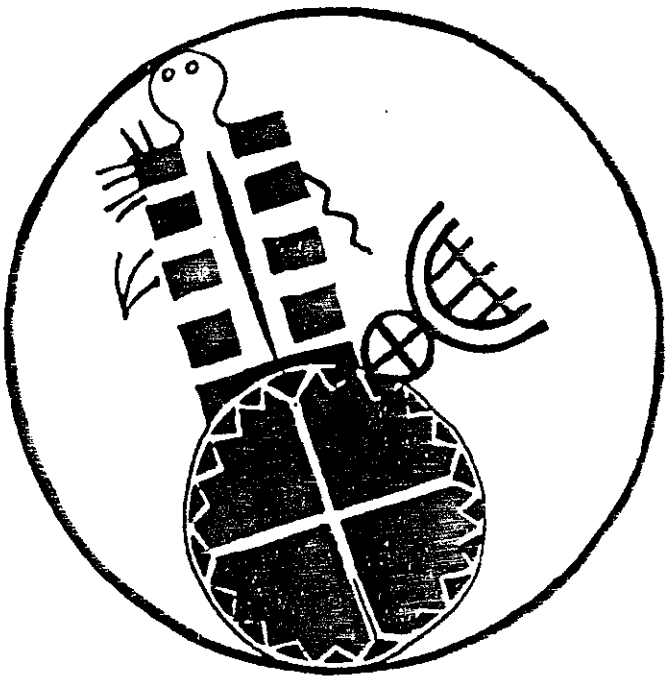
REPORT 8

SURVEY OF CALIFORNIA AND OTHER INDIAN LANGUAGES



Proceedings of the Meeting of
the Society for the Study of the
Indigenous languages of the Americas
July 2-4, 1993
and the Hokan-Penutian Workshop
July 3, 1993

both held at the 1993 Linguistic Institute at Ohio State University in Columbus, Ohio

Margaret Langdon, Volume Editor
Leanne Hinton, Series Editor

CONTENTS

GENERAL

SOUTH AMERICA

HOKAN-PENUTIAN WORKSHOP

CONSTITUENT ORDER VARIATION IN APURINÃ

(Arawakan)

Sidney Facundes

University of Oregon, Eugene, OR, USA

Museu Paraense Emílio Goeldi, Belém, Pará, Brazil

0. Introduction*

The present paper is an attempt to present some descriptive facts about clitics and constituent order variation in Apurinã,[1] with a discussion about whether Apurinã has a configurational or non-configurational constituent structure organization. This leads to discussion of issues related to the properties of configurationality, following mainly the works of Jelinek (1984) and Hale (1990).

The discussion about configurationality focuses on one main property of configurational languages, namely the fixed order of clausal constituents. Other properties generally associated with the configurationality issue (eg. hierarchical vs. flat clausal constituent structure organization, continuous vs. discontinuous expressions, and so on -- cf. Jelinek (1984) and Hale (1990)) will be left aside in this paper.

Section 1 presents the description of grammatical relations

---

* I wish to thank Doris Payne, Spike Gildea and Denny Moore for initial discussions and suggestions on preliminary versions of this paper. I wish also to thank the Inter-American Foundation, the Conselho Nacional de Pesquisas (CNPq/Brazil) and the Museu Paraense Emílio Goeldi for financial support of this research. As part of an ongoing language research project, the syntactic description of the language in focus here is rather preliminary; much of the grammar needs to be known in more detail.

[1] Apurinã belongs to the Maipuran branch of the Arawakan linguistic family (David Payne 1991). It is spoken mainly along the tributaries of the Purus River in the Western Amazonian region of Brazil. There are more than 2,000 Apurinã, and at least 50% still speak the native language (Facundes 1990); however, the Apurinã language has been increasingly replaced by Portuguese. In most of the villages Portuguese is being learned by children as their first language. Dialectal variation can be found in some of the nearly 20 Apurinã villages. The present analysis is intended to cover only the dialect spoken in the Japiim village, along the Paciá River, near Lábrea city, in the state of Amazonas.

and all the possible variations of constituent order; section 2 focuses on the syntactic status of the nominal clausal constituents in the various orders, as well as on the syntactic status of the verbal person markers; section 3 suggests a syntactic analysis of the given facts based on the notion of configurationality, and offers a brief discussion of basic constituent order; finally, section 4 is a brief conclusion.

1. Constituent Order

Previous analyses of Apurinã constituent order (Pickering 1974, Aberdoor 1985, Derbyshire & Pullum 1981, and Facundes 1992b), have considered all nominal clausal constituents (preverbal and postverbal nominals in OSV, SVO, OVS and VOS orders) as expressing arguments of the verb.[2] Moreover, the person markers which are attached to the verb were considered as verb agreement markers.

Pickering, Derbyshire & Pullum suggest an analysis based on structural facts of the Apurinã syntax to argue for OSV as the basic constituent order of the language. Aberdoor presents a study of frequency counting in which she shows that OSV is very rare in text. A reanalysis of the data may show that the role played by clausal constituent order is correlated with the role played by the person markers on the verb; that such person markers are in complementary distribution with preverbal nominals, but not with postverbal ones; and that clitics can also express verb arguments, while nominals split in argumentative and adjunctive functions depending on whether they are pre- or postverbal.

1.1 Grammatical relations

Before getting into the description of constituent order variation, one would like to understand how the verb argument structure is syntactically marked; that is, how core grammatical relations are marked in Apurinã. In a rather simplistic way, it is possible to distinguish at least the grammatical relations subject and object in this language.[3] The distinction between subject and object here is primarily based on morphosyntactic evidence, namely the person markers on the verb (cf. Table 1). As will be seen in

---

[2] The term 'argument' is used here to refer to the core grammatical relations. 'Core' is defined to mean the grammatical relations which are structurally required by the verb as part of its subcategorization frame. At this point only subject and direct object are clearly core grammatical relations in Apurinã (see also fn.3).

[3] The discussion of grammatical relations and constituent order in this paper does not consider clauses with possible trivalent verbs. There are interesting phenomena in Apurinã related to the verbs which commonly behave as trivalent verbs across many languages; although they may be important in arguing about grammatical relations and constituent order, their description will be delayed until the results of further research are available.

(1.2), such person markers may be coreferential with certain overt NOMINALS.[4]

Table 1. System of Person Markers[5]

PERSON	SUBJECT SG	SUBJECT PL	OBJECT SG	OBJECT PL
1	nɨ-/ñi-/n-	a-	-nu	-wa
2	pɨ-/pi-/p-	hi-/h-	-ɨ	-ɨ
3M	ɨ-/i-/ϕ-	...-na[6]	-rɨ	-rɨ
3F	u-/ũ-	...-na	-ru	-ru

Note that the pronominal marker system above allows the identification of person, number and gender of both subject and object. As the same set of subject pronominal markers is used to refer to both intransitive and transitive subjects while the set of object markers is only used to refer to objects, the system of grammatical relations follows the nominative-accusative pattern.

1.2 Constituent Order Variation

There are six logical possibilities for the relative order of subject, verb and object: SOV, SVO, OSV, VSO, OVS and VOS. Of the six logical possibilities only VSO has not been found. Pickering (1974) has stated that the SO sequence is simply ungrammatical; however, at least in elicitation it is possible to collect cases of SOV. An analysis in detail of the discourse-pragmatic functions of the several constituent orders given in this paper is still required and, therefore, is not discussed here.

The sentences in (1-3) show one of the nominals referring to subject or object in preverbal and another one in postverbal position, i.e. OVS and SVO orders. As the ungrammaticality of the sentences in (4-5) indicates, while the postverbal nominals ARE coreferential with the pronominal marker on the verb, preverbal

---

[4] For the time being, until their status is discussed further, pre- and postverbal nominal constituents will be simply referred to as NOMINALS.

[5] It appears that all allomorphs of person markers are morphophonologically conditioned. Not all possible allomorphs are listed here; some forms also have a nasal counterpart whose conditioning has not yet been clearly determined.

[6] The plural for the 3rd person masculine or feminine is formed by the prefixation of the subject marker of the 3rd person singular masculine or feminine plus the plural suffix -na in the end of the verb.

nominals are NOT.[7]

(1) O V S
ɨwa *u* *-mapuruka* *uwa*
it 3SG.FEM.SUBJ-root.up she
'She rooted it up'

(2) S V O
uwa *mapuruka-rɨ* *ɨwa*
she root.up -3SG.MASC.OBJ it.MASC
'She rooted it up'

(3) S V O
ɨwa *mapuruka-ru* *uwa*
he root.up -3SG.FEM.OBJ it.FEM
'He rooted it up'

(4) S V O
**ɨwa* *ɨ* *-mapuruka-ru* *uwa*
he 3SG.MASC.SUBJ-root.up -3SG.FEM.OBJ it.FEM
(he rooted it up)[8]

(5) O V S
**uwa* *ɨ* *-mapuruka-ru* *ɨwa*
it.FEM 3SG.MASC.SUBJ-root.up -3SG.FEM.OBJ he
(he rooted it up)

Moreover, the ungrammaticality of the sentence in (6), in contrast with the one in (7), indicates that once both nominals are postverbal the nominal referring to the object must precede the subject; thus, VOS is allowed, but VSO is not.

(6) V S O
**a* *-mapuruka-rɨ* *ata* *ɨwa*
1PL.SUBJ -root.up -3SG.MASC.OBJ we it.MASC
'We rooted it up'

---

[7] Abbreviations and special symbols used:

O	Free Object	S	Free Subject
o-	Bound Object	s-	Bound Subject
MASC	Masculine	FEM	Feminine
SG	Singular	PL	Plural
ɨ	High Central Unround Vowel	č	Plosive Alveo-Palatal
ñ	Nasal Palato-Alveolar	š	Fricative "

[8] Aberdoor (1985) does mention cases of preverbal nominals expressing grammatical relations which are coreferenced by person markers on the verb. Nevertheless, such examples appear to be extremely rare in her texts. This fact offers some clues about the development of pronominal clitics which, however, are beyond the scope of this paper.

(7)

	V	O	S
a	*-mapuruka-ri*	*iwa*	*ata*
1PL.SUBJ-root.up	-3SG.MASC.SUBJ	it.MASC	we

'We rooted it up'

The additional examples below confirm that preverbal nominals referring either to subject or object are not cross-referenced by person markers on the verb. In (8,11) both (OSV, SOV) nominals precede the verb, which bears no person marker. In (9-10 and 12-13) the sentence is not accepted when either of the nominals is cross-referenced on the verb.

(8)

O	S	V
iwa	*ata*	*mapuruka*
it	we	root.up

'We rooted it up'

(9)

O	S	V
**iwa*	*ata*	*mapuruka-ri*
it.MASC	we	root.up -3SG.MASC.OBJ

(we rooted it up)

(10)

O	S	V
**iwa*	*ata*	*a* *-mapuruka*
it.MASC	we	1PL.SUBJ-root.up

(we rooted it up)

(11)

S	O	V
ata	*iwa*	*mapuruka*
we	it	root.up

'We rooted it up'

(12)

S	O	V
**ata*	*iwa*	*mapuruka-ri*
we	it.MASC	root.up -3SG.MASC.OBJ

(we rooted it up)

(13)

S	O	V
**ata*	*iwa*	*a* *-mapuruka*
we	it	1PL.SUBJ-root.up

(we rooted it up)

As might be expected, nominals referring to the subjects of intransitive verbs follow the same behavior as those referring to transitive subjects. In (14) the subject precedes the verb and no pronominal marker is attached to the verb; in (15) the presence of the pronominal marker on the verb when the nominal is preverbal leads to an ungrammatical sentence. In (16) the nominal is postverbal and the pronominal marker occurs, whereas in (17) the lack of the pronominal marker with a postverbal nominal causes the sentence to be ungrammatical.

(14) S V
iwa *pičeka*
it.MASC grow
'It grew up'

(15) S V
**iwa* *i* *-pičeka*
it.MASC 3SG.MASC.SUBJ-grow
'It grew up'

(16) V S
i *-pičeka* *iwa*
3SG.MASC.SUBJ-grow it.MASC
'It grew up'

(17) V S
**pičeka* *iwa*
grow it.MASC
'It grew up'

Based on the examples given up to now, it is possible to summarize the constituent orders in Apurinã as follows:

Table 2: Apurinã Core Constituent Orders

Transitive									Intransitive			
O	S	V	O	V	S	V	O	S	S	V	V	S
S	O	V	S	V	O	*V	S	O				

Looking at the table above, one would tend to postulate that Apurinã is almost completely "free" constituent order language. In the next section, some arguments against a "free" constituent order language are discussed.

2. The Syntactic Status of Nominals and Person Markers

There is a problem for the interpretation of Table 2 above as presenting indications of a "free" constituent order language. Such an interpretation only works under the assumption that pre- and postverbal nominals have identical syntactic status. This section presents arguments against that assumption.

2.1 Postverbal Nominals as Adjuncts

In all the examples given above which show at least one postverbal nominal, such a nominal can be missing. For instance, the examples in (18-19), in contrast with those in (1-2), show that postverbal nominals occur optionally in a clause when there is a coreferential person marker on the verb.

(18) O V
ɨwa *u* *-mapuruka*
it.MASC 3SG.FEM.OBJ -root.up
'She rooted it up'

(19) S V
uwa *mapuruka-rɨ*
she root.up -3SG.OBJ.MASC
'She rooted it up'

Thus, Table 2 above can be revised as follows:

Table 3: Apurinã Core Constituent Orders

Transitive			Intransitive	
O S V S O V	O V (S) S V (O)	V (O) (S) *V (S) (O)	S V	V (S)

In addition, in (20-21) the sentences show that neither the postverbal nominal referring to object nor the postverbal nominal referring to subject can occur without the coreferential verb person marker. This fact should falsify the claim that the postverbal nominal triggers an 'agreement' marker on the verb. Also, if a clause can 'lose' postverbal nominals without affecting its propositional content, it seems strange to argue that such nominals express arguments of the verb

S V O
(20)**uwa* *mapuruka* *ɨwa*
she root.up it.MASC
(she rooted it up)

O V S
(21)**ɨwa* *mapuruka* *uwa*
it root.up she
(she rooted it up)

The label **OPTIONAL** is used here to refer to the quality that a nominal may have of OPTIONALLY occurring in a clause when it corefers to a person marker on the verb. Since only postverbal nominals can (and have to) be coreferential with a verbal person marker, a distinction can be drawn between pre- and postverbal nominals: The latter are optional whenever they can occur whereas the former are not. This distinction can be used as an evidence that these pre- and postverbal elements have a different syntactic status. The hypothesis is that while preverbal nominals express core grammatical relations, those which are postverbal are adjuntive elements. Therefore, since postverbal nominals are optional whenever they occur, it is possible to claim that they behave syntactically as oblique elements. Such oblique elements

will be labeled **ADJUNCTS**.[9]

2.2 Preverbal Nominals as Core Grammatical Relations

The evidence used here for categorizing postverbal nominals as adjuncts does not apply to the preverbal ones. Preverbal nominals are NOT optional whenever they occur. The optionality of postverbal nominals can be seen in the examples given in (2) and (19) (which are repeated for convenience in (22) and (23)). The non-optionality of preverbal nominals can be seen by contrasting the examples (22) and (23) with the one in (24): (24) is ungrammatical because the preverbal nominal referring to the subject is missing. Since preverbal nominals are not cross-referenced on the verb, once they are missing there has to be another element to express the argument of the verb (cf. the subject marker in (25)). The fact that the verb person marker shows up in the absence of a preverbal nominal (or vice-versa) is a clue to the role played by the former in a clause, as will be seen in the next subsection.

```
       S          V                      O
(22)  uwa    mapuruka-ɨ̃            ɨ̃wa
      she    root.up -3SG.MASC.OBJ  it.MASC
      'She rooted it up'

       S          V
(23)  uwa    mapuruka-ɨ̃
      she    root.up -3SG.MASC.OBJ
      'She rooted it up'

(24)      V
     *mapuruka-ɨ̃
      root.up -3SG.MASC.OBJ
      (she rooted it up)

                          V
(25)  u              -mapuruka-ɨ̃
      3SG.FEM.SUBJ-root.up -3SG.MASC.OBJ
      'She rooted it up'
```

Therefore, in a certain way it is possible to argue that free-standing nominals may or may not express verb arguments in Apurinã depending on whether they precede or follow the verb in a clause.

2.3 Person Markers as Clitic Arguments

Given the evidence that, due to their optionality, postverbal

---

[9] The syntactic categorization suggested here for postverbal nominals in Apurinã has already been suggested elsewhere (Jelinek (1984) and Hale (1990)) for nominals with similar behavior in other languages.

nominals cannot be the verb arguments, one, then, needs to ask what syntactically represents the verb arguments in sentences with no preverbal nominal. As shown above, preverbal nominals cannot be optional whenever they can occur because they are not cross-referenced by verb person markers. Thus, preverbal nominals and verb person markers are in complementary distribution. The reason to be in complementary distribution is that both accomplish the same syntactic function; that is, they play the same syntactic role in a clause. This role is to express the argument of the verb.

As a syntactic element, rather than simply a morphological affix attached to the verb, the person markers behave as PRONOMINAL CLITICS. To consider person markers as pronominal clitics does not mean simply to find another label for an atypical affix; more than that, it is to try to describe more precisely the syntactic behavior and function of such an element.

As the syntactic elements which express verb arguments, the clitics behave as pronouns which are attached to the verb; that is, clitics seem to function as normal pronouns, except that they are phonologically bound morphemes. Furthermore, the difference, for instance, between a free-standing pronominal and a clitic is that while the former has the typical syntactic distribution of a nominal which can function as subject or object grammatical relation, the latter has a distribution which is morphologically determined. The distribution of the clitics is morphologically determined in that they fill up a fixed slot in a verbal construction formed by morphological rather than syntactic operations.

Nevertheless, although free-standing nominals expressing arguments find themselves identified by means of syntactic rather than morphological rules, their syntactic function are basically the same: Both free-standing nominal arguments and clitic arguments function as the syntactic bearing elements of verb arguments.

3. The Syntactic Analysis

It has been said above that the arguments of the verb can be expressed either as clitics or as free-standing nominals; however, nominals which co-occur with (and are coreferential with) clitics are NOT arguments of the verb but, rather adjuncts. In this section, the discussion is focused on some possible syntactic implications of the stated analysis. Configurationality and constituent order are issues related to the clitic and preverbal arguments plus postverbal nominal adjunction claim.

3.1 The Theoretical Notion of Non-configurationality

As Jelinek (1984) has pointed out, the recent interest on the

non-configurational property of some languages has been motivated principally by Ken Hale's work on Australian and Native American languages. The initial discussion by Hale (1980, 1981, 1982, 1983) aimed to account for some of the characterictics usually found in non-configurational languages. Some of these supposedly common non-configurational properties would be free clausal constituent order variation, syntactically discontinuous expressions and null anaphora. In these works Hale suggested some parameters of configurationality, which have been revised by Jelinek (1984).

For a better understanding of the notion of non-configurationality, it may be helpful to look at one Warlpiri example. "In the following Warlpiri sentence, any word order is possible, with the provision that the AUX clitic sequence occur in the second position.

(26) **Ngarrka-ngku ka wawiri panti-rni.**
man-ERG AUX kangaroo spear-NONPAST
'The man is spearing the kangaroo'"
(Jelinek 1984:39-40)

Jelinek has claimed that the arguments of verbs in Warlpiri are expressed by the clitics;[10] the nominals which are coreferential with the clitics would be optional and, thus, non-argumental features. Finally, based on these ideas, she has proposed an extended configurationality parameter for languages which share some of the Warlpiri grammar features. This configurationality parameter would be as follows:

(27) "a. In a configurational language, object nominals are properly governed by the verb.
b. In a W-[Warlpiri]type non-configurational language, nominals are not verbal arguments, but are optional adjuncts to the clitic pronouns that serve as verbal arguments."
(Jelinek 1984:73)

This new analysis of the non-configurational properties in Warlpiri has been endorsed in Hale's 1990 Core Structures and Adjunctions in Warlpiri Syntax:

> "there might exist languages... whose free word order results simply from the fact that (certain or all) overt phrasal expressions are adjuncts..."
> (Hale 1990:36-7).

Considering such a theoretical approach to constituent order variation, it may be interesting to draw some attention again to the Apurinã constituent order variation decribed above.

---

10 For details and examples, see Jelinek 1984.

3.2 Configurationality in Apurinã

The summary of constituent orders given in Table 2 above suggested a system of partially free constituent order variation in Apurinã. Revising Table 2, Table 3 represents the occurrence of postverbal nominals as adjuncts. Under the analysis of pronominal clitics as arguments of the verb and postverbal nominals as adjuncts, Table 3 should be further revised as follows:

Table 4: Apurinã Core Constituent Orders

Transitive			Intransitive	
O S V S O V	O s-V S V-o	s-V-o	S V	s-V

Such a revision is due to the fact that if postverbal nominals are adjuncts, they cannot also be core grammatical relations at the same time. With respect to the analysis of the transitive verbs, the result of this revision is a three-way system of syntactic expressions of the verb argument structure: (i) Both the verb arguments can be syntactically realized as free-standing nominals, or (ii) both the arguments can be simultaneously realized as clitics, or yet (iii) one verb argument can be realized as a free-standing nominal while the other is realized as a clitic.

Analogously, in relation to the intransitive verbs, the verb argument can be syntactically expressed as either (i) a free-standing nominal or as (ii) a clitic.

Transitive and Intransitive verb sentences can be grouped into three order types based on the syntactic realization of their arguments in a sentence, as seen in Table 5 below. Type I groups the sentences with only phonologically free-standing syntactic elements expressing arguments; Type II groups only phonologically bound syntactic elements expressing arguments; and Type III groups sentences which mix Types I and II.

Table 5: Grammatical Relations Organization

Type I	Type II	Type III
O S V S O V	s-V-o	O s-V S V-o
S V	s-V	

For the purpose of analyzing configurationality in the distribution of the syntactic elements expressing arguments in Type I, only sentences with transitive verbs are relevant here. The indications are that OSV and SOV can be used interchangeably as long as their interpretation is not ambiguous. However, when

ambiguity exists, the object is required to precede the subject. For instance, the variation seen in the unambiguous examples in (28-31) cannot occur in ambiguous examples like (32-33). The second interpretation of the sentences in (32-33) is not possible when the context does not disambiguate; when this happens, the position of the syntactic elements is fixed and, therefore, the order of the constituents is configurational.

