

Some cases of regrammaticalization in Tupí-Guaraní languages

Aryon Dall'Igna Rodrigues

Resumo

Apresento uma discussão sobre algumas mudanças morfossintáticas devidas a causas linguísticas ou sociais, ou de ambas naturezas, durante o desenvolvimento histórico de algumas línguas Tupí-Guaraní.

Palavras-chave: Tupí-Guaraní, morfossintaxe, linguística histórica, mudança linguística.

Abstract

I present a discussion on some cases of morpho-syntactic changes due to linguistic or social causes, or both, during the past historical development of some Tupí-Guaraní languages.

Keywords: Tupí-Guaraní, morphosyntax, historical linguistics, linguistic change.

1. Introduction¹

On the basis of more conservative languages of the Tupí-Guaraní linguistic family (such as Tupinambá, Old Guaraní, Kamayurá, Parintintín) it is possible to reconstruct some grammatical properties of Proto-Tupí-Guaraní – the pre-historical common ancestor of this family – and to observe several cases of regrammaticalization that have contributed to the differentiation in the realm of this language family. My aim in this paper is to present and to discuss in a very preliminary way a few instances of such morphosyntactic changes.

Tupí-Guaraní is a close-knit linguistic family even though it is a very widespread group of languages. With about 40 languages, it extends from the eastern coast of Brazil to the foothills of the Bolivian Andes and from French Guyana in the north to Argentina in the south. Both lexically and morphologically its languages share a high percentage of features, and their phonologies are not strongly divergent (Rodrigues 1985, Dietrich 1990). Only a very few languages, such as Kokárna, Sirionó, and Guayakí, show strikingly different grammatical features whose nature must be especially

¹ This paper is a revised version of a manuscript written in Brasília, January 1989.

investigated. Cabral (1995) has shown that Kokáma, even though with a Tupí-Guaraní basic lexicon, has a grammar that cannot be derived from Proto-Tupí-Guaraní, what means that it cannot be a member of the Tupí-Guaraní genetic family. Guayakí and Sirionó (and probably also Horá and Yúki) need yet to be submitted to a thorough study for ascertaining the nature of their divergences.

Some of the grammatical changes to be considered here are clearly consequences of phonological changes. Others are probably due to language contact, particularly with Portuguese in the case of Amazonian Língua Geral.

2. Nominalization of animate and inanimate subjects

Chiriguano (also called Áva and Bolivian Guaraní) is a Tupí-Guaraní language spoken in Southeastern Bolivia, Paraguay, and Northern Argentina. It belongs to the same branch I of the family as Guaraní and several dialects of it may be distinguished under the same or other names (e. g. Izozó, Tapiete). Its speakers have been known as invaders in the Inca Empire. The data considered here were published by Schuchard (1979) and Dietrich (1992).

Chiriguano has systematically shifted Proto-Tupí-Guaraní (PTG) stress from the last to the penultimate syllable, has dropped final consonants and, in addition, has changed the PTG phoneme **c* [ts] into zero between vowels.² Under the accumulated effect of these three changes, there was a merger of the nominalizing suffixes for animate agent **-ár ~ -cár* and inanimate agent (or circumstance) **-áβ ~ -cáβ*, both reduced phonologically to *-a*. This merger has obscured the opposition between animate and inanimate agents and has enhanced the use of relative nominalization, but only for human referents. Relative nominalization in Tupí-Guaraní languages is a device for nominalizing 3rd person syntactic subjects independently of their pragmatic role and its marker in PTG is reconstructed as **-βaʔé* (or perhaps **βaʔé*) suffixed to the verb inflected for the third person, and its reflex in Chiriguano is *-βae*. Now Chiriguano has, for instance, *mbaépa héi ndéβe oporopoanó-βae* ‘what has the doctor said to you?’, where *oporopoanó-βae* is derived from *oporopoáno* ‘he applies medicine’ and means literally ‘the one who applies medicine’, instead of PTG **moro-pocá-nóη-ár-a* ‘the animate agent (**-ár*) of applying (**nóη*) medicine (**pocáη*) for humans (**moro-*)’ (cf. Tupinambá *moroposánóηára*, Old Guaraní *poropohánóηára*, Mbyá Guaraní *poropoanõá*,

