96)
    3-see-neg-agnt

210.a. c-ep^yia-pɨr-eĩ   (Hoeller 1932:96)
    3-see-ptnt-neg
    'one who wasn't seen'

   b. c-aɨcu-eĩ-pɨr   (Hoeller 1932:96)
    3-love-neg-ptnt
    'one who wasn't loved'

211.a. c-ep^ya-ca-eĩ   (Hoeller 1932:29)
    3-see-nom-neg
    'his not being seen'

   b. c-ep^ya-eĩ-ca   (Hoeller 1932:29)
    3-see-neg-nom

Imperatives are negated by *eme*.

212.   e-mboʔe-eme   (Hoeller 1932:96)
    2s-teach-neg
    'don't teach'

213.   pe-mbo?e-eme   (Hoeller 1932:96)
        2p-teach-neg
        'don't teach'

The future tense of the verb has its own negation: *nd-ičira*. This is because the future

tense, *-ra*, derives from one of the postverbal particles in PTG which was preceded by

**-c(o)*ʷ*e-* after a negative verb; in other words, *-iči-* is cognate with Tb *-iš̌o(e)-*, OG *-iče-*,

Gj *-zo*, and AX *-ō/-(i)ñō*.

6.4.13. Negation in PTG

Negation in TG generally follows the pattern: the ambifix *nd-i* for verbs, *-e?ím*

(or its equivalent) for nouns. We reconstruct *nd-i* as a unified morpheme for PTG

because it is in most TG languages including the oldest Tb and OG and most

conservative modern languages (Gy, Pt, GM, Ch). Where languages have *nd-* and *-i* not

always together (Pt, Gj, Kb) these are (clearly) demonstrable innovations. The negation

for imperatives was apparently either *e?ím* or a modified form of it: most use a modified

form of **e?ím*, but there is no uniform way of modifying **e?ím* into the imperative

negative (Tb *ume*, OG *eme/ime*, AT *-eme*, Gy *-eme*). Gj uses separate words *nən* and

ni?i that precede the noun to negate nouns. *ni?i* appears to be the concatentation of the

two elements of the ambifix **nd-i;* *nən* from **nd-* + *aani 'no'.[6]

---

[6] A similar construction is early modern Paraguayan Guarani *ahaniri* 'no' from *a(h)ani* 'no' + *(i)ri* 'not'
and modern Paraguayan Guarani *nahaniri* 'no' from *nd-ahaniri*.

We have discovered a heretofore undiscussed negative morpheme in PTG: *-c(o)ʷe-, used between the negative suffix -i and various modal and/or pragmatic particles outside the verb complex. Since this morpheme always followed -i (as far as we can tell) the initial c would have palatalized to č. The original meaning of *-c(o)ʷe- is unknown; its use seems purely syntactic, without contributing any meaning to either the verb complex which preceded it, or to the particles that followed it. Its reflex is attested in both Tb and OG. Only three modern languages attest it: Gy where it is used exclusively with the future tense suffix -ra; and Gj (zo) and AX (ō/(i)ñō) where it is used exclusively to negate imperatives (these are the only two languages that don't negate imperatives with some form of *eʔim).

In Tb and AX there is no functional distinction between nd-i and eʔim. Each could be used for nouns or non-imperative verbs. While this may be an independent innovation in the two languages, it does suggest that, if negative morphemes were functionally distinct, the distinction between noun and verb was not all that strict. Even if nd-i and eʔim weren't restricted to verbs and nouns respectively, this still leads to the conclusion that the distinction between nouns and verbs was not very clear in PTG: in our comparison between noun complexes and verb complexes, virtually the only consistent structural distinction between them is negation. If this was also relatively arbitrary, this leaves essentially no structural difference between noun complexes and verb complexes. We have already seen that there is no structural difference between noun root classes and verb root classes. Even the inventories of noun and verb roots

themselves are non-distinct: *-epʸák* means both 'see' and 'sight'; *-úr* means both 'come' and 'arrival'.

6.5. Comparing constituents of the verb complex

Having reconstructed verb complexes as verb phrases in PTG, we now turn to understand the constituent elements of verb complexes to determine the structure of verb phrases in PTG.

Phrase	Neg	+	Pron	+	Head	+	Post	
V	neg	+	pron	+	Noun	+	aspect	-neg
N	neg	+	pron	+	Noun	+		-neg

Piece	affix	Base	affix
V	pre-	V/N	
N		V	-suffix

When we look at the structure of the verb complex in various TG languages we see considerable similarity. What is even more significant, however, is the similarity within each language between verb complexes and noun complexes. In Gj, for example, verb and noun complexes are very parallel to one another. According to Bendor-Samuel (1972) the constituent structure of the verb complex is as follows:

Pre-verb + Pronominal prefix + Head + Post-verbal

The detailed explanation for these terms is given below. The constituent structure of the

noun complex is:

Pre-nominal + Genitive + Head + Post nominal

The pre-verbal and pre-nominal are both Bendor-Samuel's terms for verbal and nominal negation, respectively. 'Genitive' is manifest as a nominal possessor or parallel to the verb complex, by a pronominal prefix. The 'Post' slot is adverbial modifiers in the verb complex; in the noun complex it is filled by either post-nominal modifiers (modifiers used only with nouns) or post-verbals (modifiers used with either verbs or nouns) There are no modifiers used only for verbs. The 'Head' of the verb complex can be either a simple verb, a derived verb, or a noun; the 'Head' of a noun complex can be either a simple noun, a derived noun, pronoun, or demonstrative pronoun (called 'specifier' by Bendor-Samuel).

In addition to all this, both nouns and verbs have derived counterparts called 'pieces' by Bendor-Samuel. Each piece is made up of a root base plus an affix (a prefix for the verbal piece, a suffix for the nominal piece). The root for the verbal piece can be a verb root another verbal piece, or a noun root, or a demonstrative pronoun. The base of a nominal piece can only be a verb or verbal piece. These structures are diagrammed below:

Complex	V	Pre	Pron	H	Post
	N	Pre	Gen	H	Post

Piece	V	Prefix	V/Vp/N/dem	
	N		V/Vp	Suffix

We have seen that the structure of derived verbs is largely the same for nearly all TG languages: noun stems and verb stems can be prefixed with the reflexes of PTG *mbo-, *ero-, *ye-, and *yo-. And we have seen that derived nouns have a very similar structure in TG languages. In most TG languages, then, the above structure of derived verbs and nouns, and of verb and noun complexes, holds true. We can confidently then posit this structure for PTG complexes as well, or, more accurately for PTG verb and noun phrases, since we have seen in (6.2.1.) above, complexes of attested TG languages derive from PTG phrases.

The 'pre' slot is always negation except for the optative prefix t- attested in all TG languages, and because verbal negation is an ambifix n_i, which this scheme breaks up into a prefix 'Pre' and a suffix in the 'Post' slot, it seems more reasonable to leave negation in a separate slot, reanalyzing complexes as follows

V	Neg	Pron	H	Post	Neg
N	Neg	Pron	H	Post	

This reveals that negation is one of the few differences between noun complexes and verb complexes. The parallels are striking. Not only is there a remarkable parallel between the structure of verb complexes and noun complexes, but the heads of the

complexes themselves are virtually interchangeable: non-verbalized nouns can be the heads of verb phrases as well as noun phrases. Verbs, whether simple or derived, cannot head noun complexes in Gj unless nominalized by a nominalizing suffix. But verb roots could be the head of noun complexes in Tb, for example.

214. a-i-pota nde-so-awéra   (A 1595:27)
      1s-3-want 2s-go-past
      'I wanted you to go' ('I wanted your (former) going away')

Here nominal tense *awéra* is affixed directly to the verb root *só*, treating it as the head of a noun complex.

215. še-juka-rese   (A 1595:27v)
      1s-kill-about
      'because of the killing of me'

216. še-r-oriβ nde-só-rese   (A 1595:27)
      1s-cm-happy 2s-go-about
      'I'm happy about your going'

In these examples, postpositional phrases are constructed, as usual in Tb, by a postposition cliticized onto a noun complex; but the noun complexes in these examples are headed by verb roots, *yuká* 'kill' in the first example, and *só* 'go' in the second.

Similar parallels exist in other TG languages. In Kb the comparative structures look like this:

     V   neg-Pron-Head-Post-neg
     N   Gen-Head-Post-neg

Piece
      V   pre-V/N
      N   V/Vp-post

Tupinamba has a similar structure:

      V   neg-pron-Head-post-neg ~ pron-Head-post-neg
      N   Gen-Head-post-neg

The situation in Kb is more complicated because of the narrative form and focus forms of the verb. As was demonstrated in Chapter 5, these structures were actually nominalizations in PTG and are in most modern languages as well.

Just as nouns themselves have been attested as the head of a verb complex in Gj the same is true in Kb.

217.   na-way-i   (Dobson 1988:19)
              neg-tail-neg
              Neg-H-Neg
              'he doesn't have a tail'

We begin to see a common pattern emerging for the structure of verb and noun complexes in TG languages.

In Pt the structure of verb complexes also follows the pattern seen in Gj and Kb, but with even greater interchangeability of verbs and nouns.

      Neg-pre-pron-Head-post

As in Gj and Kb, the verb can be with or without a verbalizing prefix:

218.　ča-ño-mo-i?ĩ[7]　　　　　　'let's talk to each other'　(Pease 1968:16)
　　　let's-rec-caus-talk

```
          verb  N
             \/
       verbl   comp V
             \/
     pron  comp V
          \/
       verb complex
```
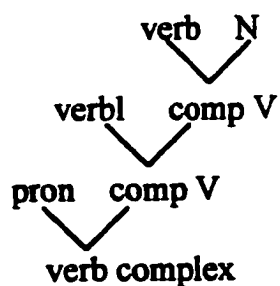
Transitive verb complexes have either a noun complex or an object pronoun before the verb stem:

219.　　　　ji-pi?a-kutukutu　　'piercing my liver'　(Pease 1968:17)
　　　　　　1s-liver-pierce

```
    noun complex  verb stem
               \/
        verb complex
```

They can also exhibit a noun incorporation:

220.　　　　o-pia-rupi?ri　　　'they held up their feet'　(Pease 1968:17)
　　　　　　3-foot-hold=up
　　　　　　s-noun-verb
　　　　　　　　stem

---

[7] *ča-* is from **t-ya-*.
　　perm-1pn

As in other TG languages, noun roots in Pt can serve as the head of a verb complex:

221. ñi-mbate   (Pease 1968:18)
    1s-baggage
    'I have baggage'

222. na-ne-mbater-i   (Pease 1968:18)
    neg-2s-baggage-neg
    'you don't have baggage'

223. nda-h-oriβ-i   (Pease 1968:3)
    neg-3-happiness-neg
    'they aren't happy'

Unlike many other TG languages however, Pt may also use verb stems as the head of noun complexes. Pease (1968) notes that one form of noun stem is a simple verb stem:

224. i-pɨ   (Pease 1968:31)
    cm-dance
    'the dance'

225. ŋã-ndur   (Pease 1968:31)
    3p-come
    'their coming'

Although this use of verbs stems as nouns seems to be rare, the conservatism of Pt makes this grammatical feature significant.

Note here the fact that 'stative' or 'descriptive' verbs in Pt are grammatically identical to nouns. Pease (1968) lists as the two simplest forms of descriptive verb stem, descriptive verb 'bases'. (root + 3poss. marker *i-* or *h-*) and noun 'bases':

226. i-katu   (Betts 1968:14)
     3-pretty

227. ñi-mbate   (Pease 1968:18)
     1s-baggage

228. ñ-akã   (Pease 1968:6)
     3-head

In AX we have less data to work with, and therefore, less information about the structure of verb and noun complexes. We do know the following:

AX does not have this nominalized form of the verb.

229. AT   a-ata we-ha 'I walk as I go'   (Nicholson 1982:13)
          1s-walk 1s-go

230. AX   a-ata a-ha 'I walk as I go'   (Nicholson 1982:13)
          1s-walk 1s-go

Apparently, these nominalizations have been reinterpreted as verbs.

This is also true of the 'focus' or 'OBTOP' nominalizations:

231. AT ipira-rehe i-ha-i 'they went for fish' (Nicholson 1982:13)
        fish-for   3-go-foc

232. AX ipira-rehe a-ha 'they went for fish'   (Nicholson 1982:13)
        fish-for   3-go

Nicholson (1982) gives the following examples of relic forms of these constructions in AX:

233.   ka-rehe itori (< i-tor-i)   (Nicholson 1982:13)
       then-about 3=come
       'she'll come then'

234.   o-čip ota   (< or-a)   (Nicholson 1982:13)
       3-descend coming
       'he came coming down'

235.   ipira-piikawa weraha imoʔayka   (Nicholson 1982:13)
       fish-net   taking tied=up-a
       'taking the fish net, he tied it up'

It is not clear whether these are frozen forms or, as in Gj, productive forms for a closed set of verbs. In any case, the general loss of the serial nominalization of verbs may be responsible for the application of both kinds of negativization to verbs. In other words, forms like *a-kohaw-iʔim* may be found in the same place syntatically where one would expect to find serial verbs; the distinction is still there, but, for the most part, only manifext when the verb is negated. Not surprisingly, AX has lost Set 3 pronominal markers at least as reflexive possissives in noun complexes.

6.6. Noun/Verb Identity

The picture of PTG that emerges suggests an identity between nouns and verbs, i.e. a language in which there is no grammatical distinction between nouns and verbs. Such a lack of a grammatical noun/verb distinction has been documented. In such cases, however, there is always some other grammmatical distinction which identifies the predicate of a clause. In Straits Salish, for example, the lack of a noun/verb distinction at the word level is 'compensated for' by marking the valence of each lexical root as either transitive or intransitive (Jelinek and Demers, 1994). Although there is one

grammatical distinction between nouns and verbs in PTG, Set 1 and Set 4 markers are only used with verbs, and when subjects are cross-referenced on verbs, only Set 1 or Set 4 markers are used; Set 3 markers are only used with nouns - This is actually similar to the situation in English where many roots may function as either verbs or nouns, the difference being marked grammatically only by the case of the pronoun(s) associated with them (and word order, not a crucial function in TG). However, PTG seems to resemble languages like Salish in having a fundamental grammatical distinction that seems to ignore the distinction between nouns and verbs. Consider the case of Kb.

In Kb we do seem to see a shift in the scope of reflexive pronouns that makes nominalized verbs seem more like verbs.

236. oroy-aʔɨr ʔar amũ kawĩpie apo-w oro-yo-upe   (Dobson 1988:86)
     1px-refl-son fall when mingau make-nar 1px-rec-to
     'when our sons were born, we made mingau for each other'

237. w-oɣ ipe ʔŋa o-y   (Dobson 1988:86)
     3r-house to 3ms go-foc
     'he went to his (own) house'

238. era oye-upe t-ur-ɨpɨ ramũ n-o-yemi-ʔuar-i   (Dobson 1988:87)
     news 3r-to poss-come-first when neg-3-eat-neg
     'they hadn't eaten when the news first came to them'

As a result of this expansion in scope, narrative verbs or serial verbs can refer to a declarative sentence far in advance of that sentence. There are examples parallel to the usual situation in other TG languages.

239.   o-o ye te-yawk-a   (Dobson 1988:103)
          1s-go 1s 1s-bathe-nar
          I'm going to take a bath

In Kb, however, every verb may be a narrative verb in a sentence.

240.   Yuʔi ʔŋa w-ipiwiɣ-amū a-w-aw a-wu-a ʔiri-piter ipe   (Dobson 1988:103)
       toad 3ms 3-sink-nar 3-go-nar 3-fluctuate-nar water-middle in
       'The toad sinking, left and appeared in the middle of the lake.'

If we interpret all narrative verbs in Kb as nouns, we should have to translate this as

'There was the sinking, going and appearing of the toad in the middle of the lake.' Such

an interpretation would be reasonable, and seems supported by the fact that these

'narrative verbs' are cross-referenced by possessive pronominal markers. Such an

interpretation becomes more strained, however, when examining passages like the

following:

241.   "E- yot ʔŋa r-esak-a ʔwe," ʔy-aw kīā ewmer-a   upe. "E- yot ʔŋa
       2s=imp-come 3ms -see-nar voc say-nar 3ms cadaver-nom to 2s=imp-come 3ms

       r- esak-a   ʔwe." Aʔe pe yer-aʔir-a kīā niapoy. Aʔe-ramū kīā o-roʔi-ramū.
       cm-see- nar voc   that in 1s-son-nom 3ms weak   that-for   3ms 3-fever-nar

       Aʔe-ramū te-yor - eʔem-a esak-a.   (Dobson 1988:93)
       that-for   1s-come-neg-nar see-nom

       '"Come see him," my husband said of the corpse. "Come see him." At this time
       my son was feeling ill. He had a fever. For that reason, I did not go see him (the
       corpse).'

If o-roʔi-ramū is a postpositional complex containing the noun complex o-roʔi and o- is

a reflexive possessor, while te-yor-eʔem-a is a noun complex and te- is a reflexive

possessor, then reflexive binding in Kb is strange indeed. The *o-* of *o-roʔi* could only refer back to *aʔir-a* 'son' while *te-* would have to refer to the *yer-* of *yer-aʔir-a*.[8]

In light of the identity between nouns and verbs, however, this may be a moot question: perhaps the use of 'noun' versus 'verb' complexes was originally a pragmatic, not grammatical distinction and the use of narrative versus declarative verbs is either a preservation of this distinction, or a revival of it. In other words, it may be that what we have understood as the grammatical categories of 'verb' and 'noun' are not the categories of 'verb' and 'noun' at all, but the categories of 'non-essential' and 'essential' information. This leads to the question of what reflexive possessives really are: while they certainly function as such much of the time, they are clearly not reflexive in the above cited examples. Understanding exactly what the true function of these pronominals is will require further study, but some promise of a solution lies, for example, in comparing the role of Set 2 and Set 3 pronominals in most modern TG languages with their role in Kb and, apparently, in PTG, and the role of cognate Tupian forms outside of TG.

	Set 2	Set 3
TG	non-reflexive (possessive)	reflexive possessive
Kayabi (?PTG)	non-essential (possessive)	essential (possessive)
Munduruku	alienable possessive	inalienable possessive

---

[8] Below is an example of the wider scope of the reflexive pronominal marker used on a narrative verb:

```
Oro-yor ore ʔyekaʔwi are.  Aʔeramü ʔŋa taytetu yuka-w.
1px-come 1px stream on   conj    3ms caititu kill-nar
'We came by a little stream.  He killed a wild pig.'
```

6.7. Conclusion

In this chapter, we have seen the parallels between noun and verb structure in PTG as exemplified by attested languages whose comparison leads to the reconstruction of even closer parallelism between noun and verb structure. The most consistent difference between noun complex and verb complex structure is in the type of negation used; and evidence in Tb, OG, and AX indicate that even this difference was not consistent in PTG. We reconstruct then a parent language in which all the grammatical indicators of whether a word was a verb or noun are absent. We further have seen evidence from Gy, OG and Tb that complexes were originally phrases. Finally we saw data in Kb as evidence to support the hypothesis that the morphology affixed to roots marking them as nouns or verbs actually functioned in PTG to mark roots as containing either essential or non-essential information, respectively, and that this in turn affects our understanding of the original function of pronominal markers. Comparisons with genetically and typologically related languages should provide new insights into this aspect of PTG grammar.