(28) O S V
yuwata *nota* *etama*
knife I see
'I see the knife'

(29) O S V
nota *yuwata* *etama*
knife I see
'I see the knife'

(30) O S V
hãkiti *kɨkɨ* *keta*
jaguar man shoot
'The man shoots the jaguar'

(31) S O V
kɨkɨ *hãkiti* *keta*
man jaguar shoot
'The man shoots the jaguar'

(32) O S V
anãpa *kɨkɨ* *etama*
dog man see
'The man sees the dog'
*'The dog sees the man'

(33) O S V
Pedro *Paulo* *keta*
shoot
'Paulo shoots Pedro'
*'Pedro shoots Paulo'

Contrasting Type I with Types II and III, and following Jelinek's configurationality parameter given in (25), the tendency would be to argue for a partial configurationality in Apurinã. This tendency follows from the configurationality parameter because preverbal nominals (as in OSV, Os-V and SV-o) ARE arguments of the verb, whereas postverbal nominals (as in SV-o(O), Os-V(S) and s-Vo(O)(S)) are NOT verbal arguments, but rather optional adjuncts to the clitic pronouns which are the verbal arguments. Therefore, while Type I would be configurational, since the clausal constituent order is syntactically relevant, and Type II would be non-configurational, since postverbal nominals are adjuncts and clitics are arguments, Type III would be both.

However, if we adopt a notion of grammatical relations which

is not necessarily defined only in terms of hierarchical structures of clausal constituents (e.g. NPs, VPs or N"s, V"s), but mainly in terms of how a nominal element can syntactically interact with the verb element in a clause, there may be an alternative way to analyse the Apurinã system summarized in Table 5.

The syntactic realization of verb arguments in Type II sentences, which consists only of clitics attached to the verb, has a fixed order. The order of the clitics is morphologically determined by the position class they occupy in the verbal construction. However, syntactically these clitics are in complementary distribution with free-standing grammatical relations; and lexically, as portmanteau morphemes, they bear grammatical information as case roles (nominative-accusative, cf. Table 1), person, and gender (feminine and masculine). Therefore, considering the syntactic behavior of the clitics, the Type II does not necessarily poses a problem to a notion of configurationality based on the functional features of the clausal constituents. That is, the only additional feature is that the arguments are phonologically realized as bound morphemes.

Finally, Type III which is constituted of a mixture of types falls out from the description of Types I and II. In a language that allows the syntactic expression of verb arguments by means of free-standing as well as bound morphemes, a hybrid kind of argument expression including both the Types I and II should be expected to occur. Rather than posing a problem, Type III reinforces the analysis suggested for the first two types.

3.3 **Basic Constituent Order**

The criteria usually used to determine the basic constituent order of languages can be grouped into three sets: Descriptive simplicity, statistical frequency, and pragmatic neutrality (following Mithum 1992). It is beyond the scope of this paper to present a study on pragmatic values of clausal constituent orders. What will be mostly considered here is the descriptive simplicity criterion and a very preliminary study on statistical frequency.

As has been mentioned earlier in this paper, in previous analyses of Apurinã constituent order (Pickering 1974, Derbyshire & Pullum 1985, and Facundes 1992b), the suggested basic constituent order was based on the descriptive simplicity criterion: Having assumed that pronominal markers on the verb were agreement triggered by the postverbal subject and object, all the constituent order patterns found could be derived from an unmarked one, namely OSV.[11]

---

[11] Aberdoor (1985) also did some work on Apurinã which included statistical frequency and discouse-pragmatic functions; however, her analysis is based on the assumption that pronominal markers are verb agreement markers and that missing nominals are result of zero anaphora.

The descriptive simplicity criterion permits one to say that OSV is more basic than SOV, since the occurrence of the latter is predictable. SOV can occur only when there is no ambiguity in a clause. Another indication in favor of OSV would be that, at least in texts, the adjuncts which are coreferential with the clitics can only occur in the sequence OS.

The decision of whether or not to consider order types like Os-V and s-V-o in the analysis of constituent order is usually related to pre-established theoretical assumptions. For instance, one assumption could be to consider as pertinent for the analysis of constituent order only the occurrence of the free-standing grammatical relations; another one could be to postulate that free-standing elements tend to be more neutral than bound morphemes.

Based on the description of the system of grammatical relations suggested here, one would tend to consider bound clitics as relevant in analysing Apurinã constituent order. If clitics behave as normal subjects and objects, except that they are phonologically bound, their role on constituent order might be as important as that of any other subject or object.

On the other hand, one of the possible consequences of the phonological attachment of clitics to verbs is that the clitics, then, follow the rules of the morphology and no longer of the syntax. What this might mean is that additional syntactic or morphological tests are required to establish the relevance of argument clitics for an analysis of basic constituent order which is based only on the descriptive simplicity criterion.

The statistical frequency criterion would lead one to choose OSV as more neutral than SOV, since SOV does not seem to occur in text but only in elicited data (cf. Facundes 1993) or in quotative clauses (cf. Aberdoor 1985). However, OSV is extremely rare in frequency (0.9%).

Os-V and s-V-o occur with equal frequency (33%) and are the most frequent order types, which makes either of them good candidates for basic order and reinforces the idea that bound grammatical relations are (statistically) relevant in defining basic constituent order. SV-o, however, occurs with low frequency (2.5%).

Based on a structural analysis including formal rules, an alternative syntactic analysis of constituent order would postulate a right-dislocation to generate postverbal nominals followed by the attachment of the person markers on the verb. If clitics are not considered, the result of this analysis would be that OSV would be the basic order of the language. This approach to the data, however, would require further details (i.e. deletion of postverbal nominals) in order to account for the optionality of such nominals. No motivation for right-dislocation has been found up to now, and, besides, the evidence that postverbal nominals are adjuncts rather

than core grammatical roles also poses problems to such an analysis.

Also based on formal rules, another alternative syntactic analysis of constituent order would be to postulate a left-deslocation of postverbal nominal(s) and person marker(s) deletion. By this analysis the basic constituent order would be VOS. Such an approach, however, only would account for the occurrence of preverbal nominal in a formal description, but would not say anything about their syntactic function, or their syntactic status in contrast with the status of the postverbal nominals.

There is no definitive hypothesis to be presented at this point in relation to a basic constituent order. Additional information about the grammar of this language might provide better clues. Nevertheless, it seems pretty clear that a synchronic description of aspects of the Apurinã syntax must involve a certain degree of complexity.

4. **Conclusion**

The ideas about the grammar of Apurinã described here were intended to show how the verb argument structure is syntactically expressed and how it corelates with constituent order. Such ideas are bound into the synchronic internal evidence of the language. Historical considerations and discourse-pragmatic functions were delayed until the results of further research are available.

The initial appearance of "free" constituent order is not actually confirmed when the grammar of the language is studied more carefully. Constituent order is relevant for Apurinã syntax, which motivates the attempt to describe its unmarked, most frequent or most neutral realization. However, this last task requires additional research; any hypothesis about basic constituent order would be more theoretically dependent than motivated by the internal language structure. Nevertheless, the preliminary description of Apurinã shows a corelation which may exist between pronominal clitics, argument roles and configurationality.

5. **References**

Aberdoor, K. 1985. **Referential Devices in Apurinã Discourse.** In Porto Velho Working Papers, ed. David Fortune. SIL, Brasília, pp.43-91.

Derbyshire, C. Desmond and Pullum Geoffrey K. 1981. **Object Initial Languages.** IJAL, vol. 47, nr. 3, July, pp. 192-214.

Facundes, Sidney da S. 1990. **Levantamento Linguístico do Rio Purus.** Relatório de pesquisa para o Conselho Nacional de

Pesquisas, Brazil.

-------. 1992a. **Split Case-Marking in Apurinã (Arawakan, Maipuran)** (ms.)

-------. 1992b. **Apurinã (Arawakan) Word Order** (ms.)

-------. 1993. **Apurinã Word Formation** (ms.)

Hale, Ken. 1980. **Remarks on Japanese Phrase Structure: Comments on the papers on Japanese Syntax**, in Y. Otsu and A. Farmer (eds.), MIT Working Papers in Linguistics, Vol. 2.

-------. 1981. **On the position of Warlpiri in a Typology of the Base.** Indiana University Linguistics Club, Bloomington.

-------. 1982. **Preliminary Remarks on Configurationality.** In J.Pustejovsky and P. Sells (eds.) NELS 12, 86-96.

-------. 1983. **Warlpiri and the Grammar of Non-Configurational Languages.** Natural Language & Linguistic Theory 1:3-45.

-------. 1990. **Core Structures and Adjunctions in Warlpiri Syntax.** Proceedings of the Tilburg Scrambling Conference, October 1990.

Jelinek, Eloise. 1984. **Empty Categories, Case, and Configurationality.** Natural Language and Linguistic Theory 2, 39 76. D. Reidel Publishing Company.

Mithun, Marianne. 1992. **Is basic Word Order Universal?** In Pragmatics of Word Order Flexibility. Ed. Doris L. Payne, pp.15-61.

Payne, David. 1991. **A Classification of Maipuran (Arawakan) Languages Based on Shared Lexical Retentions.** In Handbook of Amazonian Languages. Vol. 3, pp.355-517.

Pickering, Wilbur. 1974. **Gapping and Constituent Order in Apurinã.** Arquivo Linguístico No. 006. Brasília, D.F. SIL.

CASE, VERB TYPE AND ERGATIVITY IN TRUMAI

Raquel Guirardello

Museu Paraense Emílio Goeldi (Pará - Brazil)

1. Introduction[1]

This paper will present a short study of the Trumai verbal case-marking system.[2]

In a first approach to the data, it would be possible to say that Trumai has four types of verbs: intransitive, transitive, ditransitive, and a fourth type which is our main interest in this paper, for although it seems to be transitive (it basically has two participants), it is morphosyntactically distinct from prototype transitive verbs. We will show that (a) this fourth type is in reality intransitive, with one participant the subject and the other an oblique NP; we will label these verbs extended intransitives; (b) these four verbal types collapse into only two basic types, intransitive and transitive.

The lexical category of extended intransitive verb is motivated by both semantic features of the verbs themselves - i.e. the actions which these verbs express take a location rather than a patient for a second participant - and by pragmatic features of the second participant for those verbs which do take patients as second participants - i.e. they are stereotypical or otherwise of little importance. Such pragmatic factors are also important in the choice of when to use an extended intransitive verb versus a transitive verb when either is available, e.g. for pairs of the type ***fa*** `kill, hit' and ***disi*** `kill, hit', one of which is transitive and the other extended intransitive. In essence, the speaker chooses which verb to use depending upon the topicality of the patient/locative second participant.

2. Description: the verb types and the case-marking frames in Trumai

First, we will illustrate the four verb types of Trumai, showing how the morphosyntatic system is organized.

The object of a transitive verb (O) receives the same treatment as the subject of an intransitive verb (S),[3] that is: the same case-marking (unmarked); the same position in the clause (in adjacency with the verb); occurrence of the third person clitic **-n** when the overt nominal does not occur; and a certain control of the position of these functions in the case of lexical item deletion, through the ***i*** (or ***ii***) morpheme.[4] In contrast, the subject of a transitive verb (A) receives a distinct case-marker (the suffix **-*k***), can vary in position in the clause and, when omitted, leaves behind neither a marker on the verb nor another morpheme anaphorically (such as the ***i/ii*** of absolutive). Since S and O pattern together morphosyntactically, and since A receives its own unique morphosyntax, this language is clearly ergative-absolutive. There are other ways in which ergativity manifests itself, as for instance the imperative construction, which we will treat later. We now offer illustrative examples of the four verb types in Trumai.

Type 1 - Intransitive verbs: Absolutive S (unmarked); verbal cross-reference to Abs (V-n '3 Abs')

The intransitive subject can be expressed by a lexical NP (examples 1,2) or, if that is deleted, by means of the third person enclitic **-n** (example 3). In the case of first and second persons, the use of pronouns is obligatory, that is, these pronouns cannot be deleted. The third person enclitic also presents the allomorph **-e** for verbs which end with a consonant. Examples:[5]

S V
(1) ***ha-Ø*** ***pita***
1-Abs go.out
'I go out'

S V
(2) ***hine-Ø*** ***pita***
3-Abs go.out
'He goes out'

V-n
(3) ***pita-n***
go.out-3Abs
'He goes out'

(4) Dative marking goal of action

GOAL S V
ole-s ***ka in*** ***ha-Ø*** ***kawa***
manioc-dat ? ? 1-Abs go
'I'm going to get manioc (right now)'
(lit. `I'm going for manioc')

(5) Dative as mark of locative

S V LOC
ha-Ø ***axa'tsi*** ***ka in*** ***tehnene-ki***
1-Abs sit.down ? ? floor-dat
'I sat down in the floor'

Type 2 - Transitive verbs: Ergative A (-*k*/-*ts*), Absolutive O (-Ø;V-*n*)

The subject of a transitive verb is marked with the suffix -***k*** (cf. examples 6,8) which is preceded by epenthetic vowel ***e*** or ***a*** when attached to words which end in a consonant; the first person singular allomorph of the ergative marker is -***ts*** (cf. examples 7 a-b). The third person object (absolutive) nominal (or free pronoun) can be deleted; in this case, the clitic -***n***/-***e*** occurs (cf. 7 a-b). It is possible to find some variation in the word

order (AOV, OVA - cf. 8 a-c) but the order AVO is not allowed (cf. 8c) and the orders OAV, VAO and VOA are not attested. We see then that the sequence OV is not broken.

A O V
(6) ***hine-k*** ***atlat-Ø*** ***mapa***
3-Erg pan-Abs break
'He broke the pan'

A O V
(7) a. ***hai-ts*** ***kasoro*** ***mid-Ø*** ***husa***
1 -Erg dog neck-Abs chain
'I chained the dog (by the neck)'[6]
(lit. 'I chained the dog's neck')

V-n
b. ***hai-ts*** ***ka in husa-n***
1 -Erg ? ? chain-3Abs
'I am chaining it'

O V A
(8) a. ***wirix*** ***ma'may*** ***ka in Yaka-k***
manioc-pap mix ? ? proper name-erg
'Yaka is mixing the manioc pap'

A O V
b. ***Yaka-k ka in wirix ma'may***

A V O
c. ****Yaka-k ka in ma'may wirix***

Type 3 - Ditransitive verbs: Erg A (-*k*/-*ts*), Abs O (-*Ø*;V-*n*), Postverbal Dat IO (-*tl*/ -*ki*/ -*s*)

In the few verbs of this type, we can also see that the order OV is preserved. The indirect object (IO) is not obligatory; unlike O (but like the A), the IO nominal (or free pronoun) can be omitted with no resultant marker on the verb. For instance:

	A		O	V	IO
(9)	***kiki-k***	***ši***	***atlat-Ø***	***kiti***	***hai-tl***
	man-Erg	?	pan -Abs	give	1-Dat

`The man gave me the pan'

	A		V-n	IO	
(10)	***kiki-k***	***ši***	***kiti-n***	***ha***	***wan-ki***
	man-Erg	?	give-3Abs	1	pl -Dat

`The man gave it to us'

	A		O	V	IO
(11)	***hai-ts***	***ši***	***aros-Ø***	***kiti***	***kasoro-s***
	1 -Erg	?	rice-Abs	give	dog-Dat

`I gave rice to the dog'

	A			O	V
(12)	***hai-ts***	***ši***	***de***	***pan-Ø***	***kiti***
	1 -Erg	?	already	food-Abs	give

`I already gave food'

Type 4 - Extended Intransitive: Abs Agent (-Ø; V-n), postverbal Dat Patient (-*tl*/ -*ki*/ -*s*)

The extended intransitive verb class is interesting because semantically such verbs seem to be transitives, for they can present two participants, with one of them the agent and the other one the presumed patient of the action. The problem is that, unlike what we found for other transitive verbs, the agent here is treated as the absolutive and the 'patient' as an indirect object. See the examples below:

	AGT	V	PAT
(13)	**kiki-Ø**	**fa**	**hine-tl**
	man-Abs	hit/kill	3 - Dat

'The man hit/killed him'

	AGT	V	PAT
(14)	***kiki-Ø***	***fa***	***kodešiš-es***
	man-Abs	hit/kill	snake -Dat

'The man hit/killed the snake'

AGT V PAT
(15) ***ha-ø*** ***sone*** ***cafe-s***
1-Abs drink coffee-Dat
'I drank coffee (a lot)'

AGT V PAT
(16) ***ha-ø*** ***sone*** ***cafe-ki***
1-Abs drink coffee-Dat
`I drank coffee (a little)'

V-n PAT
(17) ***ma-n*** ***kuman-ki***
eat-3Abs bean-Dat
`He ate beans'

The dative marker varies according to the kind of NP which occurs in this position: singular pronouns and human nouns receive the suffix ***-tl*** (with an epenthetic vowel added to the forms which end in a consonant); plural pronouns and human nouns receive the suffix ***-ki***. Non-human nouns can receive two kinds of marking, ***-tl*** or ***-s***, according to the verb: for instance, ***fa*** 'hit/kill' requires the marker ***-s*** , while ***make*** 'bite' requires ***-tl***. Thus, extended intransitive verbs can be subcategorized into two classes on the basis of which dative marker they choose for non-human patients. These classes are lexically determined and are mutually exclusive, that is, verbs which require one suffix do not accept the other. Either ***-tl*** or ***-s***, however, can alternate with ***-ki***, when the patient consists of a small quantity of something (see examples 15 and 16 above). The following table summarizes the allomorphy of the dative marker:

PRON - HUMAN NOUN		NON-HUMAN NOUN	
SG	-tl	Verb Class I	-tl / -ki
PL	-ki	Verb Class II	-s / -ki

3. Analysis: Verb type 4 as morphosyntactically intransitive

When we look more carefully at the entire system of the Trumai language and its morphosyntactic elements, we are forced to conclude that verb type 4 is actually intransitive. Only one of the participants is essential, the subject; the second participant is not obligatory and can be omitted without troubles. This is prima facie evidence that we are dealing with a syntactically intransitive verb type, albeit one which can be extended by taking the 'Patient' NP as a syntactically IO. Thus the label extended intransitive, which emphasizes that the verb in question is basically intransitive, rather than transitive.

The arguments in support of this analysis are: constituent order, verbal person marking, case-marking, and the morphosyntax of the imperative construction. Observe the following table:

a) Basic Word Order

INTRANSITIVE		S	V	(LOC)	or	(LOC)	S	V
TRANSITIVE	A	O	V					
DITRANSITIVE	A	O	V	(IO)[7]				
EXTENDED INTRANSITIVE		S	V	(IO)				

We see from clause types 2 and 3 that the position of O is preverbal, whereas in type 3, the IO (and the oblique LOC in type 1) comes after the verb and is optional (it can vary its position, occuring before the subject, but not before the verb - cf. note 7). In type 4, the erstwhile patient occurs postverbaly and is optional, hence it patterns with IO and Loc rather than O.

b) Verbal person marking

INTRANSITIVE		[ ]	V-**n**	
TRANSITIVE	A	[ ]	V-**n**	
DITRANSITIVE	A	[ ]	V-**n**	IO
EXTENDED TRANSITIVE		[ ]	V-**n**	IO

As we said before, when the third person S or O nominal (or free pronoun) is ommited, the enclitic -**n**/-**e** occurs. If verb type

4 were transitive, the enclitic should refer to the O, as it does in type 2; but on the contrary, it refers to the S.

c) Case System

INTRANSITIVE		S-***Ø***	V	Goal-***s***/ -***ki***
TRANSITIVE	A-***k***	O-***Ø***	V	
DITRANSITIVE	A-***k***	O-***Ø***	V	IO-***s*** / -***tl*** / -***ki***
EXTENDED INTRANSITIVE		S-***Ø***	V	IO-***s*** / -***tl*** / -***ki***

Looking at the case system, it is clear that the second participant of the extended intransitive is not an O, for it presents exactly the same case-markers as the indirect object of a ditransitive verb.

d) Imperative

The final argument, which confirms that type 4 verbs are intransitive, is the morphosyntax of the imperative mood in Trumai. The imperative particle ***wana*** is employed to mark imperatives for intransitive verbs, while transitive and ditransitives verbs use the particle ***waki***. The extended intransitive verbs pattern with intransitive verbs, taking the particle ***wana***.

INTRANSITIVE	***wana***
TRANSITIVE	***waki***
DITRANSITIVE	***waki***
EXTENDED INTRANSITIVE	***wana***

(18) ***wana pita***
Imp go.out
'Go out!'

(19) ***wana pes hen***
Imp run then
'Run, then!'

(20) ***waki husa hen***
Imp chain then
'Chain (it), then!'

(21) ***hine-tl waki kiti***
3 - Dat Imp give
'Give (it) to him!'

(22) ***wana sone***
Imp drink
'Drink (it)!'

(23) ***wirix-ki wana sone***
manioc.pap-Dat Imp drink
'Drink the manioc pap!'

As this final test indicates most clearly, Trumai basically has not four different morphosyntatic verb types, but only two, TRANSITIVE and INTRANSITIVE, with some transitives (type 3) and some intransitives (type 4) extended by means of the adjunction of an optional indirect object. Although the morphosyntax is clear, we are left with a puzzle: why would semantically transitive (i.e. two-participant) verbs be obligatory codified in morphosyntax as intransitive?

4. A possible explanation: semantically and pragmatically reduced transitivity

The fact that an action involves two participants doesn't necessarily mean it has high transitivity. As shown by Hopper and Thompson (1980), transitivity involves not only the number of participants, but an entire set of components, such as volitionality of the agent, punctuality and telicity of the action, the degree of affectedness of the patient, individualization of the patient, etc. The relevance of these components can vary from language to language in determining how a given action will be codified in morphosyntax.

The Trumai language codifies as intransitive verbs those two-participant verbs which have inherently reduced transitivity, due to either semantic features of the verb (section 4.1) or pragmatic features of the second participant, the erstwhile patient (section 4.2).

