UNIVERSIDADE FEDERAL DO PARÁ

Reitor
Alex Bolonha Fiúza de Mello
Vice-Reitora
Marlene Medeiros Freitas
Pró-Reitor de Administração
Murilo de Souza Morhy
Pró-Reitora de Ensino de Graduação e Administração Acadêmica
Sônia de Jesus Nunes Bertolo
Pró-Reitora de Extensão
Terezinha Valim Oliver Gonçalves
Pró-Reitor de Pesquisa e Pós-Graduação
João Farias Guerreiro
Pró-Reitor de Planejamento e Desenvolvimento
Francisco de Assis Matos de Abreu

EDITORA UNIVERSITÁRIA UFPA

DIRETORA DA EDUFPA
Laïs Zumero
Divisão de Editoração
José dos Anjos Oliveira
Divisão de Distribuição e Intercâmbio
Sérgio Lima

GRÁFICA UNIVERSITÁRIA UFPA

Diretor
Manoel Gomes de Lima

CNPq

LÍNGUAS INDÍGENAS BRASILEIRAS
FONOLOGIA, GRAMÁTICA E HISTÓRIA

ATAS DO
I ENCONTRO INTERNACIONAL DO
GRUPO DE TRABALHO SOBRE LÍNGUAS INDÍGENAS
DA ANPOLL

Tomo I

Ana Suelly Arruda Câmara Cabral
Aryon Dall'Igna Rodrigues
(organizadores)

EDITORA UNIVERSITÁRIA UFPA
Belém - Pará
2002

voces en tipo diferente. Viena: Faesy & Frick / Paris: Maisonneuve, 1876 [1639/1640].
Seki, L. *Gramática Kamaiurá* - Língua Tupi-Guarani do Alto Xingu. Campinas, Editora da UNICAMP, 2000.
Weiss, H. Para um dicionário da língua Kayabí. Tese de doutorado, Universidade de São Paulo, 1998.

Pre-Proto-Tupí-Guaraní Main Clause Person-Marking*

Spike Gildea (University of Oregon)

1 INTRODUCTION

This paper surveys prior work in historical syntax in the Tupí-Guaraní (TG) family, then proposes a new reconstruction of Pre-Proto-TG main clause grammar. Jensen (1990, 1998) surveys the person-marking patterns of the various synchronic TG languages and reconstructs the virtually universal TG hierarchical (inverse, according to D. Payne, 1994) verbal person-marking pattern to Proto-TG. Jensen (1998) suggests a five-stage scenario by which that person-marking pattern was generated historically. In her scenario, Pro-Proto-TG begins with an absolutive person-marking system, identical to the absolutive person-marking pattern found in most modern TG subordinate clauses, nominalizations and topicalized oblique constructions. As a specialist in the reconstruction of grammatical patterns, I disagree with Jensen's theoretical assumptions about the nature of historical change, and about appropriate methods for diagnosing relative age of grammatical systems. Drawing on principles well-known in the literature on grammaticaliza
tion, I conclude that at no stage can we reconstruct a Pre-Proto-TG absolutive person-marking system. Rather, the modern hierarchical (inverse) system evolved from a pre-Proto-TG system that must have been originally more nominative. All the absolutive person-marking patterns in modern comparative TG grammar must be reconstructed to the grammar of nominalizations or other nonfinite inflections.

Schleicher (1998) offers a series of innovative proposals regarding Proto-TG and Pre-Proto-TG which have yet to be addressed in the published TG literature. He disagrees with Jensen both in the phonological forms that he reconstructs and in his reconstruction of the grammatical person-marking systems. I do not address his reconstruction in this paper, except to note that he disagrees with Jensen on bases other than those adduced here.

---

* Special characters
Both special characters are vowels, and are done in IPATimes.
ɪ = barred i ˊ = schwa

I am not a specialist in TG linguistics, hence I do not take a stand on the differences between Schleicher and Jensen regarding their reconstructions of phonological form. Regardless of the actual phononological content of the Proto-TG morphological forms, the reconstruction of the grammatical system of person-marking is a logically distinct problem, which can be treated independently. In the interests of consistency, I have chosen to use Jensen's (1998) proto-forms. As specialists in individual TG languages (in personal communication) have endorsed Jensen's treatment of their data, and as Both Jensen and Schleicher agree that the same system reconstructs to Proto-TG, I take their reconstruction of the Proto-TG system as my point of departure.

