

FRANÇOISE ROSE (CNRS/FRANCE)*

Borrowing of a Cariban number marker into three Tupi-Guarani languages

Abstract

This paper discusses the process of borrowing of a Cariban number marker **komo* by three Tupi-Guarani languages (Wayampi, Emerillon and Zo'é). This process had already been brought to light before (Jensen 1979: 16) and is uncontroversial. In a contact situation in which lexical borrowing has occurred, this grammatical marker is apparently the only bound morpheme to have been borrowed by the languages under study. This paper has two main objectives; the first one is to investigate the historical contact situation between the populations involved and the second to describe the specific contact-induced changes at the phonological, distributional, morphological and semantic levels, in order to facilitate hypotheses regarding the socio-historical and linguistic modalities of this loan. The detailed study of the unique case of borrowing of a bound grammatical morpheme in these languages is the starting point for discussing the present understanding of contact-induced changes in general.

1. Introduction

This paper studies a process of borrowing of a Cariban number marker into three Tupi-Guarani languages, Wayampi (1), Emerillon (2) and Zo'é (3). This process has already been brought into light before by Jensen (1979: 16). This author stated without any more details about Wayampi and Emerillon that “the two languages have borrowed the subject/object pluralizer *kō/kom* from the Cariban language family.”¹ This claim has been reasserted in the later literature (Dietrich 1990: 65, Jensen 1999: 151, Rose 2011: 116) and extended to Zo'é (Cabral p.c.). It is considered uncontroversial, even though no argument has ever been offered for it, and the socio-historical and linguistic modalities of the suggested borrowing have never been discussed.

* I would like to warmly thank many colleagues who contributed to this paper with their insightful comments: Françoise and Pierre Grenand, Denis Creissels, Peter Bakker, Sérgio Meira, Claudine Chamoreau, Odile Renault-Lescure, Wolf Dietrich, François Copin, Pieter Muysken and Frank Seifart. Vincent Monatte has kindly drawn the map for me, on the basis of the maps in Queixalós & Renault-Lescure (2000).

¹ “as duas línguas têm emprestado o pluralizador do sujeito/objeto, *kō/kom* da família lingüística Carib.” (Jensen 1979: 16, about Wayampi and Emerillon).

Emerillon, Tupi-Guarani (Rose 2011: 34)

- (1) *aʔe-nawe* *sikãĩ-we* *wãĩwĩ-kom* *aʔuwe*
 DEM-POSTP be.small-too woman-PL still
 ‘At this time, the girls were still small too.’

Wayampi, Tupi-Guarani (Grenand 1980: 58)

- (2) *e-memi-kũ*
 1SG-child-PL
 ‘my children’

Zo'é, Tupi-Guarani (Cabral, p.c., her translation)

- (3) *kubi'é* *kã*
 man COL
 ‘all the men’

The morpheme illustrated above in three Tupi-Guarani languages can indeed easily be compared to the collective particle **komo* reconstructed by Gildea (1998) for Proto-Carib, not only for its form and its meaning but also for specific aspects of its distribution.

When following a possessed noun, it indicates the number of the possessor in the Cariban languages, as exemplified in (4). A parallel is found in Emerillon (Tupi-Guarani), as illustrated in (5).

Kaxuyana, Cariban (Gildea 1998: 117)

- (4) *ki-mnayi* *kumu*
 1+2-sister COL
 ‘the sister of all of us’

Emerillon, Tupi-Guarani (Rose 2011: 267)

- (5) *i-kiwir-a-kom*
 3.II-brother-REF-PL
 ‘their brother’

The objective of this paper is to investigate, first the socio-historical conditions of such a borrowing in the light of the contact history of the populations involved, and second, the linguistic aspects of the process of borrowing at the phonological, distributional, morphological and semantic level. The final goal of this paper is to propose hypotheses about the modalities of this unique morphological loan (Why? How? When? Where? etc.). Several hypotheses will be put forward, trying to explain the discrepancy of distribution of this morpheme between the Cariban system and the recipient languages. Another interesting aspect of this particular borrowing is that, in a contact situation in which lexical borrowing has occurred, this grammatical marker is apparently the only bound morpheme to have been borrowed into the languages under study. This fact will trigger discussion on some generalizations that have been suggested in contact linguistics.

Section 2 starts by presenting the contact history of Tupian and Cariban people and puts forward hypotheses about the possible socio-historical contexts of the specific borrowing under study. Sections 3 and 4 then examine the linguistic aspects of this loan. Section 3 describes the morpheme under study in the Proto-Cariban system and in the modern version of some possible Cariban source languages at the phonological, distributional, morphological and semantic levels. Section 4 presents the borrowed morpheme also at each level in the three Tupi-Guarani recipient languages and discusses the process of borrowing. Section 5 situates this specific borrowing of a bound grammatical morpheme in the present understanding of contact-induced changes in general.

2. Hypotheses on the socio-historical context of the borrowing

Cariban groups and Tupian groups are generally considered as constituting two different phyla, both very important in the Amazonian basin. This section presents information on the relationships between on the one hand Cariban groups and on the other hand the Tupian stock, and two internal sub-groups of the stock, first the Tupi-Guarani family and within it, the most northern Tupi-Guarani groups. Each of these three relationships will be examined as a possible historical context for the borrowing of the number marker.

2.1. Common genetic heritage or ancient contacts between Cariban and Tupi

Rodrigues (1985) analyses some phonological correspondences between the lexicon of the whole Tupian stock (of which the Tupi-Guarani languages form a branch) and of the Cariban family.² Some evidence suggests that at some point, the Cariban family and the Tupian stock were related, either by a genetic relationship or by an ancient contact situation. This question boils down to one of the major challenge of Amazonian linguistics: some clear cases of correspondence between lexical items and morphosyntactic features across large Amazonian families (Cariban, Tupian and Macro-Je) may be due either to some common heritage or to inter-ethnic contact at a remote point. Hornborg (2005) insists on the importance of the trade network in Amazonia in the first millennium BC, and its consequences for ethnic identity, linguistic differentiation and cosmology. Contact phenomena in Amazonia have often been underestimated as a possible trigger for language/cultural change, not just in terms of simple borrowing but also in terms of linguistic/cultural shift or merging.

² It appears now that most of the Cariban words that Rodrigues mentions in this list come from Hoff's grammar of Carib (Hoff 1968) and that there are in fact less cognates between Tupi languages and other Cariban languages of the Guianas (Meira, p.c.).

However, the number marker under study is attested in only three languages of the Tupi stock (Wayampi, Emerillon and Zo'é), without any traces of it in any of the other languages of the stock. Therefore, the possibility that this cognate form, shared among three Tupi languages and all the Cariban languages, originates in a form shared by a common ancestor of the two families or by some ancient contact between them is very unlikely and must therefore be discarded.

2.2. Old contacts between North Amazonian Carib and Tupi-Guarani

Rodrigues (1985: 375) gives another set of cognate forms with sound correspondences this time between North-Amazonian Cariban languages only and Tupi-Guarani languages only. The Tupi-Guarani family is the most important member of the Tupi stock in terms of the number of ethno-linguistic groups and their geographical dispersion. It is also a very coherent ensemble, indicating a rather recent differentiation. According to Rodrigues, these cognate forms very likely reflect the result of contact of population rather than of genetic relationship. Rodrigues (2000) synthesizes the results of the contact periods proposed in Rodrigues (1985) and the history of Tupi-Guarani dispersion and migrations. He works out that the hypothetical contact between Tupi-Guarani and Carib people must have taken place before the split of the Sateré-Mawé away from the Tupi-Guarani branch, and therefore probably in the area between the upper Juruena and Arinos rivers (still further south of the Amazon river).

Since, in the case under study in this paper, the three Tupi languages, Emerillon, Wayampi and Zo'é, all belong specifically to the Tupi-Guarani family (Rodrigues & Cabral 2002), the shared morpheme could go back to these old contacts between North Amazonian Carib and Tupi-Guarani. However, these three languages all belong to a single sub-branch (sub-branch 8) of the Tupi-Guarani family. Since the borrowed morpheme is not attested in any other Tupi-Guarani language, the hypothesis that the actual similarity originates in such an old contact is therefore unlikely and must also be discarded.

This assertion is supported by the fact that there is in fact no widely-shared morphology for plural marking within the Tupi-Guarani family (Jensen 1999: 151–152). Despite the great lexical and morphological similarity of Tupi-Guarani languages (Jensen 1999: 128), no plural morphology has ever been reconstructed for Proto-Tupi-Guarani. The number category is not even mentioned in any of the historical works on the family (Dietrich 1990, Jensen 1998, Schleicher 1998), and it has never been investigated before at the level of the family. This is illustrated below by examining number expression in four Tupi-Guarani languages: Tapirapé, Kamaiurá, Urubu-Ka'apor and Emerillon. These examples do not constitute a comprehensive survey of number expression in the family, but serve to illustrate how heterogeneous the category is. Two remarks can be made. First, no number marker is widely shared. Most are restricted to one or a few

languages.³ Second, most of these markers result from recent grammaticalization. Their use as number markers can be derived from a reconstructed Proto-Tupi-Guarani morpheme with some related meaning.

Tapirapé (Praça 2007: 61–64) does not display any plural marker as such. However, it uses the suffix *-kwer* ~ *-wer* (probably derived from the Proto-Tupi-Guarani past nominal tense suffix)⁴ for collective on unpossessable nouns with human referents (6), and reduplication for non-humans (7). The most productive device for number marking is an associative particle *agỹ* ~ *gỹ* on animate subjects (8). Finally, the singular is marked with the particle *gã* on demonstrative pronouns in subject function (9).