4.1. The semantic factor: locative object verbs (eg. 'bite', 'hit', etc)

In some actions, the second participant (i.e. the non-agent) is not a true patient (which is completely affected by the action), but rather is a kind of location. That is, while contact is made with the second participant, the effect of the contact may or may not be transferred to it (cf. Hopper and Thompson's 1980 conclusion that 'transferral' of the action is the most basic component of transitivity). It is thus more basically a location where the action occurs than a patient affected by the action. Some languages, such as Trumai, mark the difference between patient-objects and locative-objects in surface morphosyntax. Other languages, such as English, mark this difference only in syntactic variation (cf. Fillmore 1970). Observe the following examples:

(24) (a) I hit him (he may be affected)
(b) I hit at him (he is not affected)

(c) I killed him (he is affected)
(d) * I killed at him

(25) (a) I gave him a hit
(b) * I gave him a kill

From such examples, Fillmore (1970) argues that the erstwhile patient of hit is a location (in example 25, in reality, it is a recipient, which is a kind of metaphorical location). In (24a), the blow arrives at the second participant, but the second participant may or may not be affected. In (24b) the blow is aimed at the second participant, but either the blow does not connect or it has no effect; the use of a locative preposition to mark the second participant argues for a semantic case role of location rather than patient.

The second participant of kill, in contrast, is necessarily affected; therefore it is a patient to whom the action is transferred (and, for this reason, the use of a locative preposition here is impossible). In the case of hit the actual patient is the blow which is created through the action, and the erstwhile patient him is a kind of location where the 'real' patient is created. Example (25) shows this more clearly: it is possible to give a hit (the patient which is invisible in 24a-b) to the second participant (who is recipient), but is impossible to give a kill to the second participant, precisely because he IS the patient.

Different languages treat the lexical category of locative object verbs differently. Where English uses syntactic variation, Lhasa Tibetan allows only the syntax of (25) above for such verbs (DeLancey - p.c. with Gildea), and in Trumai, type 4 verbs allow only the syntax of (24b) above. A brief list of the semantically conditioned type 4 verbs includes:

The Trumai Extended Intransitive Locative Object Verbs

make	'bite'	;	***fa***	'hit/kill'
xom	'suck'	;	***laxod***	'smell (action/perception)'
xu'tsa	'look/see'	;	***fa'tsa***	'listen/hear'
lax	'hunt'			

Verbs like 'hunt' also have a locative object, for the object need not be necessarily present to conduct the action (we can hunt

all day and at the end of hunting have never even encountered a likely patient).

4.2. The pragmatic factor: the non-topical patient

With some verbs the patients are habitual, very predictable, often indefinite and unindividuated. Most of the remaining extended intransitive verbs in Trumai are of this type. While the agent is topical, the patient has little importance or individuation and is often stereotypical. The Agents of 'eat' and 'drink', for example, are affected by these actions in a way that is likely to be more salient to a speaker than the effect on the thing eaten. Hopper and Thompson (1980) note that lowered topicality in a patient is likely to lead to less transitive morphosyntax, such as antipassives, etc. A subclass of type 4 verbs in Trumai are of this type, where stereotypically non-topical patients obligatorily occur as indirect, rather than direct, objects:

The Trumai Extended Intransitive Verbs with non-topical patients

xoxan 'wash' ; ***sone*** 'drink'
ma 'eat' ; ***maska*** 'sew'

At this point, we should point out some potential problems for our analysis: most troubling is the verb ***suda*** 'make (something)' which in Trumai is an extended intransitive. This verb is a problem for the explanation given above because the object of the action falls into neither of the categories above: it is not predictable nor it is semantically a locative, but rather a patient. It is hard to understand why the language would codify this apparently fully transitive action with intransitive morphosyntax. In another way as well, the Trumai system is not totally coherent. Like ***lax*** 'hunt', ***padi*** 'wait' is neither a telic action nor must its object be necessarily present. But while 'hunt' in Trumai is extended intransitive, 'to wait' receives fully transitive morphosyntax.

5. The pragmatic uses of transitive/extended intransitive pairs

In Trumai, some pairs of verbs are semantically equivalent (that is, they express the same action), but belong to different morphosyntatic categories. For each of these verbs it is possible to obtain a paradigm with all persons of subject and object (1st, 2nd, 3rd) and with nouns.These verbs are the followings:[8]

EXTENDED INTRANSTITIVE	TRANSITIVE	GLOSS
fa	***disi***	`hit/kill'
make	***tako***	`bite'
dama	***tuxa'tsi***	`pull'

These verbs alternate with one an other in actual use: it is possible to use either the extended intransitive or the transitive forms any time the basic concept needs expression. What is interesting to see is that speakers seem to select which category of verb to use based on the persons of agent and patient. When first person acts on third, speakers prefer the extended intransitive form; when third acts on first, the preferred verb used is the transitive.

This preference was observed in two kinds of data:

(1) In elicitation of paradigms from various consultants; e.g, when the paradigm was requested for 'hit', the consultants systematically selected the verb type depending on the person of A and O: 1A -> 3O, 1A -> 2O and 2A -> 3O were most commonly expressed with extended intransitive verbs, whereas 3A -> 1O, and 2A -> 1O were most commonly expressed with transitive verbs. When only third person was involved, the changes were not so systematic;

(2) texts. For example, Monod-Becquelin (1976) observes that in a text where a Trumai person tells about the killing of an uncle by the Kayabi indians, the transitive verb ('form ergatif' in her terminology) occurs very frequently when the consultant speaks about the actions of the aggressors (kayabi) on the victims (Trumai). While telling of the revenge (Trumai on Kayabi - the Trumai indian speaks of the Kayabi people) the intransitive

('construction Sujet-Objet' in her analysis) occurs. Here we have again the difference between NPs:

attack: Kayabi (3rd pl) on Trumai (1st pl) = transitive
revenge: Trumai (1st pl) on Kayabi (2nd pl) = intransitive

But sometimes the selection does not occur in the expected way (that is, it is a tendency rather than a rigid rule).

This new pattern can be explained by the difference in topicality in objects: first person is inherently more topical than third person, for the speaker will always consider himself to be the center of speech.[9] Morphosyntactically, first person is given a topical position, as the object. As third person has inherently lower topicality, it tends to hold the less topical position of indirect object. From this follows the selection of verbal forms:

(26) 3Agt --> 1Pat = inherently topical object uses the transitive verb.

O	V			A	
ha-Ø	***disi-tke***	***ka***	***in***	***hinak***	***wan-ek***
1-Abs	hit/kill-Desid	?	?	3	pl -Erg

'They want to kill/hit me'

(27) 1Agt --> 3Pat = inherently less topical object uses the extended intransitive verb

A	V			IO
ha-Ø	***fa-tke***	***ka***	***in***	***hine-tl***
1-Abs	hit/kill-Desid	?	?	3 -Dat

'I want to hit/kill him'

It is interesting also to observe that, in the case of ***fa*** and ***disi***, there are two possible meanings for these verbs: 'hit' and 'kill'. Since 'hit' takes a locative object and 'kill' does not, one might ask how it is that (a) a single verb can mean both and (b) one verb with both meanings ends up as transitive and another as extended intransitive. We can make the hypothesis that in old Trumai there might have been one verb to one meaning, i.e. probably the extended intransitive verb *fa* meant only 'hit' and

the transitive verb ***disi*** meant only 'kill'. We would them assume that the semantics of these historically distinct verbs evolved closer to one another (i.e. ***fa*** developed the meaning 'kill' and ***disi*** the meaning 'hit'), while the morphosyntactic category of each remained constant.

6. Conclusion

These are the results obtained by our analysis. There are other facts we intend to investigate in the future, such as the following:

* the selection when, in the use of ***fa/disi*** and other pairs, the involved NPs are third person nouns (i.e. pronoun versus common noun; human noun versus non-human noun, etc.)

* If the choice of dative markers (***-s/-tl/-ki***) is partially conditioned by other characteristics of the NP-Indirect object, e.g. degree of individuation; degree of affectedness (total/partial); animacy; human versus non-human in relation to pronouns (3rd pronoun refferring to human versus 3rd pronoun non-human), etc.

* It was suggested to us that, although the choice of verb from pairs (like those shown in section 5 above) is essentially lexical rather than morphological, the syntactic effect greatly resembles that of an antipassive: for one verb, the second participant is an O (cf. the active in a language with a morphological anti-passive); for the other verb, the second participant is an oblique, the IO (cf. antipassive construction which demotes the O to oblique). The difference is that in Trumai this is a nonproductive lexical pair rather than a morphosyntatic process (reminiscent of the distinction between so-called 'lexical causatives' like the pair die/kill and truly productive morphological causatives).

That is an interesting idea to be discussed; the question is if such pairs do function as anti-passives, and if so, how productive this system would be. As we need more information about these pairs and the entire system of the Trumai language (i.e. about the possibility of a morphological anti-passive

construction in Trumai), we prefer to reserve this discussion until another paper.

NOTES

1. My research on Trumai has been conducted since 1989 in the Xingu Reserve in Mato Grosso, Brazil. The data used here were given by the consultants Kumaru, Amati and other persons from Terra Preta Village. Past research was funded by Brazilian foundations: CNPq (Conselho Nacional de Pesquisa), FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), FAEP (Fundação de Apoio à Pesquisa), CAPES (Coordenadoria de Aperfeiçoamento ao Pessoal de Ensino Superior); current research is funded by CNPq and FINEP (Fundação Nacional de Ensino e Pesquisa). A previous analysis of the Trumai verbal case-marking system benefitted from comments by Dr. R. Dixon. An earlier version of this paper was presented at the 1993 Summer Meeting of the Society for the Study of Indigenous Languages of the Americas (SSILA); financial support for participation in this meeting was given by CNPq (Brazil), the Linguistic Society of America and the AID SUNY/Training Program. Thanks to Denny Moore for encouragement and Spike Gildea for detailed comments and discussion; some ideas presented here were suggested by Scott DeLancey. Thanks also for comments received during the SSILA meeting presentation. Remaining mistakes are my own responsibility.

2. Trumai is an isolated indigenous language spoken in the Xingu Reserve, in the central region of Brazil. The Trumai people live distributed in four different places (three villages and near the P.I. Pavuru, an administrative post). While there are between 100 and 109 inhabitants in Trumai communities, due to intermarriage with indians of other Xingu tribes, the actual number of speakers is between fifty five and sixty.

3. We adopt in this paper the abbreviations used in Dixon (1979), adding also the term IO:

S	intransitive subject
A	transitive subject
O	transitive direct object
IO	indirect object

4. As the analysis of the ***i/ii*** morpheme is complex, we will not treat it in this paper. In brief, this morpheme optionally marks overt NPs. When A or IO NPs are deleted, the ***i/ii*** morpheme is also deleted; in contrast, when the S or O NP is deleted, the ***i/ii*** morpheme can remain.

5. The following abbreviations are used in examples:

Abs	absolutive case
Dat	dative case
Erg	ergative case
AGT	agent
PAT	patient
SG	singular
PL	plural
Imp	imperative

6. The words ***kasoro***, ***aros*** and ***cafe*** are borrowings from Portuguese ('cachorro, arroz and café', respectively).

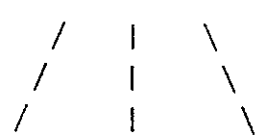
7. In reality, as LOC, IO can present variations in position:

Extended Intransitive :	(IO)	S	V	
Ditransitive :	(IO)	A	O	V

8. Perhaps there would be other pairs. At the moment (at this stage of our knowledge about the language) these are the pairs attested by us.

9. This is reminiscent of Dixon's (1979) proposed hierarchy for person-based split ergativity, and of Gildea's (to appear) discussion of a similar hierarchy for inverse systems. In both, the 1 > 3 is a basic part of the hierachy. Other kinds of NPs are considered. Dixon's hierarchy is the following:

1 > 2 > 3 pronouns > 3 proper nouns > nouns



human > animate > inanimate

While in some languages it is obligatory to follow this hierarchy, in Trumai is not, although as yet we have no evidence for the rest of hierarchy (human > animate > inanimate).

REFERENCES

Dixon, R. (1979). **"Ergativity"**. Language. vol. 55, no 1 . pp. 59-138.

Fillmore (1970). **"The grammar of hitting and breaking"**. Readings in Transformational Grammar. Ed by R. Jacobs and P. Rosenbaum, 120-133. Waltham, MA: Ginn.

Gildea, S. (to appear). **"Semantic and pragmatic inverse - 'inverse alignment' and 'inverse voice' - in Carib of Surinam"**. The pragmatics of voice: Active, Inverse, Passive. ed. by T. Givón, Typological Studies in Language, vol. 30. Amsterdam: John Benjamins.

Guirardello, R. (1992). **"Aspectos da Morfossintaxe da Língua Trumai (isolada) e de seu Sistema de Marcação de Caso"**. Master Thesis . Universidade de Campinas, Brasil.

Hopper, P. & Thompson, S. (1980). **"Transitivity in Grammar and Discourse"**. Language vol. 56, no 2. pp. 251-299

Monod-Becquelin, A. (1976). **"Classes verbales et construction ergative in Trumai"** . Amérindia 1. Paris. pp. 117-143.

NHEENGATU (LÍNGUA GERAL AMAZÔNICA), ITS HISTORY, AND THE EFFECTS OF LANGUAGE CONTACT

Denny Moore, Sidney Facundes and Nádia Pires
Museu Paraense Emílio Goeldi, Brazil

INTRODUCTION[*]

One of the most fascinating cases of a language altered by contact with other languages has remained largely unknown to English-speaking linguists--the case of Nheengatu, also called Língua Geral Amazônica. This language was once dominant throughout the settled Brazilian Amazon region and is still spoken in its modern form in some areas, especially in the region of the Upper Rio Negro.

The indigenous language which was the source of Nheengatu, Tupinambá, is known through descriptions written by Jesuit missionaries (for example, Anchieta 1595 and Figueira 1621), sources which provided the basis for the modern analysis of this now extinct language by Rodrigues (1958, 1990). Old documents in Nheengatu survive from each successive century. There are collections of texts and amateur grammatical descriptions (rigidly following European grammatical categories) from the last two centuries (Magalhães 1876, Rodrigues 1890, Silva 1945, Michaele 1951). The few modern linguistic treatments of Nheengatu include Taylor (1985, 1988), Borges (1991), Grenand and Ferreira (1989), and Rodrigues (1986: ch. 10). The latter work deals explicitly and authoritatively with the diachronic evolution of Nheengatu from Tupinambá; the others are more concerned with phonology than with grammar.

Our own research on modern Nheengatu began in Belém, Brazil, in 1987, initially as a means for teaching field methods. Rather unexpectedly, the research continued sporadically for three years, with a total of ten texts transcribed and analyzed. Emphasis was given to the syntax because of its lack of professional description.

On this basis we present a very brief description of some of the main structural features of the contemporary Nheengatu of the upper Rio Negro, noting obvious resemblances to the structure of its indigenous ancestor or to Portuguese. Unfortunately, no information is yet available on the Nheengatu of other regions and so little can be said about the important question of variation within modern Nheengatu--which may be considerable.

* We wish to thank the SSILA for creating an extra session for Brazilian Indian languages at their summer meeting. Travel to the 1993 Linguistic Institute and SSILA summer meetings was made possible by financial support from USAID and CNPq (the Brazilian National Research Council) in the case of Nádia Pires, and by support from the Inter-American Foundation in the case of Sidney Facundes. We thank our Nheengatu informant, Lenir da Silva, for her invaluable assistance with the language and for checking the examples in this paper for accuracy. Spike Gildea removed a number of errors from an earlier version. Support from the Museu Goeldi and from CNPq has been indispensable for our work.

To gain at least a superficial historical perspective on Nheengatu and its evolution, some sources on Tupinambá and on the history of Nheengatu and its relation to socio-political events in Amazonian history were consulted. (Many were not immediately available and could not be consulted.). The linguistically sophisticated work of Freire (1983) was especially useful. On this basis a quick outline of the history of Nheengatu is given, immediately below, focusing on those aspects most relevant for understanding the transmission and the modification of the language during its various phases. After the summary description of the structure of Nheengatu, some final observations are offered about the possible effects of the different types of language contact situations through which Nheengatu passed in different historical periods.

BRIEF HISTORY OF NHEENGATU

In 1500, at the time of the first Portuguese contact with what is now Brazil, the eastern coast from São Paulo to the mouth of the Amazon was occupied by native peoples speaking Tupinambá, one of the languages of the Tupi-Guaranian family (of which twenty or so still survive), the most widespread of the ten families of the Tupi linguistic stock. Since there were relatively few European women among the first colonists, many of the Portuguese men married Tupinambá women. Tupinambá was spoken in the household and the mestizo children spoke it natively (Rodrigues 1986:101).

The initial impression of the Europeans was that all the Brazilian Indians spoke the same language, and they thought that knowledge of the language would facilitate the work of conquest and conversion. The Jesuits were active with the indigenous peoples and languages, producing the descriptions by Anchieta (1595) and Figueira (1621). Figueira referred to the language of the coast as the 'Língua Brasílica'.š This name was commonly used to refer to it in the Seventeenth Century, though in the second half of that century the name 'Língua Geral' came into use, and in the latter part of the Nineteenth Century the name 'Nheengatu' became common (Rodrigues 1986: 100-103).

The colonization of the Amazon River and its tributaries lagged behind the colonization of the southern regions, where a língua franca with an indigenous base, Língua Geral Paulista, developed and then almost completly disappeared by the 18th century (Rodrigues 1986:102). The Luzo Brazilian occupation of the Amazon region began in 1616 with the establishment of Forte do Presépio in the mouth of the Amazon River.

In the Sixteenth Century two expeditions on the Amazon River had been struck by the enormous number of indigenous languages --a very different situation from the coastal uniformity. A Spanish Jesuit who traveled the Amazon River counted more than 150 different languages along the banks of the Amazon and the mouths of its principal tributaries (Acuña 1641:199, cited in Freire 1983: 42).

The European colonists (and the mestizos) depended on Indian labor to extract wealth from Amazônia. A system of slavery and 'aldeias de repartição' (resettlement villages) for 'free' Indians was established. Large numbers of indigenous people from many regions, speaking many different languages, were taken from their homes and resettled as laborers for colonists and missionaries. Língua Geral was spoken by the Europeans and mestizos to these Indian laborers.

The use of Língua Geral as a língua franca was favored by the presence of many languages of the Tupi-Guaranian family in the region and by the colonists' desire for a language to communicate with the captured labor force (as well as with their own Tupinambá allies) and by the widespread fluency in Língua Geral that had already been obtained on the coast.

Three years after the Jesuits gained control over the indigenous population through the Regimento das Missões in 1686, Língua Geral was recognized as the official language of Amazônia by the government in Portugal, which endorsed its spread. The Jesuits increased the time that the indigenous inhabitants of the resettlement villages spent in the villages, reducing the time spent in extractive activities. They systematized more the education in Língua Geral. They also increased the expeditions to subjugate and relocate native peoples from more and more remote villages.

Some census figures help understand the sociolinguistic situation at the end of the Seventeenth and the beginning of the Eighteenth Centuries. According to Baena (1831:247, cited in Freire 1983:50), in the four years 1687-1690, just from areas reached by the Tocantins, Amazon, and Negro Rivers, 184,040 Indians were seized and relocated for King and Church. By comparison, the European population was tiny. The 150 Europeans who arrived in 1616 had only grown to 1,000 by 1720, whereas only in Pará, excluding Maranhão, there were 63 resettlement villages with 54,264 Indians, as well as more than 20,000 Indian slaves and a number of mestizos (Raiol 1900:132, cited in Freire 1983:52).

Two facts are noteworthy here. One is that there were massive numbers of new speakers of Nheengatu during the phase of its expansion in the Seventeenth Century and the first half of the Eighteenth Century. Also, these new speakers were from various tribes and spoke various languages; many of which fell into disuse as the speakers' children learned Nheengatu. The existence of a multiplicity of indigenous languages among the captured Indians would favor the spread of Nheengatu, as they turned to it to communicate with each other, just as many Brazilian Indians today speak to other Indians in Portuguese.

The second fact is that there was a large community of native speakers of Nheengatu. Of the classes of people mentioned in the census of 1720, the Whites born in Brazil, the mestizos, the Indian slaves, and the more acculturated Indians in the resettlement villages spoke Nheengatu. While it was the case that the Jesuits used the language as a means of instruction, it would seem, on general grounds, that language learning in the classroom would have been much less significant as a means of transmission than was informal contact with the many native speakers during work, visiting, or religious activities.

By the middle of the Eighteenth Century Nheengatu was nearly universal in colonized Amazônia, even in the capital Belém. This success brought on its decline. Through their

knowledge of the language and control over the Indians, the Jesuits constituted a political force which rivaled that of the State. In the second half of the Eighteenth century the Jesuits were expelled, the State assumed control over Indians and attempted to introduce Portuguese influence into Amazônia. Nheengatu was persecuted and Portuguese was promoted as the language of instruction.

Instruction in Portuguese was ineffective among the Indians. Catastrophic depopulation was already decimating the resettlement villages. Between 1743 and 1750, 40,000 Indians died from diseases in the villages in Pará alone (Freire 1983: 62). In the hands of the State, the Indians continued to fare poorly. Some Portuguese settlers and African slaves were introduced into eastern Amazônia, altering the population balance somewhat in that area.

Brazil became independent in 1822. There had been native insurrections and rebellions previously, all violently surpressed. But the rebellion called the Cabanagem was a large-scale revolt by Indians, caboclos, and negros against the Europeans that lasted ten years, 1837-47, and cost 40,000 lives. The language of the Cabanos was Nheengatu. After the defeat and decimation of the Cabanos, the predominance of Nheengatu was greatly reduced, though it continued in western Amazônia, which still largely depended on Indian labor. The introduction of settlers from the Northeast in the last decades of the Nineteenth Century during the rubber boom reinforced the use of Portuguese.

Freire (1983:73) notes that Correa de Faria, in the mid-Nineteenth Century, compared the Nheengatu he had learned on the Upper Rio Negro with that of the Seventeenth Century, as described by Figueira (1621), and found it to be very different.

In this century, Portuguese has continued to replace Nheengatu, which survives, however, on the Rio Negro, on the Middle Amazon, and probably on the Solimões River.