This paper is organized as follows: Section 2 presents the methods by which the relative ages of modern morphosyntactic patterns can be determined, Section 3 presents the Proto-TG person-marking system, Section 4 discusses Jensen's 1998 internal reconstruction of the evolution of the Proto-TG system, Section 5 presents my alternative hypothesis, and section 6 discusses some implications of the new reconstruction.

2 METHODS

The comparative method is well documented as a means of reconstructing phonological systems and lexical items to proto-languages, but methods for reconstructing grammar are more recent, and as such less well-known. Much of the work on such methods has followed from recent advances in Grammaticalization Theory (cf. Heine et al 1991, Hopper and Traugott 1994, Bybee et al 1994, Harris and Campbell 1994, etc.). For a brief survey of this work, and in particular for a discussion of how it provides principles that can guide researchers in interpreting modern grammatical patterns so as to determine the type of change (reanalysis, analogical extension or contact-induced change) that has led to a given modern pattern, cf. Gildea (1998, ch. 3). The contributors to Gildea (ed. 2000) use these methods in a number of case-studies of comparative grammatical reconstruction in language families of the Americas. Further principles of internal reconstruction are provided in Givón (1979, 1995, and especially 2000), some of which are useful in determining relative ages of individual components found in a complex modern system.

In this section, I introduce the modern patterns of data that result from the three major (and perhaps the only) mechanisms by which new morphology is introduced into grammatical systems: reanalysis (2.1), analogical extension (2.2), and borrowing (2.3). I conclude the section with a list of the properties are generally associated with relatively older

morphology versus relatively younger morphology in a complex system (2.4).

2.1 REANALYSIS

Reanalysis is the process by which a given construction, with its attendant morphology and syntactic relations amongst separate components of the construction, is reinterpreted by the speakers into a new construction type. A common example of reanalysis is the creation of a progressive aspect from a locative construction (Heine et al 1991; Heine 1993, 1994; Bybee et al 1994; Gildea 1998).

> In many languages worldwide, locative constructions of the type "Peter is at home" have been used to express verbal aspects ("Peter is at / in / on working" > "Peter is working"). In such cases, a verb in some nominalized form, such as a participle, a gerundial, or an infinitive ('work-ing'), takes the place of the noun phrase ('at home')... The use of the locative construction for a verbal aspect creates a form-meaning asymmetry, at least for some time; since the form "Peter is at / in / on working" is ambiguous, it may mean something like 'Peter is at his place of work' or else 'Peter is working.' ... The transition from locative construction to aspect marking leaves three types of morphological material behind: the erstwhile auxiliary ('be' in our example), the adposition ('at'), and the nominalization marker of the main verb ('-ing'). This means that there are three forms for marking one function only.
> (Heine, Claudi and Hünnemeyer 1991.214-5)

The key properties of reanalysis are:

- Reanalysis is covert: there is no immediate change in actual utterances. Only later do changes come that show the items in the construction have taken on a new morphosyntactic relationship with each other. These changes are not a part of the reanalysis, but are later extensions that reveal the morphosyntactic consequences of the reanalysis.

- Reanalysis creates polysemy: the morphemes that occur in the reanalyzed construction take on new meanings/functions, while maintaining their etymologically prior meanings/functions in other constructions. For instance, English *-ing* is still a perfectly good nominalizer outside of the progressive construction (e.g. *I like dancing. Drinking can be fun. These boots were made for walking*, etc.).

- Reanalysis affects entire constructions: e.g., the reanalysis that created the English progressive did not simply add one more inflection to the existing verbal paradigms, but rather it brought an entirely new system into main clause grammar. Instead of a simple inflected verb, it required a tensed auxiliary, dual marking of the verb (the preposition *on* and the suffix *-ing*), and an accusative case-marker *of* (*He is on building of a house.*).