Tapirapé, Tupi-Guarani (Praça 2007: 61–64)

- (6) *akoma'e-kwer-a a-a i-ãpy-wo ka-ø*
 man-GRUP-REF 3.I-go 3.II-burn-GER field-REF
 'The men are going to burn the field.'
- (7) *ãxe'i rãka ã-pyyk ma'e-ma'e-ø confresa-pe*
 yesterday PAS.REC 1sg.I-take thing-RED-REF Confresa-LOC
 'Yesterday I bought many things in Confresa.'
- (8) *konomĩ-ø=gỹ-ø a-yj 'ot-a*
 child-REF=PL-REF 3.I-run 3.come-GER
 'The children came running.'
- (9) *ã'ẽ=gã-ø wer-or 'ãwãxi-ø*
 DEM=SG-REF 3.CAUS.SOC-come corn-REF
 'She brought the corn.'

Kamaiurá, along with possible cognates of collective *-kwer*,⁵ uses nasalization of final consonant (10) as a way to signal number on kinship terms (Seki 2000: 58–59) and a plural particle *awa* with variable position depending on whether it pluralizes a 3rd person subject, a possessor or a postposition object (11) (Seki 2000: 90).

Kamaiurá, Tupi-Guarani (Seki 2000: 59, 90).

- (10a) *i-tutyt* (10b) *i-tutyñ*
 3-uncle 3-uncle.PL
 'his/her uncle' 'his/her uncles'

³ In Tenetehára, the collective suffix *-kwer* and reduplication are found like in Tapirapé, as well as a final-particle *wa* pluralizing subjects that reminds of Kamaiurá *awa* (Duarte 2007). Guajá uses cognates of Tapirapé *-kwer* and *agỹ* (Magalhães 2007). In Guajá, nouns are unmarked for plural, but a collective suffix *-kér* ~ *-ér* is used for a coherent group (Magalhães 2007: 148), and a final-position particle *wỹ* ~ *awỹ* to pluralize a third person subject. With a subject of first or second person, or third person singular, it conveys an associative meaning (Magalhães 2007: 89).

⁴ Wolf Dietrich (p.c.) considers the plural marker and the past nominal tense suffix *kwer* a unique morpheme, linked by the concept of 'not belonging to'.

⁵ Three plural suffixes are presented, *-met*, *-het* and *-wet*, the distribution of which is not clear but is restricted to some animate nouns, mostly with human referents (Seki 2000: 58–59).

- (11) *a-raha i-upe-awa*
 1SG-carry 3-DAT-PL
 ‘I carried for them.’

Urubu-Ka'apor developed a plural marker *ta* (12) from the Tupi-Guarani descriptive noun *-eta* ‘many’, with some associative nuance (13).⁶

Urubu-Ka'apor, Tupi-Guarani (Kakumasu 1986: 328, 374).

- (12) *jawar ta*
 dog PL
 ‘the dogs’

- (13) *kaitā-ta*
 Caetano-PL
 ‘Caetano’s folks’

Finally, Emerillon displays three different plural markers (Rose 2011: chapter 5). Two of them encode the plurality of the subject: the verbal phrase clitic *-(o)ŋ* is clearly an innovation of Emerillon (14), while the *kupa* plural is also shared with Wayampi (15). Both forms are hypothetically the result of the grammaticalization of a gerundive. The lexical origin of *-(o)ŋ* is unknown while *kupa* is probably a frozen residual form of a verb **kúβ* ‘be together’ with the gerundive suffix **-a* (Jensen 1989: 104).

The third plural marker, *-kom*, results from a particular strategy: borrowing. It constitutes the central topic of this paper, and is presented in more detail in section 4. Emerillon also employs indirect strategies like reduplication expressing pluractionality, which can imply a multiplicity of arguments (16), and the use of completive aspect (17) that can indicate that the event expressed by the verb concerns the whole range of possible subjects of an intransitive verb, or objects of a transitive verb.

Emerillon, Tupi-Guarani (Rose 2011)

- (14) *o-me?eŋ-gatu sipa?a-m i-ŋupe-ŋ.*
 3.I-give-INTENS metal-TRANSL 3.II-for-PL.S
 ‘They give him metal.’
- (15) *e-re-iba, e?i baipuri-r-ie-pope-ne kupa-o.*
 1SG.I-RELN-pet 3.I.say tapir-RELN-belly-in-CONTR PL.S-CONT
 ‘“My pet”, they say together from inside the tapir’s belly.’
- (16) *iar-a-pope o-iju-ijuŋ beku.*
 canoe-REF-in 3.I-RED-put creeper.sp
 ‘They put all the creepers in the canoe.’

⁶ The element *eta* has also been a source for plural marking in Nheengatu (daCruz 2011: 377–378). Wolf Dietrich (p.c.) informed me that *-réta* is also found as a plural suffix on nouns, pronouns, and verbs in Chiriguano. In Tapiete (see González 2005: 101) it is *-reta* and *-re*. For Tupinambá, *-eta* was indicated as a plural marker by Anchieta (1595: 8).

- (17) *o-kuwa-pa* *o-manō-mãʔẽ*.
 3.I-know-COMPL 3.I-die-REL
 ‘He knows all the dead ones.’

This survey enables us to highlight the poverty and heterogeneity of number marking in Tupi-Guarani, contrasting with a very dense and homogenous marking in Cariban languages.

2.3. Recent contacts between some Tupi-Guarani groups and some Cariban groups, in the eastern Guianas and adjacent northern Brazil

Tupi-Guarani people split from the core of the Tupi family, in the State of Rondonia in south-west Brazil (Rodrigues 1985: 371–372). Very few people crossed the Amazon. Among them, members of the eighth sub-group are now geographically extremely distant from the core of the family, and lie at the very north-eastern margin of the Tupi expansion. They reached their present location in the last centuries. The Tupi-Guarani ethnic groups that later grouped together to form the Emerillon group entered the territory of present-day French Guiana probably in the late 15th century (Grenand 1982). The Wayampi probably crossed the Amazon in the early 18th century (Grenand 1982), and entered what is now French Guiana in the early 19th century (Hurault 1972: 79). Very little information is available on the history of the Zo'é people, officially contacted for the “first” time by non-Indians in 1987 (Bechara Sanchez 1998). The Emerillon now live in French Guiana, the Wayampi in French Guiana and in the Brazilian states of Amapá and Pará, and the Zo'é in Pará (cf. Map 1). They are now the most northern representatives of the Tupi-Guarani family, which spreads throughout Brazil, down to northern Argentina, Paraguay, and Bolivia and up to French Guyana. While northern Tupi-Guarani groups are the result of a migration from the south, the Cariban people followed a migration the opposite way around, from north to south. As a consequence, the borrowed plural marker is probably the result of a rather recent contact, probably in some already rather northern area. The contact involved some Cariban people (rather than a Proto-group) and either the common ancestor of these three modern Tupi-Guarani people (if this happened before the 15th century, and south of the Amazon) or the three groups individually, since they were already distinct in the last centuries.⁷

The direction of diffusion is without any doubt from Cariban to Tupi-Guarani. First of all, while **komo* is found only in three Tupi-Guarani languages, it is found in every single Cariban language and can be reconstructed in Proto-Carib by applying regular sound changes (Gildea 1998: 116–117). Second, while it cannot be integrated into any

⁷ This loan could logically have happened separately in each of them, or in one or two of them only. Contact, at least between the Emerillon and the Wayampi, may have facilitated the diffusion of the borrowing between them.

morphological Tupi-Guarani paradigm, it is part of a cross-Cariban series of number particles and third-person plural particles ending in *mo* (Meira, p.c.). Third, while number marking has little density in Tupi-Guarani languages (cf. section 2.2), Caribbean languages display a great number of encoding devices in that domain.⁸ This suggests that *komo* is originally a Caribbean item.

The contact between some Caribbean group and the three Tupi-Guarani groups under study (or their common ancestors) could have been facilitated by several factors. First, the most obvious factor is the cohabitation on the same geographical area. In what is now French Guiana, the presence of the newly arrived Tupi-Guarani groups since the 15th century resulted in different groups sharing the same living area, generally without permanent settlements. The Proto-Emerillon group entered the region as a wedge between in the north and the west Caribbean languages and in the north-east and the east Arawak languages, and some few Caribbean languages thus separated from the main block. Proto-Emerillon had therefore a more regular contact with Caribbean people of the Central Plateau (Aramisho, Aramakoto, Taripi, Kaikushian) than with those of the coast (Kali'na) (Pierre Grenand, p.c.). The Proto-Wayampi came into contact with Caribbean people about two centuries later. Emerillon have lived close to the Wayana (Caribbean) at least since the 18th century, and Wayampi got closer later on in the late 19th century (Hurault 1972: 179, 192). This has led nowadays to some multilingualism between Wayana, Emerillon and Wayampi. Métraux (1927: 34) states that since 1850, the Wayampi have been friends with the Wayana, who exerted a strong influence on them. As far as the Proto-Zo'é are concerned, they have been separated by a strong Caribbean block from the two former groups at least since the 18th century. They have very likely been in contact with some Carib people, since the region they live in (and have presumably been living in for some time) was populated by sub-groups of Proto-Tiriyo and Proto-Apalai speaking people (Bechara Sanchez 1998). In all these cases, the various populations sharing a common living area were obviously interacting. Nowadays, some north-eastern Caribbean groups such as Carib/Kali'na⁹, Wayana, Tiriyo, Apalaí and Kaxuyana live in the same area as the three Tupi-Guarani groups under study (sometimes with large empty spaces in between), that is to say in the eastern part of the Guianas and the adjacent part of northern Brazil (cf. Map 1).¹⁰ However, many of the Caribbean people that the Proto-Emerillon, Proto-Wayampi and Proto-Zo'é may have encountered have completely disappeared, as is the case for the groups that used to live in the central part of French Guiana (Hurault 1972).