PHONOLOGY

There are some modern treatments of the phonology of Nheengatu, especially the sketches in Taylor 1985 and 1988, and the thesis of Borges (1991). Some observations are offered by Grenand and Ferreira (1989: xiv-xvii). However, many aspects of the phonology are still debatable. We will limit ourselves to a brief, tentative characterization of the phonology of Nheengatu, using the above sources as a point of departure and indicating the details which are unresolved.

One complication is the existence of dialect differences. Another is the problem of separating vocabulary items according to their origins, since there are at least two phonological patterns present: words descended from Tupinambá and words borrowed, more or less recently, from Portuguese. The complexity of the question can be seen from the example of **jirimṹ** 'squash', which is listed as part of the vocabulary of Nheengatu by Grenand and Ferreira (62) and which they would consider to be a borrowing from Portuguese since the initial consonant, a

voiced palatal fricative, only occurs, according to them, in such borrowings. However, the word 'jerimum' is itself of Tupi-Guaranian origin, probably borrowed at an early date into Portuguese and then, apparently, borrowed again later into Nheengatu. It is difficult to recognize such examples or even those from other indigenous languages, for example, **dakirú** 'violin', said by Grenand and Ferreira (xi) to be of Tukanoan origin. The phonological analysis will, of course change greatly as a function of which vocabulary items it covers.

BORROWED VOCABULARY

There are some old borrowings from Portuguese which follow indigenous phonological patterns:

Nheengatu	Portuguese	English
sorára	'soldado'	'soldier'
kamišá	'camisa'	'shirt'

At least for bilingual speakers, recent Portuguese borrowings seem to follow the phonological patterns of Portuguese, with all the consonants and the seven vowels of that language:

Nheengatu	Portuguese	English
[hópʌ]	'roupa'	'clothing'
[presízu]	'preciso'	'necessary'
[uištudái]	'estuda'	'(he) studies'

NATIVE VOCABULARY

The surface phonemes of what appear to be non-borrowed words form a more restricted inventory. None of authors cited immediately above agree as to the details of this inventory, though they do agree on its basic components. The analysis adopted here (presented in the table on the left, below) also differs in its details from the others. Marginal or debatable phonemes are enclosed in parentheses. For comparison, the phonemic inventory of Tupinambá, from Lemle 1971 (109), based on Rodrigues 1958, is given below on the right. Some details of the sound system of Nheengatu are discussed and compared with Tupinambá or other languages of the Tupi-Guaranian family.

Modern Nheengatu

p	**t**	**(č)**	**k**	**(kw)**	**(?)**
mb	**nd**		**ng**		
(b)			**(g)**		
	s	**š**			
w	**r**	**y, ỹ**			
m	**n**				
i		**u**			
e					
	a				

′ (stress)
~ (vowel nasalization)

Tupinambá

p	**t**		**k**	**?**
b				
	s	**š**		
w	**r**	**y**		
m	**n**		**η**	
i	**i**		**u**	
e			**o**	
	a			

′ (stress)
~ (vowel nasalization)

Most occurrences of the palatal affricate, **č**, precede **i**, but a few examples do not:

čã́ 'no more' **č-áku- mã?ã́** 'don't know...'
not-know-what

While some **č** before **i** can optionally be **t**, indicating a palatalization rule like that of many dialects of Portuguese (eg. **kičí** ~ **kití** 'toward'), others cannot (eg. **čĩ́**, ***tĩ́** 'nose'), and some examples of **t** before **i** cannot be palatalized (eg. **ratíwa**, ***račíwa** 'uncle'). So, provisionally, **č** will be considered a phoneme, with some fluctuation with **t** before **i**, at the surface level.

The prenasalized voiced stops, **mb**, **nd**, **ng**, are common and recognized by all authors as phonemes. They occur initially and intervocalically, nasalizing the immediately preceding vowel, even across morpheme boundaries. They appear to occur only before oral vowels:

mbíra	'offspring'	**sẽ-mbíra**	'my offspring'
ã-mba?ú	'(I) eat'	**u-sẽndú**	'(he) hears'

Since in many Tupi-Guaranian languages (for example, Urubu-Kaapor, Kakumasu 1986: 401), the nasal consonant phonemes have prenasalized voiced oral stops as allophones before oral vowels, one would assume that the nasal series in Tupinambá is the diachronic source of both the nasal series and the prenasalized series in Nheengatu, though it is not clear what the conditioning factor for the split was. Interestingly, the principal informant prefers **yané-** as the first person plural prefix of the nominal series and **yãndé** as the free pronoun 'we'.

Oral voiced stops, **b** and **g**, are relatively scarce and are not recognized as phonemes by Taylor or Borges. However, they do occur in words which are not obvious borrowings, before oral or nasal vowels:

búya	'snake'	**se-búya**	'my snake'	(**móya** in Tupinambá)
bũmbáka	'a palm species' (Grenand and Ferreira: 26)			
tibíyara	'a bird species' (Grenand and Ferreira: 166)			
garapé	'creek'			
apigáwa	'man'			

The nasals **m** and **n** occur before or after oral or nasal vowels.

míra	'person'	**se-míra**	'my person'
nãmbí	'ear'		
amã́na	'rain'		

The palatal nasal is analyzed here as (the typically Tupian) **ỹ** instead of **ñ** because (1) it is usually a glide phonetically and (2) the vowels on either side are obligatorily nasalized, unlike the case of the nasals **m** and **n**. It occurs intervocalically and (rarely) initially.

ỹṹ	'alone'	**ãỹṹ**	'only'
ỹã?ã́	'that'	**kũỹã́**	'woman'

Unlike **ỹ**, the nasal labiovelar glide is rare. Whereas the Tupi-Guaranian languages generally have notable nasalization spreading, this is very marginal in Nheengatu. For example, in **yãndé** 'we', the initial glide is oral, and in **aẽtá** 'they', an oral vowel precedes a nasal vowel within the same syllable, at least on the surface.

Two oral glides are generally recognized for Nheengatu, **y** and **w**. As analyzed here, these are only slightly reduced high vowels which occur syllable initially and do not carry stress. Examples:

yaučí	'turtle'	**waimĩ́**	'old woman'
íwa	'tree'	**iwá**	'fruit'

Unlike Portuguese, Nheengatu, following the indigenous pattern, permits syllables containing two vowels. Note 'turtle' and 'old woman' above and also the following examples:

aẽtá	'they'	**pakúa**	'banana'	**múčiu**	'bellybutton'
u-ikú	'he is' (normal pronunciation, secondary stress on the first vowel)				

To avoid sequences of three vowels in one syllable in the example **apukwái** 'tie', we tentatively recognize a labiovelar stop, **kw**, which is probably derived from underlying **ku**. Some examples of **kw** (eg. **aikwé** 'there is') cannot be **ku**, though this sequence also exists (eg. **iku?ẽ́ma** 'light-colored').

There are only four vowel phonemes in modern Nheengatu, at least in the dialect studied. Nineteenth Century sources often note a fifth vowel, presumably ɨ.

Each Nheengatu morpheme has one primary stress. Within the word, the rightmost stress is maintained and the preceding stresses are successively reduced. Word boundaries can be determined on this basis. Example:

"'u-"mu-'kiri	**u-mu-kíri**	'he causes to sleep'
	3-transitivizer-sleep	

In our transcription we indicate the stress of each root morpheme with an acute accent mark, though only the rightmost is unreduced. Affixes, except the diminutive, the augmentative and the plural, are stressed on the syllable adjacent to the stem. Affix stress is not marked here.

The status of the glottal stop is not yet clear. Frequently it can occur optionally at morpheme boundaries intervocalically, even before an unstressed vowel, eg., **se-?iwá** 'my fruit'. It also occurs morpheme internally before stressed vowels, e.g. **ka?á** 'forest'. It may be fully predictable in this position, but for the time being, it will be transcribed when it is possible morpheme medially.

The syllable pattern of (C)V(V) in Nheengatu differs from that of Tupinambá, which permitted syllable final consonants morpheme finally.

In the transition from Tupinambá to Nheengatu, the principal changes in the inventory of segmental phonemes, pointed out by Rodrigues (1986: 104), were the merger of Tupinambá **b** (a bilabial fricative) with **w**, the merger of Tupinambá **o** with **u**, and the disappearance of the velar nasal η, with accompanying nasalization of the preceding vowel.

MORPHOLOGY

WORD CLASSES

Nheengatu words fall into eight word classes, approximately equivalent to those of Portuguese: nouns, verbs, adjectives, adverbs, postpositions, pronouns, demonstratives, and particles. Most words in modern Nheengatu texts are of native origin, though there are many borrowings. Almost all borrowings from Portuguese are nouns, verbs, or particles; the other categories seem to be essentially of indigenous origin.

Nouns can be distinguished from adjectives in that the former accept prefixes of the nominal series and the latter do not (though stative verbs homophonous with adjectives do accept these prefixes). Also, adjectives can modify nouns which precede them, but nouns cannot

modify preceding adjectives. Both simple nouns and derived nouns take the same person prefixes, e.g. **se-pú** 'my hand', **se-yasí-wéra** 'my cry-baby'.

Verbs follow a typically Tupi-Guaranian pattern, falling into three mutually exclusive subclasses, intransitive, transitive, and stative. All verbs are obligatorily prefixed for subject. Only verbs can constitute complete one word sentences:

intransitive:	**a-puraki**	'I work.'
transitive:	**a-mũỹã́** (object)	'I make (object).'
stative:	**se-ruri**	'I am happy.'

Note that the stative verbs use prefixes of the nominal series, while the intransitive and transitive verbs use prefixes of the verbal series (which occur with no other class). Borrowings from Portuguese seem to enter only the intransitive and transitive subclasses, not the stative subclass. All stative verbs have corresponding adjectives, for example **se-ruri** 'I am happy' and **suri** 'happy', but the converse is not true.

Adjectives can be either attributive (**maniáka akíra** 'green manioc') or predicative (**maniáka i-akíra** 'the manioc is green). Some predicate adjectives occur with the invariant prefix **i-**, which is homophonous with the third person of the nominal series. By contrast, stative verbs occur with all the prefixes of the nominal series, showing concordance with the (optional) subject. Adjectives, but not nouns or adverbs, accept the suffix ~**to** 'semi' (**purã́ngã́-to** 'almost good', ***ukã́-to** 'almost a house'). Adjectives, unlike transitive and intransitive verbs, cannot accept the prefixes of the verbal series.

Adverbs can be distinguished from nouns and verbs by their lack of person prefixes. They differ from adjectives in that they cannot modify preceding nouns. The free movement of adverbs also distinguishes them from particles and other word classes.

The pronouns are either personal or interrogative. The same set of personal pronouns is used as subject or as object of a verb, as in Tupinambá. Most of the Tupinambá pronouns survived into Nheengatu, but the pronominal system was reanalyzed, converging toward Portuguese. As analyzed by Rodrigues (1990: 420), the Tupinambá system functioned in terms of 'parameters of (a) contrast between speaker and hearer and (b) focality of the 3rd person', rather than the person and number system of today.

Nheengatu Personal Pronouns:

isé	1 sing	**yãndé**	1 plural
ĩndé	2 sing	**pẽỹẽ́**	2 plural
a?é	3 sing	**aẽtá**	3 plural

Nheengatu also has two interrogative pronouns which are used as question words, both from Tupinambá.

mã?ã́	'what, who, whom'
awá	'who, whom'

All the postpositions are of indigenous origin, remaining in Nheengatu even though the basic word order changed from SOV to SVO. Postpositions accept prefixes of the nominal series (e.g. **se-irṹ** 'with me'), but cannot occur with a free pronoun (***išé irṹ** 'with me'). Some of these postpositions are the following:

rupí	'through'	**upé**	'in'
su?í	'from'	**resé**	'in'
irṹ	'with'	**kití**	'to'

The Portuguese numerals can be used in modern Nheengatu, though at least the lower numerals still exist: **yepé** 'one', **mukṹĩ** 'two', and **musapíri** 'three'. Unlike in Portuguese, numbers, even borrowed ones, can precede (**dózi akayú** 'twelve years') or follow (**akayú dózi**) a noun.

There are two demonstratives (**kwá** 'this' and **ỹã?ã́** 'that'), which can precede or be the head element in a noun phrase. They cannot occur with pronominal prefixes but can occur with the plural suffix (e.g. **kwá-itá** 'these'). According to Rodrigues (1986: 105) these two elements are the only survivors of the rich Tupinambá system of demonstratives which included forms meaning 'this (close to the speaker)', 'that there (close to the hearer)', 'that over there (visible)', 'that over there (invisible)', 'that physically present', 'that we are talking about', etc.

Particles do not accept inflectional or derivational affixes, though some can form constructions with another free element. Examples:

ramẽ́	'when'
ĕí	NEGATIVE
nẽ́ĩ ~ nẽ́	'nor"
arã́ma ~ arã́	'for'
wa?á	RELATIVIZER
aikwé	'there is'
ãỹṹ	'just/only'
sá	'if'
ki	'that (COMPLEMENTIZER)'
presízo	'It's necessary'

Some of the particles are borrowed, such as **presízo** 'it's necessary' and **nẽ́ĩ ~ nẽ** 'nor', **sá** 'if' and **ki** 'that'.

COMPOUNDS

Tupinambá was morphologically complex, with an ample system of incorporation. Examples from Rodrigues (1990: 398-99):[1]

ya-y-namí-?ók-ukár
3-relational-ear-take.off-Caus
'CUT THE EAR OFF OF'

ya-y-pó-pwár-atã
3-relational-hand-tie-hard
'TIE UP HIS HANDS TIGHTLY'

Compounding is no longer a very productive process, but a variety of compounds do exist. Examples:

N + N > N	e.g.	**pí-puãpé** foot-nail	'toe nail'
N + N > Adj	e.g.	**sasí-ára** pain-day	'sad'
N + Adj > N	e.g.	**maniákã-mbéka** manioc-soft	'soft manioc'
V + Adv > V	e.g.	**kwá-katú** know-well	'think, believe'
N +N > N	e.g.	**nã́mbi-púra** ear-part.inside	'earring'
Ptc + Ptc > Ptc	e.g.	**čí-arã́ma** not-to	'to not'

AFFIXES

Affixation, as well as compounding, was reduced during the evolution of Nheengatu. Rodrigues (1986: 105) neatly sums up the grammatical changes, observing that (our translation): 'The greatest alterations suffered by Tupinambá in the process of becoming Língua Geral resulted from a progressive simplification of the grammatical forms, accompanied by reorganization of the construction of sentences'. For example, he points out that the Tupinambá verbal morphology, which included a system of five moods (indicative, imperative, gerund, circumstantial and subjunctive) converged to the indicative mood. The noun morphology, which included a system of conjugations in six grammatical cases (nominative, vocative, attributive and three locative cases) was lost in Nheengatu.

1 Special abbreviations:

Caus	= causativizer	Cop	= copula	V'	= complex verb
Neg	= negative	Rel	= relative	R[illegible]	= relational
Ptc	= particle	Foc	= focus	R[illegible]	= relativizer
Comp	= complement	S'	= embedded clause	[illegible]	= future

The inflectional and derivational affixes of modern Nheengatu are from Tupinambá; that is, there is no borrowing of any Portuguese affixes. Even recent Portuguese borrowings in Nheengatu can accept person prefixes. Some modern affixes seem to be the result of grammaticization of what were formerly lexical items. The plural suffix, -itá, a convergence toward Portuguese, was formerly a lexical item, etá, meaning 'many' (Rodrigues, personal communication). Some modern affixes:

INFLECTION		DERIVATION	
Verbal Series		mu-	TRANSITIVIZER
a-	1 sing	yu-	INTRANSITIVIZER /REFLEXIVIZER.
re-	2 sing	-~to	'semi, almost'
u-	3	-mã̄ỹã	'someone with tendency for...'
ya-	1 pl	-sára ~ -gára	AGENT
pe-	2 pl	-éra ~ -wéra	'habitual doer of...'
aẽtá-ú-	3 pl.	-wára	'characterized by'
		-ĩ́ma	'without'
Nominal Series			
se-	1 sing	-mirí	DIMINUTIVE
ne-	2 sing	-asú	AUGMENTATIVE
i-	3		
yané-	1 pl		
pe-	2 pl		
aẽtá-	3 pl		
-itá	PLURAL		
-ã́na ~ -wã́na	PERFECTIVE		
-rẽ́	IMPERFECTIVE		

Reduplication to indicate repetitive action has been retained as a morphological process in Nheengatu, for example, ya-yapí 'throw or shoot repeatedly', pi-píka 'drizzle'. Reduplication was present in Tupinambá, as in most Tupian languages.

SYNTAX

MATRIX CLAUSES

The matrix clauses show very few borrowings of grammatical morphemes. Their syntax shows convergence toward Portuguese in some aspects and preservation of characteristically indigenous features in other aspects.

Sentence Types

There are three sentence types in Nheengatu, in embedded as well as matrix clauses. Verbal sentences consist of an optional subject followed by one or more VP's containing verbs prefixed for subject. These verbs may be intransitive, transitive, or stative. Transitive verbs are optionally followed by an object, as in Portuguese, in contrast to the Object -Verb order of Tupinambá (and of most Tupian languages). (In the examples below, embedded rather than matrix clauses are given as illustrations if the text examples of the latter are lacking or unclear.) Examples of verbal sentences:

[ya-mũỹã́]$_{VTrans}$ čĩmbi?ú [ya-pinačika]$_{VIntr}$ [ya-mũỹã́]$_{VTrans}$ kaširí
1p-make food 1p-fish 1p-make chicha
'WE MAKE FOOD, WE FISH, WE MAKE CHICHA.'

[yã-mbúri]$_{VTrans}$ maniáka paranã́ upé [i-mẽmbéka]$_{VStat}$ arã́ma
1pl-put manioc river in 3s-be.soft to
'WE PUT THE MANIOC IN THE RIVER TO BECOME SOFT.'

There are two verbs which might be considered auxiliaries, which occur after the main verb, contrary to the order in Portuguese: **putái** 'want' and **ikú** 'be'. The former can occur without a subject prefix, forming a complex verb. The latter can be preceded by a verb, an adjective, or a postpositional phrase. Examples:

[a-yuwíri putái]$_{V'}$ se-retã́ma kiti
1s-return want 1s-land to
'I WANT TO RETURN TO MY LAND.'

yãndé [[ya-purĩngitá]$_{V}$ [ya-ikú]$_{Aux}$ ye?ẽngatú]$_{VP}$
we 1p-speak 1p-be Nheengatu
'WE ARE TALKING NHEENGATU.'

išé [se-rúka upé]$_{PP}$ a-ikú
I 1s-house in 1s-be
'I'M IN MY HOUSE.'

The copula sentence type consists of an obligatory subject followed by a predicate noun phrase or adjective phrase. There is no overt copula, unlike in Portuguese. Examples (with inverted order):

rẽ-mbe?ú aẽta-supé [purã́nga išé]$_{S'Cop}$
2s-tell 3p-for good I
'TELL THEM THAT I'M FINE.'

The third sentence type consists of a VP with no subject. The VP is composed of a predicative particle followed by a NP or by a clause with an overt subject. These resemble impersonal constructions in Portuguese except that the predicative particle shows no verbal characteristics. At least one of them, **presízu** (< Port.: É preciso...) 'It is necessary...' is borrowed, and the first syllable of **aikwé** 'there is' looks like Portuguese 'aí' 'there'. Example:

[aikwé]$_{Ptc}$ kašuéira
there.be waterfall
'THERE ARE WATERFALLS.'

[presízo]$_{Ptc}$ aẽtá u-ištudái pohtugéš upé
need they 3-study Portuguese in
'IT'S NECESSARY THAT THEY STUDY IN PORTUGUESE'

Syntactic Processes in Matrix Clauses

The major syntactic processes affecting matrix clauses look more indigenous than European.

Negation. Verb phrases can be individually negated with the particle **čí**:

čí [a-pitá]$_{VP}$ a-iwíri kwá-kití
not 1s-stay 1s-return this-toward
'I DON'T STAY, I COME BACK TO BELÉM.'

Note the structural similarity of this to multiple negation in Tupinambá (Rodrigues 1985: 399):

i-sɨ́ n i-memɨ́r-asɨ́-y na s-uwɨ́-y n i-mara?ár-i...
Rlt-mother not Rlt-son-pain-Neg not Rlt-blood-Neg not Rlt-sick-Neg
'HIS MOTHER DID NOT FEEL ANY CHILDBIRTH PAIN, DID NOT BLEED, WAS NOT SICK...'

The negative particle can occur in the beginning of the clause, negating all of it. It can also form a negative focus construction with a fronted NP:

[čí tapi?íra]$_{Foc}$ apigáwa u-yuká
not tapir man 3-kill
'IT WAS NOT THE TAPIR THAT THE MAN KILLLED.' (elicited)

Topicalization. Noun phrases can be topicalized, leaving behind third person copies:

[ỹã?ã́ yawára,]$_{Top}$ a?é u-su?ú apigáwa
that dog it 3-bite man
'THAT DOG, IT BIT THE MAN.'

Questions. Polar questions can be formed by intonation.

ĩndé re-murái apekatú kwá-su?í tetã́ma suí?
you 2s-live far this-from city from
'DO YOU LIVE FAR FROM HERE FROM THIS CITY?'

Interrogative word questions are formed using indigenous interrogative words and the particle **ta?á**.

mã?ã́ ta?á re-wasẽ́mu pušuéra?
what Q 2s-find ugly
'WHAT DO YOU FIND UGLY?

As in Portuguese, the interrogative word need not necessarily be fronted.

taína u-mã?ã́ mã?ã́?
child 3-see what
'THE CHILD SAW WHAT?'

Adverbial movement. Sentence level adverbials can be fronted or placed between phrases.

[kuši?ĩ́ma]$_{Adv}$ aikwé yepé feičiséiro a-koñeséi wa?á
formerly there.be a shaman 1s-know Relz
'FORMERLY, THERE WAS A SHAMAN WHOM I KNEW.'

Some common syntactic processes in Portuguese, such as passives or clefts, do not occur in Nheengatu.

EMBEDDED CLAUSES

Nheengatu embedded clauses are especially noteworthy in that they show three different patterns:

(1) Subordinate clauses formed on an indigenous pattern
(2) Subordinate clauses formed on a Portuguese pattern, but using indigenous morphemes
(3) Frank borrowings from Portuguese, with accompanying Portuguese grammatical morphemes.