CONCLUSION

7.0 Summary and contributions of the study

In this dissertation we have examined data from a variety of TG languages to arrive at a reconstruction of TG, more precisely a re-examined reconstruction, modifying Jensen's (1990) reconstruction of PTG. Here we may summarize what we have reconstructed for PTG, highlighting revisions in the reconstruction.

The phonemic inventory was as follows:

```
p       t       k
pʷ              kʷ
pʸ
β       r
m       n       ŋ
        c
        cʷ
            y           w

i       ɨ       u
e               o
        a
```

The most significant difference between this inventory and the ones that have been reconstructed in the past is the discovery of the fortis/lenis distinction. This distinction, apart from the obvious fact of its being a different (and more symmetrical) way of organizing TG and PTG phonologies carries with it many advantages. It combines with remarkable elegance with other seemingly unrelated facts about the PTG phonemic inventory to lead to principled internal reconstruction of Pre-TG phonemes; both the fortis/lenis categorization of PTG phonemes make possible a much more lucid

understanding of PTG morphophonology and its development into modern TG morphophonologies; the reconstructed Pre-TG makes possible tremendous progress in understanding the development of TG morphology and morphophonology; and looking beyond TG, the discovery of the fortis/lenis categorization opens up the possibility of similarly new avenues of progress in the study of other Native American languages: I have found not fewer than nine outside Tupian to have f/l phonologies, and have already used the f/l distinction to yield interesting insights into the development of Pre-IE and PIE phonology (Schleicher 1994).

Another significant difference in this phonemic system is that labialized and palatalized consonants are seen as single segments rather than as sequences of consonant + glide as with Lemle (1971).

The universally accepted reconstruction of two affricates c and $\check{c}$ is not accepted here; instead only c is reconstructed in place of c and $\check{c}$. This lends to a more detailed understanding of the relative importance of dialect borrowing in TG languages.

In addition to recognizing the autosegmental phonemes of nasality and orality, we have discovered autosegmental preglottalization with the syllable as its domain. This solves the puzzle of the so-called metathesis of glottal stop.

As stated above, analyzing the phonology of PTG as f/l makes possible a

principled account of Pre-TG phonology. The inventory reconstructed is shown below.

CONSONANTS

Stops	p	t	k
labialized	p^w	t^w	k^w
palatalized	p^y	t^y	k^y
Nasals	m	n	ŋ
Flap		r	

VOWELS   i   ɨ   u   ī   ɨ̄   ū

          e   o   ē   ō

            a   ā

AUTOSEGMENT: ʔσ

This reconstruction is itself an innovation, since no one before has reconstructed a Pre-TG phonemic inventory. Of particular interest is three complete series of stops. This reconstruction of these three series (and proposed sound change of Pre-TG **t^w and the palatalized series to PTG *c^w and *c respectively) makes it easier to find vocabulary in non-TG Tupian languages which are cognate to PTG vocabulary (see **2.2** above). The hypothesis that lenes are from PreTG stops and PTG stops are from Pre-TG pre-glottalized stops makes possible an explanation of the morphophonemics of PTG nominalization suffixes, viz. /-a/ and /áβo/. The nominalizing suffix used with vowel-final words has been reconstructed as */áβo/. Analyzing this as *-áβ-βo and making the suffix fit the grammar of PTG also makes a new contribution to this study.

In Chapter 3 we reviewed morphophonological rules in Tb and modern TG languages to arrive at some understanding of the morphophonology of PTG. This study

made the following original conclusions:

Prenasalized stops, which were oral allophones of nasals in PTG and are in historical TG, originated as nasal allophones of stops in Pre-TG. The transition from pre-nasalized stops to post-oralized nasals accompanied the transition from nasality as a segmental feature in Pre-TG to an auto-segmental feature in PTG. This ties in with the f/l phonology of PTG: stops in Pre-TG had pre-nasalized allophones while pre-glottalized stops did not. This understanding of the relationship between nasality and stops contributes to the hypothesis of the origin of root classes arrived at in Chapter 4. It was also discovered that vowels before final nasals were, at least phonetically, always nasal.

In Chapter 3 the epenthesis rules (3 and 16) given in Jensen (1984) were related to one another and to the 'nominal case' suffix. As alluded to above, the discovery that ? is autosegmental in Pt and Kb, probably was in OG and was in PTG led to a reanalysis of Jensen's metathesis rule. The f/l phonology of PTG led to a reanalysis of Jensen's rules 4, 8, 9, and 13; in particular, that rules 8 and 9 were really the same nasalization rule.

In Chapter 4 our understanding of the f/l category in PTG and its origin in Pre-TG, together with our understanding of the relationship between nasality and stops, led to the hypothesis of a phonological origin of the root classes in PTG. This hypothesis

may be tested by comparison with non-TG Tupian languages. Also discovered was the connection between locative clitic *pe* and diffusive locative *-βo* also due to our knowledge of f/l phonology. The origin of PTG *c* from Pre-TG palatalized stops shown in (2.2.) together with the f/l distinction and the lenition rule, led to recognizing that serial verb suffixes have a common origin with agentive and circumstantial markers *čár* and *čáβ* this not only supported the idea that serial verbs are really nouns, but led to the explanation of how the focus verbs were also nouns. This in turn, led to a new conclusion about subordination in PTG, which in turn led to a new and better understanding of the pronominal system; in particular that PTG had never been an ergative language.

In Chapter 5 we saw that there was not just one form, *če* for the first singular Class 2 pronominal as had been previously reconstructed, but two dialectal variants, *ye-/če-* and that the *če-* form, from *iče* /ice/ was due to adding *t-* to first person singular pronominal forms: *(t-)a-*, *(t-)ye-*, *(t-)we-*. This in turn was an extension of the alternation in the first plural inclusive *i/t-i-/ya-/t-ya-*. By showing that the original form of the first person inclusive was *i-* and the first singular was *ye-*, we can better apply internal reconstruction to PTG pronominals and compare the results with other Tupian languages, and ultimately outside Tupian (already the internal reconstruction here seems to lend support to the Tupi-Cariban comparison done by Rodrigues (1985)) The internal reconstruction of Pre-TG palatalized stops explained why there were two third person Set 2 markers: *i* and *c* as *i-* and *t-i-*. This led in turn to a clear analysis of the

morphemes *t-*, *o-*, and *i-* in Pre-TG and PTG explaining also, for example, such phenomena as two future markers for nouns: **rām* and **wām* as being ***t-am* and ***o-am*.

In Chapter 6 we saw that what has long been known about roots in PTG is true also for complexes: there is little structural difference between verbs and nouns and minimal difference in PTG. The study of negativization yielded the reconstruction of a heretofore unknown PTG morpheme: **č(o)ʷe*. The development of noun incorporation in OG as well as data in Tb and Gy pointed to verb and noun complexes having been phrases in PTG. And finally data from Kb has shown that the categories of noun and verb were perhaps better understood as the categories of essential and non-essential information.

7.1. Internal classification of TG

No attempt throughout this study has been made to determine the internal relationships within the TG family. Wolf Dietrich (1990b) has made what is the best attempt to date in achieving this. He has accomplished this by dispensing with the traditional tree model of classification. Instead he classifies TG languages according to levels of mutual affinity as seen in the chart on the following page.

▦ = conservative languages

```
                          AsX
                           |
                                  Em
   I    ▦▦▦Mb▦▦▦▦Wa▦▦▦
        ▦|▦▦AsT▦|▦
        ▦Ap▦▦▦▦WJ▦
        ▦Kw▦▦▦Te-Gj▦
   _    ▦|▦Gy▦▦▦▦▦▦        axis of morpho-
        ▦Av▦-▦Tb▦▦Pt▦       logical conserva-
        ▦▦▦▦▦Km▦             tism
        ▦CT/Ch▦▦▦Tp▦
        ▦▦▦▦▦▦▦▦▦
                          Kb

                       Ub    Si
   -------------|----------------------------
   II
              Jm          Gk  -  Co
   -------------------------------------------

   III                   Aw   SM
```
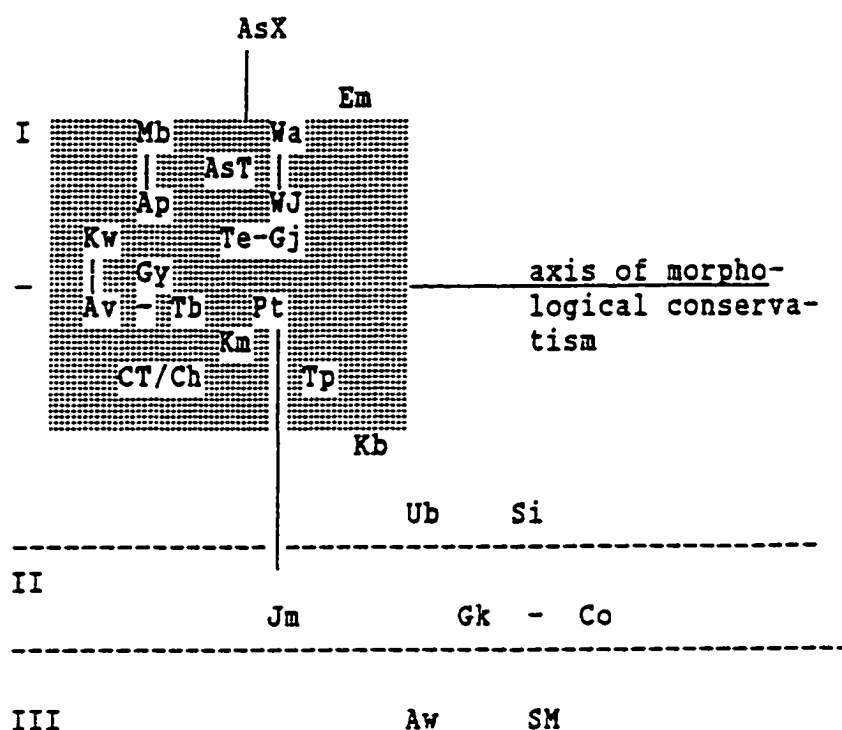
However, it is possible to use the data of this study to propose an Urheimat for PTG. To do that we can look at the geographic scope of TG and compare several isoglosses. Working in this way, we see that the Guaranian languages (and OG itself) lie in a relic area as do Gy and Pt. In some respects Kb seems to be in a relic area, but not to the same degree as, say, Gy. Gj, UK, AX, AT are squarely within the central area for most isoglosses. The way W has stretched the central area northward seems to testify to a relatively late migration. Tb and OG have not been included in the isogloss maps because their earlier date at least partly cancels out the importance of geography in their conservatism. By discerning the central innovating area, if we propose a homeland in the center of that, we have a homeland for PTG in the Amazonian interior.

While an internal classification may not be possible (or even appropriate) at this point, the principle of migration theory proposed by Dyen (1956) and exemplified by Diebold (1960) can help in proposing a homeland for PTG. By looking at the isogloss chart, the greatest mutual difference between individual languages (among those studied here) is that between Pt, Km, Kb, and Tp. These languages all have are adjacent to one another on the map and their regions taken together comprise that of greatest diversity of any set of adjacent languages. It makes sense, therefore to tentatively propose the PTG homeland as being in this area, in the Planalto do Mato Grosso. That area is not exactly where migration theory would point to as the PTG homeland, but this can be complemented by insights from dialect geography: the distribution of linguistic innovations indicates that all migrations (and linguistic innovations they brought) followed rivers downstream and most rivers from this area lead in the same direction, i.e. toward the Amazon River. Hence Pt, which found itself upriver from other languages is for that reason alone in a relic area, while many innovations are shared between AT, AX, Gj, and W at or near the mouth of the Amazon river. In this scenario, one group migrated directly south along the Paraguay River towards southern Brazil and southern Paraguay (OG), while other groups migrated north. Gy may have migrated directly east from the PTG homeland, or split off from the Southern group before OG underwent the phonological changes that make it as we see it attested. If Gy was part of an eastward migration, down the Guaporé River, Pt may or may not have followed the same route, only continuing down the Madeira River rather than stopping in Bolivia. Tupinamba, not belonging to the southern migration, represents one of the northern migrations to

the Amazon, down the Amazon to the Atlantic coast, and then south all along the eastern coast of Brazil. Early explorers did find Tupinamba speakers all along this coast at the time of Anchieta's grammar, and this explains the innovations in Tb.

The most interesting feature of this hypothesis is how it speaks to the subgrouping of Gy, Tb and the Guaranian languages according to Dietrich's internal classification: this relationship does not constitute an actual subgroup of TG, but merely reflects their strong similarity to PTG as the result of being in relic areas. The problem is somewhat similar to the proposition of a genetic subgrouping in Indo-European of Celtic, Italic, and Tocharian based on shared conservative features, really the result of their location in relic areas respective to the speech area of Indo-European as a whole.

A PTG homeland in Planalto Mato Grosso is the most reasonable proposal for a homeland because it is the only location where all the rivers used in the migration routes necessary for the current distribution of TG languages have their headwaters. It is not terribly far from the location of other branches of Tupian, which are clustered close together in the area of Rondonia. But it is far enough as to suggest that 1) the speakers of PTG migrated some distance east before TG began to differentiate into different languages; 2) PTG was itself in a 'relic' or 'lateral' area with respect to other Tupian languages and therefore probably is relatively conservative with respect to Proto-Tupian.

When we review all the isoglosses we can combine them in a single isogloss map or chart as follows:
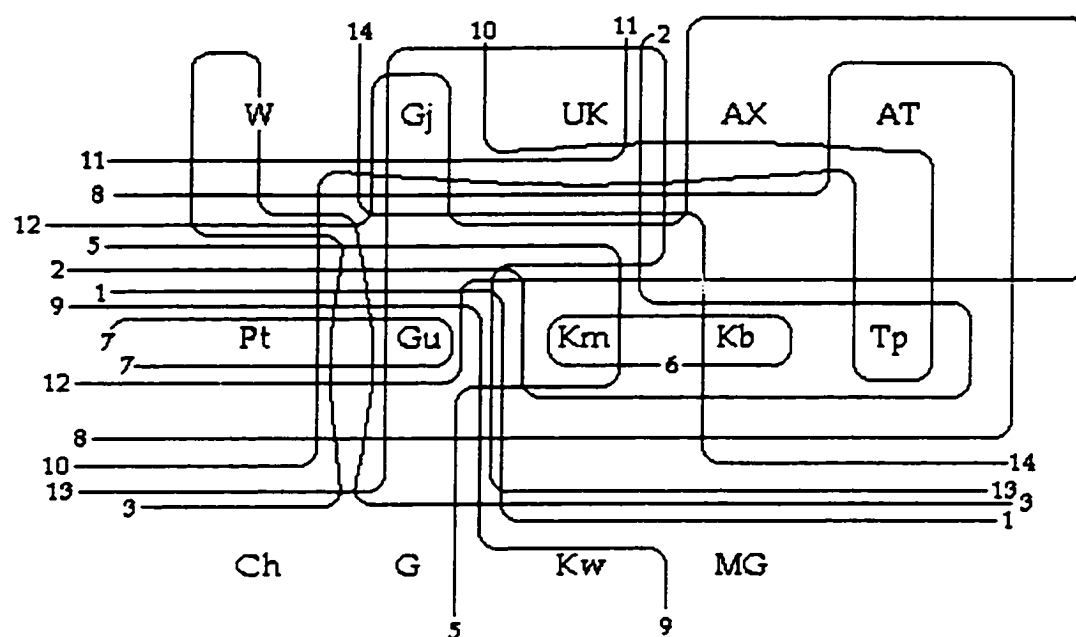
1. $\beta \rightarrow$ w (above)
2. βa?e $\rightarrow$ ma?e
3. loss of word final consonants
4. obtop or focus nominalization
5. Set 4 markers *oro-/opo-* not lost or altered by analogy
6. preservation of constrast between reflexes of *p^w and *k^w
7. preservation of p^y
8. Set 3 markers
9. [ti]
10. *ya-/ti-/tya-* alternation grammaticalized
11. *ye-/yo-* merged
12. Set 4 marker(s) *epe* (and *epeyepe*)
13. Set 1 1s *ye-/*če- (*če- inside and below line)
14. loss of autosegmental nasality

Some caution needs to be used in interpreting the significance of the isoglosses. For example, as Jensen has pointed out, palatalization rules like [ti] → [či] are so commonplace as to be nearly useless for subgrouping. The same is true for loss of autosegmental nasality (Soares 1979:105).

This chart strongly supports the grouping of Ch, G, Kw, and GM in its own subgroup, OG as well as AT and AX in a subgroup and, perhaps to a lesser extent, of W and Gj. The isogloss map also parallels Dietrich's chart of internal relationships in showing two major facts about TG: 1) there are almost no subgroups within TG — isoglosses separate almost every language into its own subgroup and no two isoglosses contain the same languages; 2) The languages we have examined have grouped into three levels of relatedness, each level set off by an isogloss bundle. Hence the Guaranian languages show closest affiliation with one another, PT, Gu, Km, Kb, and Tp show least similarity among themselves, but all have very conservative features; and the languages W, Gj, UK, AX, and AT, are more similar to one another (principally because of shared innovations rather than shared conservative features), but are overall the least like reconstructed PTG.

One of the important contributions of Dietrich's analysis is that it goes a long way toward factoring out which TG really belong within the language family and which ones are not really TG, though closely related. Among those languages heretofore considered

TG which are shown in Dietrich's study not to be TG is Kokama. This is confirmed in this study for different reasons, namely, that the pronoun system of Kokama could not have derived directly from PTG; note for example, the first person plural inclusive *ini* from **inde*, a form not reconstructed for PTG, but possible for a later stage of Pre-TG. This last detail agrees with Dietrich's statement that Kokama "should be on the very periphery of the TG family." Of Sirionó, Dietrich (1990b) concludes that it is "so defective in its grammatical structure, so peculiar in its phonological structure and vocabulary, that it is near the third category," i.e. among languages closely related to, but outside of, TG. He nevertheless deems it "a full-member of the TG language family." For reasons given in Chapter 2, I have placed Sirionó just outside TG, although admittedly the few basic elements of morphology which Sirionó shares with TG (especially first and second pronominals) can only be directly derived from PTG, with the possible exception of the first person singular subject prefix *(t)a-*. On the one hand *(t)a-* may be considered non-TG (cf. *ta-* in Kokama), but not necessarily so, given our earlier discussion of *t-* variants in first person pronouns in TG (note also *te-* in Kb, which is unique among TG languages). There does seem to be not so much a clear cut TG family distinct from all other Tupian languages, but a continuum of relatedness, part of which is sectioned off as TG. Dietrich draws the line in such a way as to include Sirionó; I draw the line in such a way as to exclude it. Actually, given the use of Set 1 markers as both object and subject markers, the paucity of shared (albeit basic) vocabulary, and the radically different phonology, Sirionó may in fact be a creole resulting from contact between PTG and either non-TG Tupian speakers or non-Tupian speakers, or both.