Early cases of reanalysis can be difficult to identify, since one is always able to maintain a conservative analysis of the morphosyntax and simply suggest that the original grammar now supports a new function. Early in the reanalysis of the English Progressive, one can observe a shift in meaning from 'subject is physically located someplace where he is in the midst of doing something' to 'Subject is in the midst of doing something' without the locative sense (cf. the detailed discussion in Bybee et al 1994.136). Nonetheless, a **syntactic** analysis does not follow automatically from a **semantic** innovation. Before the changes later wrought by analogical extension, an analyst could have insisted that the nascent English progressive remained simply a nominalization placed in a prepositional phrase, which was then made the predicate of the copula and hence attributed some property to the subject of that copula.

2.2 EXTENSION

In contrast to reanalysis, extension makes explicit changes in the form of utterances, motivated by a condition of analogy between two utterance types. Changes brought about by extension affect only one part of a construction at a time, leaving the rest essentially unchanged. In our example of the English Progressive, speakers apparently drew an analogical connection between the new progressive and other main clauses, then changed the progressive to bring it more into alignment with other main clauses. First the predicate preposition *on* was reduced phonologically (*on* > *'n* > *'-* > *Ø-*), then the accusative preposition *of* became optional and disappeared altogether. While the progressive did not become one more simple verbal inflection, two major morphemes that distinguished it from other main clauses were eliminated. Extension can also operate in the absence of construction reanalysis, as when person-markers inside an existing paradigm expand in semantic scope (e.g. a marker of A expands to mark Sa, and numerous other examples in Harris and Campbell 1994). The key properties of extension are:

- Extension causes overt changes in form.
- Extension is the mechanism that reveals prior reanalysis.

- Extension affects individual morphemes or syntactic patterns.

2.3 BORROWING

The third major mechanism for introducing new grammatical patterns or morphemes is borrowing. In cases of intensive language contact, fluent bilingual speakers may introduce patterns from one language into another. These patterns may include actual morphemes from the source language (e.g. the Portuguese complementizer *que* into Nheengatu), or speakers may simply press existing morphemes into service to form a parallel pattern in the borrowing language (the term *calque* is a label frequently associated with this process). The key properties of borrowing are:

- Borrowing causes overt changes in formal structure.
- Grammatical borrowing is usually accompanied by extensive lexical borrowing (cf. Thomason and Kaufman 1988).

2.4 INTERNAL RECONSTRUCTION: TYPICAL PROPERTIES OF OLDER MORPHOLOGY

Givón (2000.120-121) discusses a number of general principles of historical syntax that yield a list of predictions, among which I highlight the following:

- **Phonetic size:** The smaller a morpheme is, the older it is.
- **Semantic size:** The more generic, grammaticalized, or semantically opaque a morpheme is, the older it is.
- **Distance from stem:** All other things being equal, the closer a morpheme is to the stem/root of the word, the older it is.
- **Morpho-phonemic irregularity:** The more morpho-phonemically irregular or variable a morpheme is, the older it is.

The caveat "all other things being equal" is necessary because factors other than morpheme age can lead to any of these properties individually. But when they are all found together in the case of a specific morpheme, the burden of proof falls heavily on the analyst who intends to argue that the morpheme in question is relatively young. Further, very old morphemes can sometimes be found in innovative constructions, in which case the age of the morpheme in question is not a good indicator of the age of the overall

system. For example, in the Cariban family a number of old nominalizers with irregular plural forms and morphophonemic irregularities were reanalyzed into verbal tense-aspect markers, thus bringing evidence of very old morphology into very new grammatical constructions (Gildea 1998, ch. 9-14). In such cases it is important to take into account the entire constructions in question to ascertain whether the grammar is conservative or innovative.

To Givón's list of properties, I would add another:

- **Identifiable source forms:** Morphology that has no "cognates" elsewhere in the grammar to serve as possible source forms for reanalysis or extension, or that was not plausibly borrowed into the lnaguage from an identifiable source in another language, is more likely to be old.