⁸ These devices include particles (like *komo*, *tomo*, *jamo*), but also suffixes that mark number on the verb (*-të*, *-ti*, *-hki*), and for many languages there even is a clitic particle (or a suffix, depending on the language) that marks number of the argument of a postposition (*-ne*, *-nje*, *-ine*, *-xine*) (Meira, p.c.).

⁹ Carib is a language of the Caribbean family, also called Kali'na in French Guiana.

¹⁰ Kaxuyana is a less likely source, since it has come only recently (in the last 50–60 years) to the region. Before that, it was spoken along the Katxuru river, a tributary of the Nhamundá, in the Brazilian state of Pará but much to the south-west of the Surinam-French Guiana area (Meira, p.c.).

were also very active, with for example exchanges between Kali'na on the coast and Wayana and Aparais on the Jari and Paru rivers from the end of the 17th century to the early 18th century (Hurault 1972: 89). Wayana are still particularly known as traders of indigenous goods, pets and manufactured objects from Europe. According to Grenand (1982), they have had intense contact with Tupi-Guarani groups from 1830 to 1960.

Fourth, the use of pidgins as lingua franca also constitutes a likely vector for contact-induced changes. It is well-known that a Carib/Kali'na-based pidgin called *langue générale galibi* was used at least between the different Amerindian populations and the Frenchmen for more than 2 centuries.¹¹ Maurel underlines the predominant status of Kali'na or its pidginized variant, as first a lingua franca in the Guianas and second as a language used by the missionaries in French Guiana (Maurel 2000).¹² As argued in Taylor & Hoff (1980) and Renault-Lescure (1984), the description of Carib by Biet (1896[1661]) is that of a Carib/Kali'na-based pidgin. Another Cariban-based pidgin was first used between the Ndjuka and the Tiriyo, later extended to the Wayana.¹³ It has been reported in the early 20th century, and was still spoken at the end of the century, when Huttar described it (Huttar 1982). De Goeje (1990 [1937]: 100) also mentions another *jargon de traite*, supposedly a simplified version of Wayana with some Carib/Kali'na words. It is certain that the Emerillon and Wayampi living in the area may have used these pidgins when in contact with other Amerindian populations, Maroons, Creoles or Europeans even though they were not primarily involved in the constitution of these pidgins.

The contact between Cariban and Tupi-Guarani people of the area triggered interferences among their languages, although they are essentially limited to lexical borrowings (cf. Rodrigues 1985: 392, Rose & Renault-Lescure 2008). Jensen (1979) discusses how language contact led to the borrowing of Cariban words on the one hand into Emerillon and Wayampi (first two lines of Table 1) and on the other hand just into Emerillon (last two lines of Table 1). This fits with the fact that the Emerillon have had a longer period of contact with Cariban populations in French Guiana. Some European words were also borrowed into Wayampi and Emerillon indirectly via a Cariban language and through Maroon creoles (*arakapusa* 'gun' from Fr. *arquebuse* or Sp. *arcabuz* (Grenand 1989: 128), *kakusa* 'needle' from Sp. *aguja*, *sautu* 'salt' from Dutch *zout*). Rodrigues (1985: 392) gives a list of possible borrowings from Wayana into Wayampi.

¹¹ There is also some historical connection between this pidgin and the male speech of Island Carib (Taylor & Hoff 1980).

¹² Jesuit missions were active from 1711 to 1765, mainly along the coast and on the Oyapock (Hurault 1972: 119–136).

¹³ Historical information is however too scarce to know in which missions and with which ethnic groups this pidgin was used.

Proto-Tupi-Guarani	Wayampi	Emerillon	Carib/Kali'na	Translation
<i>ayuru</i>	<i>kure</i>	<i>kure</i>	<i>kurekure</i>	~ parrot sp.
---	<i>mariya, kise</i>	<i>baridža</i>	<i>mariya</i>	~ knife
<i>ok</i>	<i>oka, tapui</i>	<i>tapiđž</i>	<i>tapii</i>	~ house
<i>yacitata</i>	<i>yaıtata, silike</i>	<i>sirikə</i>	<i>siriko</i>	~ star/Venus

Table 1: Words borrowed from the Carib/Kali'na language into Emerillon (and Wayampi for some) (adapted from Jensen 1979: 13–14)¹⁴

In brief, the hypothesis according to which the morphemes *-kom*, *-kũ*, *-kã* of Emerillon, Wayampi and Zo'é were borrowed from a Cariban language is supported by the contact history of the languages of the area. In the following sections, several hypotheses will be put forward on the socio-historical contexts and source languages for the borrowing.

2.4. Hypotheses on socio-historical contexts and source languages for the borrowing

Taking into account what is known of the history of the recipient languages, the following hypotheses can be put forward:

Hypothesis 1: The morpheme was borrowed in some remote time into the common ancestor of these three languages, therefore long before their entering the French Guiana area.

1a. The source language could have been Proto-Tiriyo, Proto-Apalai or any other then existing Cariban language.

1b. The source language could have been a Cariban-based pidgin (primarily used to deal with Frenchmen).

This is theoretically less likely, since the type of contact involved with traders would be less intense, or in other words low on Thomason & Kaufman's scale of borrowings (1988), and therefore would rather facilitate borrowing of lexical item rather than of grammatical item.¹⁵

Hypothesis 2: The morpheme was borrowed when the three people were already differentiated.

2a. The three Tupi-Guarani newcomers arrived in some area where Cariban population was important. Any close local Cariban language could then have been a source.

2b. Population fusion with survivors of Cariban people could also be a good context for the borrowing, at a time when indigenous populations were pushed away by the colonizers and their slave expeditions, grouped into missions and then decimated by

¹⁴ The Wayampi data has been updated with the help of F. Grenand, the Carib/Kali'na data with the help of O. Renault-Lescure and the Emerillon data with my own corpus.

¹⁵ Yet the plural marker could have entered the recipient languages along with nouns borrowed both in the singular and plural form, and therefore have been easily analyzed and extended to native nouns.

epidemics. This situation would correspond to intense contact for the Tupi-Guarani speakers in Thomason & Kaufman's scale, leading to more structural borrowing. This would also imply some substratum interference in the speech of the Cariban speakers joining the Tupi-Guarani ones. From their perspective, this means language shifting. Consequently the source languages would now be extinct.

2c. Again a pidgin would be a less likely source for the borrowing.

The major problem raised by Hypothesis 2 is why then would this same and unique grammatical borrowing have happened exactly to these three languages? Chance is very unlikely, since we deal with three languages that are among them the closest linguistically and geographically. Other Tupi-Guarani people have also been in contact with Cariban populations and the Kamaiuras still are (Pierre Grenand, p.c.), but yet their languages were not influenced as far as number marking is concerned (cf. section 2.2). This common borrowing cannot be explained by contact among them, except for Emerillon and Wayampi of French Guiana. The Zo'é have been geographically separated from the other two groups at least since the 18th century, and the Wayampi who live on the Amapari in Brazil did not have contacts with the Emerillon. An acceptable explanation would be a common socio-historical context. Hypothesis 2b is interesting in that respect.

The following sections focus on the linguistic aspects of the borrowing process. Two caveats must be put forward. First, there is no description of many languages that could have been the source for this borrowing. Some may even remain unknown to Western scholars. Second, of all the languages for which a description is available, it is their modern version that has been described. There are some older descriptions only for Carib/Kali'na, some of which are suspected to actually describe a pidgin used with Frenchmen (Renault-Lescure 1984). This means there is little chance to find in the modern version of the languages some strong clues about the most likely source and the chronology of the borrowing process. It is however still worth studying this case of borrowing on the linguistic side. First, it could still tell us much regarding the socio-historical contexts of borrowing, for example whether it is a single process or three parallel ones. Second, it is interesting for the general discussion on borrowability to examine the linguistic modalities of a borrowing process when it is the unique case of a grammatical item being borrowed in the recipient languages. Third, as we will see, this morpheme has a very particular distribution in the two sets of languages, offering another challenge on how this change in distribution can be explained by the borrowing process. These points will be discussed in section 4, where the morpheme under study will be presented in the Tupi-Guarani languages (section 4) and compared with the suggested Proto-Carib reconstruction and the contemporary version of the local Cariban languages presented in section 3.