Aside from these details, Dietrich finds two major groups within TG: 'Southern' and 'Amazonian'. Again, this grouping is not based on isoglosses directly, but on rates of mutual similarity in phonology and morphology. Dietrich found a high rate of similarity between what we have termed Guaranian languages and Gy followed by less mutual similarity, not only between these languages and other TG languages, but also among the others. This correlates well with the isogloss map above. As the isogloss map indicated, a simple tree diagram cannot be drawn for TG, but certain lines of relatedness can be seen.

All of the significant innovations in Tb are shared with some of the innovating languages far from the speech area of Tb: the use of 'nominal case', substituting *ramo* for another word for when/if (Pt and Tb both use *reme*); and *-ca* instead of *-ta* for the nominalization of *r*-final verb stems. In short, since Tb was spoken in a relic area and was therefore conservative in its own day, we might stipulate that these differences between Tb and OG are among the most basic among TG languages. This is tantalizingly suggestive of the possibility that four centuries ago, when Tb was spoken, the ancestors of the 'Amazonian languages' were so much closer to Tb than they are today, that we may have to radically rethink the chronology of PTG. The only date so far proposed for PTG is 5500 to 6000 years ago (Swadesh 1959). This was based on glottochronology. Whether or not one accepts the claims of glottochronology, they may not hold for TG. If TG languages have changed so rapidly that a language in the central innovating area four hundred years ago looked like Tb, it would mean that TG have

changed much faster than glottochronology would predict. In fact, such rapid change has been observed in Guarani.[1] At such rapid rates of change, we might imagine the break-up of PTG as late as the thirteenth century AD. Such a late date may seem radical at first glance, but not when one considers that most of the modern languages with all their innovations were only known in the latter half of the twentieth century, and that those recorded a century ago reveal half their innovations in the last 100 years. Dietrich himself comments on the conservatism of TG in contrast to the posited great age of the language family:

It is amazing to see that TG languages are not only related to one another by a common lexicon and a reconstructible phonologic structure, but that they have also inherited a common basic grammatical shape, which has been preserved by the majority of these languages for centuries, even by the smallest groups scattered all over a great part of the continent. (Dietrich 1990:112-13)

It might not be so amazing to see this if it should happen that TG began to diversify into separate languages only seven or eight centuries ago.

---

[1]While studying OG at the University of Kansas, I spent a total of three hours translating a single page of OG text, written by a native speaker, with a literate, educated, native speaker of Paraguayan Guarani. Most of her attempts at identifying lexical items were incorrect, in spite of the explicit Latin title of the text. It seemed comparable to a native English speaker's attempt to translate Piers Ploughman, or even Beowulf.

Obviously, these are currently open questions. Hopefully, this work will make possible new and fruitful avenues of inquiry into the relationships among language families of the Amazon Basin.

Appendix I
Comparative Word List

1. ʔá 'fruit'
   iʔβa

   Km a; Pt iʔβa; UK maʔeiwa; W ʔa; Kb iʔwa; GM ʔa; AT iʔa, iwaa; AX iʔa; Kw iβa; Gy a, iβa

2. ʔáβ 'hair'

   Km ʔap; Pt ʔaβ; UK iʔa 'his hair', ha 'fur, hair of body, bird feather'; W apira; Kb ʔap; GM ʔa ('to cut hair', ʔigwe); AT ʔáwa; AX awa; Kw ʔáβi; Tp ham; Gy a, aɣʷer

3. ʔáβ/ʔáw 'lie down'

   Km ʔaw; Pt -ʔaw; UK upe; W ʔaw; GM tui; AX ʔap; Tp ham

4. ʔám 'to stand'
   poʔam 'to stand up'
   Km -ʔam; Pt -ʔam; UK puʔam; W puʔã; AX poʔom

5. ʔapik 'to sit'
   Km -ʔapik; Pt ʔapiɣ; AX apik; Kw -gwapi

6. ʔapitī 'tie up'
   Pt upatī, ʔapitī; W apasī; AX či; Kw apitī; Tb apitī

7. ʔár 'fall (human)'
   Pt ʔar; UK -ʔar; W ʔa; GM -a; AT ʔan; AX ʔat; Kw ʔa; Tp han; Gy ʔa

8. ʔé 'say'
   Km ʔi (iʔi); Pt ʔe; UK ye (is said); Wj ʔi; GM ʔe; Gy e

9. ʔirū 'other', 'companion'[1]

o-yo-ʔirū-ndɨk
mo-yo-ʔirū } 'four'
ʔirū-katú
ʔirū-paβ

Km irū 'husband', moyoʔirū; Pt -irū 'other', 'one more'; Kb irupāwe 'four'; GM irū 'husband', irundɨ; AT naʔiroihi 'three', iroŋatoete 'four'; AX iroma'e 'three', ožeroik 'four'; Kw irundɨ; Tp irõ 'companion'; Tb oyoirundɨk 'four'; Gy irū 'addition' 'increase', iruŋgatu 'four'

10. ʔɨβ 'tree'

Pt ʔɨβ, ɨβa; UK -ʔɨ (iʔɨ); W ʔɨ; Kb ʔɨp; AT ɨwa; Tp hɨwɨra; Gy ɨ 'stick' 'stem' 'tree'

11. ʔɨčár 'canoe'

Km ʔɨat; Pt ɨhar; UK yarusu; Wj ɨa, Wa ɨar; Kb ɨat; GM ɨa, ɨɨa; AT ɨhara; AX ɨhara; Tp ɨan; Tb isár

12. ʔɨtáβ 'swim'

Km ɨtap; Pt -itaβ; W ʔita; GM -ʔita; Tp hɨɨtam; Gy ita

13. ʔŋā 'they'

Pt ʔŋā, ʔŋahā 'they' ʔɣa, ʔɣaha 'he'; UK ŋā; Kb ʔŋā 'they', ʔŋa 'he'; Tp aheŋɨ 'they', aheŋā 'he'

14. ʔók 'dig'

Km yoʔok; Pt ɨβɨkoɨ; UK pɨkūy; GM joʔo 'hollow out'; Kw joʔo; Gy -o

15. ʔú 'eat (trans.)'

Km ʔu; Pt ʔu; UK ʔu; W ʔu; GM ʔu; AT ʔo; AX -ʔo; Tp ho; Tb ʔú; Gy ʔu

16. aʔé '3rd person'

Km aʔe; UK aʔe; W aʔe; GM haʔe, haʔevaʔe; AX aʔe; Kw haʔe; Tp aheŋɨ 'they', aheŋā 'he'; Gy ae

---

[1] See 1.1.

17. aʔɨy 'seed'

    Km -aʔɨy; Pt -aʔɨ̃n; UK -aʔĩ; W aʔɨy; Kb aʔɨy; GM -aʔɨ̃ĩ; AX -aʔɨpa; Tp āhɨy; Gy -aɨy

18. aβá 'person'

    Km awa; Pt aβaʔŋa (pl); UK awa; Wj awa, Wa aβa; Kw ava; Tp āwā; Tb aβá; Gy aβa

19. aβatí 'corn'

    Km awaci; Pt awati; UK awaši; Wj awasi, Wa aβasi; Kb awaci; GM aβaši; AT awačia; AX awači; Kw aβati; Tp hāwāči; Tb aβati; Gy aβači

20. aíβ 'bad'

    W ai; GM βai; Kw βai; Tp āip; Gy ai

21. aimbé 'sharp'

    Km aime; Pt aimbe; W ayme; GM -aĩmbe; AX aimbe; Kw -hāimbeʔe; Tp āip

22. akʷaimbaʔé (?< *kuyāʔĩ-mbaʔé) 'man'

    Km akʷamaʔe; Pt akʷaimbeʔe; AT akomaʔe; AX kužamaʔe; Kw kʷimbaʔe

23. akáŋ 'head'

    Km ʔakaŋ; Pt akaŋ; UK āka (i-āka 'his head'); Wj akā, Wa akaŋ; Kb akaŋ; GM akā; AT akɨŋ(a); AX akɨŋ; Kw akā; Tp ākɨŋ; Tb akáŋ

24. akɨm 'humid, wet'

    Km -ʔakɨm; Pt -akɨm; UK ākɨm; GM akɨ; AT -akɨm; AX -akɨm; Kw -akɨ; Tp ākɨm

25. -akúβ 'hot'

    Km hakup; Pt -akuβ; UK haku; W aku; Kb akup, akuwa; GM -aku, -aku βeβuy; AT hakom; AX hakop, hakow; Kw akɹ, haku; Tp ākop

26. amán 'rain'
    Km aman; Pt aman; UK aman; Wj amā, Wa aman; Kb amān; GM amā;
    AT amɨna; AX amɨna; Kw ama; amɨn, 'rain, rainy season'; Tb amán; Gy amar,
    aman-

27. amō 'other'
    amo-aʔe
    amo-e
    amo-te
    Km amo; UK amō; W amū; GM amboae, mboae; AT amoa, amote; AX amote,
    amo; Kw ambue; Gy amo 'some-'

28. amō-ité 'far'
    Km amoete; AX mo-ite, awa-ite

29. amōy 'old man, grandfather'
    W amūy; GM amōy; Tp amōy

30. anī 'no'
    Km anite; Pt anī; UK anī; GM anɨ; Kw ani, ani, naháni; Gy ani

31. áŋ 'this'
    Km ʔaŋa; Pt aŋa 'this here'; Tp aŋā

32. apé 'back'
    Km ape; W ape; AT iápee; AX ape; Tp āpe; Tb ape; Gy ape 'outside''surface'
    'bark'

33. apé ~ peé 'road'
    Km tape; Pt pehe; UK pe (hape 'his road'); W pee ~ ape; GM -ape, tape pia;
    Kw tape; Tp āpe; Gy ape 'here, there'

34. apɨ 'burn'
    Pt -apɨ; W apɨ; GM -apɨ; Tp āpɨ

35. apūy 'nose'
>Pt -apiñ; UK āpūy; GM apɨy; Tp āpīwam

36. apó 'root'
>Km -apo; Pt -apo, -apoʔa; UK -apo; W apo; Kb apo; GM -apo; AT -apa;
AX ɨwɨpa (<*iβɨ-apo); Kw hapo; Gy capo

37. apoʔá/apuʔá 'short'
>Pt apoʔa; UK yupuʔa; GM apuʔa

38. apuʔá 'round'
>Km ahuʔa; UK puʔa; W apuʔa; GM apuʔa; Kw apuʔa; Tp āpayā; Gy apua
'something round'

39. ár 'day'
>Km at; Pt ar; UK ʔar; Wj arɨwo, Wa arɨβo; GM ʔara; AT ara; AX ara; Kw árɨ;
Tp han; Tb ár 'day, sun, light, time'; Gy ar, arɨ 'sun, day, time, weather'

40. atá 'walk'
>Km ata; Pt ʔatauhu; UK wata; W ata; GM gwata; AT -ata; AX -ata; Kw -gwata;
Tp ātā; Tb -atá; Gy γʷata 'to go, wander, stride'

41. (t)atá 'fire'
>Km ata; Pt -ata; UK tata; W tata; GM -ata; AT tata; AX tata; Kw tata; Tp ātā;
Tb tatá; Gy tata

42. (t)atá-tíŋ 'smoke'
>Km tatacɨŋ; Pt -atatiŋ; UK tata tadāšī; Wj tatasī, Wa tatasɨŋ; GM atašī, tata
ratašī; AT tatasɨŋa; AX tatašɨŋ 'smoke from the earth', 'thick fog'; Kw hatatī,
tatatī; Gy tatačī

43. atɨr 'mountain'
iβɨ-atɨr 'mound/mountain of earth'
>Km atɨt; GM iβɨtɨ; AT ɨwɨtɨra; AX ɨwɨtɨra; Kw iβɨatɨ gwasu; Tp ɨwɨtɨt; Gy atɨ
'much'

44. atĩ 'horn'

Km hacĩ; Pt atĩ, UK haxĩ 'is pointed, sharpened'; W namirasĩ; Kb acĩ; AT háčia; AX hači; Kw hatĩ, hatĩgwe; Tp āčĩ; Gy ači

45. -āy 'tooth'

Km hãy; Pt añ, hahĩ; UK -āi (hāi 'his tooth'); W āy; GM aĩ; AT hóča; AX rĩža; Kw -āi, tāi; Tp ĩy; Tb -āy; Gy -āy

46. (mbo)ayán 'push'

Pt añan; W moayã; GM moaña; Kw -moaña; ? Tp āčim; Gy moañani 'push forwards'

47. ayurú 'parrot'

Km ayuru; Pt ajuruʔi 'blue parrot'; Kb ayuru; AT ačorohoa; AX aižora; Tp āčoro; Gy ayuru 'talking parrot'

48. βeβé 'fly'

Km -wewe; Pt -βeβe; UK wewe; Wj wewe, Wa βeβe; GM -βeβe; AT -wewe; AX βeβe/wewe; Kw -βeβe; Tp wewe; Tb βeβé

49. βeβúy 'float'

Km -wewɨy; Pt mbovevuy 'make float'; UK wewi 'is light'; Wj wewɨy, Wa βeβɨy; AX -vevoy; Gy βeβɨy 'float, be light'

50. βók 'crack, split'

Km -mowok; Pt -βoβoɣ; UK -wok; Kw -mboβu; Tp mawāk, 'go around something', wawāk, 'to turn around'; Gy βo

51. βúr 'swell'

Km -wuwut; Pt -βur; Tp wowot

52. βúr 'float'

Pt -βur; GM -βu; Tp wot; Gy βu 'fluctuate, bubble, run aground, it goes over, cooks'

53. čám 'cord'

    tupa-čam 'hammock-cord'

    Km tupaham; Pt ham, iβɨrapaham; UK tupāham; W ā; GM šā; AT očon;
    AX topam; Kw isā; Tb sám; Gy cā, tupācā; Gj həm

54. -čaráy 'play, amuse'

    UK yumusaray, yumaray; W yimaray; AT-semoaray; AX (d)žembaray;
    Kw -ñembosaray

55. čeβoʔí 'worm'

    Km ewoʔi; Pt -eβoʔi; UK sowoʔi; Wj ewoʔi; Wa eβoʔi; Kb ewōʔi; GM ʔeβoʔi
    (minhoca); Gy ceβoʔi

56. čéy 'wash'

    Pt -hey; GM -ei, -joi; AT -pohey; AX -po(h)ey; Kw -johey, -hey; Tp -ey, -pohey;
    Tb séy; Gy cey, yocey

57. čɨ 'mother'

    Km ɨ; Pt ɨ; UK -hɨ; W ɨ; Kb ɨ; GM šɨ; AT hɨke; AX hɨ; Kw -sɨ; Tp ɨ; Tb sɨ;
    Gy cɨ

58. čɨβ 'clean'

    Pt -hiβ; Kb ʔɨpia; GM -mbojeɨ; AT -hɨm; AX (h)ɨp; Gy cɨ, cɨy, yocɨy

59. čɨβ 'rub'

    Pt -hiβ; AT oseehɨm; AX žeaɨp

60. čɨm 'smooth'

    Km -īm; Pt -hɨm; UK -hɨm; GM išī; Kw -sɨy

61. čɨrɨ 'run (water)'

    W ɨrɨ; GM -šɨrɨ; Kw -sɨrɨ; Gy cɨrɨ 'move along'

62. čók 'pull off'
   Pt hoγ, poanoŋ; UK -hok; Tp ak 'break'; ?Gy coka 'der Stossel'

63. čuʔú 'to bite'
   Km -ʔuʔu; Pt -huʔu; UK suʔu; W suʔu; GM šuʔu; AT -oʔo; Kw -isuʔu; Tp oho;
   Gy cu

64. čún 'black'
   pi-čún 'black'
   Km picun; Pt -un; UK -un, -ū, pihun; W pi(y)ū; Kb ūnai; GM -ū, huū; AT hon;
   AX -on; Kw -ū, hū; Tp un; Tb sún; Gy cū

65. eʔím 'not'
   Pt im; UK im; GM ani, eme, eʔĩ; Kw ani; Tp ehimamo 'because not'; Gy eĩ

66. eʔïy 'scratch'
   Km eʔïy; Pt aβekyi 'scratch the hair'; Kw -ñeheʔiy, -heʔiy; Tp ehĩy; Gy -eïy

67. eβék 'belly'
   Pt eβeγ; Kb -ewek; GM eβi atā 'hard belly'; AT -ewéŋa; AX -rawera; Tp ewek

68. ečá 'eye'
   Km -ea; Pt eakʷar, eaʔir; UK eha; W ea; Kb -rea; GM -eša, -eša raʔiï; AT -éha;
   AX -reha; Kw -esa; Tp ẽã; Tb -esá; Gy -eca

69. ekíy 'to pull'
   Km ekiy; UK hikiy; W ekiy

70. ekó 'to live'
   Km -kowe 'to live', -iko 'to be'; Pt -kóji; UK -ko, šuwe 'alive'; W eko 'to be in
   movement'; GM -iko; Tp ekoete 'alive'; Gy -eko 'be'

71. embi-rekó 'wife'
   Km -emireko; Pt -embireko-; W emireko 'possessed thing'; Kb emireko;
   GM embireko; AX -mirika; Kw -embireko; Gy -embireko

72. endé 'you'
Km ene; Pt nde, ndehe; UK nde; W ene; Kb ene; GM ndee; AX ende; Kw nde;
Tp ane; Tb ende; Gy nde

73. endɨ 'saliva'
Km henɨ; Pt -e-ndɨ; UK -endɨ; W -eni; GM -endɨrɨ; AT hénɨa; AX -rendɨ;
Kw -endɨ; Tp enɨ; Gy -endɨ

74. endúβ 'hear'
Km anup; Pt -enduβ 'pay attention, listen with respect'; UK -endu; W enu;
GM -endu; AT -enom; AX endop; Kw -hendu; Tb -endúβ

75. enɨpɨʔã 'knee'
(?< endɨpɨʔã)
Pt -enɨpɨʔã; W enɨpiã; Kb -renɨpɨʔa; GM -enapɨʔã; Kw -etɨpɨʔã; Tb -enɨpɨʔá;
Gy -enɨpia

76. enõy 'call'
W enũy; GM -enoĩ; AX enoi

77. epʸák 'see'
Km -ecak; Pt epiaγ; UK -sak; W esa 'meet'; GM -eša; AT -eča(ŋ), -ečaŋokan
('to show'); AX esak; Kw -heše; Tb -epʸák; Gy -epia

78. -ér 'name'
Km -et; Pt -er; UK -er; Wj e, Wa er; GM -erɨ; AT -era; AX -era; Kw hérɨ;
Tp et; Tb -ér; Gy -er

79. -etá 'much'
Pt hete; UK heta; GM -eta; AT heta; AX eta 'many'; Kw heta 'many'; Gy -eta

80. etɨmã 'leg'
Km -etɨmakaŋ; Pt -etɨma; UK itɨmã; W etɨmã; Kb -up; GM -etɨma; AT -etɨmoa
'below the knee'; AX tamakaŋ; Kw -etɨma; Tp etɨmɨkɨnapipem 'top part of the
leg'; Gy -etɨma

81. -etún 'smell'
   Km wetun; Pt etun; UK hētū; W etū; GM etū; AT eton; AX eton; Kw -hetī;
   Tp eton