² The dropping of final consonants and the change **c > ∅* are shared with other languages of branch I, but the shift of stress to the penultimate syllable is probably due to contact with speakers of Bolivian Kechua.

all for 'doctor'). For non human referents, animals or things, Chiriguano has the nominalizer of agent *-a* without the addition of the relative nominalizer *-bae*, as in the following examples: *yandúti oisúwa* 'a spider's sting' (*oisúwa* < PTG **o-i-cu?w-áβ-a* 'action or effect of biting'), *ché aiasói asóya-pe* 'I cover myself with a blanket' (*asóya* < **aco?j-áβ-a* 'thing for covering').

3. From gerundial towards serial verbs and the loss of switch-reference

In the verbal system of the more conservative Tupí-Guaraní languages the verb of the predicate of a simple sentence or of the first predicate of a complex sentence occurs either in the indicative or in the imperative mood. In these moods, which take no inflectional suffix, the subject is marked by personal prefixes of two sets, one with six forms for the indicative mood and the other with two forms for the imperative. Such forms refer to the grammatical persons or combinations of persons of the subject: 1 'the speaker', 12(3^f)/3^f 'the speaker and the addressee with or without a non focal third person', 123^f/3^f 'the speaker and the addressee with one or more focal third persons or only one or more focal third persons', 13^f 'the speaker and one or more focal third persons', 2 'the addressee', 23^f 'the addressee and one or more focal third persons'. For transitive verbs, after the personal prefix for the subject and immediately before the verbal stem, there is a relational prefix referring to the object. While the paradigm of the indicative mood has prefixes for all those six "persons", namely 1 **a-*, 12(3^f)/3^f **ja-*, 123^f/3^f *o-*, 13^f **oro-*, 2 **ere-*, 23^f **pe-*, the paradigm of the imperative has only those for 2 **e-* and 23^f **pe-*. Besides the indicative and imperative there are also a co-referential and a non co-referential mood. This latter, morphologically marked by the suffix reconstructed as **-eme* ~ *-me*, has been called subjunctive and the co-referential, with the suffix **-áβo* ~ *-βo* ~ *-a* ~ (*-r* → ∅), has been called gerund by the first grammarians of these languages (Anchieta in the 16th century (1595) for Tupinambá and Ruiz de Montoya in the 17th century (1640) for Old Guaraní). The Tupí-Guaraní gerund implies the same subject as that of the indicative in the first or main clause, whereas the subjunctive implies a different subject, so that we can consider these two moods as the two opposite members of a switch-reference system. The gerund has a particular set of personal subjective prefixes in intransitive verbs, but no such prefixes in the transitive ones; these take only referential prefixes making reference to their objects. The subjunctive takes only relational prefixes in both transitive and intransitive verbs, but in the latter they make reference to the subjects.

The reconstructed Proto-Tupí-Guaraní forms for the personal prefixes of the gerund are: 1 **wi-*, 12(3^f)/3^f **ja-*, 123^{+f}/3^{+f} *o-*, 13^{+f} **oro-*, 2 **e-*, 23^{+f} **pe-*. With the exception of 1 and 2 these forms are the same as those of the indicative mood, but in the indicative they occur not only in the intransitive, but also in the transitive verbs. For the relational prefixes the reconstructed forms are: **r-* ~ \emptyset ‘the referent is the immediately preceding noun’, **c-* ~ *t-* ~ *i-* ‘the referent is a not immediately preceding noun’.