3. The collective markers of Cariban languages

In this section, the systems constituting likely sources for the borrowing process under study are presented. It describes the phonological form, the distribution, the morphological status and the semantics of Proto-Carib **komo* and its reflexes in modern Carib/ Kali'na, Wayana, Tiriyo, Apalaí and Kaxuyana (the five Cariban languages presently spoken closer to the three Tupi-Guarani languages concerned by the diffusion). Table 2 below summarizes the information detailed in this section. The Cariban-based pidgins were excluded from this presentation, since no mention of a reflex of **komo* (nor any kind of regular morphological number marker) has been found in the limited available literature.¹⁶

Languages	Proto-Carib	Carib/Kali'na	Wayana	Tiriyo	Apalaí	Kaxuyana
References	Gildea 1998	Hoff 1968/ Renault- Lescure p.c.	DaSilva Tavares 2005	Meira 1999	Koehn & Koehn 1986	Gildea 1998
Phonological forms	<i>*komo</i>	<i>koŋ</i> ¹⁷ / <i>kon</i>	<i>kom</i>	<i>kom</i>	<i>komo</i>	<i>kumu</i>
Distribution	Possessor	+	+	+	+	+
	Noun	–	animate and inanimate nouns	on some animate nouns	–	–
	Absolutive argument of verb	–	on some independent nominalized verbs	–	–	–
	Object of postposition	–	–	–	–	–
Morphological status	particle	suffix	suffix	suffix	particle	particle
Semantics	collective	plural /collective	collective	collective	collective	collective

Table 2: Summary on the collective marker **komo* in some Cariban languages

3.1. Phonological form

As presented in Table 2 above, all the Cariban languages we are interested in present a reflex of **komo*, without important phonological changes. It can be noted that Kaxuya-

¹⁶ Only in the male speech of Island Carib, apparently closely related to the Carib/Kali'na pidgin, the plural/collectivizer *-kul-gu* is attested as a productive morpheme (Taylor 1954).

¹⁷ In Hoff's transcription of Carib, the symbol *ŋ* stands for advanced velar nasal.

na uses the vowel /u/ rather than /o/, and that Carib/Kali'na realizes the nasal as an advanced velar rather than a bilabial nasal. The form *komo* as such is preserved only in Apalaí. It is a frequent variant in Wayana, a rare one in Tiriyo and is even rarer in Carib/Kali'na (Meira, p.c.).

3.2. Distribution

**komo* is one among various particles (most with phonological and suppletive allomorphs) used to express collective number. They are distributed in a complex manner that is moreover not systematically consistent within the family. **komo* is used only in some part of the distributional range of the collective particles. In the specific positions where it may be used, it is essentially an allomorph among others of the other collective particles, while on possessed nouns it is the only collective particle attested.

- i) In most languages of the family, only animate nouns have a collective form. The suffixes *-*yamo*, *-*tomo* or *-*nomo* are used. They are idiosyncratic to individual nouns.

However, in Wayana (DaSilva Tavares 2005: 151–152), the use of **komo* has extended to this context. Even though most underived nouns take *-tom(o)* (20), most pronouns take *-kom(o)* as in (19), and some underived nouns as well (18). The marking of the collective category is extended to inanimate nouns (20).

Wayana (DaSilva Tavares 2005: 151–152)

- (18) *eluwa-kom*

man-COL
'the men'

- (19) *sin-kom*

DEM.INANIM.PROX-COL
'these'

- (20) *pakolo-tom*

house-COL
'the houses'

In the Carib/Kali'na language described by Hoff (1968: 227–229), the reflex of **komo* also displays a more important use on nominals. *-koŋ* ~ *-goŋ* is used as the major allomorph of the plural suffix on underived nouns (21), along a few rarely used other allomorphs.¹⁸ It is also used on inanimate demonstratives (22) (Hoff 1968: 277). Finally, *-koŋ* is found on nominalized forms (30). Since derivation of adverbs, verbs, postpositions and numerals is crucial in that language, *-koŋ* results as being the most important collective marker (Renault-Lescure, p.c.).

¹⁸ For instance, *-yan* or *-san* are restricted to kinship terms and 'man' and 'woman' (Renault-Lescure, p.c.).

Carib, Cariban (Hoff 1968: 228, 277)

(21) *wo:to-koŋ*
fish-PL
'fishes'

(22) *e:ro-koŋ*
DEM-PL
'these (things)'

- ii) The number of the possessor is indicated by a collective particle following the possessed noun. Its form is extremely consistently **komo* across the family.

This exact configuration is found in Kaxuyana (see (23) below), Tiriyo (Meira 1999: 140) and Apalaí (Koehn & Koehn 1986: 88–89).

Kaxuyana, Cariban (Gildea 1998: 117)

- (23a) *minayi* 'my sister'
 (23b) *minayi-tomo* 'my sisters'
 (23c) *ki-mnayi kumu* 'the sister of all of us'
 (23d) *ki-mnayi-nomo* 'the sisters of we two'
 (23e) *ki-mnayi-nomo kumu* 'the sisters of all of us'

The Carib/Kali'na language does not conform to the above generalization, in that number marking of the head noun and of the possessor is not as much distinct (Hoff 1968: 227–229). As in Proto-Carib, the suffix *-koŋ* can pluralize the possessor (24) on a possessed noun, with the exception of a first person possessor as exemplified in (25)¹⁹ (Hoff 1968: 245). But since the distribution of *-koŋ* has extended to head nouns, *-koŋ* on possessed nouns can therefore ambiguously refer to the possessor or the possessee as in (24).

Carib, Cariban (Hoff 1968: 245–246)

- (24) *aye:mari:-koŋ*
2.path-PL
'your (PL) path'/'your (SG or PL) paths'
 (25) *ye:mari:-koŋ*
1.path-PL
'my paths', *our path

- iii) Some languages of the family developed innovative verbal systems, due to the reanalysis of nominalized forms as finite verbs. As a consequence, the collective possessor morpheme that was used to mark the possessor of nominalized forms (their absolutive) is now used to mark the absolutive argument of finite verbs. In (26), the collective suffix modifies the unique argument of an intransitive verb, while in (27) it modifies the P argument of a transitive verb.

¹⁹ This restriction applies to verbs, as well as to nouns and postpositions.

Macushi, Cariban (Gildea 1998: 165, repeated from Abbott 1991: 116)

- (26) *aminke a-tti-'pi-kon*
 far 2-go-PAST-COL
 'You all went far.'

Kuikúro, Cariban (Gildea 1998: 167)

- (27) *tujá-héke léfa Ø-ate-lə-ko léha*
 water-ERG ASP 3-encircle-PUNCTUAL-PL ASP
 'The water encircled them.'

In the area under scrutiny, Carib/Kali'na uses some nominalized verbs in independent clauses with a TAM value (28).²⁰ On these verbs, *-koŋ* marks the plural of S or P (Renault-Lescure p.c.). In Wayana (DaSilva Tavares 2005: 210), Tiriyo (Meira 1999: 293–296), Kaxuyana and Apalaí (Meira, p.c.), the number marker used on finite verbs is never **-komo*.

Carib/Kali'na, Cariban (Renault-Lescure, p.c.)

- (28) *tamelo loten politique ta i-womitoki-li-kon*
 diversely only politics in invest-NOMN-NB
 'They invested in politics only in diverse ways.'

A last use not discussed in Gildea's work is the presence of a collective marker on a postpositional phrase. De Goeje (1990: 42) states that **komo* is used on postpositions carrying a personal prefix in several Cariban languages, all of which (Pemon, Kapon, Makushi, etc.) are spoken outside of the Guianan area (Meira, p.c.). In Carib/Kali'na (Hoff 1968: 253–254), in Wayana (DaSilva Tavares 2005: 295) and in Tiriyo (Meira 1999: 372) (29), the collective suffix found on postpositions is *-(i)ne*, *-he* and *:-ne* respectively. In Carib/Kali'na, *-koŋ* can occur on postpositions only when they are nominalized (30).

Tiriyo, Cariban (Meira 1999: 374)

- (29) *i-pëe-ne*
 3-ABL-NB
 'from them'

Carib/Kali'na, Cariban (Hoff 1968: 253–254)

- (30) *po:-no-koŋ*
 at-NOMN-NB
 'inhabitants of'

²⁰ This is the case with the nominalizers *-li* (sometimes expressing visibility) and *-po* (expressing past or completed) (Renault-Lescure, p.c.).

3.3. Morphological status

In terms of morphological status, Gildea reconstructs the collective marker on nouns as a suffix and the one for possessor (or argument of a reanalyzed nominalized clause) as a particle. He specifies that the modern reflexes of the collective possessor particle **komo* are often described as suffixes rather than particles. As a suffix, they cannot receive independent stress, but can condition syllable reduction in preceding morphemes, and consonant assimilation phenomena. The reflex of **komo* is described as a suffix in Carib/Kali'na, Wayana and Tiriyo, and as a particle in Apalaí and Kaxuyana.

3.4. Semantics

At the semantic level, it is important to note that:

plurality is not a grammatical category in most Cariban languages and is not obligatorily marked in those languages that ostensibly do mark the plural. Rather than distinguishing between *one* (singular) and *more than one* (plural), the more common distinction is between *all* (collective) and *fewer than all* (non-collective). Thus, a noun that is not marked for number might be singular or plural, but it is not all of a perceived group; a noun that is marked for collective number will certainly be plural (usually more than two or three), and in the mind of the speaker, the referents taken together constitute the entirety of some group. (Gildea 1998: 116–117)

Meira indicates that, despite the Caribanist tradition of using the 'collective' terminology, 'totalitative' would be more appropriate, since the semantics of number opposes 'all' to 'less than all' (Meira 1999: 138). However, this author chooses to conform to the use of the term 'collective'.²¹ This traditional gloss may also hide the fact that some uses of number markers in Cariban languages look closer to those of a 'definite plural' (Meira, p.c.).²²

4. The borrowed form in Emerillon, Wayampi and Zo'é (Tupi-Guarani) and hypotheses on the process of borrowing

In this section, the phonological form, the distribution, the morphological status and the semantics of the borrowed item is examined in the three recipient languages: Emerillon,

²¹ In an important but far earlier study, Hoff (1968: 47) uses the term 'plural'. In the Apalaí description, the so-called 'plural' markers are said to be 'group markers' and are glossed 'collective' (Koehn & Koehn 1986).