82. itá 'stone'
   Km ita; Pt itaky; UK ita; W takuru; GM ?ita; AX ita; Kw ita; Tp itã; Tb itá;
   Gy ita

83. iye, iče 'I'
   Km iye; Pt jihi; UK ihē; W ie; Kb ye; GM šee; AX (i)dže; Kw še; Tp ie; Tb iše;
   Gy če

84. i 'water'
   Km i, Pt ihi, i; UK i; W i; Kb ?i; GM i; AT ia; AX i; Kw i; Tp hi; Tb í; Gy i
   'water', 'river'

85. i-upá 'lake'
   Km iupawape; Pt ipia?i, ipiahu; UK ipa; W ipa; GM ii rupa, iupa; Kw iupa;
   Gy iupa

86. i?itíŋ 'sand'
   Km iciŋ; Pt i?itiŋ; Tp ičiŋ 'ajuntar'

87. i?ú 'drink'
   Km i?u; Pt -i?u; W i?u; GM -i?u; AX -i?o, -i?o; Kw -i?u; Tp iho; Gy iu

88. iβí 'earth'
   íy
   Km iy, iwi; Pt ij, iβi; UK iwi; Wj iwi, Wa iβi; Kb iwi; GM iβi; AT iča; AX iwi;
   Kw iβi; Tp iwi; Tb iβí; Gy iβi

89. iβák 'sky'
   Km iwaka; Pt iβaγ; Wj iwa, Wa iwaŋ; Kb iwak; GM iβa; AT iwaŋa; AX iwak;
   Kw ári; Tp iwāk; Tb iβák; Gy iβa

90. iβák-tíŋ 'cloud'
    Km iwiciŋ; UK iwatakāšī; Wj iwasī, Wa iβasī; Kb iwasiŋ;
    AT iwaŋa-siŋ; Gy iβĭčīŋ/iβičīŋ

91. iβirá 'tree'
    Km iwira; Wj iwira, Wa iβira; GM iβira; AT iwira; AX iwira; Kw iβira;
    Tp iwirā; Tb iβira; Gy iβira; wira 'wood'

92. iβitú 'wind'
    Km iwitu; Pt iβitu; UK iwitu; Wj iwitu, Wa iβitu; Kb iwitu;
    GM iβitu; AT iwitoa; AX iwito; Kw iβitu; Tp iwito; Tb iβitu; Gy iβitu

93. ié 'belly'
    ié-pʷér 'intestines'
    Km iepo; Pt -ihe 'large intestines' -iepohun 'small intestine', -iepojuβ; Wj eikʷe
    'stomach', Wa eikʷer; Kb -riepuku; GM ie, iekʷe poʔi, iekʷa 'bowels'; AT hieʔia,
    hieawóa; Kw -ie, -iekʷe; Tp ie; Tb ié, iepʷér; Gy cie

94. (e)imbá 'domestic animal'
    W eima, Wa eima ~ ima; GM -imba; Kw ṃimba; Tp eimam

95. ipé 'bark'
    Km ipe; Pt ape, apir, pe, ipe ('of a tree'); Kb ipe, ipeϕet; GM pekue; Tp ipe;
    Gy ape 'bark, surface'

96. ipitún 'night'
    Km ipitun; Pt ipitun; UK pitun; W pitū, pia; Kb ipitūn; GM pitū; AT ipitonoho,
    ipiton; AX ipiton, ipytonoho; Kw pyhare; Tp hipiton

97. irapár/iβirapár 'bow'
    Km ʔiwirapat; Pt iβirapar; UK wirapar (irapat); AT iwirápara; AX iwirapára;
    Kw gwirapa; Tp iwirāpan; Gy -apar (Class II)

98. -iti?mbór 'dust, powder'

(iβiti?mbór) 'dust of the earth'

Km iwicimot; Pt -moiti?mboruhu 'make powder'; Gy čimbo 'smoke, dust', iβičimbo

99. kʷaáβ 'know'

Km -kʷahap; Pt kʷahaβ; UK -kʷa; W kua; GM -kʷaa; AT -kʷaham; AX kohap; Kw -kʷaa; Tp kʷaam, akʷaam, 'known'; Tb kʷaáβ; Gy kʷaa

100. kʷár 'sun'

kʷar-či 'Mother Sun'

Km kʷat; Pt kʷara, kʷarahiuhu 'the sun after the rain'; UK warahi, kʷarahi; W kʷarai; Kb kʷat; GM kʷarai; AT kʷarahia; AX kʷarahi; Kw kʷarahi; Tp kʷan 'sun, day, era'; Tb kʷár, kʷarasí

101. ka?á 'scrubland, forest'

Km ka?a; UK ka?a; W ka?a; GM ka?a; AX ka?a; Kw ka?a; Tp ka?a; Gy kaa

102. ka?api?í 'grass, weeds'

UK kapī; W ka?api?i; AX ka?ápi?í; Kw kapi?i

103. ka?í 'monkey'

Km ka?i; Pt ka?i; UK ka?ijarar 'caiarara monkey'; W ka?i; Kb ka?i; GM ka?i; AT ka?ia; AX ka?i; Kw ka?i; Tp kã?i

104. ka?mbú 'suck'

Pt ka?mbu; GM kambu; Tp kamo; Gy kambu

105. káβ 'fat'

Km kap; Pt -kaβ; UK ma?e ka (iša); W ka; AT -kam; Tp kam

106. kám 'breast'

kam-í 'milk'

Pt -kam ('chest'), -ka?maŋuhū, -kambipi; W kā; Kb kā; GM kā; Tp kim 'chest', 'breast'; Gy kā 'breast', kambi 'milk'

107. káŋ 'bone'

Km -kaŋ; Pt kaŋ; UK kaŋgwer; Wj kaŋgwe, Wa kaŋgwer; Kb -kaŋ; GM kā, kāgwe;
AT íkɨŋ(a); AX kɨŋ; Kw -kāɨ, -kāgwe; Tp kɨŋ; Tb káŋ; Gy kaŋᵂer

108. karãy 'scrape'

UK karãy; W karãy; Kw ñekarãy; Gy kãrãy

109. karú 'eat (intrans.)'

GM karu; AT karo; AX karo; Kw karu; Tp karo; Tb karú; Gy karu

110. katú 'good'

Km katu; Pt katu; UK katu; W katu; AT kato, katoete; AX kato; Tp kāto;
Gy katu; Gj katu

111. káy 'get burned'

Km -kay; Pt -kay; UK -kᵂay, -kay; W -kay; GM -kay; AT -kay, ɨwɨkay 'dig';
AX ɨwɨkay; Kw -kay; Tp ɨwɨkay, 'dig'; Tb -káy; Gy -kay

112. kér 'sleep'

Km -ket; Pt kir; UK kᵂer, ker; W ke; GM ke; AT -ken; AX kit; Kw -ke; Tp ket;
Tb kér

113. kɨʔá 'dirty'

Km Pt -kɨʔa, apiŋgɨʔa; UK šiʔa (3rd sing); W kɨʔa; GM kɨʔa; Kw -kɨʔa; Tp kɨʔā

114. kɨβ 'louse'

Km kɨp; Pt -kuβ; UK kɨ; W kɨ; Kb kɨp; GM kɨ; AT kɨwa; AX kɨwa; Kw kɨβɨ;
Tp kɨp; Tb kɨβ;

115. kɨčé 'knife'

Km kɨe; Pt kɨheʔi; UK kɨse; W kɨse; GM kɨše; AT kɨeʔia; AX kɨʔe, kɨhe;
Kw kɨse; Tp kɨče; Tb kɨsé; Gy kɨce

116. čikiyé 'fear'
Km -kiye; Pt -kihiji; UK kiye; W kiye; Kb akiiye ʔŋa 'he is scared'; GM -kije;
AT kiiče; AX kiiže; Kw -kihije; Tb sikiyé; Tp kiiče; Gy cikiye

117. (a)kir 'green'
Pt yukiri; -akir, -kir; UK -akir; GM aki, ki; AX ižokira

118. (pi)kir 'fat'
Pt -pikir; GM kira; Tp kirã

119. kitíŋ-ʔók 'clean'
Km kiciŋoki; Pt -kitiʔŋoɣ 'wash'

120. kitī/kiti 'to cut'
Km kici; Pt kiti; UK kitik 'to grate'; W kisī; GM kixī; AX kiči (serrar); Kw -ikītī;
Tp kiči

121. kū 'tongue'
Km kō; Pt -kū; W apekū; Kb kū; GM apekū; AT íkoa; AX ko; Kw -apekū;
Tp kō; Tb kū; Gy kū

122. ko 'this'
UK komeʔē; GM ko, kovaʔe; Kw kóvaʔe, kóa; Tp kō; Gy ko

123. kočúβ 'ashes'
Km ikohup; Pt kuhuβ; Gy kocuɣʷer

124. kuʔí 'sand'
UK iwiku?i; GM kuʔi; Kw itakuʔi; Tp kohi; Gy kuʔi

125. kupé 'back'
Pt kupe; UK kupe; GM kupe; Gy kupe 'back, shoulder'

126. kurumí 'boy'
     UK kurumi ra?ɨr, kurumi; AX konomi; Kw kunumi; Tp konomi; Gy kunumi 'child, lad, boy'

127. kutúk 'bore, perforate'
     Km -kutuk; Pt -?apɨtekutuɣ, -kutuɣ; UK kutuk; Wa kutu; AX pakutuk; Tp kotok;

     Gy kutu

128. kúy 'fall (object)'
     Km -kuy; Pt -kuy (involuntary); UK kukuy (fruit or leaf); GM -kuy (-jaikuy, 'begin to fall'); Tp koy; Gy kui

129. kuyã 'woman'
     kuyã-taí 'girl'
     Km kuyã 'woman', kuyatã imet 'girl'; Pt kuña; UK kũyã 'woman', kũyãtãi ra?ɨr 'girl'; Kb kũyatãy 'girl'; W kuyã 'man's sister'; GM kuña; AT kosoa; AX kunɨ; Kw kuñã 'woman', kuñatãí 'girl'; Tp kočɨ; Tb kuñã; Gy kuña 'woman' kuñataí 'girl'

130. mandí 'manioc plant'
     mandi?ók 'manioc'
     mandi?ɨβakãŋ 'manioc plant'
     Pt mandi?oɣ, mandi?ɨβakã; UK mandi?i, mandi?ite, mandi?ok, mani?ok; W mani?o; Kb mani?ok; GM mandi: AT mani?aŋa; AX mani?ak(a); Kw mandi?o; Tp manihak, manihɨp 'manioc plant', manihɨwakɨŋ 'foliage of the manioc plant'; Tb mandi?ok; Gy mandi?o

131. manõ 'die'
     Km mano; Pt mano; UK manõ; W manũ; GM -mano; AT -mano; AX mano; Kw mano; Tb mano; Gy mano; Gj məno

132. mba?é reče 'why'
     Km ma?are; UK ma?e rehe, GM mba?e re;

133. mbaraká 'sing'
    Km -maraka; Pt mbaraka; UK maraka 'maraca used by shamans'; AX marakanïŋ;

134. mboβúk/mboβɨk 'sew'
    UK muwɨk; Kb momɨ; GM -mboβɨβɨ; AX -mombok, -mombak; Kw mboβɨβɨ;
    Tp momok, 'to bore'; Gy mboβo 'to divide in pieces'

135. mbočapɨr 'few'[2]
    Km moʔapɨt; Pt mbohapɨr 'three or more'; UK mahapɨr; Wj moapɨ, Wa mosapɨ;
    Kb muapɨt; GM mboapɨ; Kw mbohapɨ; Tp maãpɨt; Tb mosapɨr 'three';
    Gy mbocapɨ

136. (mbo)mbór 'throw'
    Km -momot; Pt -mombor; UK mbor 'throw away', mbor ombor 'throw ball';
    W -moma; AT -maman; AX -mamat; Kw mombo; Tp mamat; Gy mombo 'throw'

137. mbóy 'snake'
    moï; Pt mboja; Uk mɨya; Wj moi, Wa moe; Kb moy; GM mboi;
    AT mača; AX mbaiʔa; Kw mboy; Tp may; Tb mboy; Gy mboi 'snake' 'rainbow'

138. -mbukú 'dust'
    Pt itanimbukujuhu, -timbuɣuhu, ɨβɨmuʔmbuɣ 'dust'; UK mukuʔi 'sprinkle powder'

139. meʔéŋ 'give'
    Km meʔeŋ; Pt kʷavaʔeŋ; UK meʔē; W meʔē; GM -meʔē; AT -mon;
    AX mut; Kw -mēʔē; Tp meheŋ; meʔeŋ; Gy mēē

140. membɨr 'child'
    AX membɨra; Tp memɨt; Tb membɨr; Gy membɨr

141. mén 'husband'
    Wj mē, Wa men; Kb men; GM me; AT -mena; Kw -me; Tp men

---

[2] See 1.1.

142. miyár 'animal'
   Km miyat; UK miyarai; Tp mian 'game'; Gj miar 'animal'

143. mokõy 'two'[3]
   Km mokõy; Pt mokoñ; UK mokõy; W mokũy; Kb mukũy; GM mokõy;
   AT mokoy; AX mokoy, mokoiñe; Kw mokõy; Tp mokõy; Tb mokõy;

144. mombe?ú 'recount'
   Pt mombe?u; GM -mombe?u; Tb mombe?u; Gy mombe?u

145. nambí 'ear'
   Km -nami; Pt -nambi; UK -nambi; W nami; Kb -nami; GM nambi;
   AT -námia; AX nambi; Kw nambi; Tp nami; Tb nambi; Gy nambi

146. nupã 'beat'
   Km -nupã; Pt nupã; UK nupã; W nupã; GM nupã 'to thrash', 'tread on';
   AT -nopo; AX nopiŋ; Kw nũpã; Tp nopɨ; Gy nũpã

147. o?ó 'flesh'
   Km -a?o; Pt a?o; UK so?o 'game', so?oruk^war 'flesh of game'; W o?o, so?o 'game';
   GM -o?o; AT -ra?a; AX mama?ea?a, ma?eta?a, -ra?a; Kw so?o; Tp a?a; Tb so?ó

148. oβɨ 'green'
   Km -cowɨ; PT oβɨ 'blue'; UK -owɨ te; Wj sowɨ, Wa soβɨ; GM oβɨ; Kw -oβɨ, oβi;

149. óβ 'leaf'
   Pt -aβa, -oβa; UK -o, (ho, -ro); W o; Kb op; GM -ogue, -oβana 'with budding
   leaves', -oky 'new leaf'; AT hawa; Kw hogwe, hogwek^we; Tp ap; Tb óβ; Gy -o

---

[3] See 1.1.

150. ók 'house'
    Km ok; Pt onga, tuju?oguhu, okay; UK ok; W oka 'house', 'village'); Kb ok;
    GM -o, oguy; AT aŋa, aŋohoa; AX aka; Kw oy, óya; Tp akan 'house of men';
    Tb ók; Gy o, oɨɣ

151. ore 'we (exclusive)'
    Km ore; Pt ore; W ore; Kb ore; GM ore; AT ore; AX ore; Kw ore, orekʷérɨ;
    Tp are; Gy ore

152. páβ 'all'
    W pa; GM pa, opa, paβē; Gy pa

153. papár '(re)count'
    Km papat; GM papa; Kw -papa; Gy papa

154. paranã 'river'
    Km parana; Pt paranaŋ 'flexible, movable'; UK parana 'Gurupi R.'; W parana;
    Kb parana; AT paranoa; AX paranɨ; Gy para 'sea'

155. pé 'one'[4]
    mo-ye-
    o-ye-
    -te-ĩ
    -ʔĩ
    -yo-e
    Km moyepet; Pt ojipéji; UK peteĩ; W peʔĩ; Kb aype(y)tee; GM peteĩ;
    AT očepečowe; AX moiñepen; Kw peteĩ; Tp ãčape, ãčapeče 'more'; Tb oyepé;
    Gy ñepeĩ

156. pēē 'you all'
    Km pehē; Pt pe, pehe; UK pehē; W peyē; Kb pe; GM peē; AX pende;
    Kw pēē, pēēkʷérɨ; Tp peē; Gy pēē

---

[4] See 1.1.

157. pepó 'wing'
    Km pepo; Pt pepo, pepoatā; UK -pɨpo; Wj pepokā, Wa pepūkā; Kb pepo;
    GM pepo; AT pepa; AX pepa; Kw pepo; Tp pepa; Gy pepo 'feather' 'wing'

158. petɨm 'tobacco'
    Km petɨm; Pt mohatatīhaβ (mohatatiŋ = 'to smoke'); UK pitɨm; W petī;
    Kb pɨytem; GM petī; AT petɨma; AX petɨm; Kw petī; Tp petɨm; Gy pētī

159. peyú 'blow'
    Pt -peju; UK peyu; W peyu; GM -peju; Kw -peju; Tp pečo; Gy peyu

160. pín 'rub'
    Km -pin; Wj pipī 'scratch', Wa pī

161. pír 'skin'
    Km -piret; Pt -pir; pirwer; UK pirer; Wj pire, Wa pirer; Kb pit; GM pire;
    AT -piréra; AX pirera; Kw -pire; Tp pit; Tb pir/mbir; Gy pirer 'hide' 'skin' 'shell'

162. pirá 'fish'
    Km -pira; Pt pira; Uk pira; W pira; Kb ipira; GM pira; AT ipira; AX ipira;
    Kw pira 'big fish'; Tp hipirā; Tb pira; Gy pira

163. pirér 'bark'
    GM pire; AX ɨwɨra-pirera; Kw -pire; Gy pirer 'hide, skin, shell'

164. pitáŋ/mitáŋ 'child'
    Km pitaŋ; Pt pitaʔŋi 'baby'; Kw mitā; Tp pitɨŋ 'baby'; Gy pitani

165. pɨ 'foot'
    Km -pɨ; Pt -pɨ; UK -pɨ; W pɨ; Kb pɨ; GM pɨ; AT ípɨa; AX pɨ; Kw pɨ; Tp mɨ;
    Tb pɨ/mbɨ; Gy pɨ

166. pɨʔá 'liver'
    Pt pɨʔa; UK pɨʔa; W pɨʔa, Wj pɨʔak*e, Wa pɨakʷer; Kb pɨʔa; GM pɨʔa; AT pɨʔa;
    AX pɨʔa; Kw pɨʔakʷe; Tp mɨhã; Tb pɨʔá/mbɨʔá; Gy pɨʔa 'heart, stomach, innards'

167. pɨčačú 'new'
    Km pɨau; Pt -pɨahu; UK pɨahu; W pɨau; GM pɨau; AT iʔɨahoa; AX ʔɨaho;
    Kw pɨahu; Tp hɨão; Tb pɨsasu; Gy pɨacu

168. pɨčɨ́k 'catch'
    Km pɨhɨk; Pt pɨhɨγ 'take'; UK pɨhɨk; GM jopɨ, pɨ; AT pɨhɨ; AX pɨhɨk; Kw pɨhɨ;
    Tp pɨik, 'take, keep, buy'; Gy pɨcɨ