Exemplifying: Tupinambá, for instance, has *a-só kunumĩ pişik-a* [1.ind-go boy take-ger] ‘I went and took the boy’, *a-só wi-porasěj-a* [1.ind-go 1.ger-dance-ger] ‘I went and danced’, *ere-só e-porasěj-a* [2.ind-go 2.ger-dance-ger] ‘you (sg.) went and danced’; but *kunumĩ o-só o-porasěj-a* [boy 3.ind-go 3.ger-dance-ger] ‘the boy went and danced’, *oro-só oro-porasěj-a* [13.ind-go 13.ger-dance-ger] ‘we went and danced’, *pe-só pe-porasěj-a* [23.ind-go 23.ger-dance-ger] ‘you (pl.) went and danced’. The gerundial mood is used not only for additive clauses, but also for final or purposive ones, so that the first example above may also mean ‘I went in order to take the boy’ and the second example may equally mean ‘I went in order to dance’.

3.1 Chiriguano

In Chiriguano the changes of stress placement and the loss of final consonants have concurred to the loss of originally unstressed final syllables (**kutúk* > *kútu* ‘to stab’, *tetáma* > *těta* ‘place’) and thus have affected also the gerundial forms (**kutúka* > *kútu* ‘stabbing’, *ocóbo* > *óho* ‘(he) going’), thereby erasing the phonological difference between the gerundial and the indicative forms for all persons but the 1st and 2nd singular of the intransitive verbs, which had different personal prefixes. But analogy has then replaced the gerundial correferential prefixes for 1 and 2 of the intransitives by the corresponding prefixes of the indicative, so that any difference between the two verbal moods disappeared for these verbs. A further analogical step replaced the gerundial forms of the transitive verbs, which had no subject prefixes, by those of the indicative with their own obligatory prefixes. A consequence of all these changes is that Chiriguano has no gerundial form anymore. This notwithstanding, the syntax has only partially changed. For final clauses a new construction is now being employed, namely the prospective form of the relative nominalization (*oyápo... ogwápi-wãe-ra* ‘he made it... in order to sit down’, where *-wãe-ra* < **-baŕe* ‘norninalizer’ + **-ram* ‘nominal prospective state’. For additive clauses, no change has occurred, so that now the successive verbs occur in the indicative without any special marking, as in the following examples: *i-fi*

ô-e ó-ho o-gwáta ‘his mother came out and went walking’, litt. ‘his mother came.out went walked’), *kwápe á-iko-ma fe a-kwakwáa* ‘here I have lived and I have grown up’, *hãe-ma gwé-ru háwo júpe, hãe-ma o-mboyakio, hãe-ma o-ãpi, ópa o-ãpi* ‘he brought (*gwé-ru*) him (*júpe*) soap (*háwo*), he loused him (*o-mboyakio*), he cut him the hair (*o-ãpi*), he cut him (*o-ãpi*) all the hair’ (examples from text in Dietrich 1972). As we see, instead of the gerundial constructions, Chiriguano has now a serial chaining of verbs with the same subject, instead of the common Tupi-Guaranian switch-reference system.

4. Changes in the pronominal systems

4.1 Amazonian Língua Geral

Amazonian Língua Geral (LGA, also called Nheengatú) evolved from 17th century Tupinambá (T) as the language of descendants of Tupinambá mothers with Portuguese men in northern Brazil. The offspring of such mixed marriages was in part bilingual in LGA and Portuguese (P) and, although LGA came to be the dominant language of a large population for about 300 years, there has always been bilingual speakers with Portuguese either as a first or as a second language (Rodrigues, 1996).

4.1.1 Tupinambá and Língua Geral Amazônica

Tupinambá had three pronouns and three agreeing subject markers for the ‘first person plural’, one of them excluding the addressee and including a focal third person, the other two including the addressee, and one of these including also a focal third person: *oré, oro-* ‘1+3^{+f}’, *jané, ja-* ‘1+2’, *asé, o-* ‘1+2+3^{+f}’. The prefixes *ja-* and *o-* agreed also with non-focal third persons (3^{-f}) and with focal third persons (3^{+f}), respectively. This fact has been interpreted as reflecting a basic distinction between the existence and the absence of a speaker-addressee contrast in T (Rodrigues 1990). In LGA, only one of these pronouns, *jané*, survived, having become a universal ‘we’, just as Portuguese *nós*. The prefix *ja-* continued to agree with *jané*, with the same new universal meaning ‘we’; the prefix *o-*, now in some dialects of LGA *u-*, specialized for agreeing only with third person subjects, whereas the prefix *oro-* disappeared totally together with the corresponding pronoun *oré*. The end-product of these changes was a personal system more similar to the Portuguese one, with only one pronoun and only one agreeing marker for a universal ‘we’ (*nós* and *-mos*, respectively).