²² Renault-Lescure's analysis is more precise: she says it can express delimitation of a sub-group within a group, and diversity within a group (Renault-Lescure, p.c.).

Wayampi and Zo'é. The information available is summarized in Table 3 below. Much information is unfortunately lacking for Zo'é.

Languages		Emerillon	Wayampi	Zo'é
References		Rose 2008, 2011, fieldnotes	Grenand 1980, 1989; Jensen 1993; Copin p.c.	Cabral 2000, p.c.
Phonological forms		<i>-kom</i>	<i>-kũ/kõ</i>	<i>kã</i>
Distribution	Possessor	+	–	?
	Nouns	+	+	+
	Absolutive argument of verb	+	+	?
	Object of post-position	+	+	?
	Coordination	+	–	–
Morphological status		clitic	suffix/clitic/noun?	particle
Semantics		plural	plural (Grenand)/ collective (Copin p.c.)	collective

Table 3: Summary on the plural/collective marker of Emerillon, Wayampi and Zo'é

At each level, changes involved in the process of borrowing will be discussed.

4.1. Phonological form

The form of the marker is *kom* in Emerillon, *kõ* in the Amapari dialect of Wayampi spoken in Brazil and *kũ* in the Wayampi dialect spoken in French Guiana, *kã* in Zo'é. Lexical nasality is an exception among the Emerillon affixes, which usually undergo (de)nasalization depending on the nasality of the root they attach to. This fact constitutes additional evidence for the borrowing process.

Comparing the phonetic realization of the segments with modern versions of Cariban languages may not make much sense if the borrowing is really old. However, the syllabic constituency is interesting to examine.²³ In Cariban languages, it is either CVCV or CVC. In Tupi-Guarani languages, it is either CVC or CV. It seems likely that the morpheme was monosyllabic (CVC) in the source language(s). The internal phonological evolution of Tupi-Guarani languages does not involve syllable reduction, nor final vowel deletion. It would be unlikely that a CVCV form would have been reduced to CVC or CV. This almost certainly excludes Kaxuyana as the source language. Loss of final consonants is nevertheless a regular diachronic evolution in Wayampi, and would fit a CVC form in the source language. Whether this occurred before or after the loss of

²³ While it is imaginable that Carib/Kali'na *-koŋ* had once been realized *-kom*, it is highly improbable that Kaxuyana *kumu* was once realized *ku* or *kum*, the Proto-Cariban reconstruction being **komo*.

final consonants, the loanword had to adapt to the current phonological system of Wayampi, maintaining the nasality on the vowel only. This does not help to date the borrowing. In Zo'é, the absence of the nasal consonant cannot be explained by the loss of final consonants, since only final **b* was systematically lost in that language.

The variation of the vowel quality in Wayampi (/u/ in the dialects of French Guiana, /o/ in Brazil) is unexplained. The mutation of the vowel in Zo'é is less of a problem, since [ã] and [õ] are in free variation in Zo'é nasal roots (Cabral 2000: 584).

4.2. Distribution

This section will essentially focus on Emerillon, because very little information on the distribution of Wayampi and Zo'é forms is given in the literature.

First, as mentioned above, *-kom* maintained in Emerillon the very special distribution as a modifier of the possessor on a possessed noun (5). Second, the *-kom* marker can also be found on an unpossessed noun (31), or any other kind of nominal phrases (personal pronoun (32), demonstrative pronoun (33), nominalization (34), relative clause (35)).

Emerillon, Tupi-Guarani (Rose 2011: 112)

- (31) *wãĩwĩ-kom-enam* *o-wikirog* *o-wikirog*.
 woman-PL-TOP.SWITCH 3.I-empty 3.I-empty
 'The women gutted the fish.'
- (32) *k^{wi}-ena* *orone-kom-uwe* *pe-r-upi* *oro-ho-tar-uwe*.
 once-DUB PRO1EXCL-PL-too path-RELN-on 1EXCL.I-go-FUT-too
 'One day maybe, we too, we will go on this path.'
- (33) *wijn-a-kom-ãhã* *o-wir* *o-ho*.
 DEM-REF-PL-only 3.I-go.up 3.I-go
 'Only they go up.'
- (34) *o-me?eŋ* *karakuri* [*iar-a-ma?ẽ-har*]-*a-kom-a-pe*.
 3.I-give money canoe-REF-look-NOMN-REF-PL-REF-to
 'He pays those who watch the canoes.'
- (35) ... [*o-?i-wi-wi* *o-?ur-a-ma?ẽ]-kom* *amõ*...
 3COREF-river-ABL-RED 3.I-come-REF-REL-PL other
 '...others, who come from their river...'

As a consequence, *-kom* on a possessed noun can pluralize the possessee, the possessor or both (36).

Emerillon, Tupi-Guarani (Rose 2011: 114)

- (36) *i-kiwir-a-kom*
 3.II-brother-REF-PL
 a) 'his brothers'; b) 'their brother'; c) 'their brothers'

Third, *-kom* is used on postpositions to modify their 3rd person object (37).

- (37) *o-wig* *Ø-ehe-kom*.
 3.I-arrive 3.II-to-PL
 ‘He arrives next to them.’

Fourth, *-kom* is sometimes found on a transitive verb to modify a 3rd person object (38) and more rarely on an intransitive verb to modify its subject, as in (39).

- (38) *wij-a-te* *o-zika* *kom*.
 DEM-REF-FOC 3.I-kill PL
 ‘It is him who killed them.’

- (39) *aʔe-kom* *nōde-kom* *se-iko-kom*.
 DEM-PL PRO1INCL-PL 1INCL.I-be-PL
 ‘These people (who survived the wars), we are them.’

A final and marginal use of *-kom* is its function as a noun phrase coordinator, when adjoined to the last element of a string of nouns, as illustrated in (40) and (41).

- (40) *uruwi,* *tareʔir,* *pakupitaŋ-a-kom* *ō-wem*.
 fish.sp aïmara pacu-REF-PL 3.I-go.out
 ‘The uluwi, the aïmaras and the pacus came out.’

- (41) *pureru* *zawar-a-kom-a-r-ehe* *o-maʔẽ-maʔẽ* *kiʔi*.
 toad dog-REF-PL-REF-RELN-POSTP 3.I-RED-watch then
 ‘Then he watches the toad and the dog. (‘The toad then watches the dogs’, with the normal plural interpretation, would be possible had the story several dogs in it).’

Descriptions of Wayampi contain little information about the distribution of *kō/kũ*. It occurs on nouns, as in (42) or (43), on pronouns (apparently in variation with *kupa*), and also on relativized clauses. Possessor pluralization is not one of its function (Copin, p.c.), so that the equivalent of (36) cannot have the b) interpretation.

Wayampi, Tupi-Guarani (Grenand 1980: 58, 100)

- (42) *taiwe-kũ*
 ancestor.spirit-PL
 ‘the spirit of the ancestors’
- (43) *e-memi-kũ*
 1SG-child-PL
 ‘my children’
- (44) *ɔlɔ-kũ*
 PRO.1INCL-PL
 ‘us all’

- (45) *o-kaʔu-mãʔẽ-kũ*²⁴
 3.I-drink.manioc.beer-REL-PL
 ‘those who drink manioc beer’

The Wayampi morpheme *kõ/kũ* is also presented as a subject pluralizer in Grenand’s work, with the same position and function as *kupa* (46) (Grenand 1980: 68). In Jensen’s 1993 paper on negation in Wayampi, *kupa* is found in several examples to pluralize the subject. Copin (p.c.) notes that *-kõ* occurs on transitive verbs to pluralize a third person object when it is not expressed with an overt noun phrase (47). It thus seems that *kõ/kũ* refers to the absolutive argument as in Carib/Kali’na.

Wayampi, Tupi-Guarani (Grenand 1980: 68; Jensen 1993: 354; Copin, p.c.)

- (46) *ɔ-wãʔẽ-ma-kupa*
 3.I-arrive-TAM-PL.S
 ‘They arrived.’
- (47) *o-su'u-kõ*
 3.I-bite-PL
 ‘He bit them.’

Jensen’s data also show that *kõ* can be used on a postposition and then pluralizes its object represented by a person index. Lastly, nothing is mentioned about a coordinating function such as found in Emerillon.

Wayampi, Tupi-Guarani (Jensen 1993: 354)

- (48) *n-o-kyje-’ãi* *apã* *kõ* *i-wyi* *kõ* *amẽ* *mijã*.
 NEG-3.I-fear-FUT.NEG enemy PL 3.II-of PL then UNR
 ‘Their enemies wouldn’t have been afraid of them.’

Publications on Zo’é essentially focus for the time being on phonology. However, it can at least be said that *kã* is found on unpossessed nouns (49) and on pronouns (50), but is not used as a coordinator as in Emerillon (Cabral, p.c.).

Zo’é, Tupi-Guarani (Cabral, p.c.)