In the first pattern, the clause contains a subordinating particle immediately after the head of the VP, that is, after the main verb, after the predicate nominal or adjectival, or after the predicating particle, according to the type of the VP. These particles include **wa?á** RELATIVIZER, **ramẽ́** TIME, **arã́ma** PURPOSE and **či-arã́ma** NEGATIVE PURPOSE.

(This last particle occurs clause initially). The relative clauses can have an external head and a corresponding empty internal extraction site:

a-yururé se-mãỹã u-pitá arãma iane-rẽndá upé [se-ratíwa u-šári wa?á yãndé arã]$_{S'Rel}$
1s-ask 1s-mother 3-stay Purpose 1p-farm in 1s-grandpa 3-leave Relz us Purpose
'I ASKED MY MOTHER TO STAY IN OUR FARM THAT MY GRANPA LEFT FOR US'

(Note that the relative clause modifying 'farm'; has been extraposed from inside the postpositional phrase to the end of the sentence.)

Or they may be headless, with one missing argument:

aẽtá u-kõtái [∅ u-akõteséi wa?á garapé apíra kití]$_{S'Rel}$
they 3-tell ∅ 3-happened Relz stream headwaters toward
'THEY WOULD TELL US WHAT HAPPENED ON THE HEADWATERS OF THE STREAM.'

The time, purpose and negative purpose clauses formed by **ramẽ**, **arãma**, and **čí-arãma**, respectively, distribute like adverbials or adjectivals:

aẽtá u-pisíka pa?á yãndé [ya-ú ramẽ čĩmbi?ú irusãnga]$_{S'Adv}$
they 3p-catch they.say us 1p-eat Time food cold
'THEY WOULD CATCH US WHEN WE ATE COLD FOOD.'

yã-búri maniáka paranã upé [i-mẽmbéka arãma]$_{S'Adv}$
1p-put manioc river in 3-be.soft Purpose
'WE PUT THE MANIOC IN THE RIVER IN ORDER FOR IT TO BECOME SOFT.'

ya-ú čĩmbi?ú sakú [čí-arãma kurupira-itá u-rasú yãndé]$_{S'Adv}$
1p-eat food hot Neg-Purpose kurupira-Pl 3-take us
'WE WOULD EAT HOT FOOD FOR THE KURUPIRA NOT TO TAKE US AWAY.'

ya-mãỹã čĩmbi?ú [apigáwa u-ú arãma]$_{S'Adj}$ (elicited)
1p-see food man 3s-eat Purpose
'WE SAW THE FOOD FOR THE MAN TO EAT.'

In the second pattern, a subset of the Nheengatu WH words (MA words) are used in embedded clauses in a manner similar to that of Portuguese. The MA words are **awá** 'who(m)', **mã?ã** 'which, that', **mairamẽ** 'when', **marãma** 'because', **mamẽ** 'where', and **mayé** 'as'. The relative clauses with **awá** and **mã?ã** cannot have external heads:

[mã?ã́ u-yururé i-tupã́na u-yũmbu?é tupã́na supé...]$_{S'Rel}$
what 3-asked 3s-god 3-pray god to
'WHAT HE ASKED (FROM) HIS GOD, PRAYING TO HIS GOD...'

*apigáwa [mã?ã́ u-yururé i-tupã́na]$_{S'Rel}$...
man who 3-asked 3-god
('THE MAN WHO/THAT ASKED HIS GOD...')

The clauses formed by the other MA words distribute as adverbials or adjectivals:

aẽtá u-mã?ã́ úka [mamẽ́ a-murái]$_{S'Adj}$
they 3-see house where 1s-live
'THEY SAW THE HOUSE WHERE I LIVE.'

išé čí a-sasá i-puší [mayé aẽtá ũ-mbe?ú]$_{S'Adv}$
I not 1s-pass 3-bad how they 3-say
'I'M NOT HAVING A BAD TIME LIKE THEY SAY.'

Embedded questions also follow the Portuguese pattern, but using indigenous MA words:

...čí aẽtá u-kwá [mã?ã́ kurupíra-itá u-mũỹã́ yane-irṹ]$_{S'Q}$
not they 3-know what kurupira-Pl 3-do 1p-with
'...THEY DIDN'T KNOW WHAT THE KURUPIRA DO TO US.'

Some transitive verbs can take unmarked sentential complements:

...nẽ́ išé a-mãndu?ái [a-mã?ã́ ĩndé]$_{S'Comp}$
nor I 1s-think 1s-see you
'...NOR I THINK OF SEEING YOU.'

In the third pattern, obvious borrowings from Portuguese include:

function:	Nheengatu	Portuguese	English
complementizer	kí	'que'	'that'
conjunction	i	'e'	'and'
disjunction	u	'ou'	'or'
negative disjunction	nẽ́	'nem'	'neither/nor'
conditional	sá	'se'	'if'

...[re-murái iké]$_{VP}$ u [re-murái ĩterió kičí]$_{VP}$
2s-live here or 2s-live interior toward
'...YOU LIVE HERE OR IN THE INTERIOR...'

PHRASES

The structure of phrases is rather conservative. Two major changes from the indigenous pattern are the order Verb - Object in the VP and the greater elaboration of adjectival and adverbial phrases as incorporation within the verb declined. Noun phrases retain the order Genitive - Noun:

[yane-ye?ẽ́nga]$_{NP}$ 'OUR LANGUAGE'
1p-language

[kariwa ye?ẽ́nga]$_{NP}$ 'WHITE MAN'S LANGUAGE'
White.man language

Also the order Noun - Adjective:

[ć́í ya-pudéi [[[ya-ú [ć́ĩmbi?ú irusã́nga]$_{NP}$]$_{VP}$]$_{S'Comp}$]$_{VP}$
not 1p-can 1p-eat food cold
'WE CANNOT EAT COLD FOOD...'

And Demonstrative - NP:

[[ku?á [se-awá-itá purã́nga]$_{NP}$]$_{NP}$...
this 1s-hair-Pl pretty
'THIS PRETTY HAIR OF MINE...'

There is a position after the head of the VP which contains aspectual suffixes, subordinating particles, and auxiliaries:

a-mũỹã́ pãỹẽ́ mã?ã́ mamẽ́ [a-purakí wa?á a-ikú]$_{VP}$
1s-do all what where 1s-work Relz 1s-be
'I DO EVERYTHING WHERE I'M WORKING.'

[[a-síka ramẽ́]$_{VP}$]$_{S'}$ sẽ-mbíra-itá [u-kíri-ãna u-ikú]$_{VP}$
1s-arrive Time 1s-child-Pl 3-sleep-already 3-be
'WHEN I ARRIVE, MY CHILDREN ARE ALREADY SLEEPING.'

Nheengatu retains postpositions, in contrast to the prepositions of Portuguese, which occur as the head of postpositional phrases which, as is characteristic of Tupian languages, have a strictly adverbial distribution, never modifying nouns.

a-morái ramẽ́ [[se-páya]$_{NP}$ irṹ]$_{PP}$...
1s-live Time 1s-father with
'WHEN I LIVED WITH MY FATHER...'

TEXT FRAGMENT OF MODERN NHEENGATU

Conversation in Belém between Two People from the Upper Rio Negro

This is the beginning of a text which was recorded and transcribed in 1988, in the Museu Goeldi in Belém. The two speakers are Lenir da Silva, a young woman in her thirties from the region of the Upper Rio Negro, trilingual in Nheengatu, Portuguese, and Spanish, and Gerson, a somewhat younger man from a Baniwa community who is bilingual in Nheengatu and Portuguese, and who lives in the city of San Gabriel da Cachoeira.

Gerson: __

1. ĩndé mu?í akayú ta?á re-morái iké kwá sidádi upé
   you how.many years Q 2s-live here this city in
'FOR HOW LONG HAVE YOU LIVED HERE IN THIS CITY.?'

Lenir:__

2. išé akayú nóvi akayú-ana a-yuwíri se-retã́ma su? í
   I year nine year-already 1s-return 1s-city from
'IT HAS BEEN NINE YEARS THAT I LIVE IN THIS CITY.'

3. išé a-yupukwá iké
   I 1s-accustom here
'I GOT USED TO THIS PLACE.'

4. išé cí a-mãndu?ái a-yuwíri se-família-itá rúka kití
   I not 1s-think 1s-return 1s-family-Pl house to
'I DON'T THINK OF RETURNING TO MY FAMILY'S HOUSE'

5. a-kwakatú išé čí a-yupukwá a-kití
   1s-believe I not 1s-accustom there-to
'I THINK I CANNOT ACCUSTOM MYSELF TO THAT PLACE ANYMORE.'

6. a-pitá kurí iké até kumairamẽ́ Tupãna-itá kurí u-kwá
   1s-stay Fut here until when God-Pl Fut. 3-know
'ONLY GOD KNOWS HOW LONG I'M GOING TO STAY HERE.'

7. mayé ta?á a-sú ãỹṹ a-watá se-retã́ma kití a-mã?ã́ arã́ se-anã́ma-itá
   how Q 1s-go only 1s-walk 1s-city to 1s-see Purpose 1-family-Pl
'HOW CAN I GO BACK TO THAT CITY ONLY TO SEE MY FAMILY?'

8. išé čã a-mãndu?ái a-yuwíri a-kití
   I not 1s-think 1s-return there-to
'I DON'T THINK OF GOING BACK THERE.'

9. a-yuwíri kurí ãỹṹ a-mã?ã́ arã́ se-anã́ma-itá
1s-return Fut only 1s-see for 1s-family-Pl
'I WILL GO BACK THERE JUST TO VISIT MY FAMILY.'

Gerson:__

10. kuši?ĩ́ma re-yúwi ramẽ́ kwá kití mayé-ta re-yúwi arã́
formerly 2s-come Time that-to how-Q 2s-come Purpose
'FORMERLY, WHEN YOU CAME HERE, HOW DID YOU COME?'

11. aikwé áwa u-rúi ĩndé o re-yúwi putái te ne-rupí ...
there.be who 3-bring you or 2s-come want even 2s-by
'WAS THERE ANYBODY TO BRING YOU OR DID YOU YOURSELF WANT TO COME?'

TEXT FROM THE NINETEENTH CENTURY

This text is from *Poranduba* (Rodrigues, 1890:87-88). It is reproduced as it was written, in a transcription based on the Portuguese orthography, with no morpheme boundaries indicated and prefixes often written separately. This myth is from the Rio Solimões, about the origin of a bird species, Tinkuan (Cocculus cornutus L.), held to be an omen. The leaves of the carayuru plant produce a red dye. There are few, if any, Portugese borrowings in the text. The translation is ours.

UIRA-PAYÉ NHEENGAREÇARA
The Spirit Bird Sings

Uirá payé paá, mocoin tayra tuichaua aitá cuchiyma maarupiara, arecé
bird shaman they.say two sons chiefs they formerly happy for.this

cuité aitá tutyra u mutara ima
therefore them uncle 3 hate

'THEY SAY THAT THE SPIRIT BIRDS WERE, FORMERLY, TWO SONS OF A CHIEF, VERY HAPPY, FOR WHICH AN UNCLE HATED THEM.'

U cenõe, paá, aitá, u ayuri u itêca muirá u munhan arama cupichuaa,
3 called they.say them 3 invite 3 cut.down trees 3 make to field

u mucáo i cunhambira etá. Aé uana, paá, u iucá.
3 got.drunk 3 nephew plural Then, they.say 3 killed

'HE CALLED THEM AND INVITED THEM TO CUT TREES, TO MAKE A FIELD AND THEN GOT HIS NEPHEWS DRUNK. THEY SAY THAT THEN HE KILLED THEM.'

Aé uana aitá uiuire i aria pére, aitá anga iunto ana.
Then they returned 3 grandma with they soul only already

Aitá u purandú imu çupé:
They 3 asked brother to

'THEN THEY RETURNED TO THEIR GRANDMOTHER WHEN THEY WERE ONLY SOULS. ONE BROTHER ASKED THE OTHER:'

-- Mahy taá ne querpe?
How question 2 dream

'WHAT DID YOU DREAM?'

-- Ce querpe racói, cha yá çuca carayuru irumo.
I dreamed in.this.way I we washed carayuru with

'I DREAMED THAT WE WASHED WITH CARAYURU.'

-- Yaué tenhen racói iché ce mu.
that.manner also that.way I my brother

'I DREAMED THE SAME.'

Aintá aria uité u moacó aitá remiú. U neeng cuité aitá:
Their grandma then 3 heated their food 3 speak then they:

'WHEN THEIR GRANDMOTHER HEATED THEIR FOOD THEY SAID:'

-- Ah! ce aría, inti uana yá icó mira arama, yaué anga iunto ana.
Ah! my grandma not already we are people in yes soul only already

'AH! GRANDMOTHER, WE ARE NO LONGER PEOPLE, BUT ONLY SOULS.'

Eré ce aría, cha çu ana ne chii, re cenoe ramé cha neengare,
well my grandma I go already 2 from 2 hear when I sing

cha munhan ramé: "Tincuan! Tincuan!..."
I make when "Tincuan! Tincuan!..."

'SO, GRANDMOTHER, WE WILL LEAVE YOU AND WHEN YOU HEAR ME SING "TINCUAN! TINCUAN!..."'

re iauáo oca queté, cha neengare ramé cuité "Titi..ti..ti.." aramé re icuáo.
2 flee house for I sing when then "Titi..ti..ti.." then 2 will.recognize

'FLEE FOR YOUR HOUSE, AND WHEN I SING "TITI..TI..TI.." THEN YOU WILL RECOGNIZE ME.'

Nhaan piranga uaá ceçá recó çôui cuêra
That red that eyes in blood past.thing

'THE RED IN THEIR EYES WAS BLOOD.'

NHEENGATU AND THE EFFECTS OF LANGUAGE CONTACT

For an adequate account of the modifications in Nheengatu/Língua Geral induced by language contact over the last half millennium, it would be necessary to have a detailed account of its sociopolitical context in each historical period as well as an analysis of the language structure and lexicon as these evolved. The task is not impossible, since relevant documents do exist. Of course, for each of the linguistic descriptions which have been made, it is not immediately obvious what the relation is between that description and the speech of the community of speakers, given the possibility of regional or social dialects, of a prescriptivist attitude on the part of the person making the description, or of common errors and misinterpretations.

One fact is clear: the language called today Nheengatu has changed at a rapid rate: the contemporary form would not be mutually intelligible with its form of 400 years ago. Other Tupi-Guaranian languages have not shown the same changes or the same rate of change. More than natural language change was at work to produce the changes in Nheengatu. At the same time, Nheengatu is far from mutually intelligible with Portuguese, with which it has coexisted for centuries.

A second fact is that there was always a sizable community which spoke Nheengatu or its precursors as a first language; it was never a pidgin. There is a belief among some traditional authors on the subject that Nheengatu was a product of the Jesuits. Rodrigues (1887: x-xi) goes so far as to say that changes occurred in Língua Geral in the Amazon Valley because (our translation), 'There it was great the number of missionaries, all with different accents, who taught the languages to Nheengaiba [non-Tupi-Guaranian speaking] tribes, planting degenerate seeds in terrains of different natures, which resulted in a general corruption, not only in pronunciation, but also in meaning'. No evidence is given that this was the real cause of change, and the patterning of the changes observed points to other processes.

Assuming that a good-sized native speaker population was the main source of transmission of the language, we may look at the historical phases of Nheengatu development and see if the types of sociolinguistic effects one would predict do, in fact, agree with the linguistic record, in so far as it is known to us.

In the first century of Portuguese contact with the Tupinambá on the coast there would have been few Tupinambá who spoke Portuguese, in relation to the large numbers who did not speak it. But intermarriage would increase the proportion of Europeans who spoke the indigenous language, as well as create a group of mestizos who spoke the indigenous language but did not have an indigenous social identity.

Rodrigues (1887: viii) notes differences between the descriptions of Anchieta (1595), who lived in Bahia and Espirito Santo, and that of Figueira (1621) who lived in Maranhão. According to him, 'Anchieta wrote the speech which he learned from the Guayanazes, Tamoyos, and Tupis; Figueira that of the Tabayaras, Potiguaras, and Tupinambás properly speaking; and Montoya that of the Guaranians, Payaguás, Charruas, etc.' (Rodrigues 1887: ix). In this picture it is difficult to separate language change from dialect differences. There were relatively few borrowings from Portuguese in the early period, which is what would be expected if Portuguese was not much used by the indigenous and mestizo populations.

During the period of the expansion of Nheengatu, the Seventeenth Century and the first half of the Eighteenth Century, bilingualism with Portuguese continued at a rather low level. The major factor was, rather, the incorporation of enormous numbers of new speakers into the speech community through slavery and resettlement villages. One would expect extensive substratum effects from speakers of many different indigenous languages undergoing language shift as they are absorbed into the Nheengatu-speaking colonial system.

In fact, in the Eighteenth Century Nheengatu was already recognized as distinct from Tupinambá. It was the language of Amazonian colonial society, not the language of an indigenous tribal group. As would be expected, borrowings from Portuguese were limited, but the grammar was altered by so many new speakers. The simplification of the morphology described above was underway at this time (Aryon Rodrigues, personal communication), though the exact sequence of grammatical and phonological changes during this phase are not yet known to us. It is clear from the Nheengatu documents of the Nineteenth Century (see text above) that the reduction of the morphology had already occurred by then.

After Nheengatu was officially discouraged and many of its speakers killed during the Cabanagem, the proportion of Portuguese speakers in Amazônia increased, as well as bilingualism in Portuguese among those who spoke Nheengatu. Texts and commentaries on Nheengatu from the second half of the Nineteenth Century are readily available. These show increased Portuguese influence, with the speech of Pará, according to Barbosa Rodrigues (xii-xiii) being the most 'corrupt' He notes the addition of vowels to eliminate closed syllables. As noted above, Correa de Faria was struck by the difference between the Seventeenth and the Nineteenth Century forms of the language.

Still, even in the latter half of the Nineteenth Century, the lexical borrowings one finds (e.g. **papéru** (<papel) 'paper', **muratú** (<mulato) 'mulato', **kabarú** (<cavalo) 'horse') are phonologically marked as older acquisitions. The obvious Portuguese borrowings are lexical items. Only a few grammatical words, such as **será** 'interrogative', were borrowed. Alongside this very limited lexical diffusion is a far more extensive and more subtle influence from

Portuguese: the many examples in the syntax of what Thomason and Kaufman (1988:351) refer to as structural diffusion without the diffusion of native morphemes. Note in the following example that a native interrogative word, **mã?ã́** (**maan** in the old transcription, retained here) 'what' occurs as the external head of a relative clause formed by the native relativizing particle **wa?á** (**uaá**) (Rodrigues 1890:37):

Cuere tenhê re u maan [çacu uaá]$_{S'Rel}$
now not 2 eat what hot relativizer
'NOW YOU DON'T EAT THAT WHICH IS HOT.'

In modern Nheengatu this relativizer is usually deleted, as in the examples in the syntax section above. This origin explains why such interrogative word relatives in Nheengatu cannot have an external head, which they can have in Portuguese: the 'what' word entered into the relatives as an external head, not as a relative pronoun.

The grammaticization of **etá** 'many' to become the plural suffix, **-itá**, was already complete in the Nineteenth Century.

Thomason and Kaufman observe that such cases of structural diffusion are only attested from situations of sustained language contact over centuries. That was the case with Portuguese and Nheengatu. In spite of the limited lexical borrowings, the constant interface with Portuguese produced structural diffusion as shown, for example, in the embedded clauses and also in the reanalysis of the pronominal system.

At the present time most Nheengatu speakers in Brazil also speak Portuguese. There is heavy lexical borrowing from Portuguese, and borrowed words accept native inflectional morphology. As expected, it was only after this extensive bilingualism that syntactic patterns using borrowed morphemes appeared. These are now noticeable in Nheengatu. For example the complementizer **ki** (<Port. 'que') now appears, as well as conjunction with **i** (<Port. 'e') and disjunction with **u** (<Port. 'ou'). A number of affixes from the last century listed in Stradelli (1929) are no longer in use.

In the region of the Upper Rio Negro Nheengatu is generally considered by tribal Indians to be a language of the Non-Indians, while among Portuguese speakers Nheengatu is often considered to be an indigenous language. It is certainly a remarkable language, whose further study will enrich our knowledge of language contact processes.

REFERENCES

Acuña, Cristobal de. 1641. Novo Descubrimento do Grande Rio das Amazonas. *Descubrimentos do Rio das Amazonas*, by Carnaval et alii, 1941, São Paulo and Rio de Janeiro: Cia Editora Nacional,.

Anchieta, Joseph de. 1595. *Arte de Grammatica da Lingua mais Usada na Costa do Brasil.* Coimbra: Antonio Mariz.

Baena, Antonio Ladislau Monteiro. 1831. Representação ao Conselho Geral da Provincia do Pará Sobre a Especial Necessidade de um Novo Regulamento Promotor da Civilização dos Índios da Mesma Provincia, *Annaes da Bibliotheca e Arquivo Público do Pará - Belém* (ABAPP), vol. 2, pp.241-292.

Borges, Luiz C. 1991. *A lingua geral amazônica: aspectos de sua fonêmica.* Master's thesis, Universidade Estadual de Campinas, Campinas, São Paulo.

Figueira, Luís. 1621. *Arte da Lingua Brasilica.* Lisboa, Manoel da Silva, . New edition in 1687, by Miguel Deslandes: *Arte de Grammatica da Lingua Brasilica.* Facsimile by J. Platzmann in 1878: *Grammatica da Lingua do Brasil*, Leipzig B. G. Teubner. More recent reedition by Emílio Allain: *Arte de Grammatica da Lingua Brasilica*, Lombaerts & C., Rio de Janeiro, s.d.

Freire, José Bessa. 1983. Da "fala boa" ao português na amazônia brasileira. *Amérindia 8.* Paris.

Grenand, Françoise and Epaminondas Henrique Ferreira. 1989. *Pequeno Dicionário da Lingua Geral.* Série Amazonas-Cultura Regional, 6. Manaus: SEDUC.

Kakumasu, James. 1986. Urubu-Kaapor. *Handbook of Amazon languages*, ed. by Desmond C. Derbyshire and Geoffrey K Pullum, vol 1, pp. 326-403. Berlin: Mouton de Gruyter.