The similarity of the LGA system with that of P was enhanced by another change that affected the deictic system of the language. T had a relatively complex system of demonstratives, with several intertwined features, such as proximity to the speaker and proximity to the addressee, visibility, definiteness, anaphoric reference to what was said or meant, and had nominalized forms for each demonstrative. LGA has radically simplified this system, reducing it to a binary opposition between proximity to the speaker-and-addressee and distance from both. In a dialect of LGA fairly well documented (Magalhães 1876) the two demonstratives are *kwaʔá* ‘this’ and *jaʔã* ‘that’. *kwaʔá* corresponds to T *koβaʔé*, the nominalized form of *kó* ‘this (near the speaker)’ (*koβaʔé* > **kowaʔé* > *kwaʔé* > *kwaʔá*); in the Upper Rio Negro dialect, which has lost the glottal stop and contracted adjacent identical vowels, this word is *kwa*). The other LGA demonstrative *jaʔã* (Upper Rio Negro *jã*) stems probably from a parallel nominalization of T *iʔãŋ* ‘this (near speaker and addressee)’, but with a change in its reference. Besides these two demonstratives, LGA has also retained the anaphoric T *aʔé* ‘that one spoken or thought about’ as a personal pronoun for the third person, quite parallel in usage to Portuguese *ele/ela* ‘he/she’ (Magalhães *aʔé*, Upper Rio Negro *aé*). The parallelism with Portuguese was increased in LGA by the creation of a plural form *aʔitá* (Upper Rio Negro *aitá*) ‘they’ (Portuguese ‘eles/elas’) with the addition of LGA *-itá* ‘plural marker’ (from T *-etá* ‘(to be) many’). In T there was no personal pronoun for third person. The consequences of the changes in LGA in comparison with Tupinambá and Portuguese may be overviewed in the table below:

	Tupinambá		LGA		Portuguese	
1	<i>ifé</i>	<i>a-</i>	<i>ifé</i>	<i>a-</i>	<i>eu</i>	<i>-o</i>
12(3 ^{-f})	<i>jané</i>	<i>ja-</i>	<i>jané</i>	<i>ja-</i>	<i>nós</i>	<i>-mos</i>
123 ^{+f} (3 ^{+f})	<i>asé</i>	<i>o-</i>	<i>jané</i>	<i>ja-</i>	<i>nós</i>	<i>-mos</i>
13 ^{+f}	<i>ore</i>	<i>oro-</i>	<i>jané</i>	<i>ja-</i>	<i>nós</i>	<i>-mos</i>
2	<i>ené</i>	<i>ere-, e-</i>	<i>iné</i>	<i>re-</i>	<i>tu</i>	<i>-s</i>
23 ^{+f}	<i>peʔẽ</i>	<i>pe-</i>	<i>pejẽ</i>	<i>pe-</i>	<i>vós</i>	<i>-is</i>
3			<i>aʔé/aé</i>	<i>u- ~ w-</i>	<i>ele/ela</i>	<i>-Ø</i>
33			<i>aʔitá</i>	<i>u- ~ w-</i>	<i>eles/elas</i>	<i>-N</i>

Besides having generalized the use of *jané* and *ja-* for all occurrences of 1 with any other person (and consequently having eliminated the pronouns *asé* and *oré*), LGA has extended the use of a particular item of the deictic system of PT, **aʔé* ‘that invisible person or thing’, for referring to any person in the same way as *ele* and *ela* are used in Portuguese (recall that P *ele* and *ela* represent also the extension of use of a piece of the deictic system of the Latin language, which had no personal pronoun for 3). On the other hand, in LGA the T stem *-etá* ‘many’ came to be used as a marker of plural of the nouns under the influence of the systematic marking of plural in Portuguese nouns. This marker, LGA *-itá*, was also added to *a'é* or *aé* resulting in *aʔitá* or *aitá* corresponding to P ‘eles’ and ‘elas’. The result in LGA is a system wholly parallel to that of Portuguese, notwithstanding the contrast of a prefixal with a suffixal inflection.