One final note: when reconstructing grammar, one is limited to reconstructing morphology and morphosyntactic systems on the basis of the surviving morphology and morphosyntactic systems. One can never be confident in reconstructing a complete picture of all the grammatical resources of a Proto-language because one never knows what pieces might have disappeared altogether, leaving no reflexes in any modern language. We can only expand our vision of the original system by expanding the database, either from new descriptions of languages inside the family, or by comparison to data from other, plausibly related language families.

3. THE PROTO-TG MAIN CLAUSE PERSON-MARKING SYSTEM

Jensen (1998) reconstructs the following person-marking system to Proto-TG.

A / O	1	1+3	1+2	2	2Pl	3	So
1				*ìe (r-)*	*ìe (r-)*	*ìe (r-)*	*ìe (r-)*
1+3				*ore (r-)*	*ore (r-)*	*ore (r-)*	*ore (r-)*
1+2						*jane (r-)*	*jane (r-)*
2	*oro-*	*oro-*				*ne (r-)*	*ne (r-)*
2Pl	*opo-*	*opo-*				*pe (r-)*	*pe (r-)*
3	*a-i-*	*oro-i-*	*ja-i-*	*ere-i-*	*pe-i-*	*o-i-*	*i-*
Sa	*a-*	*oro-*	*ja-*	*ere-*	*pe-*	*o-*	

Table 1. Reconstructed person-marking on the Proto-TG main verb

While the organization of Table 1 highlights the hierarchical nature of Proto-TG verb agreement, the organization of Table 2 highlights the distinct morphemes that can be analytically separated out, assigning most to a simple

contrast between marking A/Sa as opposed to O/So. The two forms that cannot be accommodated in such an analysis, *oro* and *opo*, appear to the right.

	A	Sa	So	O
1Sg	*a-*	*a-*	*ìe* ***r-***	*ìe* ***r-***
2Sg	*ere-*	*ere-*	*ne* ***r-***	*ne* ***r-***
3Sg	*o-*	*o-*	*i-*	*i-*
1+2	*ja-*	*ja-*	*jane* ***r-***	*jane* ***r-***
1+3	*oro-*	*oro-*	*ore* ***r-***	*ore* ***r-***
2Pl	*pe-*	*pe-*	*pe* ***r-***	*pe* ***r-***

1A2SgO	*oro-*
1A2PlO	*opo-*

Table 2. An analytic view of the person markers from Table 1

An important detail about this system is that the O/So forms are not an internally consistent class: the 3O/So form reconstructs to a bound prefix with suppletive allomorphy, whereas the Speech Act Participant (SAP) O/So forms reconstruct to a series of free pronouns, linked to the verb by means of the relator prefix *r-* (which also reconstructs with suppletive allomorphy). While Schleicher (1998) considers it likely that the SAP forms were already cliticized to the verb by Proto-TG, he agrees that they must have been free forms in Pre-Proto-TG.

A second important detail: The SAP O/So forms are clearly reduced forms of a series of free pronouns that are still attested in modern languages. While the 3O prefix is found in other parts of the grammar, there is no corresponding third person pronoun, nor is there a plausible source for the suppletive allomorphy (although Rodrigues ms., 1992) has developed an explanation of this apparently anomalous fact). The SAP A/Sa forms and the two anomalous forms *oro* and *opo* are attested nowhere else in the grammar, either marking other parts of speech or as components in free pronouns.

Hypotheses about the Pre-Proto-TG verbal system must then follow from an internal reconstruction of how these sets of forms came to be combined into the complex system that we see here.

4. JENSEN'S (1998) RECONSTRUCTION OF THE PRE-PROTO-TG PERSON MARKERS

Jensen proposes a scenario with five stages: first was a stage with only absolutive person-marking, then the Sa prefix set was added, then the Sa prefixes were extended to transitive verbs, and finally the two anomalous forms, *oro* and *opo*, were added.