Lemle, Miriam. 1971. Internal classification of the Tupi-Guaranian linguistic family. *Tupi Studies I*, ed. by David Bendor-Samuel, pp.107-129. Norman, Oklahoma: SIL.

Magalhães, Couto de. 1876. *O Selvagem.* Rio de Janeiro. Typographia da Reforma.

Michaele, Faris Antonio S. 1951. *Manual de Conversação da Lingua Tupi.* Ponta Grossa, Paraná, Brazil: Edições Euclidianas (Biblioteca Brasílica), Centro Cultural Euclides da Cunha.

Raiol, Domingos Antonio, barão de Guajará. 1900. Catechese de Indios do Pará. *ABAPP*, tomo II, pp.117-183.

Rodrigues, Aryon. 1958. *Phonologie der Tupinambá Sprache*. PhD thesis, University of Hamburg.

Rodrigues, Aryon. 1986. *Línguas brasileiras; para o conhecimento das línguas indígenas*. São Paulo: Edições Loyola.

Rodrigues, Aryon. 1990. You and I = neither you nor I: the personal system of Tupinambá (Tupi-Guaraní). *Amazonian linguistics; studies in lowland South American languages*, ed. by Doris L. Payne, pp.393-405. Austin: University of Texas Press.

Rodrigues, João Barbosa de. 1890. Poranduba Amazonense. *Annaes da Bibliotheca Nacional do Rio de Janeiro*, vol. XIV, fascicle 2, Rio de Janeiro, pp. i-xv, 1-334.

Silva, Protasio I. R. da. 1945. *Tupi ou Nheengatú e Português -- Gramática Umbueçaua Mucameençáuêra*. Manaus, Amazonas, Brazil.

Stradelli, Ermana. 1929. *Vocabularios da Lingua Geral Portuguez-Nheêngatu e Nheêngatu-Portuguez*. Revista do Instituto Histórico e Geographico Brasileiro, tomo 104, vol. 158, pp. 5-768. Rio de Janeiro.

Taylor, Gerald. 1985. *Apontamentos sobre o nheengatu falado no Rio Negro, Brasil*. Amérindia 10:5-23. Paris.

Taylor, Gerald. 1988. *Ortografia do nheengatu; proposta de um sistema gráfico para transcrever a Lingua Geral do Rio Negro*, Amazonas, Brazil. Unpublished manuscript, 11 pp. Centre National de la Récherche Scientifique, Paris.

Thomason, Sarah Grey and Kaufman, Terrence. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

RECONSTRUCTION OF PROTO-TUPARI CONSONANTS AND VOWELS[1]

Denny Moore and Ana Vilacy Galúcio

Museu Paraense Emílio Goeldi, Belém, Brazil

INTRODUCTION

According to the most recent classification by Rodrigues (1984/85), there are ten linguistic families within the large Tupi stock: Tupi-Guarani, Munduruku, Mawé, Juruna, Aweti, Mondé, Ramarama, Arikém, Puruborá, and Tupari. Languages of the last five families are spoken in the Brazilian state of Rondônia (north of Bolivia). Published linguistic descriptions have mainly concerned themselves with the first two, Tupi-Guarani and Munduruku; the other families have come under study more recently. Until now the only proto-language within the Tupi linguistic stock which has been the object of detailed reconstructive work is Proto-Tupi-Guarani (e.g. Lemle 1971, Leite & Facó 1991, Rodrigues 1984/85, Jensen 1984). As part of a long-term comparative Tupi project, linguists linked to the Museu Goeldi in Belém, Brazil, are currently working in all the Tupian families in Rondônia and also in the Juruna family.

This article presents part of the initial results of some of the research which has been initiated on languages of the Tupari family during the last five years. It presents a preliminary reconstruction of the sound segments of Proto-Tupari, the mother language of the four modern languages of the Tupari family: Ayuru (Wayoró, Wayru, Ajuru), Makurap, Mekens (Mequém, Mequens), and Tupari. While the preliminary reconstruction presented here is a modest effort, undertaken mainly to guide further research on the Tupari languages, it is the only study of the four languages or of their prehistory written in English. The reconstruction of Proto-Tupari is important for Tupi comparative studies because Tupari is one of the three Tupian families with enough surviving members (four) sufficiently diverged (the languages are not mutually intelligible) to permit a reliable reconstruction at a considerable time depth, using the comparative method. The other two such Tupian families are Tupi-Guarani and Mondé.

In what follows, information about the speakers of the Tupari languages is given. Data sources and limitations are explained. The sound systems of the four languages are briefly summarized. Then the systematic sound correspondences are presented, along with a tentatively reconstructed proto-segment for each. The reconstructions are justified and the diachronic processes leading to the modern languages are summarized. Lastly, the cognates and their reconstructed forms are presented.

THE AYURU, MAKURAP, MEKENS AND TUPARI

The peoples speaking the languages of the Tupari family lived traditionally on the headwaters of various rivers, most of which drained south into the Guaporé River, which is the boundary between Brazil and Bolivia. The Makurap (and also the Aruá, of the Mondé family) lived on the headwaters of the Rio Branco, the Ayuru on the Rio Colorado, and the Mekens on the Rio Mequens. The Tupari lived on the headwaters of tributaries of the Rio Machado (Ji-Paraná). Also on this more northerly watershed were the Kepikiriwat, whose language, now apparently extinct, is the fifth language of the Tupari family. From the surviving wordlist it appears to be more remote from the other four languages than they are from each other.

According to Meireles (1989), the existence of the Mekens was reported in the Eighteenth Century, in the region of the Rio Mequens. There were two groups (Amniapá and Guaratágaja), whose speech was very similar.

Sustained contact with national society began in the second quarter of this century for most of these groups as rubber gatherers entered the region. The results of contact were usually economic exploitation and decimation through disease. Descriptions of the indigenous people of southern Rondônia include Caspar 1956, Métraux 1948, Lévi-Strauss 1948, Becker-Donner 1962, and Scolnik 1955.

The survivors of the various tribes were placed on posts of the old Serviço de Proteção do Índio (SPI), precursor of the present Fundação Nacional do Índio (FUNAI). In 1988-90, according to Braga (1992), the Posto Indígena (P.I.) Guaporé was home to approximately seventy-five Makurap, forty-one Ayuru, twenty Tupari, and one Mekens. Accordingly to her, the number of Makurap who actually spoke the language was forty-five, and Ayuru and Tupari only had eight speakers each on this post. The tendency is for young speakers to learn Portuguese as a first language. The largest concentration of Tupari is on the P.I. Rio Branco; there are also a number of Makurap there. The Mekens, with at least two dialect groups, are concentrated on the P.I. Mekens.

On the P.I. Guaporé some members of the older generation still retain the traditional knowledge of their culture. Shamanism is still practiced, sometimes involving hallucinogenic snuff, called 'rapé' in Portuguese, which is consumed in group sessions.

THE DATA

As part of an attempt to secure at least some tape documentation of the many languages of southern Rondônia, Moore tape-recorded a standardized list of lexical items in various languages during a field visit to the P.I. Guaporé in 1988. The list, recorded in Dolby stereo using an external microphone, included the Swadesh 200-word list and supplementary lists of animals, plants and material culture items common to the region.

The tapes of the four languages under study here were transcribed by Moore and Galucio independently and then compared. Other sources of data include Moore's field transcriptions of Ayuru and the Master's thesis of Braga (1992) for Makurap and that of Alves (1991) for Tupari (which was based on several hours of tapes recorded by Moore). There is an unpublished description of Tupari by Aryon Rodrigues, as well as an early attempt by Hanke, Swadesh, and Rodrigues (1958) to sketch the phonology of Mekens and relate it to other Tupian languages. These two works were not used as sources of data for the present reconstruction, however.

The transcription and analysis of Makurap by Braga generally agree well with that which is presented here. The analysis of Tupari by Alves differs from ours in several respects, but principally in the labiovelar consonants, which she does not recognize as phonemic.

Because of the very limited nature of the data, tone, length, and other subtle phonetic distinctions cannot be established with certainty. More difficult still is the very preliminary nature of our knowledge of the morphology and morphophonology of the languages.

The informants who furnished the data are the following:

- **Ayuru**: Paulina Macurap, a woman about thirty-two years old in 1988. She was raised by the Ayuru.
- **Makurap**: Sebastião Macurap, a man about twenty-two years old in 1988.
- **Tupari**: Alzira Tupari, a woman about twenty-five years old at the time of recording.
- **Mekens**: Otaviano Mequém, a man about seventy-four years old, perhaps the informant of Hanke.

SKETCH OF THE SOUND SYSTEMS OF THE TUPARI LANGUAGES

The approximate segmental phonemic inventories of the four languages are summarized in the table below. Segments whose status is still uncertain are indicated by angled brackets. Significant allophones are indicated in parentheses.

TABLE 1: PHONEMIC INVENTORY OF CONSONANTS

AYURU (Ay)

p	t	c	k	kw
	<d>		g	gw
β				
	r	y		
		(ñ~ỹ)		
m	n	<nj>	ŋ	ŋw
(mb)	(nd)		(ŋg)	(ŋgw)

MAKURAP (Ma)

p	t	c	k
			g
β			
	r	y	
		(ñ~ỹ)	
m	n	<nj>	ŋ
(mb)	(nd)		(ŋg)

MEKENS (Me)

p	t		k	kw
<b>	<d>		g	<gw>
	s			h
	.(ts)			
β				
	r	y		
		(ñ~ỹ)		
m	n		ŋ	ŋw

TUPARI (Tu)

p	t	<c>	k	kw	'
(ps)	(ts)				
(Φ)					
<b>	<d>		g	<gw>	
	s				h
β					
	r	y			
		(ñ~ỹ)			
m	n	<nj>	ŋ		

TABLE 2: PHONEMIC INVENTORY OF VOWELS - in all four languages

ORAL:	i	ɨ	u(o)	NASAL:	ĩ	ɨ̃	ũ(õ)
	e				ẽ		
		a				ã	

There are five contrasting vowels in each of the four languages. These show remarkable stability over time. Each of the five vowels can be oral or nasal. The nasality may be autonomous or may be acquired through nasalization spread from some nasal segment. Nasalization spread, e.g. Ay: *kwãβã* 'partridge', appears to occur in all of the four languages, but the exact conditions for its spread in each of the languages cannot be specified at this time. Nasalization spread is a complex phenomenon in the Tupi-Guarani linguistic family (Harrison & Taylor 1971, Lunt 1973) and in the Mondé family (Moore 1984).

The syllable canon is generally (C)V(V)(C+), where C+ represents a morpheme-final consonant. The exceptions to this are that at least Ayuru and Makurap permit a syllable-final morpheme-medial palatal glide, *y*, and Tupari permits ʼ (glottal stop) and *h* in the same position. Syllables with two vowels occur, though many of these seem to span morpheme boundaries or to be the result of diachronic consonant deletion. Braga (1992) reports phonemic vowel length for Makurap.

There is no evidence of contrastive stress in the four languages. The question of tone is unresolved. In Ayuru there are two pitch levels in ascending sequences, but at least three or four levels in descending sequences. Both the Makurap and the Tupari Indians use whistled speech to communicate in the forest. However, it is difficult to find evidence of tone contrasts.

The consonantal inventory is similar in many aspects in the four languages. Each language has a series of voiceless stops and a corresponding series of nasals. In Ayuru and Makurap the nasals have post-oralized allophones before oral vowels (e.g. Ay and Ma: [mbo] 'hand') and full nasal ones before nasal vowels (Ay: [õmẽnt] 'my husband', Ma: [nõõnt] 'other'). We will refer to the post-oralized allophones as prenasalized stops.

The prenasalized voiced palatal stop *nj* is a problem. It cannot be an allophone of the palatal nasal *ñ* since this is itself a variant of *ỹ*. Since *nj* is rare and does not occur in the correspondences, its status will be left unsolved for now.

The oral voiced stop series is marginal except for the velars and the labiovelars. The voiced bilabial stop, *b*, appears to be always derived from an underlying morpheme-final voiceless bilabial stop, *p*, before vowels. (See the table of morpheme-final consonant alternations below.) Likewise many examples of *g* are from underlying *k* morpheme finally before vowels. There are, however, some examples of *g* and *gw* which cannot be explained in this manner, e.g. Ay: *o-gotkɨp* 'my neck', *gwago* 'sweet potato'. The oral voiced dental, *d*, is very rare and does not appear in the cognates.

In all the languages the palatal glide, *y*, can optionally be pronounced as a voiced palatal slit fricative. It acquires nasalization from adjacent nasal vowels, in which case it optionally can be a palatal nasal, *ñ*, which is the normal pronunciation morpheme initially. All the languages have the voiced bilabial fricative, β, which might also be analyzed as a glide. A flap *r* occurs in the four languages. It is optionally *l* in Makurap.

The four languages differ in the points of articulation which are distinguished in the voiceless, voiced, and nasal series. Makurap lacks all labiovelar consonants. Labiovelar consonants are recognized in one or more of the series for the other three languages. They are not analyzed as a sequence of stop plus glide because they occur syllable initially, where the syllable canon does not permit consonant clusters. Further, these consonants show very regular correspondences.

Tupari has both the glottal stop, ', and the glottal fricative, *h*; Mekens has only the latter. The voiceless palatal stop, *c*, is phonemic in Ayuru and Makurap. The Tupari voiceless bilabial stop, *p*, is an affricate, *ps*, before *i*, and a bilabial fricative, *Φ*, before *o* -- allophony similar to that of the neighboring language, Jeoromitxi. Also the Tupari dental stop, *t*, is optionally an affricate, *ts*, before *i*. In Mekens the dental fricative, *s*, is optionally an affricate, *ts*.

In morpheme-final position the contrast between the voiceless, the voiced, and the nasal series is neutralized. After oral vowels only the voiceless stops *p, t, k* and the palatal glide, *y* [y^{c}], occur in word-final position. These regularly alternate with their homorganic voiced counterparts (e.g. Ma: *kɨp* 'tree', *kɨβ+ot* 'fruit') when a vowel follows the morpheme boundary:

All four languages	Ayuru & Makurap	Mekens & Tupari
/__##	/V__+V	/V__+V
p	β	b
t	r	r
k	g	g
y	y	y

After nasal vowels only strongly prenasalized oral stops, [*mp, nt, ŋk*] and the nasal palatal glide [$\tilde{y}$] occur word finally. We will analyze these as nasalized allophones of *p, t, k* and *y*, respectively, and transcribe the stop phonemes without the prenasalization. For example, [ŋẽmp] 'breast' is transcribed as *ŋẽp*. At least in Ayuru and Makurap, word-final *p, t, k*, and *y* after nasal vowels alternate with β, *n, ŋ* and $\tilde{y}$, before vowels, for example, Ay: *mẽkẽt* 'I vomit', *mẽkẽn-ẽti* 'I feel like vomiting'.

The data available are insufficient to determine the morpheme-initial morphophonemic alternations, which are more complicated. Some alternations involving dental consonants are worth noting since these help explain one of the sound correspondences, *nd:c:t:h*. This correspondence will be reconstructed as a dental consonant **D* in complementary distribution with **r*. At this point we only wish to point out the existence of morphophonemic alternations involving *r, c, t*, and *h*:

Ayuru:

ek tere	'on top of the house'
gia-rere	'up in the sky'

Makurap:

teret	'name'
o-ceret	'my name'

Tupari:

het	'name'
e-ret	'your name'

There are also morpheme-initial alternations involving these sounds in cognate words in Tupi-Guarani and in Mawé. Without going into detail, we suggest that the corresponding alternations in the Tupari, Tupi-Guarani and Mawé families will eventually be shown to have a common ancestry. Examples from Gregores and Suáres (1967:223) and Graham, Graham and Harrison (1984:189):

Guarani:		*Mawé:*	
tera	'name'	-ha	'eye'
se-rera	'my name'	u-heha	'my eye'
NP rera	'NP's name'	NP eha	'NP's eye'
h-era	'his name'	t-eha	'his own eye'
		i-ha	'his eye'

TRANSCRIPTION

The transcription adopted is basically phonemic, but with certain specified sub-phonemic variation also written. This is the case for the nasal consonants. The prenasalized allophones (*mb*, *nd* and *ŋg*) of the nasal phonemes *m*, *n*, and *ŋ* are written as such to better illustrate the diachronic process of denasalization. Similarly, the oral and nasal palatal glides, *y* and *ỹ*, and the palatal nasal, *ñ*, are distinguished in the transcription.

A few other distinctions which appear to be subphonemic are also written in case they should eventually turn out to be significant: *o/u* in all the languages, *l/r* in Makurap, *s/ts* in Mekens, and *t/s* before *i* in Tupari. Syllable break is indicated by a period (e.g. Ma: *βa.i* 'stone'), and vowel length (to the small extent to which it can be determined) is indicated by two identical vowels.

SOUND CORRESPONDENCES:

The systematic sound correspondences among the consonants of the four languages are tabulated below, organized according to the mode of production. Hypothesized reconstructed segments are shown on the left, marked with an asterisk. Conditioning environments hypothesized for the proto-language are listed on the right, when relevant, along with the numbers of the cognate sets in which the correspondence is found. Conditioning environments

for individual languages, when relevant, are given after the sound which occurs in that language, for example in the velar correspondences for cognate set (70), *g g:--g:k (h__).

Consonants clusters spanning morpheme boundaries are often maintained in the daughter languages. When one consonant is lost, as happened in cognate sets (17), (29), (35), (51), (52), (85) and (106), it is always the initial consonant which is lost, except for the Ayuru form for 'knife', *ŋgɨte* (52). Metathesis may have occurred in (10), (11) and (80). Rather than list a separate correspondence for each of these deletions, they are simply mentioned now and the cognate set in which each occurs is included as an example of the correspondence which would obtain had not the deletion occurred. For example (29) is included as an example of *y:y:y:y* although the *y* has been deleted in Mekens.

TABLE 3: SYSTEMATIC SOUND CORRESPONDENCES

Proto Tupari	Ay	Ma	Me	Tu	Cognate sets
*p	p	p	p	p	7, 14, 15, 18, 19, 20, 27, 30, 31, 32, 33, 38, 40, 41, 46, 48, 49, 52, 55, 57,58, 62, 67, 69, 70, 73, 76, 79, 90, 93, 94, 103, 112, 114,120, 122, 123,
*p	β(V+__)	p	--	p	34
*p	Ø	--	b	p	(__+'V) 31, 122
*t	t	t	t	t	4, 8, 19, 40, 50, 51, 52, 53, 68, 69, 74, 88, 91, 92, 98, 102, 106, 112, 115, 124
*t	r +V	l +V	--	t ##	73
*t	t	Ø	--	r+V	118
*t	r	r	r	t	(__+'V) 3, 5, 58, 61
*k	k	k	k	k	1, 8, 15, 22, 25, 26, 27, 29, 32, 33, 35, 37, 44, 48, 49, 56, 60, 62, 69, 71 72, 74, 83, 89, 92, 96, 99, 102, 103, 111, 114, 115, 118, 119, 123
*k	k	Ø##	--	k	9
*k	g	--	g	k	(__+') 116
*kw	kw	Ø	kw	Ø	78, 87, 107, 109,
*b	Ø	β	b	b	(V__+V) 110
*g	g	--	k	k	32, 33, 117, 121
*g	g	--	g	k (h__)	70
*g	g	--	g	--	38
*gw	gw	β	kw	β	(__Voral) 2, 5, 6, 24, 80, 105
*gw	g	β	k	Ø	(__o) 69, 105
*gw	β	--	kw	--	(#__Vnasal) 77
*ts	t	t	ts, s	t, s(__i)	23, 38, 46,59, 81, 82, 84, 96,

*(n)dz	nd	nd	s	t, s(__ɨ)	64, 66
*β	β	β	β	β	8, 20 76
*β	β	Ø	Ø	Ø	(ɨ__)12, 119
*h	Ø	Ø	Ø	h	(V__C) 34, 49, 62, 70, 79, 103, 123
*'	Ø	Ø	Ø	'	3, 5, 10, 19, 28, 30, 31, 44, 58, 61, 91, 101, 116, 122
*r	r	r	r	r	(V__V, __+V) 10, 21, 22, 37, 62, 63 77, 80, 86, 108
*r	n	l	—	—	(Vnas__Vnas) 104
*D	(n)d	c	t	h	(#__Voral) 3, 41, 54, 56, 68, 81, 82, 98
*D	(n)d	Ø/c	h	h	54
*D	—	—	s	s/h	(__i) 52, 72, 95,
*y	y	y	y	y	(__#) 29, 107, 109
*y/ñ	—	ñ(__Vnas)	—	y(__Vor)	115
*ñ	ñ	ñ	ñ	ñ	(+__Vnas) 28, 35, 39, 61, 65, 85, 113,
*ỹ	ỹ	ỹ	ỹ	ỹ	(__+, Vnas__Vnas) 1, 44, 83, 97, 99
*m	m	m	m	m	(__Vnas) 47, 51, 63, 86, 100
*m	m	p	—	m	(__Vnas) 50
*mb	mb	mb	p	p	(#__Voral) 36, 43, 67, 116
*mb	mb	—	mb	—	(Vnas+__(?)) 77
*n	n	n	n	n	(__Vnas) 13, 17, 42, 45, 47, 75, 77, 89, 104
*n	n	Ø	—	Ø	50
*n	n	t	—	—	111
*nd	nd	t	t	—	(__Voral) 4, 71
*nd	nd	—	nd	—	(Vnas+__(?)) 87
*ŋ	ŋ	ŋ	ŋ	k	(#__Vnas) 14, 89
*ŋ	ŋ Vnas	g/k Voral	k Vnas	—	17, 71
*ŋg	ŋg	ŋg	k	k	(__Voral) 52, 57, 88, 103, 110, 120, 123
*ŋg	ŋg	k	k	—	117
*ŋw	ŋgw	β	kw	β	(#__Voral) 10, 101
*ŋw	ŋgw Vor	m Vnas	m Vnas	—	1
*ŋw	β	m	m	m	(Vnas__Vnas) 25, 72, 78

The consonants of the voiceless series, *p, t,* and *k,* show near-perfect stability in all positions and are reconstructed as such. There are bilabial correspondences in (31) and (122), dental correspondences in (3), (5), (58), (61), (73), and (118), and a velar correspondence in (116) where exceptionality is due to the morpheme-final consonant alternations before vowels in the four languages. The voiceless labiovelar, *kw,* is unchanged in two languages, Ayuru and

Mekens, but disappeared in the other two. The modern examples of *kw* in Tupari are presumably from some other source.