169. pɨpɨ́r 'wide'
    Km pɨpɨt; Pt pɨpir; GM py; Kw -py; Tp pɨpit

170. pɨtér 'suck'
    Km pɨtet; Pt pɨter; UK pɨter; GM pɨte; Kw -ipete; Tp pɨtet

171. pɨtú 'breath'
    pɨtu-čem 'breathe'
    Km -yepɨtuerut; Pt -pɨtu; AT ipytohem; AX pɨtohem; Kw pɨtũhē; Gy pɨtu 'brisk',
    'quick, buoyant'

172. pó/mbó 'hand'
    Km -hwã; Pt -po; UK po; W po; Kb po; GM po; AT ípaa; AX -pa; Kw po;
    Tp ma; Tb pó/mbó; Gy po

173. poʔí 'thin'
    Km poʔi; UK puʔi; W poʔi; GM poʔi, pɨʔɨʔi; Kw -poʔi; Gy poʔi/pohi

174. po-apē 'fingernail'
    pɨ-apē 'toenail'
    Km ɨhwapē; Pt -puapē, -pypē; UK poāpē, pyāpē; W puápē, pɨápē; Kb pāpē;
    GM apē 'nail', poapē 'fingernail'; AT koape; AX poāpe; Kw -pɨape; Tp mɨāpe;
    Tb apē; Gy poāpē

175. poán (?< poayán) 'twist'
    Pt poanoŋ; UK puwan, puwā

176. počíy 'heavy'
    Km -powɨy; Pt -pohɨy; UK puhɨy; Wj poɨy, Wa po(w)ɨy; Kb -poɨyoo, puɨ(y)uu;
    GM poɨy; AT -pohoy; AX pohoy; Kw pohɨy; Tp pooy; Tb posíy; Gy pocɨy

177. poračéy 'sing, dance'
    GM porae; AT porahay 'dance'; AX porahay 'dance'; Kw porahey; Tb poraséy;
    Gy poracey 'song, to sing'

178. poti?á 'chest'
    Km -poci?a; UK piši?a; W posi?a; GM poši?a; AT -póčia; AX poči?a;
    Kw -pɨti?a; Gy počia

179. potír 'flower'
    ?ɨβ-potɨr 'blossom'
    Km potɨt; Pt iβatɨ?ri, iβɨtɨr; UK putɨr; Wj potɨ, Wa potɨr; Kb ipotɨt; GM potɨ;
    ?iβotɨ; AT ɨwotera, ɨpa-potɨra 'tree blossom'; AX ɨwɨra, potɨra; Kw iβotɨ;
    Tp hɨwatɨt; Tb potír, iβotír; Gy potɨr

180. potuká (?< po-kutuk) 'clean'
    Km potuka; Pt -potuka; UK kutuk; Tp patokā 'wash clothes'; Gy potuka

181. puká 'laugh'
    Km -huka; Pt -puka, -pukaβa?ē; UK pukʷa, puka; W puka; Kb ɸukaíta 'shout';
    GM -puka; AT -puka; AX poka; Kw pukay; Tp hãpokãy 'smile'; Gy puka

182. pukú 'long'
    Km huku; Pt puku; UK puku; W poko, -puku; Kb ɸuku; GM -puku, -βuku;
    AT poko; AX poko; Kw puku; Tp poko; Tb puku; Gy puku

183. pʷár 'tie (up)'
    Km -hʷat; Pt -kʷar; UK pukʷar (pukar); W -(o)kʷa; Kb -ɸar; GM jokʷa, jojokʷa,
    -kʷa; Tp kʷan; Tb pʷár

184. pʷeráβ 'cure'

    W puera; GM -kʷera; Tb pʷeráβ; Kb φerap 'recover'; Gy kʷera

185. ro?ɨ 'cold'

    ro?ɨ-tiáŋ

    Km ero?icaŋ; Pt -ro?ɨ, roɨtiaŋ; UK irɨsá ipe 'he is cold'; W ro?ɨ; Kb iroɨcaŋaɨ; GM ro?ɨ, ro?ɨ šá, ɨro?ɨ; AT -ro?ihɨŋ (object), -ro?ɨ (body); AX -rowɨŋ, -rohɨŋ; Kw ro?ɨ; Tp ő?ɨɨŋ; Tb ro?ɨ; Gy ro?ɨ

186. rurúk 'swell'

    UK jupururuk; W ruru; GM ruru (swelling), -mboruru 'make swell'; Kw ruru (swollen); Gy ruru

187. tanimbúk (?< tatá imbúk), 'ashes'

    UK tatimbuk; Wj tanimu, Wa tanemu; Kw tanimbu; Gy tanimbu

188. tapi?ír 'tapir'

    Km tapi?it; Pt tapi?ir; UK tapi?ir; Wj tapi?i, Wa tapi?ir; Kb tapi?it; GM tapi?i; AT tapi?íra; AX tapi?ira; Tp tápihit

189. tíŋ 'white'

    Km -ciŋ; Pt tiŋ, kɨtiŋ, roɨtiaŋ; UK šī; Wj sī, Wa siŋ; Kb ciŋaɨ; GM šīī; AT siŋ; AX čiŋ; Kw morotī; Tb tiŋ; Gy čī, moročī

190. tī 'nose'

    Km cī; Pt -tī, -ti?a?mbi; W sī; Kb cī; GM šī; AT íčia; AX či; Kw tī; Tp čī; Tb tī; Gy čī

191. -tɨk 'pull'

    mbondɨk, 'to blow'

    Km -monɨk; Pt -ojaɨtɨγ, mondɨγ 'light a fire'; AT -monɨŋ; AX mondɨk 'pull', 'blow'; Tp manɨik 'stack up, bring together'

192. túβ 'father'
   Km tup; Pt -uβ; UK -ru; W u; Kb tup; GM -u; AT tope ('his father'); AX -roa;
   Kw túβɨ; Tp op; Tb -úβ; Gy -u

193. tuβiyáβ 'big'
   Km tuyap; UK tiha; Kb -tuwioóye; GM tuβiša; Kw tuβiša; Gy tuβiča

194. tuyá 'old'
   GM tuja; Kw tuja; Gy tuya

195. uʔíβ 'arrow'
   Km iʔip; Pt -uʔiβ, -iʔipakʷar, -iʔiβ; UK uʔɨ; GM -uʔi, huʔɨ; AT oʔiwa; AX oʔiwa;
   Kw huʔɨ; Tp oʔip

196. úβ 'leg'
   Kb -up; Pt -kupy, -py; AT owákowa; AX uwa; Tp hɨp, 'leg, end', kopy, 'leg';
   Gy -u

197. upiʔá 'egg'
   Km -upiʔa; Pt upiʔa; UK ipiʔa; W upiʔa; Kb ʔupiʔa; GM -upiʔa; AT -opiʔa;
   AX opiʔa; Kw -upiʔa; Tp opihá; Tb -upiʔá

198. úr 'come'
   yor 'come'
   Km -ʔut; Pt -ur, jor; UK -wɨr, tur; W u; GM -uβɨ, -ju, -juβɨ, turɨ; AT on; AX -ut;
   Kw -u; Tp čar, 'come', čat 'arrive'; Tb úr; Gy -u/-yu

199. uwɨ́ 'blood'
   Pt -gwɨ; UK -uwɨ; W uwɨ; Kb -rɨ; GM -ugwɨ; AT wia, -owia; AX ohɨ; Kw -ugwɨ;
   Tp owɨ; Tb -uwɨ́; Gy -uɣʷɨ

200. uwáy 'tail'
   Km way; Pt gwaja; UK -uway; W uway; Kb way; GM ugway; AT wača, way;
   AX way; Kw -ugwáy; Tp owãy; Gy -uɣʷay

201. waiwī 'old (woman)'

   UK aʔi; W waimi 'woman'; GM gwaimi; Gy γʷāyβī 'old age (for women)', 'old woman'

202. -wáŋ 'red'
   pɨtáŋ
   pír-(w)áŋ 'red skin'
   Km -waŋ; Pt ŋwaŋ 'red', -piraŋ/-miraŋ/-βiraŋ ('to have rough, red skin'); UK -pirā; W pirā, pɨtā; Kb piraŋaɨ; GM pɨtā; AT -piroŋ; AX pirūŋ, pirāŋ, pirā; Kw pɨtā; Tp pirɨŋ; Tb pɨtáŋ; Gy pītā 'brown, red-brown'

203. weʔén 'vomit'

   UK weʔen, weʔē; W weʔē; AT -weʔeŋ; AX oʔen; Kw ŋwēʔē; γʷēē

204. wɨrá 'bird'

   Km wɨra; Pt gwɨra; UK wɨrahu te 'eagle'; W wɨra; Kb wɨra; GM gwɨra; AT wɨra; AX wɨra; Kw gwɨra; Tp wɨrāhi, wɨrakáy, 'hen'; Tb wɨrá; Gy γʷɨra, γʷɨraī

205. yɨ 'ax'

   Km yɨ; Pt ñumbere; UK yi; W yɨ; AT čɨa; AX džɨ; Tp čɨ

206. yaʔwár 'jaguar'

   Km yawat 'jaguar', wararuyap 'dog'; Pt jaʔγwar 'jaguar', ñaŋwatiŋ, ingaruruʔī 'dog'; UK yaŋgwate, yawarete 'jaguar', yawar, yawaruwir 'dog', yawan 'wild dog, fox'; Wj yawa, Wa yawar; Kb yaʔwat; GM jaγʷa 'dog'; AT čawára 'jaguar, dog'; AX džawara 'jaguar, dog'; Kw jaγʷarete 'jaguar' jaγʷa 'dog'; Tp čawan 'jaguar, dog'; Tb yawár 'jaguar'; Gy yaγʷar 'jaguar, fat-eater, falling star'

207. yačɨ 'moon'

   Km yaɨ; Pt jahɨ; UK yahɨ; W yaɨ; Kb yaɨ; GM jaŠɨ; AT čahɨa; AX džahɨ; Kw jasɨ; Tp čāɨ; Tb yasɨ; Gy yacɨ 'moon' 'month'

208. yačɨ-tatá 'star'

   Km yaɨtataʔi; Pt jaɨtataʔi; UK yahɨ rata; W yaɨtata; Kb yaɨtata; GM jaŠɨ tata; AT čahɨtata; AX džahɨtɨta; Kw jasɨtata; Tp čāɨtātāhi; Gy yacɨtata

209. yakaré 'crocodile'
   Km yakaré; Pt jakare, jakareti?ŋgi; UK yakare; W yakare; Kb yakare; GM jakare;
   AT čakaré; AX džakaré; Kw jakare; Tp čãkãre; Gy yakare 'caiman'

210. yán 'run'
   Pt ñan; UK yākʷen, yan; GM -ña; AX ñɨn; Tp čan, 'to gather', 'to pick'; Gy ña
   'leap'

211. yande 'we (inclusive)'
   Km yene; Pt ñande; UK yande 'we'; W yane; Kb yane; GM ñande; AT čane;
   AX žande; Kw ñande; Tp čane; Tb yande/ñande; Gy ñande

212. (mbo)yár 'tighten'
   Pt -jar, -mbojar; UK muyar; GM mbopija; Gy mboya, -ya 'fasten, stick to
   something'

213. -yarú 'to play'
   Km -puyaru; GM -mbojaru; Gy yaru 'play', mboyaru 'make fun of someone'

214. yáy 'laugh'
   Pt -jay; Kw -jay; Tp čayka 'cry'; Gy yay

215. ye?éŋ 'speak'
   Km -ye?eŋ; Pt ñi?iŋ; UK ye?ẽ; W ye?ẽ; GM ñe?ẽ 'speech'; AT -če?eŋ; AX že?eŋ;
   Kw -ñẽ?ẽ; Tp čɛheŋ; Gy ñẽẽ 'speech' 'word' 'sound' 'voice'

216. ye-βɨr 'return, come back'
   Km -yewɨt; Pt -jiβir, -βɨr 'come back from the grave'; UK yɨwɨr; GM -jeβɨ;
   AT -čewɨn; Kw -jeβɨ; Tp čewɨt; Gy yeβɨr, yeɣʷɨr 'back' 'again'

217. yu?ū 'grass, weeds'
   Pt ñuŋgwaβ; GM ñu?ū; AT čoowia; Tp čõ

218. yúβ 'yellow'

Km yup; Pt juβ 'yellow with fever'; UK -ju; W yu; GM apeju, ju; Kw -ju, saʔɨju; Gy yu 'yellow, whitish, pale'

219. yúk 'rotten'

Pt -juɣ, -kujuɣ; UK -yuk

220. yuká 'kill'

Km -yuka; Pt -juka; UK yukʷa, yuka; W yuka; GM -juka; AT -čoka; AX džoka; Kw juka; Tp čokā; Tb yuka; Gy yuka

221. yukɨrɨ 'yellow'

Pt jukɨrɨ; Kb yukɨriai; AT ičokɨri(a); AX ɨžókɨra; Gy yukɨri ɨ 'a tree whose seed when mixed with lemon turns red'

222. yukɨr 'salt'

Km yukɨt; Pt jukɨr, ñukɨpan; UK yukɨr; Wa yukɨr; GM jukɨ; AT čokɨra; Kw jukɨ; Gy yukɨr

223. yúr 'neck'

Km ayut; Pt -jur, -jurupɨ; UK yurupɨ; AT ičóroa; AX žora; Kw -ajúra, -ajurapɨ; Tp čot

224. yurú 'mouth'

Km yuru; Pt juru; UK yuru; W yuru; Kb yuru; GM juru; AT čóroa; AX žóro; Kw juru; Tp čoro; Tb yurú; Gy yuru