Stage 1: verb has only absolutive forms, all but 3O as free forms preceding the verb; A forms are all free pronouns as well, not in constituency.
Stage 2: Sa verbs add new Sa prefix set, distinguishing them from So
Stage 3: Analogical extension of Sa prefixes to 3O transitive verbs; loss of internal 3O prefix in some languages, as it is now redundant.
Stage 4: Development of 1 > 2 > 3 hierarchy
Stage 5: New portmanteau forms *oro* (< *oro* '1+3'?) and *opo* 'are developed' to fill the gap

Given the principles described in section 3, the first stage of this scenario is problematic:

- The oldest forms ought to be the most reduced, but the SAP O/So markers (Stage 1) are reconstructed as free forms, whereas all the more recent person markers (Stages 2, 3, 5) are bound affixes

- The forms with no obvious sources elsewhere in the grammar ought to be the oldest, but the SAP O/So prefixes have clear cognates in free pronouns and person-markers throughout the grammar; further, the SAP forms require *r-* 'Relator', a sign of a phrasal — as opposed to paradigmatic morphological — relationship. In contrast, none of the more recent person markers have cognates outside the verbal prefix sets.

While the 3O/So prefix fits all the criteria for an old morpheme, including that it is closer to the verb when it co-occurs with SAP A prefixes, the remaining O/So prefixes do not appear to fit into this scenario. It would be better if we could imagine a scenario in which the 3O/So prefix is quite old, but in which the remainder of the O/So prefixes are not.

5. AN ALTERNATIVE HYPOTHESIS

Given all the signs of age attending the 3O prefix, I presume it was in place on the verb before any other prefixes became bound phonologically. Any other person-marking that might have co-existed with this prefix has been lost, leaving us unable to reconstruct a larger system. In essence, I reconstruct the 3O/So prefix as a relic from an unrecoverable prior system. From this starting point, Stage 2 is to posit the accretion of the A/Sa prefixes, presumably from a set of free pronouns that was already lost prior to Proto-TG, replaced by the pronouns that gave rise to the O/So person-markers. At this point, the system is essentially nominative-accusative, with the possible complication of the 3So form, which may have remained in

place. The forms *oro* and *opo* remain mysterious, but as they are bound forms and as their source is unknown, I place them in Stage 3, prior to the obviously younger SAP O/So forms. There is no necessary reason that they could not have evolved prior to — or at the same time as — the A/Sa forms, but their inclusion in Stage 2 would further muddy the otherwise relatively consistent nominative-accusative stage, so for aesthetic purposes, I separate them here. Finally, in Stage 4, the SAP O/So forms cliticize to the verb, creating the system that that has been reconstructed for Proto-TG.

Stage 1: The innermost 3O prefix *i- /c-* '3O/So'
Stage 2: The A/Sa Prefixes (jointly, or in either order)

1	> 3	*a-i-*	*a-c-*	1S	*a-*
1+3	> 3	*oro-i-*	*oro-c-*	1+3S	*oro-*
1+2	> 3	*ja-i-*	*ja-c-*	1+2	*ja-*
2	> 3	*ere-i-*	*ere-c-*	2	*ere-*
2Pl	> 3	*pe-i-*	*pe-c-*	2Pl	*pe-*
3	> 3	*o-i-*	*o-c-*	3	*o-* ~ *i-/c-*

Stage 3: The 1Sg/Pl —> 2 forms are added
Stage 4: The SAP O/So Prefixes/pro-forms are added

At this point, the internal evidence is exhausted. What remains is to discuss the implications that this hypothesis might have for further research, especially in related language families.

6. DISCUSSION

This hypothesis suggests that cognates to the oldest form, *i- / c-* '3O/So', ought to be found elsewhere in the Tupían stock, and potentially in the Cariban and Macro-Gê language families as well. Rodrigues (ms., 1992) has encountered evidence for a third person possessive morpheme *i-* (which he considers to be another form of a relator prefix) in all three; this form shows suppletive allomorphy as well, but it is not as strong as a potentially cognate third person **verbal** prefix might be. Interestingly, Sérgio Meira and Petronila Tavares (in personal communication) have posited a prefix *i-* '3O' in Cariban languages Carib of Surinam, Tiriyó and Wayana:

	Carib of Surinam	Tiriyó	Wayana
1A3O	s-i-	w-i-	w-i-
2A3O	m-i-	m-i-	m-i-
3A3O	n-i-	n-i-	n-i-
1+2A3O	kïs-i-	---	---