4.2 Ka'apór

Ka'apór (Urubú, Urubú-Ka'apór) is a Tupí-Guaraní language now spoken in western Maranhão state in northern Brazil. Its speakers have formerly lived more to the west, in the Tocantins river valley (Ballée 1994:30ff). In their migration they contacted people speaking LGA and their language has loanwords and other traces of the influence of that language (Corrêa da Silva 1987). The elimination of the distinction in the expression of the combinations of the first with the other persons in the pronouns is probably due to the contact with speakers of LGA. The set of personal pronouns and the corresponding personal prefixes of Ka'apór is the following:

	Personal pronouns	Personal prefixes
1	<i>ihẽ</i>	<i>a-</i>
12(3)	<i>jané</i>	<i>ja-</i>
2	<i>ne</i>	<i>re-</i>
23	<i>pehẽ</i>	<i>pe-</i>
3	<i>aʔé</i>	<i>o- ~ u- ~ w-</i>
33	<i>aʔé ta</i>	<i>o- ~ u- ~ w-</i>

As may be easily seen in this set, the pronoun *jané* and the corresponding prefix *ja-* have had its meaning extended to all situations

in which 1 is associated to another person, be it 2 or 3 or a combination of 2 and 3 or of 3 and 3. This is essentially what has happened in LGA, as shown in the preceding section.

Analogously to LGA Ka'apór employs the deictic *aʔé* as a personal pronoun for 3, namely *aʔé*, and uses a reduced form of the stem *-etá* 'to be many' for pluralizing this pronoun: *aʔé ta* 'they'.

5. The reduction of the case system

5.1 Tupinambá

Tupinambá had a system of five nominal cases: (a) argumentative (also called onomatic, nominal or nominative) for arguments in a clause, (b) transitive (or predicative) for being in or getting into a state, (c) punctual locative for well defined location, (d) diffuse locative for diffuse or spread location, and (e) situational (also called partitive) locative for situation relative to a part of a whole, as in the following paradigms:

	C-ending stems	V-ending stems
Stems:	<i>ajúr</i> 'neck'	<i>kuʔá</i> 'waist'
argumentative	<i>ajúr-a</i>	<i>kuʔá-∅</i>
transitive	<i>ajúr-amo</i>	<i>kuʔá-ramo</i>
punctual locative	<i>ajúr-ipe</i>	<i>kuʔá-pe</i>
diffuse locative	<i>ajúr-iβo</i>	<i>kuʔá-βo</i>
situational locative	<i>ajúr-i</i>	<i>kuʔá-j</i>

5.2 LGA

LGA has lost the case system of T as the result of a conspiracy of quite different changes. In T the caseless stems worked as existential or possessive predicates, as in *tapiʔír* 'it is a tapir' or 'there is a tapir', *ne memír* 'you have a child' or 'there is a child of yours'. The argumentative case marked the arguments of a predicate, not only the main ones, such as subject and direct object, but also the objects of postpositions, as in *ne memír-a tapiʔír-a ojuká iβírásáŋ-a pupé* 'your son killed (*ojuká*) a tapir with a club (*iβírásáŋ-a*)'. In LGA the predicative use of nouns was replaced by the existential verb of

T *-ikó* ~ *-ekó* ‘to be in movement’ and its comitative-causative derivation *-erekó* ~ *-rekó* ‘to cause something/somebody to be (in movement) with oneself’, so that now *-ikó* (Upper Rio Negro *-ikú*) is used as a copula and *-rekó* (Upper Rio Negro *-rikú*) as a possessive verb ‘to have’ with two arguments: *iné rerekó ne memira* ‘you have your child’. The caseless forms of Tupinambá disappeared and the argumentative forms became the stem forms in LGA: *memira* is no more analyzable as *memir* plus *-a*, having become the unsegmentable word *memira*.