Appendix II
Isogloss Maps
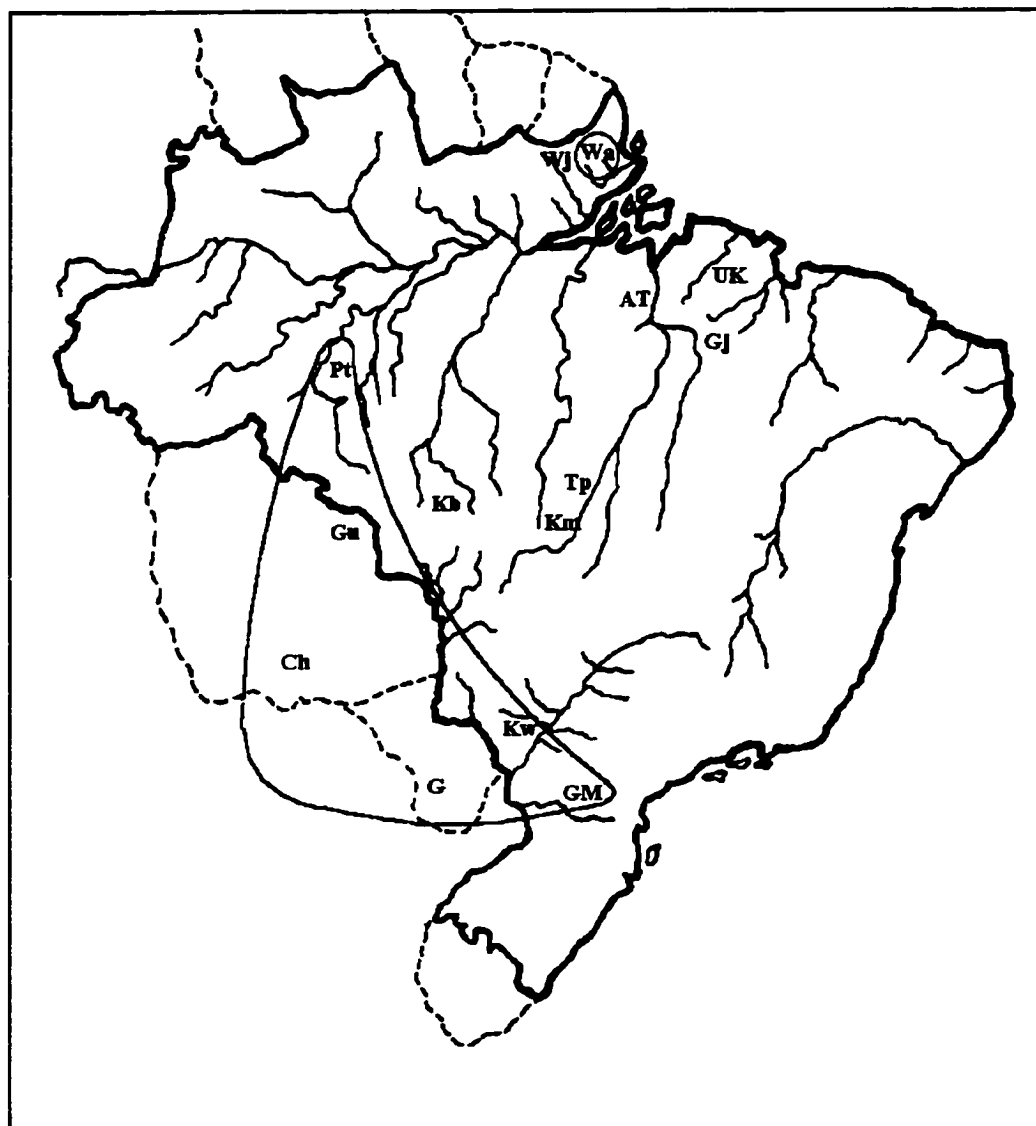
preservation of *β
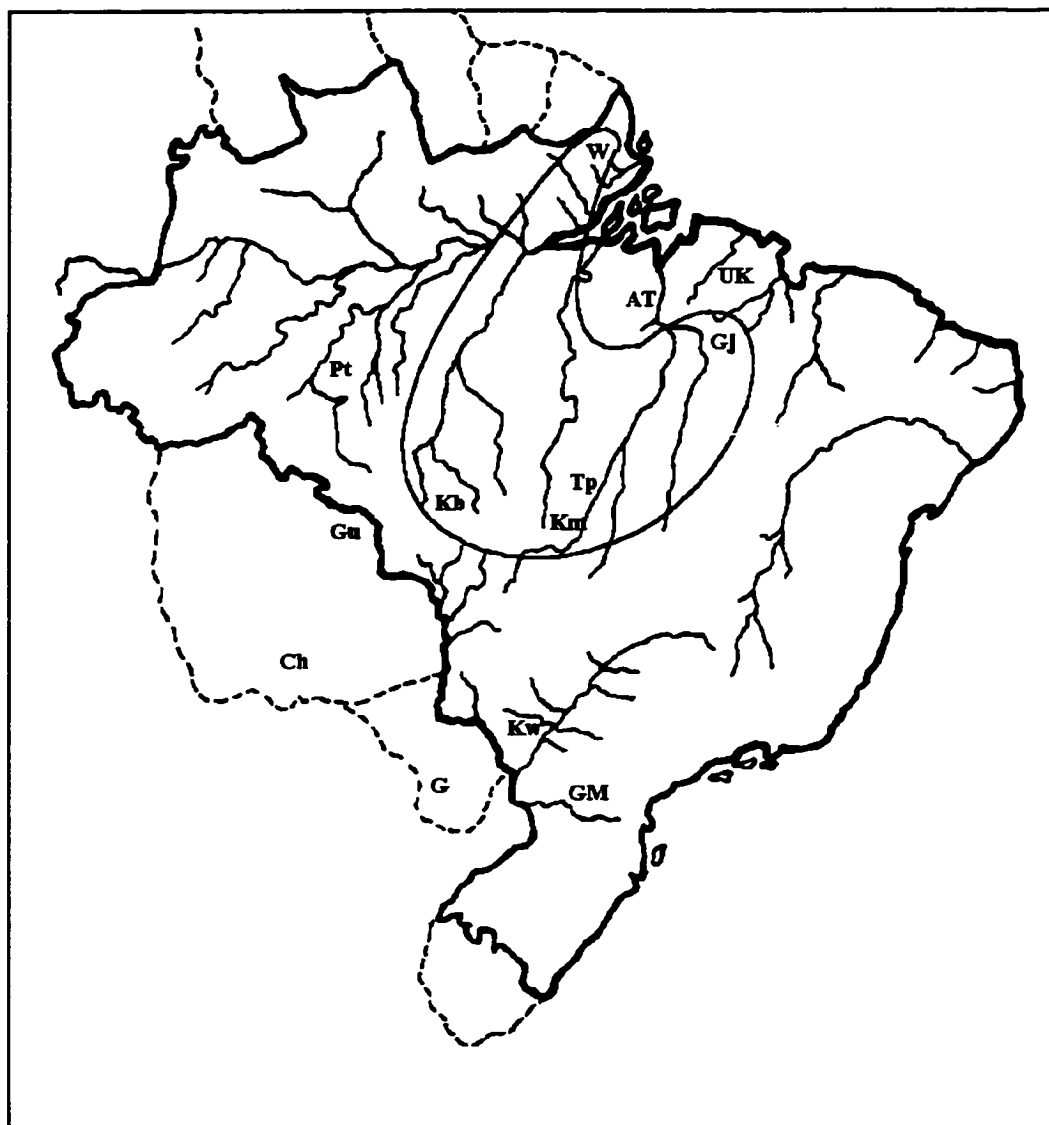
βaʔe > maʔe
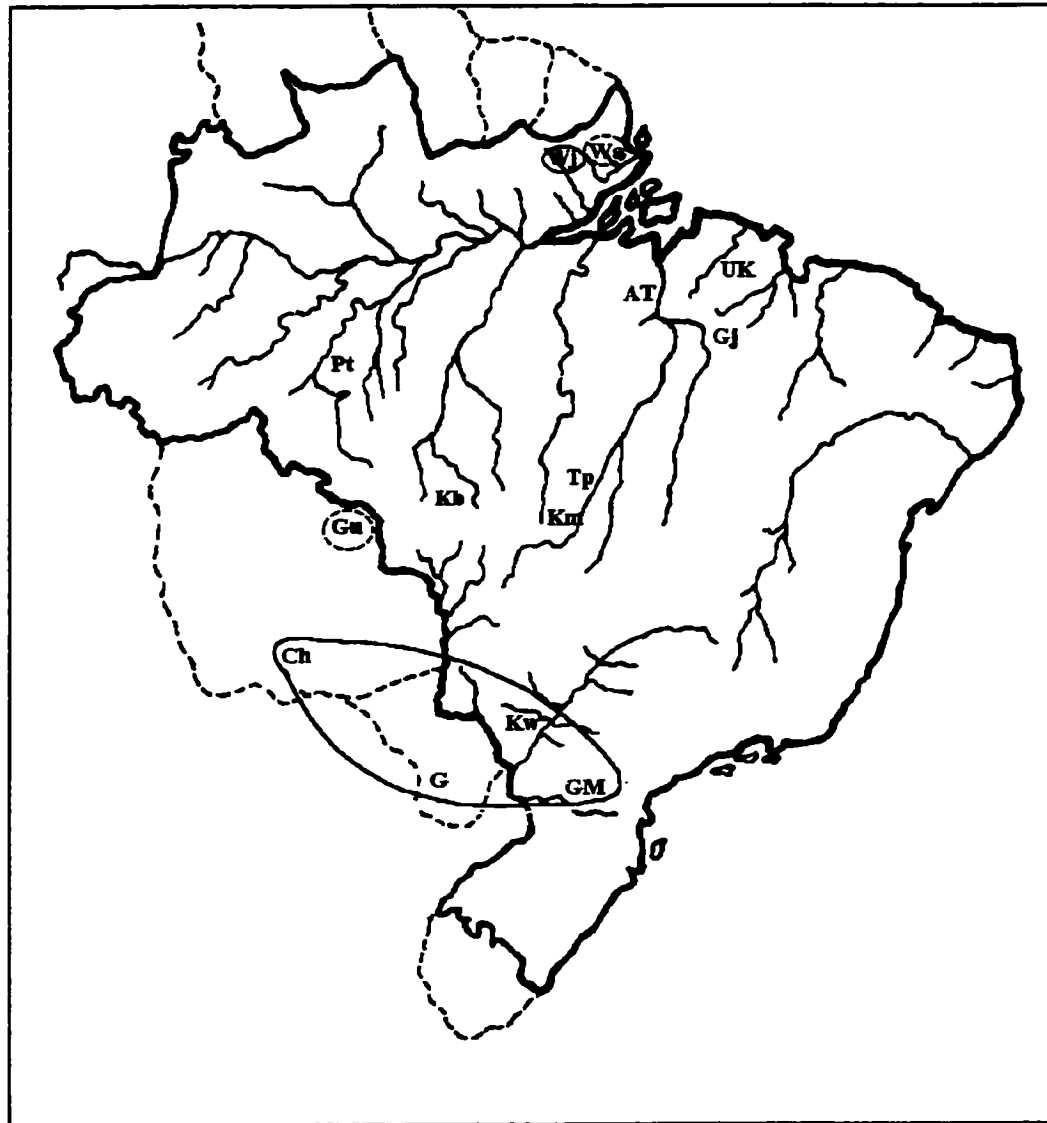
loss of final consonants
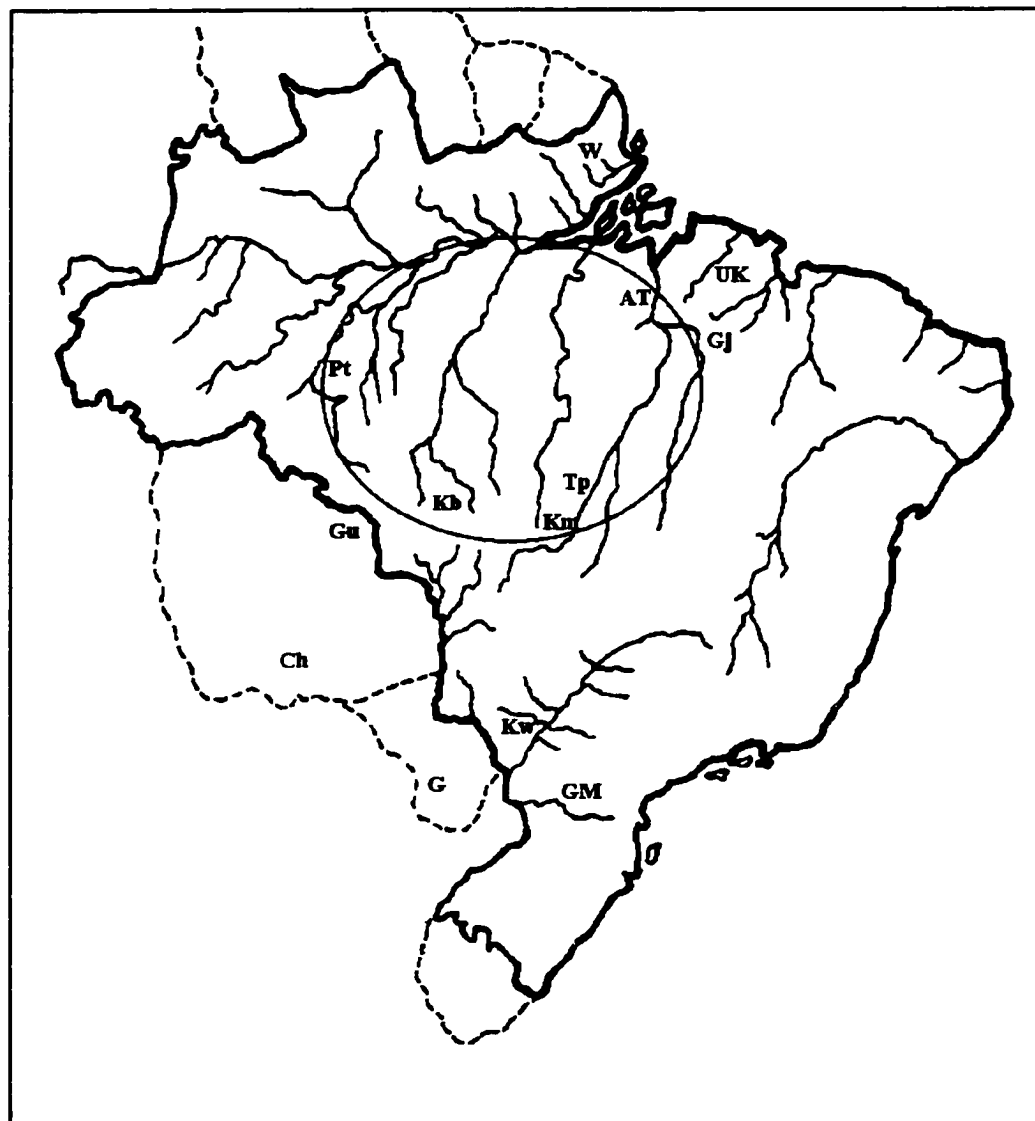
having the obtop/focus nominalizer *-*i*
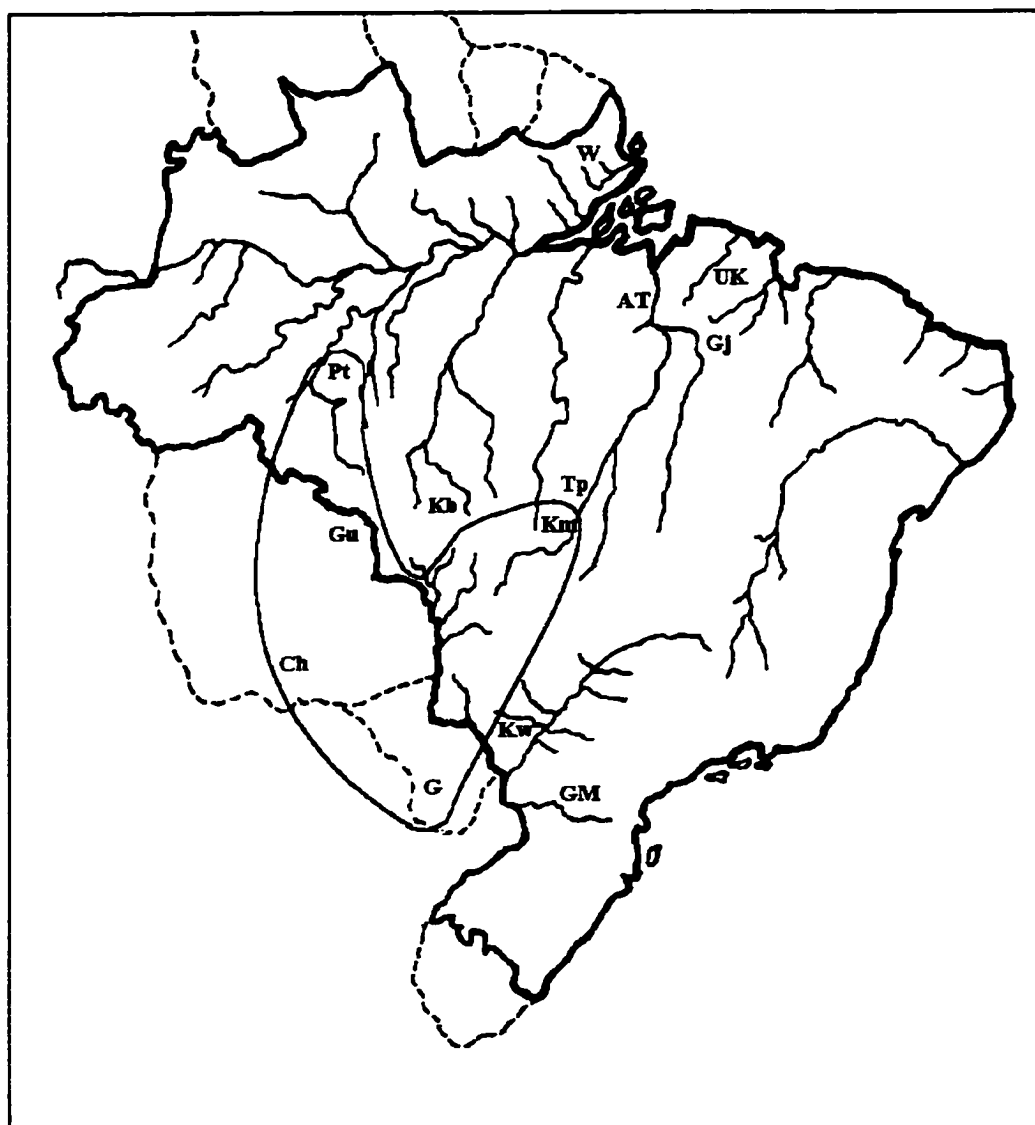
*oro-/*opo- not lost or altered by analogy
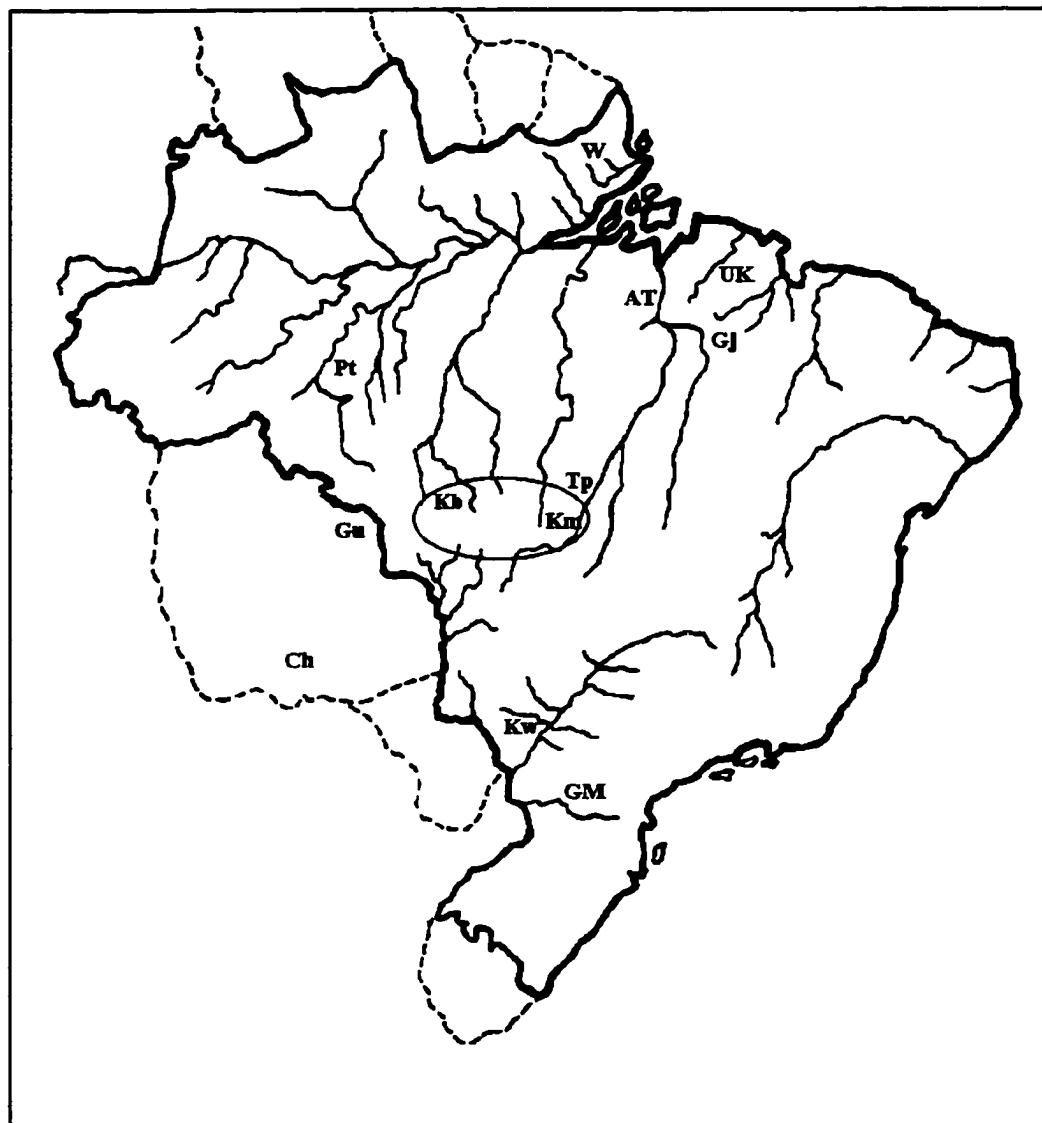
distinction between original $*p^w$ and $*k^w$
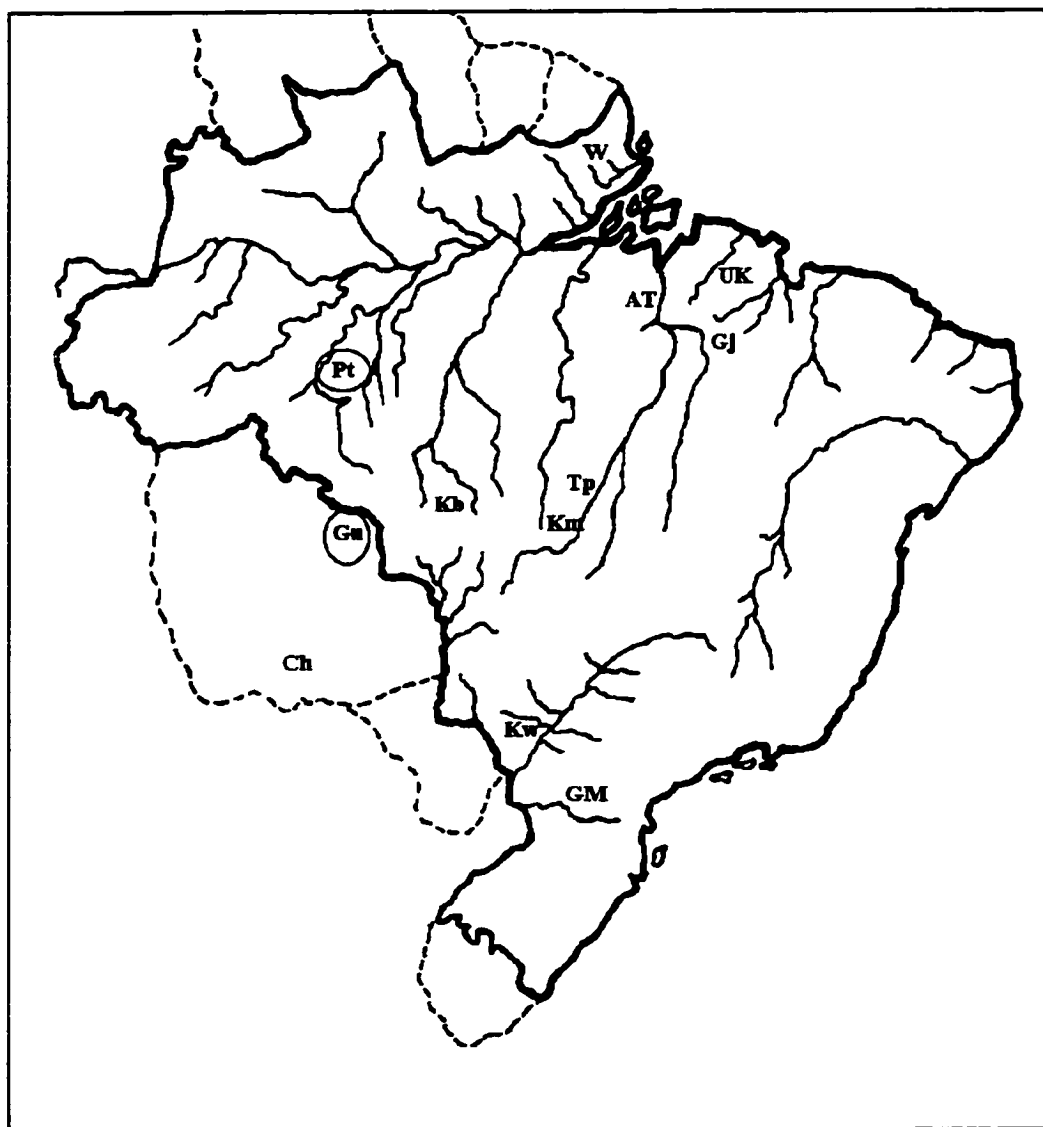