$1S_A$	wï-	t-/s-/k-/w-	w-
$2S_A$	mï-	mï-	mï-
3S	nï-	nï-	nï-
$1+2S_A$	kïtï-	---	---

In my (Gildea 1998) reconstruction of the Proto-Carib 3A3O form, I did not even consider this pattern. I now intend to return to my comparative Cariban materials to see if such an analysis can be sustained for other languages in the family, and even perhaps reconstructed to Proto-Carib. In addition to this potential cognate, in several Cariban languages certain verb tenses condition the 3A3O prefix *i-* instead of the expected *nI-*; I have been treating each case of this *i-* as an independent innovation, in which the form comes from an extension of the third person possessive *i-*. I will re-evaluate this hypothesis as well in future work, especially in light of more recent evidence for the robustness of third person *i-* in Southern Cariban language Ikpéng (Pacheco 2001).

Continuing to focus on comparisons to Proto-Carib, it is worth mentioning an asssertion by Derbyshire (1998): Given Jensen's reconstruction of Pre-Proto-TG as ergative, and given the strong likelihood that there is a genetic relationship between the Cariban and Tupí-Guaraní families, it is reasonable to presume that Proto-Cariban might also have been ergative. While the conclusion does not necessarily follow from the premises (e.g., closely related languages like Tiriyó and Carijona differ with regard to ergativity, and nobody has suggested that the relationship between Proto-TG and Proto-Carib is particularly close), it is nonetheless true that this new reconstruction of Pre-Proto-TG actually aligns much better with the existing reconstruction of verbal person-marking in Proto-Carib. Now neither shows main clause ergativity, and more, the two person-marking systems are heavily isomorphic (albeit with almost no actual cognates). In both, ergativity remains a property of subordinate clauses (cf. Gildea 1994, Derbyshire 1994).

Parallel to the investigations that this reconstruction will stimulate in the Cariban family, one might hope to look with new eyes on the comparison of Pre-Proto-TG grammar with the grammar of other languages in the Tupían stock, or of Macro-Gê. The oldest form is most likely to have reflexes extending the farthest away: are there cognates to the old *i- / c-* '3O/So' prefix to be found in verbal systems? If we are to find cognates to the younger forms, those added to the Pre-Proto-TG system in Stages 2 and 3, it will most likely be in Tupían languages. Could it be that any of the

presumed pronominal sources of these forms still exist (perhaps even still as pronouns) in other Tupían families?

Acknowledgement

The main ideas in this paper arose during my participation in the *Rencontre sur le grammaire des langues Tupi-Guarani*, held in Cayenne, French Guiana in January 2000, sponsored by the Institut de Recherche pour le Developpement (IRD) and the Centre Nacional de la Recherche Scientifique (CNRS). I presented a preliminary version of this reconstruction of Pre-Proto-TG person-marking there, and I thank all the participants for helping to stimulate these ideas and for their useful comments on this reconstruction. I am responsible for the errors that remain.

REFERENCES

Bybee, Joan, Revere PERKINS, and William PAGLIUCA. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago: University of Chicago Press.

Derbyshire, Desmond C. 1999. The Cariban Language Family. In R.M.W. Dixon and Alexandra Y. Aikhenvald (eds),*The Amazonian languages.* New York: Cambridge University Press.

Gildea, Spike. 1998. On Reconstructing Grammar: Comparative Cariban Morphosyntax. Oxford: University of Oxford Press.

Gildea, Spike, ed. 2000. *Reconstructing Grammar: Comparative Linguistics and Grammaticalization Theory.* Amsterdam: John Benjamins

Givón, T. 1979. *On understanding grammar*. New York: Academic Press.

Givón, T. 1995. *Functionalism and grammar*. Amsterdam: John Benjamins.

Givón, T. 2000. Internal reconstruction: As method, as theory. In Spike Gildea (ed), *Reconstructing Grammar: Comparative Linguistics and Grammaticalization Theory,* 107-159. Amsterdam: John Benjamins

Harris, Alice C. and Lyle Campbell. 1994. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. New York: Oxford University Press.