- (49) *kubi’é* *kã*
 man COL
 ‘all the men’
- (50) *nonẽ* *kã*
 PRO.1EXCL COL
 ‘us all’

Table 4 summarizes the distribution of the reflexes of **komo* in the three recipient languages as compared to the supposed distribution of *komo* in Proto-Carib and in Carib/Kali’na, the modern Cariban language with the largest distribution for *komo*.

²⁴ Following a comment by Copin (p.c.), the *o-* prefix was added to make the utterance grammatical.

Distribution	Proto-Carib	Carib/Kali'na	Emerillon	Wayampi	Zo'é
Possessor	+	+	+	–	?
Noun	– ²⁵	+	+	+	+
Absolutive argument of verb	–	±	+	+	?
Object of postposition	– ²⁶	–	+	+	?
Coordination	–	–	+	–	–

Table 4: Comparative distribution of *komo* in Cariban and Tupi-Guarani languages

The syntactic distribution of the reflexes of **komo* in Emerillon and Wayampi (given that little information is available on Zo'é) is wider than the distribution of **komo* in Proto-Carib or in any daughter language. Wayampi yet also shows some loss in the distribution.

The typical function of the collective particle **komo* in Proto-Carib as plural of the possessor of a possessed noun has been maintained in Emerillon but lost in Wayampi. Both languages use **komo* in contexts where particles other than *komo* are normally used in most Cariban languages (on an unpossessed NP, on verbs, on postpositions), but where it is sometimes found in Carib/Kali'na and Wayana. Moreover, Emerillon innovated with the use of *kom* as a coordinator. No comparable use is attested in Cariban languages.

Three hypotheses can be put forward.

Hypothesis A: The languages may have borrowed **komo* with its reconstructed distribution in Proto-Carib (in the case of Emerillon) and/or as a noun pluralizer (on the model of its use on nouns and nominalizations, a position where it happens very frequently in the languages that allow it). Then the recipient languages later extended its use to verbal and postpositional phrases (and coordinator function). For instance, the use on postpositions in Emerillon could have been innovated in analogy with the use of *komo* on a possessed noun. Postpositional phrases and genitive phrases show the same order of elements and the same indexation system with a Set II index for possessor or object of postposition²⁷ (compare (51) and (52)), so that the analogy is easy.

Emerillon, Tupi-Guarani

- (51) *zawar, pureru, i-dʒar-a-kom o-k^wa-ŋ.*
 dog toad 3.II-owner-REF-PL 3.I-go-CONT
 'The dog, the toad and their owner go away.'
- (52) *o-wig Ø-ehe-kom.*
 3.I-arrive 3.II-to-PL
 'He arrives next to them.'

²⁵ In this context, another collective particle is used.

²⁶ In this context, another collective particle is used.

²⁷ The Emerillon postpositions derive directly from Proto-Tupi-Guarani, with no change such as borrowing or grammaticalization. It is therefore excluded that these postpositions derive from possessed nouns that would have grammaticalized into postpositions and so would still allow the possessed noun collective marker.

Another example of an expected extension concerns the absence of restriction of **komo* to animate nouns, as was the case in Proto-Carib but less clearly so in the daughter languages. The extension to animate entities of a category previously restricted to animate nouns is common in the languages of the world. This extension could therefore have been anterior to the borrowing in the source languages or posterior to it in the recipient languages. However, if **komo* had been borrowed as a possessor and/or noun pluralizer only, the specific pattern of number marking on Emerillon verbs would be left unexplained: why is the marker referring to the absolutive argument only?

Hypothesis B: Proto-Carib and the plausible source languages have been described as using other forms of the collective marker than **komo* on verbs and postpositions, and generally on nouns as well. In Emerillon and Wayampi, **komo* covers the overall distribution of the various particles in Cariban languages. The recipient languages could have picked up a single form among the collective markers (i.e. **komo*) and extended it as the unique marker for all the functions of collective markers in the source language(s). This would mean, with Matras & Sakel's terminology (Matras & Sakel 2007), that both MATTER and PATTERN would have been borrowed, but not with a perfect match of the form/meaning set: the category of number has been borrowed in its full extension, but only one marker of the category was borrowed. This type of borrowing process would probably require a very high degree of bilingualism.

Hypothesis C: A third hypothesis is that the morpheme could have been borrowed into Tupi-Guarani languages from a pidgin, which would have already simplified the Cariban system by using a single particle (**komo*) to cover the whole range of contexts normally covered by the various particles. The pidgin would have unified the very complex suppletive/allomorphic system of Cariban collective particles. This is plausible. Taylor & Hoff (1980) have indeed shown the reanalysis of several morphemes in the Carib-based pidgin. However, as was said earlier, none of the known pidgins of the area under study have a marker for plural, and none of them have been in intense contact with the recipient languages.

4.3. Morphological status

The morphological status of the cognate form *kom ~ kũ ~ kã* is not a simple issue.²⁸

The Emerillon form *-kom* generally attaches to the right of nouns and postpositions. It can affect the preceding root. On a non-predicative nominal element, *-kom* triggers the presence of the referential suffix *-a*, a suffix that is overt only on a consonant-final root when some other element (a suffix, a discourse particle, a postposition or the possessed noun in a genitive phrase ...) follows within the same constituent, as in (36) and

²⁸ It is often difficult, in the languages of the world, to decide on clitic vs. suffix status for certain forms. The difference in the descriptions of the different languages (Cariban or Tupi-Guarani) may simply reflect different analytical preferences of the respective linguists.

(51).²⁹ I consider *-kom* to be a clitic for two reasons. First, it attaches to various parts of speech. Second, in some examples, it is not so obviously bound. When *kom* follows a verb (to modify its object), it is a separate stressed word, as in (38). When used as a coordinator, it is also stressed. However, since it follows a nominal root, the referential *-a* suffix appears in between as in (40) and (41). Queixalós (2001) suggests that the Emerillon *kom* should be analyzed as a noun with a collective meaning such as ‘group, set’. When following a noun, *kom* would be the head and the other noun its genitive modifier. Thus *wãĩwĩ-kom* ‘women’ should be analyzed literally as ‘the group of women’. This proposal agrees with the prosodic autonomy of *kom* after a verb and the fact that it can be simultaneously preceded and followed by the *-a* referential suffix such as in (34) and (41). However, this analysis does not explain first why the supposed nominal *kom* cannot take modifiers other than nouns, second why this noun can have scope on the possessor of the preceding noun rather than on this noun itself, and third why, when it stands by itself, as a supposed subject or object of a verb, it always follows the verb while the normal position of the core arguments is preverbal.

The Zo'é form *kã* is considered a particle with a nouny behavior by Cabral (p.c.). It is always written as a separate word and does not seem to show any phonetic interaction with the preceding word.

The Wayampi form *-kũ* is written as a bound morpheme in Grenand's work. It is considered to be a collective noun in Copin's analysis (Copin, p.c.), following Queixalós (2001). For the Amapari dialect, *kõ* is always written as a separate word by Jensen.

To summarize, the morphological status of the borrowed element in the recipient languages (as within the source family) varies from a free stressed word to a suffix. In all cases, it is situated at the right margin of the phrase containing the element it categorizes as plural or collective, illustrating one of the few so far widely accepted generalizations on borrowings: “It is true that the borrowing of function words tends to be accompanied by an adoption of the rules for the linear ordering of those words.” (Moravcsik 1978).

4.4. Semantics

The semantics of the borrowed item are now presented as analyzed in the three recipient languages.

²⁹ Following Queixalós' diachronic hypotheses, the suffix *-a* was previously found on any referential term. In Emerillon, it got partially lost depending on the phonological and syntactic environment (Queixalós 2001, 2006). It is still present on referential non-predicative element, but only when these are consonant final and followed by some other elements (like postpositions, suffixes, particles, *-kom* or the possessed noun in a genitive phrase). This gives us the diachronic indication that the borrowing of the *-kom* suffix probably occurred before the decay of the *-a* (or else the *-a* suffix would have been inserted before *-kom* by analogy). In Wayampi, the *-a* suffix got entirely lost since all final consonant were lost, so that *e-memi-kũ* (1SG-child-PL) ‘my children’ in Wayampi corresponds to *e-mëbír-a-kom* in Emerillon.

For Zo'é, Cabral (p.c.) states that *kã* is a collective marker rather than a plural marker (53b), just as in the source languages. It occurs only on referential nouns. Thus a bare noun can be singular or plural (53a). The use of the marker *kã* on first person plural pronouns is said to convey an associative meaning (50).

Zo'é, Tupi-Guarani (Cabral, p.c., her translation)

(53a) *kubi'é*

man

'one man/several men'

(53b) *kubi'é* ***kã***

man **COL**

'all the men'

For Wayampi, Grenand (1989: 23) defines *-kũ* as a plural marker, and Jensen (1989) glosses *kõ* as plural as well. Grenand specifies that the plural of nominal lexemes is generally unmarked (54). The plural marker is contrasted with some derivation morphemes (*-kwε*, *-ngε*, *-lε*)³⁰ that can be used to convey an associative meaning as in (55) (Grenand 1989: 23). The translation of some examples leaves open the possibility of a collective meaning induced by *-kũ* (56).³¹ Finally, as example (56) shows, *kũ* is not restricted to animate entities as in most of the Cariban languages (Wayana and Carib/Kali'na being excepted).