preservation of p^y
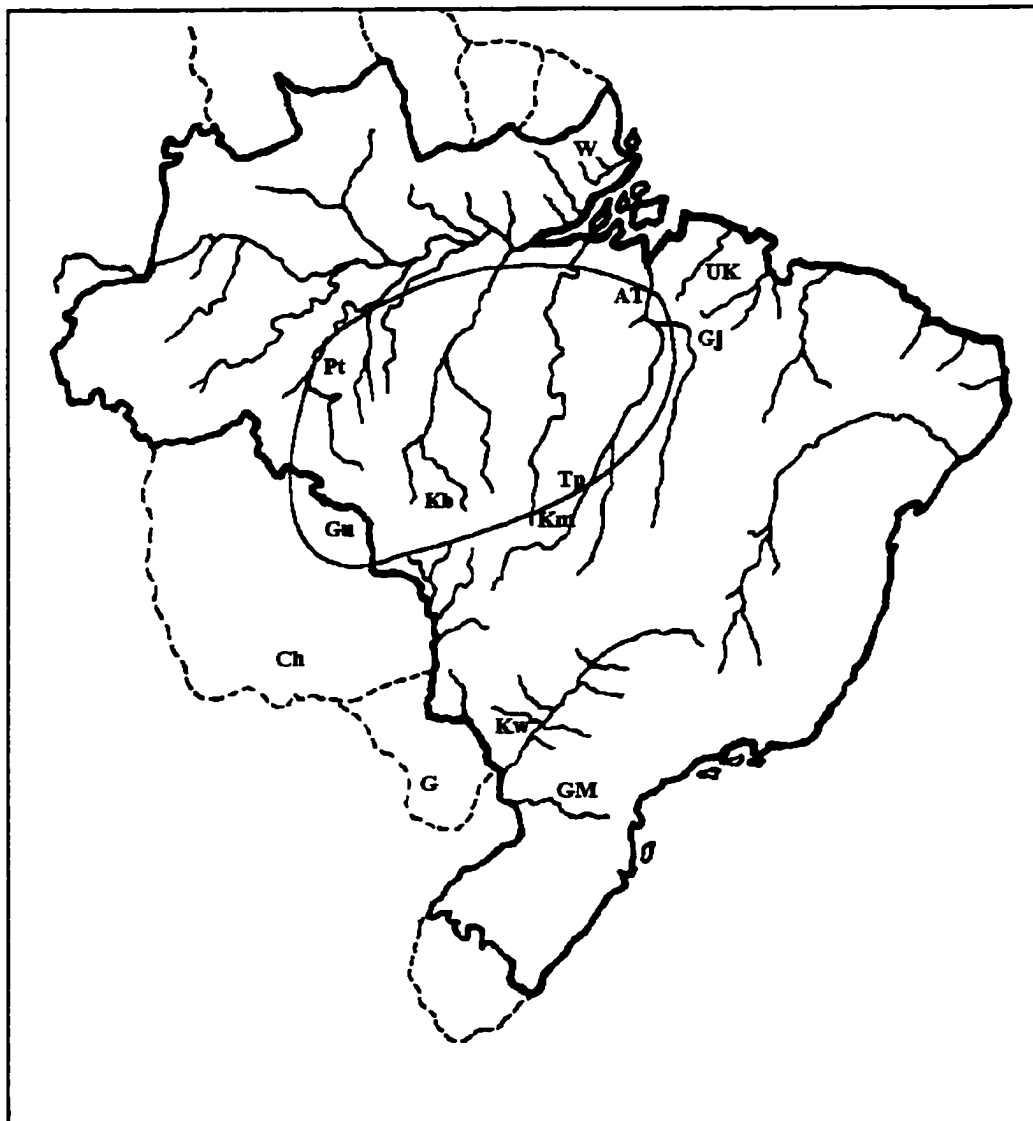
having Set 3 pronominal markers
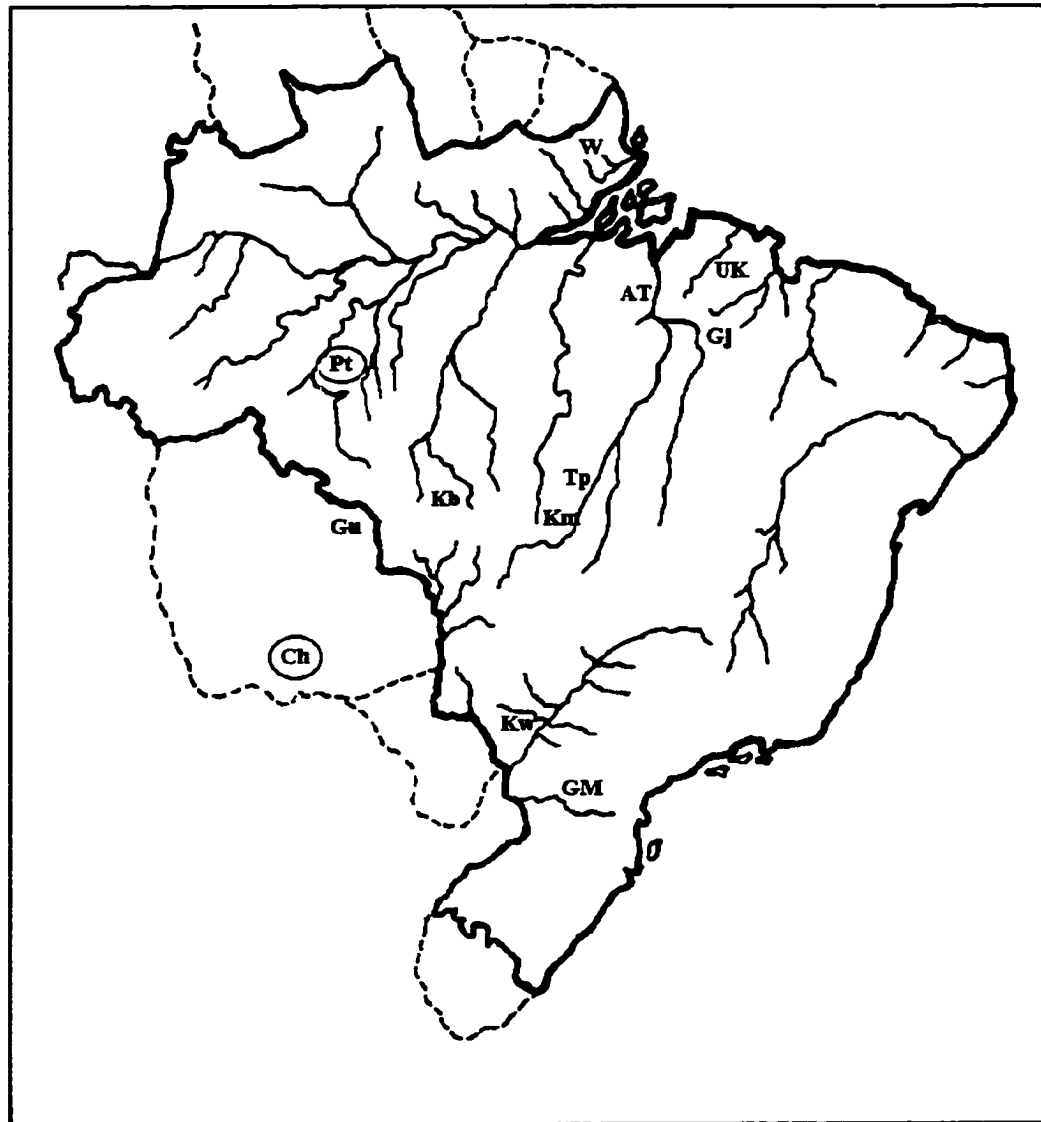
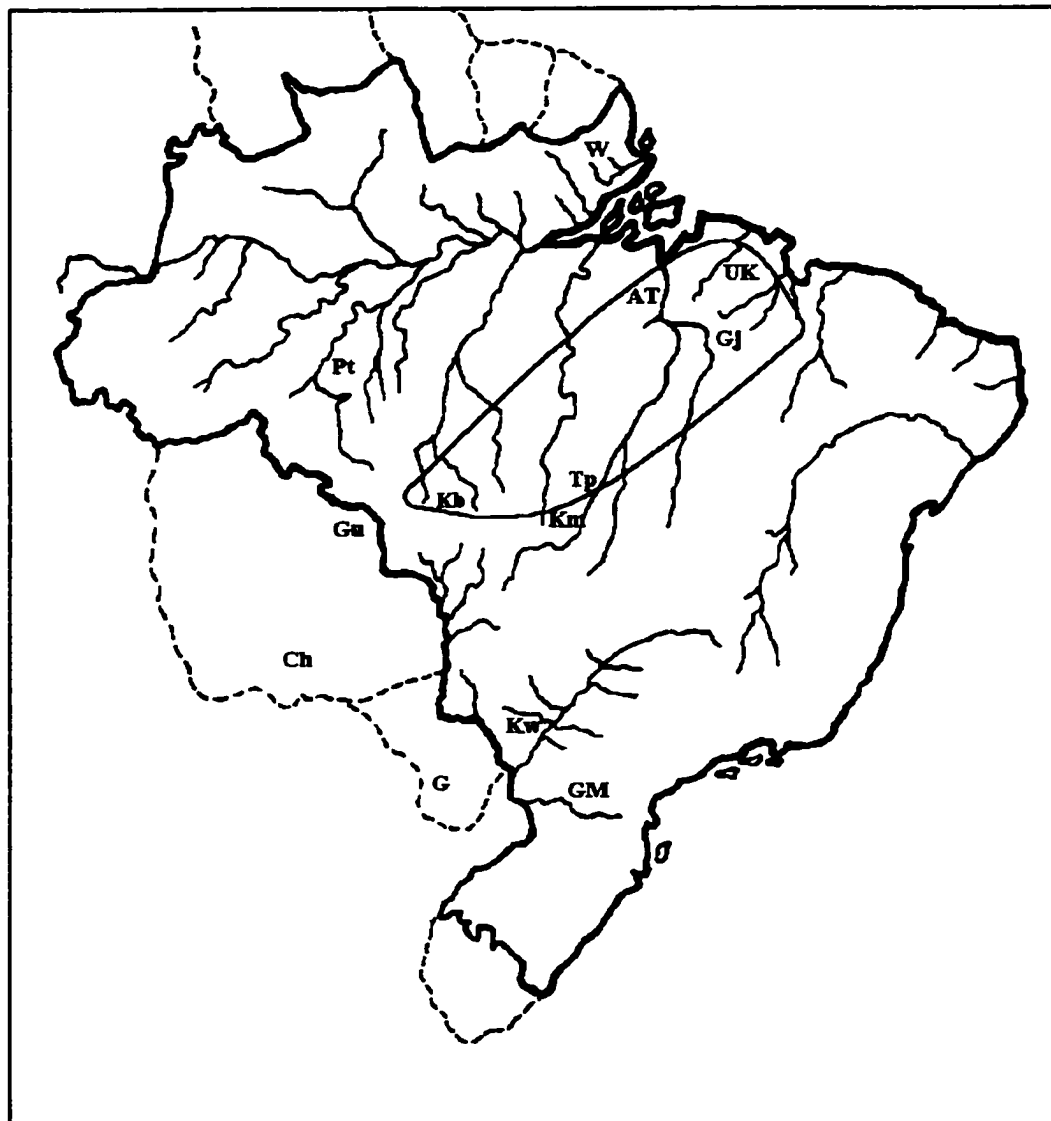preservation of [ti]
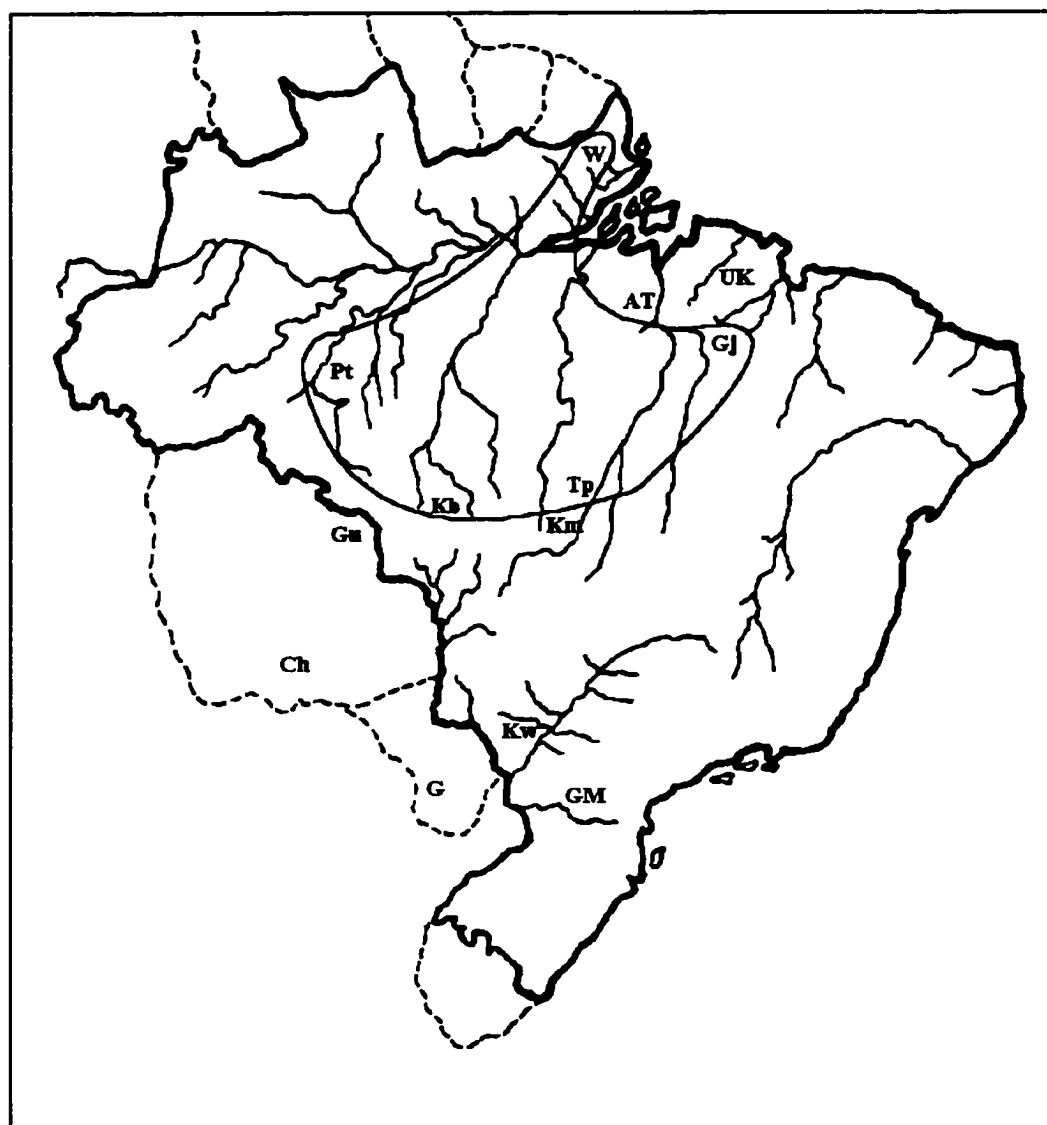
*ya-/*ti-/*tya- alternation grammaticalized
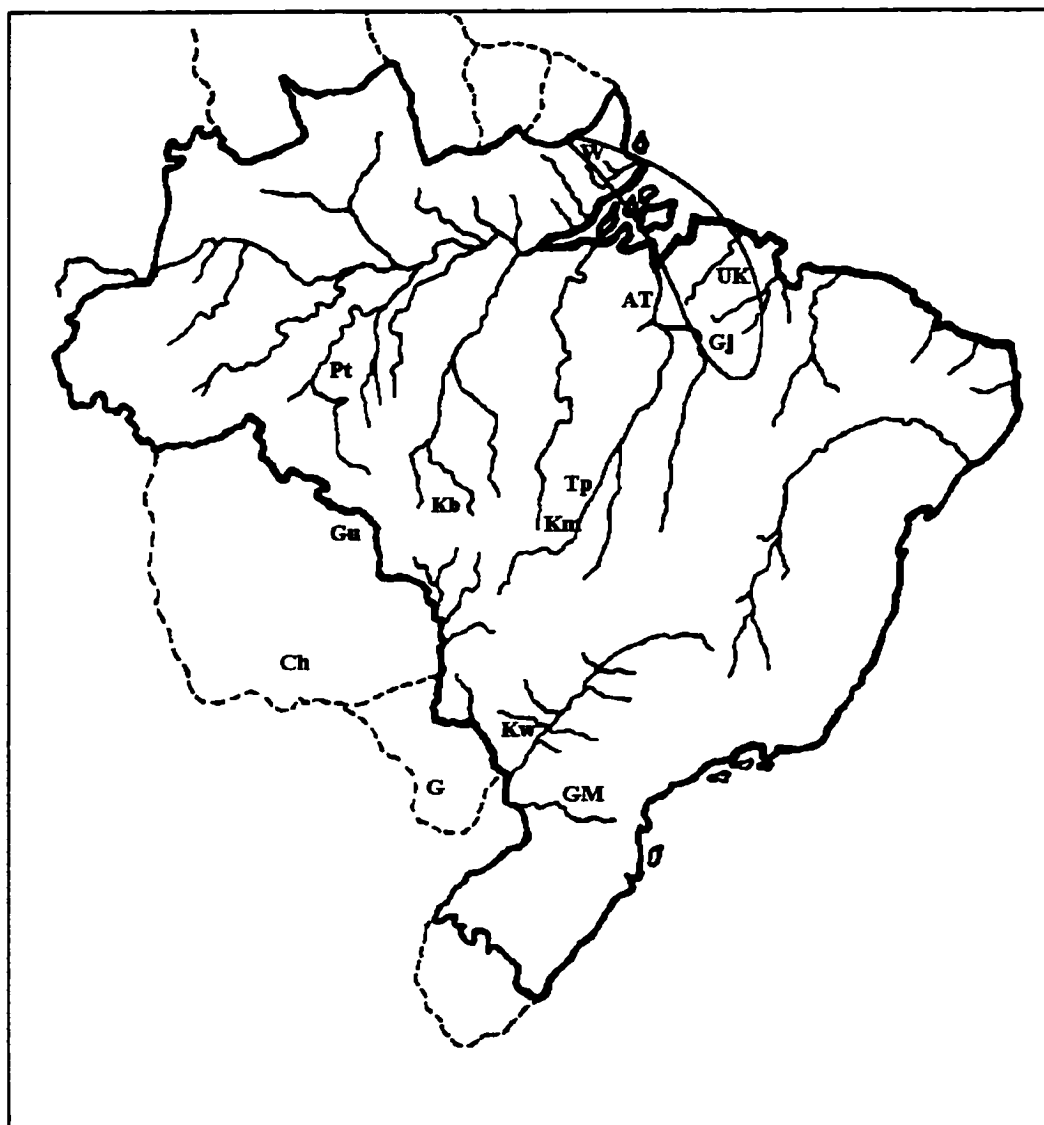
ye- and *yo-* merge
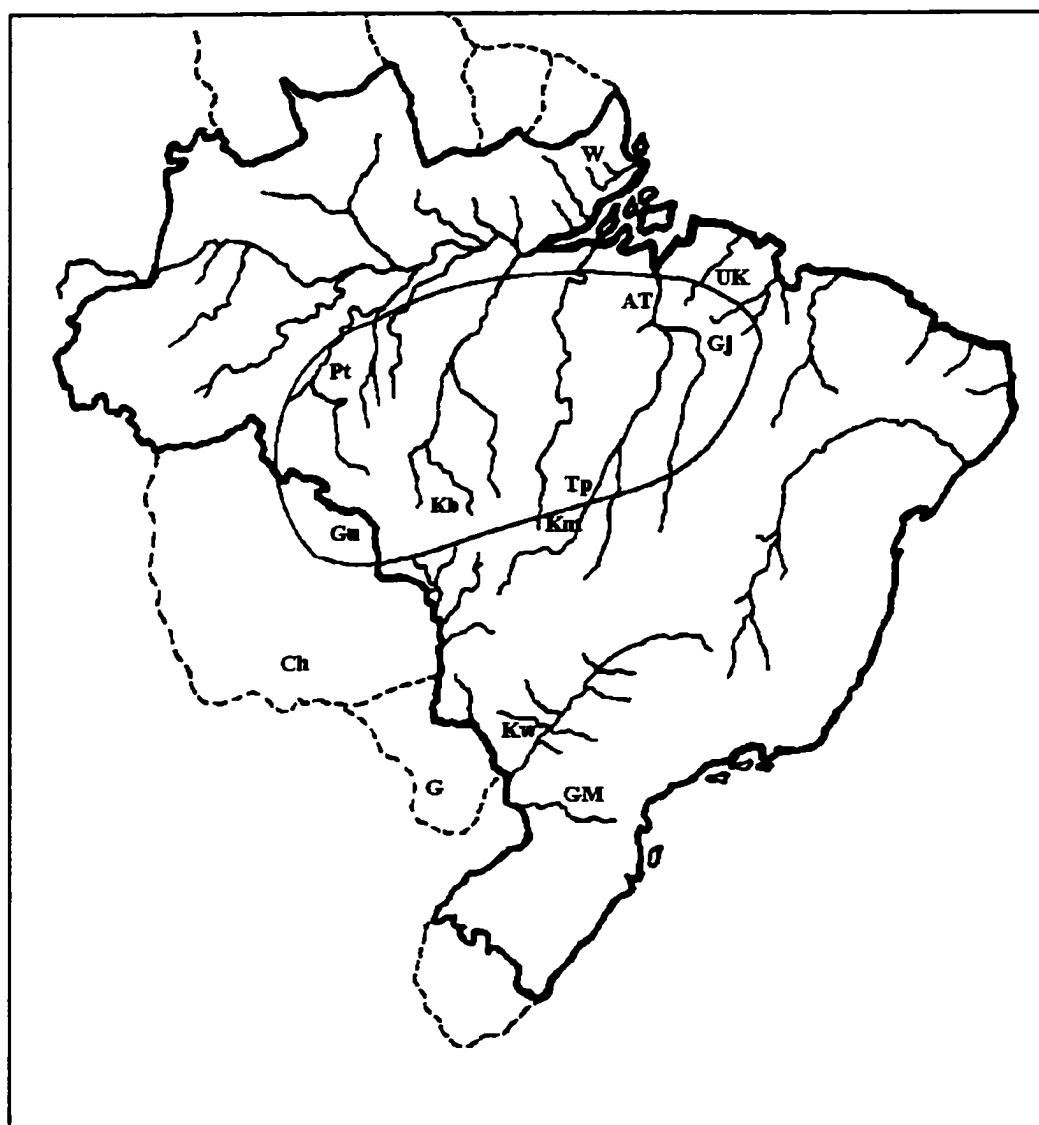
having reflex of PTG *epe*
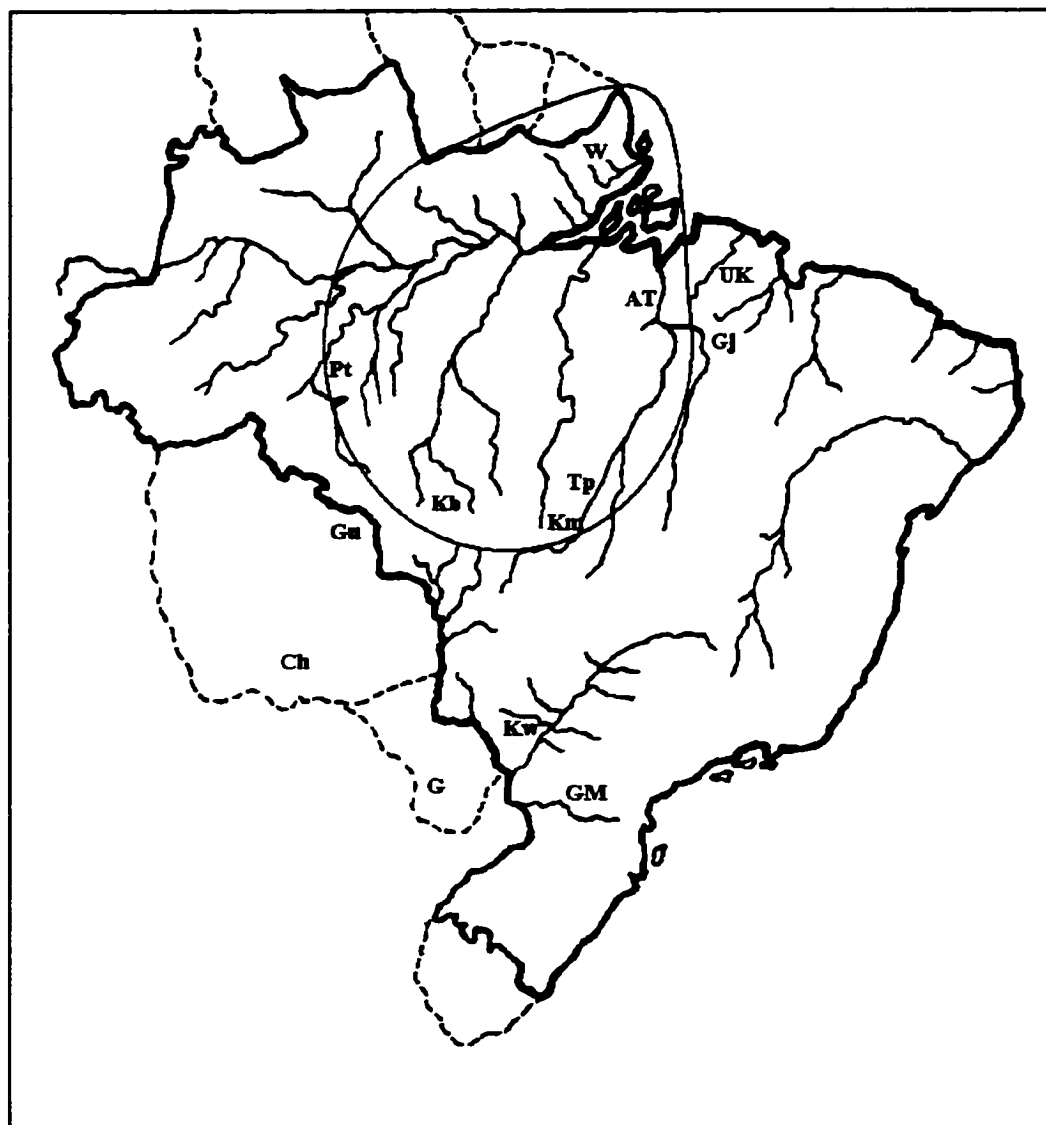
having reflex of PTG 1s *ye