Heine, Bernd. 1994. Grammaticalization as an explanatory parameter. In William Pagliuca (ed), *Perspectives on Grammaticalization,* 255-287. Amsterdam: John Benjamins.

Heine, Bernd, Ulrike CLAUDI, and Friederike HÜNNEMAYER. 1991. *Grammaticalization: a Conceptual Framework.* Chicago: University of Chicago Press.

Hopper, Paul, and Elizabeth TRAUGOTT. 1993. *Grammaticalization.* Cambridge: Cambridge University Press.

Jensen, Cheryl. 1990. Cross Referencing Changes in the Tupí-Guaraní Family. In Doris L. Payne (ed), *Amazonian linguistics: Studies in lowland South American languages.* Austin: University of Texas Press

Jensen, Cheryl. 1998. Comparative study: Tupí-Guaraní. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds), *Handbook of Amazonian languages, v. 4*

Pacheco, Frantomé. 2001. Morfossintaxe do verbo Ikpéng (Karíb). Ph.D. Dissertation, Universidade Estadual de Campinas.

Payne, Doris L. 1994. The Tupí-Guaraní inverse. In Barbara Fox and Paul Hopper (eds) *Voice: Form and Function,* 313-40. Amsterdam: John Benjamins.

Rodrigues, A. D. 1992. Gammatical affinities among Tupí, Karíb and Macro-Jê, 1992 (ms.).

Schleicher, Charles. 1998. *Comparative and internal reconstruction of the Tupi-Guarani language family.* Ph.D. Dissertation, University of Wisconsin/Madison.

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

Revendo a classificação interna da família Tupí-Guaraní

Aryon Dall'Igna Rodrigues (Laboratório de Línguas Indígenas - UnB)
Ana Suelly Arruda Câmara Cabral (Universidade Federal do Pará)

1. INTRODUÇÃO

Com base no conhecimento sobre as línguas Tupí, que se tornou disponível a partir dos anos 60, Rodrigues (1985) apresentou a sua segunda proposta de subdivisão interna da família Tupí-Guaraní. Nessa nova classificação, diferentemente da proposta classificatória de 1958, tanto o Mawé, quanto o Awetí e o Mundurukú agora vistos como outras famílias dentro do tronco Tupí, foram excluídos da família Tupí-Guaraní. Esta, com aproximadamente 40 línguas, foi subdividida em oito subconjuntos, três dos quais constituem um ramo meridional e os outros cinco um ramo setentrional (cf. Dietrich 1990, Rodrigues, 2000). Para essa classificação foram consideradas basicamente propriedades fonológicas, porque a insuficiência de documentação não permitia o uso de critérios gramaticais e mesmo lexicais para um conjunto maior de línguas. Naquela ocasião, as informações existentes sobre as línguas de povos recém-contactados, como os Araweté, os Guajá e os Parakanã, restringiam-se a amostras lexicais mínimas e a existência de línguas como o Jo'é ainda era desconhecida. Agora, com o avanço considerável na documentação das línguas da família, inclusive com a divulgação das primeiras informações sobre o léxico e sobre a gramática de línguas faladas por povos contactados nos últimos vinte anos, tem-se tornado viável e necessária uma revisão da classificação interna publicada em 1985, de modo que seja mostrada, com mais detalhes e por meio de evidências adicionais, a complexa ramificação dessa grande família lingüística.

A PROPOSTA DE 1985

Rodrigues (1985) selecionou as seguintes propriedades lexicais e estruturais como diagnósticas tanto para inclusão de línguas na família Tupí-Guaraní, quanto para exclusão de línguas geneticamente aparentadas em nível mais remoto:

(a) Prefixos marcadores de sujeito comuns aos verbos intransitivos e transitivos em orações independentes, incluindo formas iguais a, ou deriváveis fonologicamente de: *a-* 'eu', *ere-* 'você', *ja-* 'eu e você', *oro-* 'eu e ele', *pe-* 'você e ele', *o-* 'ele, eles' (também 'eu, você e ele').