Of the correspondences reconstructed as voiced oral stops, two, *Ø*:β:*b*:*b* and *g*:–:*g*:*k*, are due to morpheme-final consonant alternations before vowels--probably the only source of *b*. The bilabial is reconstructed as **b* instead of β *since b*>β is a more natural change than the reverse. There is no clear explanation for the deletion of β in Ayuru, though this maybe due to a following glottal stop which was lost.

The velar correspondences indicate the existence of an oral **g* and **gw* in Proto-Tupari. The correspondence *g*:*g*:*k*:*k* must be different from *k*:*k*:*k*:*k*, and is reconstructed as **g*. This is in harmony with the very general pattern of devoicing in Mekens and Tupari. There are two exceptions to this. The correspondence *g*:–:*g*:*hk* in cognate set (70) may be from a medial sequence **hg* which blocked devoicing of *g* in Mekens before the *h* disappeared. The correspondence *g*:–:*g*:– cannot be explained at this time.

The correspondence *gw*:β:*kw*:β (which does not occur before *o*) occurs morpheme initially and medially. It is reconstructed as **gw* because (1) labiovelars are more likely to go to bilabial fricatives or glides than the reverse, and (2) **kw* was already seen to have different, though parallel reflexes. Before *o*, **gw* seems to have lost its labialization. The correspondence Ay: β:Me: *kw* is unclear since the forms for the other two languages are missing.

Dental affricates **(n)dz* and **ts* are reconstructed because no conditioning factor could be found for **(n)d* and **t* to become affricates. It is not clear from the data whether the voiced affricate is prenasalized or not, which is an important question. The variation *t/s* in Tupari is perhaps conditioned by the following vowel.

Of the fricatives, the correspondence β:β:β:β is reconstructed as *β. The correspondence β:*Ø*:*Ø*:*Ø* is reconstructed as *β. It only occurs after *i*, and the sequence *i*β does not appear in any of the Makurap, Mekens, or Tupari forms. Tupari seems to have retained a syllable-final preconsonantal **h*, as well as a prevocalic glottal stop, *'.

The correspondences reconstructed as **r* and as **D* are particularly interesting. The phoneme *r* occurs only morpheme medially and finally in the middle of words in the four languages, and shows highly regular correspondences. The correspondence *n*:*l*:–:– in (104) is perhaps from a nasalized **r*. The correspondence *(n)d*:*c*:*t*:*h* is reconstructed as a dental segment **D* whose exact phonetic shape is unknown and which is in complementary distribution with **r*, which never occurs word-initially, whereas **D* only occurs in that position. On the basis of this complementarity and also the morpheme-initial morphophonemic alternations given above, we suggest that **D* was a desonorantized variant of **r* at some point in the past, perhaps in Proto-Tupi, since the characteristic alternations occur in several different Tupian families. The reflex of **D* in Surui, a language of the Mondé Family, is *l* (for example *let* 'name'), which also argues for an original liquid source. The correspondence –:–:*s*:*s*/*h* is reconstructed as **D* before *i*. The correspondence *nd*:*Ø*/*c*:*h*:*h* in (54) has no explanation at the present.

The oral palatal glide is stable in final position and reconstructs as *y. However it is rare and unstable morpheme-medially. There are three cognate sets, (2), (11), and (100), in which *y* or *ỹ* occur in Ayuru or Makurap corresponding to Ø in Mekens and to a glottal stop or syllable break in Tupari. For these correspondences, not listed in Table 3, a cover symbol, **Y*, is given, but no plausible reconstructions can be offered. The palatal nasal is stable, apparently occurring only morpheme initially, and is reconstructed as *ñ. The nasal palatal glide occurs elsewhere and is reconstructed as such.

Looking at the nasalized segments, the simplest and most natural overall explanation for the correspondences observed is that original nasal sonorants were progressively denasalized before oral vowels , and then the denasalized stops were devoiced by the general devoicing change in Mekens and Tupari. This implies a lack of rightward nasalization spread from nasal consonants in Proto-Tupari--otherwise there could not have been oral vowels after nasals.

An alternative which must be rejected is that original voiceless stops before nasal vowels were retained as such in Mekens and Tupari and became nasal sonorants in Ayuru and Makurap. This could not have happened because there are a number of examples of the all voiceless stop correspondences (*k:k:k:k*, etc.) before nasal vowels (e.g. 13, 44, 78, 83, 99, 106, 115, and 118), and no conditioning factor to explain why these would not also have turned into nasal sonorants in Ayuru and Mekens. The denasalization hypothesis is supported by the existence of similar denasalization in the Gavião language of the Mondé family. Compare, for example, Suruí: *mẽt*, Gavião: *mẽt* 'husband' and Suruí: *mebe*, Gavião: *bebe* 'peccary'.

There are some irregularities in this picture. Makurap sometimes has voiceless stops instead of the expected prenasalized stops. There appears to be fluctuation in the language in this regard, for example, 'wasp' may be either *ŋgap* or *kap*. In Tupari the velar nasal seems to have been eliminated altogether.

Another irregularity is that some nasal vowel correspondences are oddly sporadic. See the discussion of nasal correspondences below.

The correspondence *ŋgw*:β:*kw*:β reconstructs neatly as **ŋgw*, paralleling **gw*. The last two correspondences, before nasal vowels, are less clear. One suspects the source to be **ŋw* since that would otherwise be missing from the pattern.

VOWEL CORRESPONDENCES

Since the vowel correspondences are so regular (*a:a:a:a*, etc.) we will only list the correspondences which are NOT regular. Two irregular correspondences are not included because they probably are due to transcription errors: in cognate set (26), the Mekens form *i-kaa* should probably be *ɨ-kaa* ('water-drink'), and in set (123) the Tupari form should probably be *ahkop*, as in set (103).

TABLE 4: IRREGULAR VOWEL CORRESPONDENCES

Proto-Tupari	Ayuru	Makurap	Mekens	Tupari	Cognate Sets
*ɨ	ɨ	ɨ	i	ɨ	(#__pe/βe) 27, 119
*u	ɨ	o	u	o	(__pi,bi) 30, 110
*i	i	ɨ	i	--	metathesis ? 80
*ɨ	ɨ	Ø	Ø	--	metathesis ? 80
*a/e	a	e	a	e	11
*e/a	e	a	a	--	119
*ĩ/ẽ	ĩ	ẽ	ẽ	--	42

The correspondence *ɨ:ɨ:i:ɨ* is reconstructed as **ɨ* on the hypothesis of neutralization of *ɨ/i* in the specified environment in Mekens. The correspondence *ɨ:u:u:u* is reconstructed as **u* on a similar hypothesis of **u*>Ay: *ɨ* in the specified environment. The next two correspondences, *i:ɨ :i:--* and *i:Ø:Ø:--* are perhaps explicable by postulating the metathesis of the *ɨ* in **araɨgwi* to after the following consonant in Makurap and its deletion in Mekens. For the last three correspondences, *a:e:a:e*, *e:a:a:--*, and *ĩ:ẽ:ẽ:--*, there is no basis for positing the proto-vowel and these are given as **a/e*, **e/a*, and **ĩ/ẽ*, respectively.

NASAL VOWEL CORRESPONDENCES

Nasal vowels regularly correspond to nasal vowels and are reconstructed as such, for example, 'husband' **mẽt*, Ay: *-mẽt*, Ma: *-mẽ-picop*, Me:*-mẽt*, Tu:*mẽẽt*. However, there are some irregularities. At least one of the irregularities is probably due to a transcription error in cognate set (25) 'dog' (cf. 72). A number of irregularities appear to be due to nasal spread after consonant addition or deletion (cognate sets (84) and (85)) or to differing conditions on nasalization spread (cognate set (90) and (97)). In these cases the oral form is regarded as the original form, later affected by nasalization spread.

Some of the other irregularities show a certain degree of systematicity. They are listed below in Table 5.

Table 5: IRREGULARITIES INVOLVING NASAL VOWELS

Proto-Tupari	Ayuru	Makurap	Mekens	Tupari	Cognate Sets
*Vnasal	Vnasal	Voral	Vnasal	------	17, 53, 71
*Vnasal	Vnasal	Voral	------	Voral	9
*Vnasal	Voral	Vnasal	Vnasal	Vnasal	1, 74
*Vnasal/oral	------	Vnasal	------	Voral	115

The first irregular correspondence, with three examples (17, 53, 71), seems to be due to denasalization in Makurap, since the forms in the other two languages are nasalized. In the second irregular correspondence, we will assume the second and third syllables were oral, but there was nasality on the first which shifted in Makurap. In cognate sets (1) and (74) the Ayuru

forms appear to have been denasalized. For the last correspondence,in (115), there is no clear basis for deciding the reconstruction for 'tucan'.

The instability of the irregular nasal correspondences listed above might be explained if nasalization in Proto-Tupari was like that reported for the Tupi-Guaranian language Kaiwá by Harrison and Taylor (1971). In Kaiwá, morphemes are either nasal or oral, but it cannot be predicted which syllable(s) will receive the nasality in the nasal morphemes: both *tupã* and *tũpa* are possible.

SUMMARY

The charts of reconstructed consonant and vowel segments is given below in Table 6 and 7. The palatal stops are excluded from the picture because of lack of evidence about their origins, but they should not be forgotten.

Table 6: PROTO-TUPARI CONSONANTAL SEGMENTS

p	t		k	kw	'
(b)			g	gw	
	ts				
	(n)dz				
β					h
	r	y			
	(D)	ỹ~ñ)			
m	n		ŋ	ŋw	
(mb)	(nd)		(ŋg)	(ŋgw)	

Table 7: PROTO-TUPARI VOCALIC SEGMENTS

Oral:	i	ɨ	u(o)
	e		
		a	
Nasal:	ĩ	ɨ̃	ũ(õ)
	ẽ		
		ã	

Table 8: UNDERLYING MORPHOPHONEMES OF PROTO-TUPARI

p	t		k	kw	'
			g	gw	
	ts				
	(n)dz				
β					h
	r	y			
m	n		ŋ	ŋw	

What would appear to have been the underlying system in Proto-Tupari is presented in Table 8. In this table **D* is considered a variant of **r*, with which it is in complementary distribution. The prenasalized consonants are subsumed under the nasals as allophones. The palatal glide includes its variants. The oral series includes only **g* and **gw*, the bilabial being only derived from **p* morpheme finally before vowels.

There are many details to be verified or altered in this picture. Assuming that this preliminary reconstruction is essentially correct, the major changes operating to produce the daughter languages have been (no ordering implied):

- Denasalization of nasal sonorants before oral vowels, a process perhaps already underway in Proto-Tupari.
- Devoicing of obstruents, mainly in Mekens and Tupari.
- Attrition of the original labiovelars by loss, delabialization, or loss of the velar.
- Loss of preconsonantal **h* and of glottal stop, except in Tupari.
- Despirantization of dental affricates in Ayuru and Makurap.
- Desonorantization of **r*, a process probably initiated long before Proto-Tupari.

Because of the considerable time depth of the Tupari family of languages, the preliminary reconstruction of Proto-Tupari presented here should eventually help cast some light on Proto-Tupi. More data and more phonological and morphological analysis, as well as data from other Tupian families, are needed to refine and broaden the tentative reconstruction presented here.

It is not possible at this time to do a thorough comparison of Proto-Tupari with the languages (or proto-languages) of the other nine Tupian families. Note however some obvious cognates:

Family:	**Tupari**	**Tupi-Guarani**	**Ramarama**	**Arikém**	**Munduruku**	**Juruna**	**Mawé**
Language:	**Proto-Tupari**	**Proto-T-Guarani**	**Karo**	**Karitiana**	**Munduruku**	**Xipaya**	**Mawé**
Armadillo	*ndayto	*tatu	yayo	sosi	day^{3}do^{2}	dusa	sa'ho
Peccary	*Daotse	------	yate	soytsa	da^{3}je^{2}	uza	—

Some of the reconstructed items give a small sample of Proto-Tupari material culture: 'ax', 'basin', 'basket', 'canoe', 'hammock', 'knife', 'salt', and 'seat'. ('Clothing' is an extension of 'skin'.) Domesticated plants include 'cotton', 'maize', 'pepper', 'sweet potato' and perhaps 'tobacco'. ('Banana' probably refers to a wild species which is similar in appearance.)

LIST OF COGNATES AND RECONSTRUCTED FORMS

In the following list, some forms are included, in parentheses, even though they are doubtful as cognates. They are included since some part of them may eventually prove to be cognate or to at least be useful for clarifying the segmentation of the cognates. Note for example, that in (35) 'flea', the Mekens form, *ĩp-tsap*, supports the segmentation of the Tupari form, *ñõ-tap*.

Extraneous segments may be included without being separated by hyphens if the segmentation is obvious, as in, for example, (29) 'earth'. Where it is useful to indicate

segmentation, as in (34) 'fish', this is done by hyphens, which do not necessarily indicate morpheme boundaries. Morpheme boundaries are indicated (by a +) only when they are relevant to reconstruction and there are strong reasons to posit them, especially, (i) when a known morpheme (such as the prefix *ki-* 'first person plural' in Mekens) is involved, (ii) when considerations of canonical form indicate a morpheme boundary (such as between most consonant clusters), or (iii) when the sound correspondence is what would be predicted by well-attested morpheme-final or initial morphophonemic alternations (such as in (58) 'macaw'). Note in (58) that morpheme boundaries are indicated in the Tupari form, *pet+'a*, and in the reconstructed form, **pet+'a*, but not in the forms for the other three languages, *pera*, since fusion may have rendered the morpheme boundary undetectable in these three languages. There are several morphemes of the form *'V* word-finally in Tupari which may be classifiers, e.g. pep+'o 'wing, feather'.

The reconstructions provided aim at accounting for the forms in the daughter languages as much as possible. Inevitably, there are cases such as (24) 'distant' in which there is some irregularity which cannot be reliably distinguished from transcription errors at this point. In these cases a degree of arbitrariness in the reconstruction is unavoidable.

	English	Proto-Tupari	Ayuru	Makurap	Mekens	Tupari
1.	Agouti	*ŋwãkɨ̃ỹã	ŋgwakɨ̃ỹã	mãkɨ̃ỹã	mãkɨ̃ỹã	——
2.	Alligator	*gwaYto	gwayco	βato	kwato	βa.o
3.	Ant, big	*Dat+'a	ndara	----	----	hat+'a
4.	Armadillo	*ndayto	ndato	tayto	tato	------
5.	Assai (palm)	*gwit+'i	gwiri	βirica	kwiri	βit+'i
6.	Ax	*gwi	———	βi	kwi	βii
7.	Banana	*ehpiip	epiip	------	----	ehpiip
8.	Basin	*βãẽkɨt	βãẽkɨt	———	———	βãĩkɨt
9.	Basket, big	*ãŋgerek	ãŋgerek	akẽnẽ	———	ɨp-akerek
10.	Bat	*ŋwari+'a	ŋgwaria	βa-ca-ria-y	kwari-sa	βari+'a
11.	Blood	*a/eYɨ	o+yaɨ	c+eyɨ	ki+aɨ	e.ɨ
12.	Blow	*ɨβa	y+ɨβa	β-ɨɨ-ka	s-eb-ɨɨ.a	ɨa
13.	Brazil nut tree	*kãnã, *arao	kãnã	araokiee	kãnã	arao.a'
14.	Breast	*ŋẽp	ŋẽp	ŋẽp	———	kẽp
15.	Canoe	*kɨp-pe 'tree-skin'	kɨpe	kɨpe	———	kɨpe
16.	Capibara	(loan)	caβi	caβi	———	———
17.	Cicada	*ŋõtŋõnã	ŋõŋõnã	koko.ĩ	kõtkõnã	----
18.	Clothing	*pe	pe	———	ki+pe	pee
19.	Coati	*pi'it	piit	------	piit	pi'it
20.	Cockroach	*a/eβape	aβape	———	eβape	(paba'pairu)
21.	Cotton	*ororo	ororo	ororo	ororo	ororo
22.	Crab	*kera	(koro)	———	kera	kera.a
23.	Deer	*ɨtsɨɨ	ɨtɨɨ	ɨtɨɨ	ɨsɨɨ	------
24.	Distant	*gwetsok	gweeto	βetok	kwesop	(tog-o)

25.	Dog	*ãŋwẽko	ãβẽko	amẽko	amẽko	amẽko
26.	Drink	*ka	kap	ka	i-kaa	ɨh-kaa
27.	Duck	*ɨpek	ɨpek	------	ipek	ɨpek
28.	Dust	*ñõ'õ	kɨyt-ñõõp	-------	-------	ñõ'õ
29.	Earth	*kɨy	kɨy	kɨy	kɨmãkãy	kɨy
30.	Egg	*upi+'a	ɨpia	c+upia	s+upia	upi+'a
31.	Feather	*pep+'o	peo	-------	pebo	pep+'o
32.	Fire	*agopkap	agokap (?)	----------	----------	kopkaap
33.	Firewood	*agopkap	agopkap	(ocatpot)	(otat)	kopkaap
34.	Fish	*pot	ɨɨ-βoy	pot-kap	-------	ɨh-pot
35.	Flea	*ñõk	ñõk	ñõk	(ĩp-tsap)	ñõ-tap
36.	Foot	*mbi	mbi	mbi	pi	pi
37.	Fowl	*õkɨra	ñ•õkɨra	m+õkɨra	õkɨra	-------
38.	Genipap	*tsigaap	tigaap	(mẽẽncaap)	sigaap	------
39.	Give	*ñũã	ñõã	ñũã	ñõã	ñũã
40.	Good	*poat	poatẽp	-------	(isãmẽp)	poat
41.	Hair	*Dap	ndap	-------	onẽ-tap	hap
42.	Hammock	*ẽ/ĩnĩ	ĩnĩ	ẽnĩ	ẽnĩ	------
43.	Hand	*mbo	mbo	mbo	ki+po-pi	po
44.	Hawk	*kẽỹ+'ã	------	------	kẽỹã	kẽỹ+'ã'
45.	Heart	*ãnõã	m+ãnõã	-------	ki+ãnõã	ãnõã
46.	Heavy	*potsi	poti	poti	i-potsi	posi
47.	Honey Marten	*ãmãnã	ãmãnã	ãmãnã	------	------
48.	Horn	*apikɨp	------	apikɨp	-------	apikɨp
49.	Hot	*ahkop	y+akop	-------	s+akop	ahkop
50.	Humming bird	*mĩnĩt	mĩnĩt	pĩĩt	------	mĩĩt
51.	Husband	*mẽt	õ+mẽt	mẽ-picop	õ+mẽt	mẽẽt
52.	Knife	*ŋgɨtpe	ŋgɨte	------	kɨpe sĩt	putpe sĩĩt
53.	Know	*toã	------	kiua toa	opoe toã	-------
54.	Leaf	*Dep	kãnã-nde	ep/ cep	hep	hep
55.	Liver	*pia	pia	piat	o+pia	pia
56.	Lizard	*Dako	-------	cako	tako	haaku
57.	Louse	*ãŋgɨp	ãŋgɨp	ãŋgɨp	kɨp	kɨp
58.	Macaw	*pet+'a	pera	pera	pera	pet+'a
59.	Maize	*atsitsi	atiti	atiti	asisi	------
60.	Mandi (fish)	*mõkoa	------	mõkoa	õkoa	------
61.	Meat	*ñẽt+'ã	ñẽrã	ñẽrã	õ-ñẽrã	ñẽt+'ã
62.	Monkey, Capuchin	*sahkɨrap	------	------	sakɨrap	ahkɨrap
63.	Monkey, Spider	*ãrĩmẽ	----	(alẽbo)	ãrĩmẽ	ãrĩmẽ

64.	Mortar	*ẽndzɨ	ẽndɨ	———	ẽsɨ	sɨ-ka'
65.	Mother	*ñã	ñã	ñã	ñãã	ñã
66.	Mountain	*(n)dzo	ndoo	ndɔa	soo	tuh-tet
67.	Nail	*mbo-ape	———	mbo-ape	o+po-ape	———
68.	Name	*Det	ndet	o+cer-et	———	het
69.	Neck	*gwotkɨp (kɨp='tree' ?)	o+gotkɨp	βotkɨp	o+kotkɨp	otkɨp
70.	New	*pahgop	pagop	———	i-pagop	pahkop
71.	Night	*ŋĩndak	ŋĩndak	gitak	———	———
72.	Ocelot	*ãŋwẽko Dĩĩt	———	———	ãmẽko sĩĩt	ãmẽko hĩĩt
73.	Old	*poot	poor+ia	puul+e	———	poot
74.	One	*kiẽt	kiet	———	———	kiẽt
75.	Other	*nõõ	nõ	nõõt	———	oasĩ-nõõt
76.	Owl	*popoβa	(ɨβao)	popoβa	popoβa	———
77.	Paca	*gwãnãmbiro (*gwãnã+mbiro ?)	βãnãmbiro	———	kwãnãmbiro	———
78.	Partridge	*kwãŋwã	kwãβã	———	kwãmã	———
79.	Path	*pee	pee	pee	pe	ahpe
80.	Peanut	*araɨgwi	araɨgwi	araβɨɨk	arakwi	———
81.	Peccary	*Daotse	———	caote	tause	———
82.	Peccary, collar	*Daotsey	———	caotey	tause	aote'iri
83.	Pepper	*kõỹ	———	kõỹõ	pe-kõõỹ	———
84.	Person	*aotse	aotẽ-nãp	———	aose	———
85.	Piranha	*ipñãỹ (*ip+ñãỹ ?)	ipñãỹ	ipñãỹ	———	ĩñãỹ
86.	Push	*mõrã	õ+mõrã	———	———	i+mõrã
87.	Rotten	*ãnde, *ãkwĩ (*ã+nde, *ã+kwĩ)	(ñ+ãnde)	ãĩt	(s+ããnde)	ã.ĩ
88.	Salt	*ŋgɨɨt	ŋgɨɨt	———	kɨɨt	kɨɨt
89.	Scorpion	*kɨtnĩŋã	kɨnĩŋã	———	kɨtnĩŋã	———
90.	Seat	*ãβõ-pe *ñãp-pe	ãβõpe	ñãpe	ããβo	ñããp
91.	See	*to'a	toa	toa	———	to'a
92.	Seed	*kit	(aβi)	(tambiit)	ikit	kit
93.	Shell	*ape'	y+ape	ape	———	s+ape'
94.	Skin	*pe	pe	pe-et	ki+pe	pee
95.	Small	*Dĩĩt	———	———	sĩĩt	sĩĩt
96.	Smooth	*atsik	y-atik	———	s-asik	asik
97.	Snail	*ɨỹã	ɨyã	ãmẽỹã	———	———
98.	Snake	*Dat	ndat	cat	———	hat
99.	Sour	*kãỹ	kãỹ	———	pe-kãỹ	———