Wayampi, Tupi-Guarani (Grenand 1989: 23, 254)

(54) *ε-po*

1SG-hand

'my hand(s)'

(55) *kiki-lε*

young-PL.SOC

'seedlings'

(56) *ia-kũ* ɔ-ɔ

canoe-PL 3.I-go

'All the canoes left.'

In Emerillon, *-kom* is quasi exclusively found on referential nouns, generally animate ones (see for instance examples (31) to (36)). Animate nouns that are semantically plural but unmarked for number are either non-referential (with generic reference) as in (57) or referential, but some with generic or collective reference (without salient individuals) as in (58) and (59). As a consequence, *-kom* seems to indicate plural rather than collec-

³⁰ According to Copin (p.c.), the mid vowel *-e* alone would be a plural suffix, limited to a few nouns. It got lexicalized with them, as shown by the maintenance of the root-final consonant as in *ja'yl-e* 'children'.

³¹ According to Copin (p.c.), 'all' is not implied in this sentence.

tive in this language. Inanimate nouns can be marked for plural when referential (60) (41), but can be unmarked if non identifiable (61).

- (57) *ʔi-b d-o-zaug-i teko wāĩwĩ.*
 river-in NEG-3.I-bathe-NEG Emerillon woman
 ‘(When they have their periods) the Emerillon women do not bathe in the river.’
- (58) *apam-a-ně nōde-apisi-taně.*
 foreigner-REF-CONTR 1INCL.II-massacre-DESID
 ‘The foreigners wanted to massacre us.’
- (59) *sə sə pira-maʔě o-bo-sale.*
 big big fish-REL 3.I-CAUS-be.salted
 ‘They salted the very big fish.’
- (60) *tʃalupug-a-kom-a-we masakala-kom-a-we i-puli.*
 stewpot-REF-PL-REF-too chicken-PL-REF-too 3.II-next.to
 ‘And next to it, there are stewpots and chickens.’
- (61) *ʔn-a-wi-te arakapusa-uhu o-mōdur-oŋ³² bal.*
 DEM-REF-ABL-FOC gun-big 3-send-PL.S bullet
 ‘From there, guns were firing bullets.’

The use of *kom* on a verb to pluralize the object is restricted to animate participants, as shown in the possible interpretations of (62). Otherwise, reduplication is used, as in (16).

- (62) *za-idu-katu-kom.*
 INDET.I-hear-well-PL
 ‘People hear them well (other people, wild pigs, *engines, *chain saws).’

To conclude on the semantics of the borrowed form/meaning set in Emerillon, Wayampi and Zo'é, first of all there is no evidence in Emerillon that the collective meaning was maintained. A minor meaning shift occurred, from collective to plural. Second, while collective marking was restricted to animate nouns in Proto-Carib (and usually marked with particles other than **komo*), in at least Emerillon and Wayampi, number marking is found on inanimate nouns as well, as is the case in Wayana and Carib/Kali'na. However, the restriction to inanimate referents was maintained when *-kom* is used on Emerillon verbs to pluralize their objects.

Summary of section 4: When comparing the realization of the cognate form in the three Tupi-Guarani languages (Table 4) and the local Cariban languages and their suggested Proto-system (Table 3), it varies above all at the distributional level (cf. Table 5). The variability of the phonological form, the semantics and the morphological status of the morpheme is surprising neither within the Cariban family or the Tupi-Guarani subgroup nor between the two families. They are explainable by common internal changes. It is yet not clear whether the three languages underwent independent borrowing proc-

³² As explained in 1.1, this marker *-oŋ* indicates plurality of the subject.

esses or if one (or two) of them were intermediaries in the contact situation. The variability between the three languages could be explained by restriction and extension of use posterior to the borrowing.

5. Theoretical discussion on the borrowing process

5.1. An unlikely borrowing

Since there is no evidence to compel us to believe that the morpheme under study was first borrowed attached to a host element, it constitutes the unique borrowing of a grammatical and somewhat bound morpheme in the three languages under study. This leads us to two central issues in the field of contact linguistics: first, language-internal factors for the likelihood of transfer of such a morpheme; second, the necessary socio-linguistic factors to such a type of borrowing.

On the one hand, the likelihood of transfer is usually associated with certain internal characteristics of the source and recipient languages. Firstly, the likelihood of transfer is correlated with morphological autonomy, phonological robustness and invariance of a morpheme (Weinreich 1953), or identifiability of form-meaning sets (Field 2002). In this respect, the reconstructed **komo* is not an ideal morpheme for transfer, since the collective category of Cariban languages does not display a one-to-one mapping of meaning and form. Cariban collective particles are numerous, and their complex distribution depends on the phonological, syntactic and sometimes idiosyncratic environment. Secondly, bound morphemes are often said to be less borrowable than free items (Winford 2005: 386). Curnow (2001: 429) asserts that there are relatively few examples of bound grammatical forms and meanings transferred from one language to another (classifiers, case affixes, plural suffixes, verbal aspect markers). In the case under study, Cariban collective particles are somewhat bound. Nevertheless, **komo* was, at least in Proto-Carib, an independent particle. It may be the case that, out of the set of collective particle, the most autonomous one has been borrowed. Thirdly, it is said that the more frequent an item, the more likely it is to be borrowed (i.e. Field 2002: 5). In Cariban languages, the frequency of the collective particle **komo* is obviously restricted due to its allomorphic distribution, often based on idiosyncratic grounds. However, it is still very frequent due to its use on nominalizations, a crucial structure of Cariban languages. Moreover, collective particles as a whole are fairly frequent in Cariban connected discourse. If a pidgin had been the intermediary, the major obstacle to the borrowing of such a morpheme would be cleared, since it is quite expected for the form of a number marker in a pidgin to be invariant and frequent, with a one-to-one mapping of form and function. Still it remains to explain why a pidgin would have picked up this specific form in the first place.

On the other hand, the literature on language contact suggests that certain types of borrowings are more likely to result in certain types of sociolinguistic situations. On

Thomason & Kaufman's scale of borrowings (1998), Stage 1 of contact (referring to 'casual contact') may lead to lexical borrowing of non-basic vocabulary, while structural borrowings start only with Stage 2 ('slightly more intense contact'). The situation of Emerillon is taken as a basis for the discussion, since the contact history of this language has already been partly studied beforehand (cf. Rose & Renault-Lescure 2008). Emerillon has been borrowing lexical items from various European languages, Creoles and other Native American languages. Other loans from Cariban languages (Carib/Kali'na and possibly Wayana) are nominal items, for instance words for 'knife', 'gourd', 'beads'. The degree of contact of Emerillon with Carib/Kali'na and Wayana is around Stage 1 (labeled 'casual contact'),³³ since only lexical borrowing of non-basic vocabulary is attested. The present contact of Emerillon with French and French Creole may be between Stage 1 and Stage 2 (labeled 'slightly more intense contact') since somewhat more structural borrowings is attested (subordinators and numerals for instance) and basic vocabulary is involved. Emerillon must have been, in the past, in a sociolinguistic situation with more intense contact, a situation that enabled the borrowing of a grammatical item. Unfortunately, information on older contact situations is scarce or inexistent. The Emerillon group was first contacted in 1766 (Hurault 1972: 179).

5.2. Possible factors facilitating this borrowing

This section provides factors that may have facilitated such a borrowing, for which neither the morphological characteristics nor the socio-historical type of contact were optimal.

First, in contrast with what was said in general about bound morphology in 5.1, the borrowing of a number marker is not so unlikely. Borrowed bound morphemes are attested (Harris & Campbell 1995: 134; Stolz 2008: 21–22, Gardani 2008). It is often alleged that derivational morphemes are more easily borrowed than inflectional morphemes (see for example Moravcsik 1978: 112, Thomason & Kaufman 1988: 74–75). Harris & Campbell (1995: 135–136) claim that this generalization can better be explained on semantic grounds : affixes with clear semantic content (most derivational affixes, but some inflections such as number) are more easily borrowed than semantically weak or redundant affixes (such as a verb agreement). It is also Matras' stance concerning plurality (Matras 2007: 43):

Plural markers can be said to occupy a position between derivation and inflection markers. On the one hand, they are potentially linked to the expression of plural agreement elsewhere in the sentence, and so they operate at the level of the sentence rather than just the word. On the other hand, they indicate clear semantic opposition to singulars at the word level. Morphological plural marking meets the

³³ However, in the western part of French Guiana, the Emerillon live in the same villages than the Wayana people. Many families are now bi-ethnic.

criteria for semantic transparency which is so often noted as a factor facilitating morphological borrowing.

Gardani's specific studies on the borrowing of inflectional morphemes (Gardani 2008) and plural more specifically (Gardani 2011) explain the ease of borrowing of number markers with the following assertion: plural is an inherent inflection³⁴ (Booij 2005), and as such, is more similar to derivation than contextual inflection. Hence it is easier to borrow. In conclusion, several authors converge in saying that within bound morphology, number markers are rather likely to be borrowed. An illustrative example is the borrowing of Spanish *-s* plural into Bolivian Quechua (Harris & Campbell 1995: 135). Gardani (2008, 2011) provides a fair number of other examples.

A second point that could have facilitated this borrowing is the lack of typological constraints. Cariban and Tupi-Guarani languages are quite comparable in terms of morphological typology, so that there is no constraint on the possibility of borrowing from one language to the other (Field 2002). Languages of both families tend to be agglutinating. They may therefore be considered to respect Field's Principles of System Compatibility, principles stating that the borrowing language's morphological typology constrains the possibility of borrowing from another language (Field 2002). Typological similarity is of course only facilitating the borrowing, not triggering it necessarily. Thus the Tupi-Guarani people Tapirape and Kamaiura have also been in contact with Cariban populations and the Kamaiuras still are (Pierre Grenand, p.c.), but yet their languages were not influenced as far as number marking is concerned (cf. section 2.2).