loss of autosegmental nasality

REFERENCES

Almeida, Antonio, Irmãzinhas de Jesus, and Luíz Gouvéa de Paula. 1983. A língua Tapirapé. Rio de Janeiro: Xerox.

Anchieta, Joseph de. 1595. Arte de grammatica da lingoa mais usada na costa do Brasil. Coimbra. New edition by Julio Platzmann, Leipzig 1874.

Bendor-Samuel, David (ed.). 1971. Tupi studies I. Norman: Summer Institute of Linguistics and the University of Oklahoma.

Bendor-Samuel, David. 1972. Hierarchical Structures in Guajajara. Norman: Summer Institute of Linguistics and the University of Oklahoma.

Betts, La Vera D. 1981. Diccionário Parintintin-Portugues Portugues-Parintintin. Brasília: Summer Institute of Linguistics.

Cabral, Aan Suelly de Arruda Camara. 1995. Contact-Induced Language Change in the Western Amazon: The Non-Genetic Origin of the Kokama Language. Doctoral dissertation. University of Pittsburgh.

Derbyshire, D.C. and G.K.Pullum (eds.). 1986. Handbook of Amazonian Languages, Vol. 1. Berlin: Mouton de Gruyter.

Diebold, A. Richard, Jr. 1960. Determining the centers of dispersal of language groups. IJAL 26:1-10.

Dietrich, Wolf. 1986. El Idioma Chiriguano. Madrid: Instituto de Cooperación Iberoamericana.

Dietrich, Wolf. 1990a. Chiriguano and Guarayo Word Formation. Amazonian Linguistics: Studies in Lowland South American Languages. Ed. by Payne, 117-158. Austin: University of Texas.

Dietrich, Wolf. 1990b. More evidence for an internal classification of Tupi-Guarani languages. Indiana 12. Berlin: Gebr. Mann.

Dobson, Rose M. 1988. Aspectos da lengua Kayabí. Série lingüística 12. Brasília: Summer Institute of Linguistics.

Dooley, Robert A. 1982. Vocabulario do Guarani (Dialeto Mbyá do Brasil). Brasília: Summer Institute of Linguistics.

Dooley, Robert A. 1990. The Positioning of Non-pronominal Clitics and Particles in Lowland South American Languages. Amazonian Linguistics: Studies in Lowland South American Languages. Ed. by Payne, 457-494. Austin: University of Texas.

Dubinsky, Stanley. 1989. Antipassive in Hebrew: The Case for Extended Valence (ms.).

Dyen, Isidore. 1956. Language distribution and migration theory. Language 32.4:611-626.

Figueira, Luiz. 1687. Grammatica da lingua do Brasil. Facsimile of the 1687 a̅m republished by Julio Platzmann, 1878. Leipzig: B. G. Teubner.

Granado, Juan (ed.). 1891. Pequeño Ensayo del Idioma Guaraní. Asunción: El Paraguayo.

Harrison, Carl H. 1975. Gramática asurini. (Série Linguística IV). Brasília: Summer Institute of Linguistics.

Harrison, Carl H. 1986. Verb prominence, verb initialness, ergativity and typological disharmony in Guajajara. Handbook of Amazonian Languages, Vol. 1. Ed. by Derbyshire and Pullum, 326-403. Berlin: Mouton de Gruyter.

Harrison, Carl H. and John M. Taylor. 1971. Nasalization in Kaiwá. Tupi studies I. Ed. by Bendor-Samuel 1-14. Norman: Summer Institute of Linguistics and the University of Oklahoma.

Hoeller, Alfredo. 1932. Grammatik der Guarayo-Sprache. Guarayos (Bolivia) and Hall in Tirol (Austria): Verlag der Missionsprokura der P.P. Franziskaner.

Jensen, Cheryl. 1990a. O desinvolvimento histórico da língua Wayampi. Master's Thesis. Campinas: Universidade Estadual de Campinas.

Jensen, Cheryl. 1990b. Cross-referencing changes in some Tupí-Guaraní Languages. Amazonian Linguistics: Studies in Lowland South American Languages. Ed. by Payne, 117-158. Austin: University of Texas.

Kakumasu, James. 1986. Urubu-Kaapor. Handbook of Amazonian languages, Vol. 1. Ed. by Derbyshire and Pullum, 326-403. Berlin: Mouton de Gruyter.

Kakumasu, James and Kiyoko Kakumasu. 1990. 15 textos Urubu-Kaapor. Brasília: SIL.

Leite, Yonne de Freitas. 1977. Aspectos da fonologia e morfofonologia Tapirapé. Série Linguística VII. Rio de Janeiro: Museu Nacional.

Leite, Yonne de Freitas. 1982. A classificação do Tapirapé na família Tupi-Guarani. Ensaios de Lingüística 7.25-32. Belo Horzonte: Universidade Federal de Minas Gerais.

Lemle, Miriam. 1971. Internal classification of the Tupí-Guaraní linguistic family. Tupi studies I. Ed. by David Bendor-Samuel, 107-129. Norman: Summer Institute of Linguistics.

Mithun, Marianne. 1984. The Evolution of Noun Incorporation. Language. 60:4.846-94.

Nicholson, Velda. 1978. Aspectos da língua Assuriní. Brasília: Summer Institute of Linguistics.

Nicholson, Velda. 1982. Breve estudo da língua Asurini do Xingu. Brasília: Summer Institute of Linguistics.

Payne, Doris L. 1990. (ed.). Amazonian Linguistics: Studies in Lowland South American Languages. Austin: University of Texas.

Pease, Helen. 1968. Parintintin Grammar. Arquivo Lingüísico No. 83. Brasília: Summer Institute of Linguistics.

Pease, Helen, and LaVera Betts. 1971. Parintintin Phonology, Tupi Studies I. Ed. by Bendor-Samuel 1-14. Norman: Summer Institute of Linguistics and the University of Oklahoma.

Pederson, Donald. 1977. Grammatical Structures of Guarani. Doctoral Dissertation. University ofy Southern California.

Restivo, Paulo. 1724. Linguae Guarani grammatica. Reprint by Wilhelm Kohlhammer, Stuttgart, 1892.

Rodrigues, Aryon Dall'Igna. 1953. Morfologia do verbo Tupi. Letras, 1:121-152. Curitiba.

Rodrigues, Aryon Dall'Igna. 1958. Classification of Tupí-Guaraní. IJAL 24.231-234.

Rodrigues, Aryon Dall'Igna. 1981. Estrutura do Tupinamba. unpublished manuscript.

Rodrigues, Aryon Dall'Igna. 1984/1985. Relações internas na família linguística tupi-guarani. Revista de Antropologia 27/28/:33-53. São Paulo.

Rodrigues, Aryon Dall'Igna. 1985. Evidence for Tupi-Carib Relationship. South Amercian Indian Languages. Ed. by Klein and Stark 371-404. Austin: University of Texas.

Rodrigues, Aryon Dall'Igna. 1986. Lenguas Brasileiras. São Paulo: Edições Loyola.

Rodrigues, Aryon Dall'Igna. 1990. You and I = Neither You nor I: The Personal System of Tupinambá. Amazonian Linguistics: Studies in Lowland South American Languages. Ed. by Payne, 117-158. Austin: University of Texas.

Ruiz de Montoya. 1639. Tesoro de la lengua Guaraní. Madrid: Juan Sanchez. Reimpresió facsímile preparada por don Julio Platzmann. Leipzig: Teubner, 1876.

Ruiz de Montoya. 1640a. Arte de la lengua Guaraní. Madrid: Juan Sanchez. Reimpresió facsímile preparada por don Julio Platzmann. Leipzig: Teubner, 1876.

Ruíz de Montoya. 1640b. Catecismo en la lengua Guaraní. Madrid: Juan Sanchez. Reimpresió facsímile preparada por don Julio Platzmann. Leipzig: Teubner, 1876.

Saelzer, Meinke. 1976. Fonologia Provisória da língua Kamayurá. Série lingüística 5. Ed. by Lorraine Bridgemann. Brasília: Summer Institute of Linguistics.

Sapir, Edward. 1916. Time perspective in aboriginal American culture. Memoir 90, Anthropological Series No. 13. Geological Survey, department of Mines. Ottawa. Reprinted in Selected Writings of Edward Sapir (ed. by David G. Mandelbaum), 389-462, 1949. Berkeley and Los Angeles: University of California.

Schleicher, Charles. 1989. The Development of the Verb System in Guarani. Master's Thesis. University of Kansas.

Schleicher, Charles, 1994. A Chronology of PIE Obstruents. Indogermanische Forschungen.

Seki, Lucy. 1990. Kamaiurá (Tupí-Guaraní) as an Active-Stative Language. Amazonian Linguistics: Studies in Lowland South American Languages. Ed. by Payne, 117-158. Austin: University of Texas.

Soares, Marilia Facó. 1978. A classificação interna do Kayabi na família Tupi-Guarani. Manuscript.

Soares, Marilia Facó. 1979. A perda da nasalidade e outras mutações vocálicas em Kokama, Asurini e Guajajara. Master's thesis. Rio de Janeiro: Universidade Federal do Rio de Janeiro.

Soares, Marilia Facó, and Yonne de Freitas Leite. 1991. Vowel shift in the Tupi-Guarani language family: a typological approach. in Language Change in South American Indian Languages, Mary Ritchie Key (ed.), Philadelphia: University of Pennsylvania.

Swadesh, Morris. 1959. Mapas de clasificación lingüística de México y las Américas. México: UNAM.

Velazquez, Maura. 1986. The Lexical Approach to Syntax and Noun Incorporation in Guarani. Master's Thesis. University of Kansas.

Weiss, Helga. 1972. Kayabi verbs. Archive of the SIL (Brasília), of the Fundação Nacional do índio (Brasília, and of the Universidade Estadual de Campinas (UNICAMP).