100.	Speak	*mãYã	mãỹã	---------	----------	mã'ã
101.	Stone	*ŋwa+'i	ŋgwai	βa.i	kwai	βa+'i
102.	Straight	*kɨɨt	kɨɨt	kɨɨt	--------	--------
103.	Sun	*ŋgiahkop	ŋgiakop	---------	--------	kiahkop
104.	Surubim (fish	*ãnõrẽ	ãnõnẽ	ãnõlẽ	------	------
105.	Sweet Potato	*gwagwo	gwago	βaβo	kwako	βa.o'
106.	Swim	*tĩptĩpnã	tĩptĩpnãã	tĩtĩnã	--------	--------
107.	Tail	*okway	okway	c+oay	s+okway	oay
108.	Take	*ara	--------	--------	y+ara	ts+ara
109.	Tapir	*ɨkwaay	ɨkwaay	ɨay	ɨkwaay	--------
110.	Termite	*ŋgub+i	ŋgɨi	ŋguβa	kubi	kubi
111.	Timbo	*nĩk	nĩk	tĩk	(kĩkĩt)	------
112.	Tobacco	*pitoa	pitoa	(bɨtea)	pitoa	(kɨpea)
113.	Tooth	*ñããỹ	ñããỹ	ñããỹ	ki+ñãỹ	i+ñãỹ
114.	Tree	*kɨp	kɨp	kɨp	kɨp	kɨp
115.	Tucan	*yo/ñõkãt	------	ñõkãt	------	yokãt
116.	Turtle	*mbok+'a	mboga	------	poga	pok+'a
117.	Urucum	*ŋgop	ŋgop-gaap	iko	kob+a kaap	----
118.	Vomit	*ẽkẽt	m+ẽkẽt	n+ẽkẽ	--------	ẽkẽr+ã
119.	Vulture	*ɨβe/ako	ɨβeko	ɨako	iako	------
120.	Wasp	*ŋgap	ŋgap	ŋgap	kap	kap
121.	Water	*ɨgɨ	ɨgɨ	ɨ	ɨkɨ	ɨ-u-k'a (?)
122.	Wing	*pep+'o	peo	--------	ipebo	pep+'o
123.	Year	*ŋgiahkop	ŋgiakop	--------	----------	ɨhkop
124.	You	*ẽt	ẽt	ẽt	---------	ẽt

1. We wish to thank Paulina Macurap, Sebastião Macurap, Otaviano Mequém, Alzira Tupari, and other helpful people of the P.I. Guaporé for the information they provided about their languages. The SSILA kindly added an extra session for papers on Brazilian Indian languages to their summer meetings. Financial support for the participation of Ana Vilacy Galucio in the SSILA meetings and in the 1993 LSA Linguistic Institute was provided by a tuition grant from the Linguistic Society of America, by a grant for short-term study abroad from USAID, and by travel support from the organization People-to-People. The 1988 field visit was supported by National Science Foundation grant BNS-8619117. The comparative Tupi project is currently supported by Wenner-Gren research grant no. 5578 and by the Brazilian Financiadora de Estudos e Projetos (FINEP) grant no. 63.92.0468.00. The field research was conducted with authorization from the Brazilian National Council for the Development of Science and Technology (CNPq) and from the Fundação Nacional do Índio (FUNAI). Both the regional FUNAI staff in Guajará-Mirim and the personnel of the P.I. Guaporé were very helpful. Our work has only been possible through the excellent cooperation offered by CNPq and the Museu Paraense Emílio Goeldi. Spike Gildea and Francisco Queixalós offered useful comments on an early draft of this article; remaining errors are our own.

REFERENCES

ALVES, Poliana Maria. 1991. *Análise fonológica preliminar da língua Tuparí.* M.A. Dissertation, University of Brasília.

BECKER-DONNER, E.. 1962. *Guaporé-Gebeit.* Akten des 34. Internationalen Americanisten Kongresses (Wien 1960), 124-128.

BRAGA, A. 1992. *A fonologia segmental e aspectos morfofonológicos da língua Makurap (Tupi).* M.A. Dissertation, State University of Campinas.

CASPAR, F.. 1956. *Tuparí.* Trans. E. Northcott. London: G.Bell and Sons, Ltd.

GRAHAM, A., Sue Graham and Carl H. Harrison. 1984. Prefixos pessoais e numerais da língua Sateré-Mawé. *Estudos sobre línguas Tupi do Brasil*, ed. by Robert A. Dooley, *Série Linguística*, nº 11, pp. 175-205.

GREGORES, E. and J. Suáres. 1967. *A description of colloquial Guarani.* The Hague: Mouton & Cia.

HANKE, W., M. Swadesh, and A. Rodrigues. 1958. Notas da fonologia Mekens. *Miscellanea Paul Rivet Octogenario. Dicata*, vol.2:187-217. Mexico.

HARRISON, Carl H. and John M. Taylor. 1971. Nasalization in Kaiwá. *Tupi Studies I*, ed. by D. Bendor-Samuel, pp. 15-20, Norman, Oklahoma: SIL.

JENSEN, C. 1984. *O desenvolvimento histórico da língua Wayampi.* M.A. Dissertation, State University of Campinas.

LEITE, Yonne and Marília Facó Soares. 1991. Vowel shift in the Tupi-Guarani language family: A typological approach. *Language change in south american indian languages*, ed. by Mary Ritchie Key, pp. 36-53. Philadelphia: University Press.

LEMLE, M. 1971. Internal classification of the Tupi-Guarani linguistic family. *Tupi Studies I*, ed. by D. Bendor-Samuel, pp. 107-129. Norman, Oklahoma: SIL.

LÉVI-STRAUSS, C. 1948. Tribes of the right bank of the Guaporé River. *Handbook of South American Indians*, ed. by J. Steward, vol.3:371-379. Washington, DC: Smithsonian Institution.

LOUKOTKA, C. 1950. La parenté des langues du bassin de la Madeira. *Lingua Posnaniensis* (Poznan) 2:123-144

LUNT, Horace G. 1973. Remarks on nasality: The case of Guarani. *A festschrift for Morris Halle*, ed. by Stephen R. Anderson and Paul Kiparsky, pp. 131-139. New York: Holt, Rinehart and Winston, Inc.

MEIRELES, Denise Maldi. 1989. *Guardiães da fronteira: Rio Guaporé, século XVIII.* Petrópolis, RJ: Ed.Vozes.

MÉTRAUX, Alfred. 1948. The tribes of the upper Guaporé River. *Handbook of South American Indians*, Vol 3: *The tropical forest tribes*, ed. by Julian Steward, pp. 143-253. Washington, DC: Smithsonian Institution.

MOORE, Denny. 1984. *Syntax of the language of the Gavião Indians of Rondônia, Brazil.* Doctoral dissertation, City University of New York.

RODRIGUES, A. 1984/85. Relações internas na família linguística Tupi-Guarani. *Revista de Antropologia.* 27/28:33-53. São Paulo.

SCOLNIK, R. 1955. Observaciones sobre el estado sanitario de algunas tribus indígenas del Brasil Central. *America Indígena XV* (2):89-95.

Basic word order in Karitiana (Arikem family, Tupi Stock)

Luciana R. Storto
Pennsylvania State University

Introduction

The aim of this paper is to propose a preliminary analysis of the phenomenon of constituent order variation in Karitiana, a language spoken today by 154 people who live in "Area Indigena Karitiana", in the state of Rondonia, Brazil. The language has previously been studied by David and Rachel Landin, missionaries of the Summer Institute of Linguistics who lived among the Karitiana for five years. In an article based on his masters thesis D. Landin poses SVO as the basic word order of Karitiana (1984:221). However, no structural argument was ever given to support this hypothesis. D. Landin bases his choice on Greenberg's suggestion that the basic word order in a language should be elicited from the most frequent order used in declarative transitive sentences with overt nonpronominal arguments. Considering that SVO is only one of the 6 word orders found in Karitiana and since it is not the most common one, we have no reason to accept D. Landin's description[1]. Also, SVO is the basic word order in Portuguese, the Brazilian national language, spoken fluently by most of the Karitiana for at least half a century. It is not clear that the use of this word order in translations of transitive sentences out of context is free from influences from Portuguese.

Another problem with posing SVO as the basic constituent order in Karitiana has to do with the parametric characteristics of the language, which are consistent with OV and not VO order: noun-postposition, genitive-noun, noun-adjective. Even though the latter is slightly inconsistent with OV order, it follows a pattern found in genetically related Tupi-Guarani languages, which are OV (Moore 1991:1).

Furthermore, most other Tupian languages of which reliable studies are available are consistently OV. The word order in some of the best studied languages of the Tupi-Guarani family (by far the larger of the Tupian families) are: Kaapor: SOV; Kamayura: SOV; Tupinamba: SOV; Asurini of the Trocara: OVS; Guajajara: SOV in embedded clauses and VSO in matrix clauses. Other Tupian families show the same pattern: Munduruku (Munduruku family): SOV; Gaviao (Monde family): OV; Xipaya (Juruna family): possibly nonconfigurational, but presents OV constructions; Karo (Ramarama family): SOV; Ayuru (Tupari family): SOV. According to Moore, these facts "tend to confirm some earlier speculation that in their past stages the Tupian languages had the basic word order

[1]It is worth mentioning that my definition of word order is the underlying order of constituents in a language. This idea is certainly closer to Chomsky's notion of D-Structure than to Greenberg's criteria of frequency of occurrence.

Object-Verb (OV) rather than Verb-Object (VO)". He points out that Landin's analysis of Karitiana as an SVO language poses the most serious claim to VO order for a Tupian language (Moore 1991:2-3).

Methodology

In an attempt to solve the problem, I elicited and analyzed from a corpus of texts all sentences which presented at least a two place predicate with overt nonpronominal arguments. Pronominal arguments were not considered because I am not completely certain about the rules that regulate their distribution. However, whenever needed in the explanation, I used data which exemplifies the occurrence of free pronouns and personal verb prefixes. The texts utilized consist of mythological tales, historical narratives and dialogues.

Results

From a total of 62 sentences analyzed, the proportion of word order variation found was:

VOS: 27
OVS: 13
SVO: 9
VSO: 9
SOV: 2
OSV: 2

VOS and OVS: The majority of the texts analyzed are mythological tales or reported stories. This fact seems to have influenced the results above, since transitive subject final clauses are much more frequent in narratives than in conversations. A more careful analysis of the data might prove that the sentence final position has a semantic function of agentivity in story telling. R. Landin (1982:3-8) suggested that this position is reserved for the discourse theme. Her evidence, however, is not conclusive since what she describes as the discourse theme almost always coincides with the subject of the sentence[2].

1. Na-pisorok-Ø mijo Botỹj
   erg-gather-nf nut Botỹj
'Botỹj gathered the nuts'

2. Ga Y-ti-m-'a-t
   field 1p-top-caus-make-nf
'I made (prepared) the field'

3. Sal na-pitan-ta'ãt opok
   salt erg-share-evid white man
'The white man shared the salt'

---

[2]The abbreviations used on this paper are: nf:non-future tense; f: future tense; top: topicalizer; erg: ergative case; abs: absolutive case; 1p: 1st person singular prefix; 1p.pron: 1st person singular free pronoun; caus: causative; evid: evidential marker;co.3p.poss.: 3rd person singular anaphoric possessive pronoun; asp: aspect; asp.sup: aspect (supine); neg: negation; progr: progressive.

VOS and VSO are very common word orders in texts, in contexts where the action is being emphasized. Sentences 4-7 are extracted from a myth where the activities performed by the main character are being described:

4. Na-petet-Ø Botỹj mijo ket
   erg-cook-nf Botỹj nut unripe
'Botỹj cooked the unripe nuts'

5. Dok Byjyty
   seat (ideophone) Byjyty
'Byjyty sat'

6. Na-ambo-t kendo ohyn Byjyty
   erg-seat-nf coconut above Byjyty
'Byjyty sat on the coconut'

7. Na-mynira-t ta-iso Byjyty
   erg-lit-nf co.3p.poss.-fire Byjyty
'Byjyty lit his fire'

OVS and OSV word orders sometimes present the verb prefix **ti-** which Rachel Landin describes (1982:15) as a marker of topicalization of the object. Whenever this prefix occurs on a verb, the object is the first constituent in the sentence. However, not every instance of topicalization of the object is marked by this morpheme, as first noted by Rachel and David Landin. R. Landin correctly noted that this topicalizer occupies the same position in the verb that the ergative/absolutive markers do. She also noted that **ti-** never occurs in the narrative portion of a text, being restricted to "monologues, conversation, and speech quotes in narratives" (1982:11). D. Landin did not make use of this information in his thesis, posing an optional rule (1984:233) that deletes the case prefixes and inserts the topicalizer **ti-** when there is uncertainty as to the syntactic functions of the arguments in a sentence. However, from the examples below it is clear that **ti-** is not a disambiguator of syntactic function, since it is present even when the subject is dropped:

8. Moramon a-ti-m-'a-tykat, y-ta'it
   what 2p-top-caus-do-asp 1p-uncle
'What are you doing, uncle?'

9. Tykỹ ti-m-'a-tykat, y-saka'et
   palm heart top-caus-do-asp 1p-nephew
'I am taking (gathering) palm hearts, my nephew'

10. Pom ememo ti-m-'a-t
    nambu black top-caus-make-nf
'(He=pro) created the black nambu'

11. Ese i-ti-m'-a-t Ora
    water ?-top-caus-make-nf Ora
'Ora created the water'

12. Y'it kyry y-ti-'y-tysypak?
    1p-son liver 1p-top-eat-asp.sup.
'Am I eating my son's liver?

13. Atykiri naka-sot-Ø [esety Ora ti-m'-a]
    then erg-exist-nf [water big Ora top-caus-do
'Then there was the river which Ora created'

Also, the characterization of this morpheme as a topicalizer seems to be inadequate because it marks the fronting of elements whose referent is "unknown" such as WH words generated in object position. It is clear that in WH sentences movement is not motivated by pragmatic factors similar to the ones that characterize topicalizations.

Furthermore, **ti-** is restricted to a certain sentential type used to express direct speech or embedded sentences and it indeed occupies the same structural position filled by the morphemes which indicate ergative/absolutive case in the narrative sentential type. These morphemes (**na(ka)/ta(ka)**) were shown by D. Landin (1984:223-227) to be in ergative/absolutive distribution, where **ta(ka)** marks transitive verbs preceded by objects or intransitive verbs preceded by subjects while **na(ka)** marks everything else[3]. In a parallel fashion **ti-** marks movement of accusative, as opposed to nominative arguments (which are unmarked) to sentence initial position in direct speech and embedded sentences. That is, the presence of the morpheme **ti-** reflects a nominative/accusative system in the direct speech sentential type while the morphemes **na(ka)/ta(ka)** reflect an ergative/absolutive system in narrative sentences. In face of this evidence I suggest that Karitiana has a split ergative system.

14. Sosy i-ti-oky-t põrãsi
armadillo 3p-top-kill-nf trap
'The trap killed the armadillo'

15. Põrãsi i-oky-t sosy
trap 3p-kill-nf armadillo
'The trap killed the armadillo'

Data recently elicited in the field shows that the ergative/absolutive markers also occur in contrast in identical environments:

16. Taso na-oky-t ombaky
man erg-kill-nf jaguar
'The man killed the jaguar'

17. Taso ta-oky-t ombaky
man abs-kill-nf jaguar
'The man killed the jaguar'

According to two fairly sophisticated informants tested independently, sentences 16 and 17 convey the same meaning, but the latter is used as "a warning, when you know something will happen as a consequence of the action". In my opinion the absolutive marking is demoting the subject in its characteristics of agentivity (control and intentionality) in order to emphasize the action. The process above could perhaps be described as semantic ergativity.

The obvious conclusion concerning the consituent orders OSV and OVS is that since they get marked for object movement, they must be deviations from the basic word order.

---

[3]The allomorphs **naka** and **taka** prefix stress-initial verbs while **na** and **ta** prefix verbs with all other stress patterns.

SVO: This is the typical word order used in translations of transitive declarative sentences from Portuguese:

18. Taso na-oky-j ombaky
man erg-kill-f jaguar
'The man will kill the jaguar'

19. Y-hay naka-kip-Ø sosy
1p-old brother erg-open-nf armadillo
'My older brother opened the armadillo'

This order also occurs in both the narrative and the direct speech sentential types. Semantically, it appears to be unmarked.

The most striking fact we observed about the word order variation, is that the **SOV and OSV** word orders were found to occur only in subordinate clauses. It is a widely accepted fact that the word order in embedded clauses has a tendency to be more conservative than the one in main clauses, since the former are less subject to influences of pragmatic nature than the latter. This generalization could be used as an argument for the hypothesis that a sequence with the form NP NP V is the basic word order in Karitiana. I am inclined to think that SOV is the basic word order in the language due to the fact that it is morphologically less marked or simpler than OSV, where **ti-** sometimes indicates movement of the object (as in example 13). However, the latter seems to be by far the most used word order in subordinate clauses, SOV being more common in mythological tales. Other examples of embedded clauses are given below:

20. [ombaky taso oky tykiri] y-taka-hyryp-Ø yn
[jaguar man kill when] 1p-abs-cry-nf 1p.pron.
'When the man killed the jaguar I cried'

21. [Ahoy byhip tyki-oot] na-pa'ira-t jonso
[arroz cook when-progr] erg-angry-nf woman
'The woman was angry when she was cooking rice'

22. Atykiri Botyj naka-m-'a-'ot hyryp [Ora ta-'it
then Botyj case-caus-do-first cry [Ora co.3.poss.-son
byhot tykiri]
transform when]
'Then Botyj cried first when Ora transformed his son'

23. [Ambi Joana ama tykiri] naka-tat-Ø Maria
[house Joana buy when] case-go-nf Maria
'When Joana bought the house, Maria left'

24. [Dinheiro y-ahit-iki tykiri] y-taka-tat-Ø yn
[money 1p-get-neg when] 1ps-case-go-nf 1ps
'[When I did not receive the money] I left'

In an analogy with the analysis proposed for German and all other Germanic languages other than English, I suggest that Karitiana might be a verb second (V2) language. It is a widely accepted fact that the word order in German is that of the embedded sentences (SOV) and that the tensed verb (verb/Infl) has to raise to the COMP position in main clauses (the second structural position in the sentence)in case there is no lexical item occupying that position for the sentence to be well-formed. The COMP node in V2 languages such as German is interpreted as the head of the sentence - an inherently tensed position which needs to be lexically realized in order to assign nominative case to the subject (Platzack 1986).

A consequence of the analysis proposed for constituent order in Karitiana indicates that verbs will always be in final position in embedded sentences and never so in main clauses. This is indeed the distribution I have found in my preliminary analysis of the data.

The V2 hypothesis for Karitiana will not be fully developed in this paper because I do not completely understand the distribution of certain crucial morphemes such as pronominals and case, topicalization and tense markers. However, an interesting fact which seems to support this hypothesis is the absence of several inflectional morphemes in subordinate sentences. These clauses present either no tense marker at all or the unmarked verb suffix which indicates present or past tense. The presence of a future tense suffix in subordinate clauses is considered ungrammatical, which could indicate that COMP is the structural position where future tense is assigned to the verb since narrative main clauses always require a tense marker. Unlike main clauses, embedded sentences do not present aspect markers suffixed to the verb. Subordinate sentences which semantically bear aspect have it expressed in VP internal adverbials as in example 21.

If the V2 phenomenon is a reality in Karitiana, then the structural description of sentences will be such that the SPEC of CP position can be occupied by WH words or any argument noun phrase, while the COMP position is always occupied by either the verb or the tensed auxiliary in main clauses. The morpheme **ti-** seems to mark exactly this fronting of arguments generated in object position to SPEC of CP. A pre-sentential position has to be posed, where conjunctions which function at a level above the clause and ideophone (onomatopoetic) phrases occur. Also, there has to be a clause-final position where the subject can move in certain discourse environments.

Although I am not in a position to give compelling evidence for the V2 phenomenon in Karitiana, the data seems to point to SOV as the basic constituent order in the language. Finally, I hope to have persuaded the reader that at least further analysis is needed before we accept David Landin's claim that SVO is the basic constituent order in Karitiana.

AKNOWLEDGMENTS

I thank Wenner-Gren Foundation for Anthropological Research and The Pennsylvania State University for their financial support of my graduate studies and the Museu Emilio Goeldi/FINEP (Brazil) for travel grants which enabled my field research among the Karitiana. I also thank Samuel Cruz from FUNAI (The Brazilian Indian Affairs Bureau) for logistic support. I am grateful to Daniel Everett, Denny Moore, Jeffrey Pullum, David Wible, and my fellow junior researchers at the Goeldi Museum for helpful commentaries and discussion of earlier versions of this paper. Finally, I thank the Karitiana people for their cooperation with this research, specially Nelson and Meireles Karitiana who have been particularly interested and dedicated.

REFERENCES

LANDIN, D., 1984. An Outline of the Syntactic Structure of Karitiana Sentences. In Estudos Sobre Linguas Tupi do Brasil. Serie Linguistica 11. SIL.

LANDIN, R., 1982. Word Order Variation in Karitiana. Arquivo Linguistico no. 149. Brasilia, D.F. SIL.

MOORE, D., 1991. A Few Aspects of Comparative Tupi Syntax. 47th International Congress of Americanists. Manuscript.

PLATZACK, C., 1986. Comp. Infl and Germanic Word Order. In L. Hellan and K.K. Christensen (eds.), Topics in Scandinavian Syntax 185-234.