The translative case (Tupinambá *memiramo*) was generally replaced by the new basic form (*memira*) and partly replaced by the prospective form of the noun (Tupinambá *memirama* ‘future child’, *kuʔaráma* ‘future waist’), a replacement favored by the similarity of the endings; but in LGA the prospective form was on its turn reanalyzed as a sequence of a noun and a postposition *aráma* ‘for, in order to’ (this postposition having probably originated in the future form of the verbal nominalization for animate agent, ...-*ár-ám-a* ‘one that will ...’). T *sje rúβ-a tuβifáβ-amo oikó* ‘my father is (in the state of) a chief’, LGA *se pája uikú tui fáwa* ‘my father is a chief’; T *ja-s-enōj jané r-uβisáβ-amo* ‘we called him for being our chief’, LGA *ja-senūj aʔé jané tui fáwa aráma* ‘we called him for our chief’.

The punctual locative replaced the diffuse and the situational ones, except for a very few instances that crystallized as adverbial words. On the other hand, in most words the (punctual) locative suffix evolved into a postposition, *upé*, having as its object the noun in the new unanalyzable form originated in the argumentative case of T. Instead of T *ókipe* ‘in the house’ LGA has *úka upé*. But the suffix *-pe* has been retained in words that ended in *-β* in T, a segment that dropped before the suffix: T arg. *kisáβ-a* ‘hammock’, p. loc. *kisá-pe* ‘in the hammock’; LGA *kisáwa, kisápe*. It has been retained also in some words that took an adverbial character, such as *árpe* ‘on’ (T *ʔári-pe* ‘on the surface’), *wírpe* ‘under’ (T *wírpe* ‘in the bottom’). The retention of the locative instead of its replacement by the postposition *upé* varies according to the dialects of LGA, but the other cases of Tupinambá appear to have totally disappeared from all dialects.

Even though Portuguese is a prepositional language as opposed to the Tupián languages that are postpositional, the strong reduction of the case system in LGA may be seen as the result of a structural influence of the Portuguese language, in the sense of a change from a more synthetic to a more analytic grammatical system.

References

- Anchieta, Joseph de. 1595. *Arte de grammatica da lingua mais usada na costa do Brasil*. Coimbra.
- Balée, William. 1994. *Footprints of the forest: Ka'apor ethnobotany – the historical ecology of plant utilization by an Amazonian people*. New York: Columbia University Press.
- Cabral, Ana Suely A. C. 1995. *Contact-induced language change in the western Amazon: The non-genetic origin of the Kokama language*. Ph.D. dissertation, University of Pittsburgh.
- Corrêa da Silva, Beatriz. 1997. *Urubú-Ka'apor; da gramática à história: a trajetória de um povo*. Master Thesis, Universidade de Brasília.
- Dietrich, Wolf. 1990. *More evidence for an internal classification of Tupí-Guaraní languages*. Indiana, supplement 12. Berlin: Gebr. Mann.
- Dietrich, Wolf. 1992. *El idioma Chiriguano: gramática, textos, vocabulario*. Madrid: Instituto de Cooperación Iberoamericana.
- Rodrigues, Aryon D. 1985. Relações internas na família linguística Tupí-Guaraní. *Revista de Antropologia* 27/28:33-53. São Paulo.
- Rodrigues, Aryon D. 1990. You and I = neither you nor I: the personal system of Tupinambá. In: Doris L. Payne (ed.), *Amazonian linguistics: studies in lowland South American languages*. Austin: University of Texas Press, pp. 393-405.
- Rodrigues, Aryon D. 1996. *Argumento e predicado em Tupinambá*. Boletim da Associação Brasileira de Linguística 19:57-66.
- Ruiz, Antonio. 1640. *Arte, bocabulario de la lengua Guaraní*. Madrid.
- Schuchardt 1979. *Ñande ñe: gramática Guaraní para Castellano hablantes*. Santa Cruz de la Sierra.