Third, as presented in the section 2.2, Tupi-Guarani languages do not share a number marker. The gap in the system of plural encoding and the semantic transparency of the plural/collective marker (Matras 2007: 44) must have strongly contributed to this unique borrowing process. Moreover, a strong pressure may have been exerted by the very dense number marking system of Cariban languages on the 'poor' number marking systems of Tupi-Guarani languages. The borrowing of **komo* therefore does not constitute a case of morphological renewal but a step in the acquisition of a new category (number) along the creation of other number markers via grammaticalization (see section 2.2).

To sum up, the developing but still tentative literature on contact linguistics could make the borrowing studied here a counterexample to the definition of likelihood of transfer and the expectations on the correlations of socio-historical contact situations and stage of borrowing. On the other hand, the occurrence itself of this borrowing is an interesting case for other aspects of this literature, such as recent discussions on likely morphological borrowings, the question of typological compatibility and the assumption that a gap in the system can be a trigger for borrowing.

³⁴ The use of inherent inflection is context-free, determined by what information the speaker wants to convey, while the use of contextual inflection is context-sensitive, determined by the syntactic context in which it occurs (Booij 2005: 103–104).

6. Conclusion

This paper discussed the modalities of a borrowing process that had been previously suggested, that of a number marker **komo* of Carib origin into three Tupi-Guarani languages: Wayampi, Emerillon and Zo'é. In these modern languages, it constitutes a unique morphological loan, among lexical loans.

Firstly, the contact history of Tupi and Cariban languages have been examined, showing that this loan must be rather recent and localized. Several hypotheses on the socio-historical context of the borrowing and the source languages have been proposed. It could have occurred before, or after, that the three recipient languages separated. The source languages could have been Cariban neighbors, remnants of Cariban groups that agglutinated to Tupi-Guarani groups, or a Cariban-based pidgin. Until more historical information is available, all these hypotheses remain valid.

Secondly, a careful look was given at the phonological, morphological, syntactic and semantic characteristics of this form both in some Cariban languages (and Proto-Carib) and in the Tupi-Guarani languages concerned by the borrowing. The recipient languages use **komo* in contexts where particles other than *komo* are normally used in most Cariban languages. Two alternative explanations were given for this. On the one hand, the collective marker may have been borrowed into Tupi-Guarani languages with its distribution in Proto-Carib, and its function later extended to verbal and postpositional phrases. On the other hand, the category of number may have been borrowed into Tupi-Guarani languages in all the contexts where it is present in Cariban languages, but only one of the particles was borrowed and is used in all contexts rather than just in its original distribution. This constitutes an interesting case for the general understanding of borrowing.

Third, this paper discussed how this loan contributes to the general discussion of contact linguistics, namely in terms of sociolinguistic factors as well as linguistic features enhancing the likelihood of transfer. It will hopefully contribute to the development of a comprehensive account of language contact.

Abbreviations

3.II Set II	third person index	INDET	indeterminate
ABL	ablative	INTENS	intensifier
ASP	aspect	LOC	locative
CAUS	causative	NB	number
CAUS.SOC	sociative causative	NEG	negation
COMPL	completive	NOMN	nominalization
CONT	continuous	PAS.REC	recent past
CONTR	contrastive	PL	plural
COL	collective	PL.S	plural of subject
COREF	coreferential	PL.SOC	sociative plural

DAT	dative	POSTP	postposition
DEM	demonstrative	PRO	pronoun
DESID	desiderative	PROX	proximate
DUB	dubitative	RED	reduplication
ERG	ergative	REF	referential
EXCL	exclusive	REL	relativizer
FOC	focus	RELN	relational marker
FUT	future	SG	singular
GER	gerund	TAM	tense, aspect, mood
GRUP	group	TOP.SWITCH	topic switch
INANIM	inanimate	TRANSL	translational
INCL	inclusive	UNR	unreal

References

- Abbott, Miriam (1991): Macushi, in: Derbyshire, Desmond & Pullum, Geoffrey (eds.), *Handbook of Amazonian languages*, Vol. 3. Berlin: Mouton de Gruyter, 23–120.
- de Anchieta, Joseph (1595): *Arte de Gramatica da Lingua mais Usada na Costa do Brasil*. Coimbra: Antonio Mariz.
- Bechara Sanchez, Fabio José (1998): A ocupação do interflúvio Erepecurú/Curuá (Pará). *XXII Encontro Anual da ANPOCS*. Caxambu, Mato Grosso.
- Biet, Antoine (1896): Les Galibi: Tableau véritable de leurs mœurs avec un vocabulaire de leur langue (1661), in: *Revue de Linguistique et de Philologie Comparée*, juillet 1896, octobre 1896, 220–248, 261–338.
- Booij, Geert (2005): *The grammar of words: an introduction to linguistic morphology*. Oxford, New York: Oxford University Press.
- Cabral, Ana Suely (2000): Fonologia da língua Jo'é, in: *Universa* 8(3), 571–596.
- Couchili, Ti'iwan; Maurel, Didier & Queixalós, Francesc (2002): Classes de lexèmes en émérillon, in: *Amerindia* 26/27, 173–208.
- Curnow, Timothy (2001): What language features can be 'borrowed'?, in: Aikhenvald, Alexandra & Dixon, R.M.W. (eds.), *Areal diffusion and genetic inheritance*. Oxford: Oxford University Press, 412–436.
- daCruz, Aline (2011): Fonologia e gramática do Nheengatú. A língua geral falada pelos povos Baré, Warekena e Baniwa. PhD dissertation, Vrije Universiteit, Amsterdam.
- Da Silva Tavares, Petronila (2005): *A grammar of Wayâna*. PhD Dissertation, Linguistics Department. Houston: Rice University.
- De Goeje, Claudius (1990 [1944, 1937]): *Etudes Linguistiques Caribes*. Vaduz/Lichtenstein: Sändig Reprints Verlag.
- Dietrich, Wolf (1990): More evidence for an internal classification of tupi-guarani languages, in: *Indiana* (Supplement 12).
- Duarte, Fábio (2007): Expressão da quantificação em Tenetehára, in: Rodrigues, Aryon & Cabral, Ana Suely (eds.), *Línguas e Culturas Tupí*, Vol. I. Campinas: Curt Nimuendajú, 333–340.
- Field, Fredric (2002): *Linguistic borrowing in bilingual contexts*. Amsterdam/Philadelphia: John Benjamins.
- Gardani, Francesco (2008): *Borrowing of inflectional morphemes in language contact*. Frankfurt: Peter Lang.

- Gardani, Francesco (forthcoming): Plural across inflection and derivation, fusion and agglutination, in: Johanson, Lars & Robbeets, Martine (eds.), *The origins of bound morphology*. Leiden: Brill.
- Gildea, Spike (1998): *On reconstructing grammar: comparative Cariban morphosyntax*. Oxford: Oxford University Press.
- González, Hebe (2005): *A grammar of Tapiete (Tupi-Guarani)*. PhD dissertation, University of Pittsburgh.
- Grenand, Françoise (1980): *La langue wayãpi (Guyane française). Phonologie et grammaire*. Paris: SELAF.
- Grenand, Françoise (1989): *Dictionnaire wayãpi-français*. Paris: Peeters/SELAF.
- Grenand, Pierre (1982): *Ainsi parlaient nos ancêtres. Essai d'ethnohistoire «wayãpi»*. Paris: ORSTOM.
- Harris, Alice & Campbell, Lyle (1995): *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Hoff, Berend (1968): *The Carib language*. The Hague: Martinus Nijhoff.
- Hornborg, Alf (2005): Ethnogenesis, regional integration, and ecology in prehistoric Amazonia: toward a system perspective, in: *Current Anthropology* 46(4), 589–620.
- Hurault, Jean-Marcel (1972): *Français et Indiens en Guyane*. Paris: Union Générale d'Éditions.
- Huttar, George (1982): A Creole-Amerindian Pidgin of Suriname, in: *Society for Caribbean Linguistics Occasional Paper* 15, 1–13.
- Jensen, Allen (1979): Comparação preliminar das línguas Emérrillon e Oiapí no seu desenvolvimento do Proto-Tupi-Guarani, in: *Arquivo Lingüístico* 135, *SIL*, 22.
- Jensen, Allen (1993): Wayampi, in: Kahrel, Peter & Van den Berg, René (eds.), *Typological studies in negation*. Amsterdam/Philadelphia: John Benjamins, 343–364.
- Jensen, Cheryl (1989): *O desenvolvimento histórico da língua Wayampi*. Campinas, Brasil: Editora da Unicamp.
- Jensen, Cheryl (1998): Comparative Tupí-Guaraní morpho-syntax, in: Derbyshire, Desmond & Pullum, Geoffrey (eds.), *Handbook of Amazonian languages*, Vol. IV. Berlin: Mouton de Gruyter.
- Jensen, Cheryl (1999): Tupí-Guaraní, in: Dixon, R.M.W. & Aikhenvald, Alexandra (eds.), *The Amazonian languages*, Vol. Chapitre 5. Cambridge: Cambridge University Press.
- Kakumasu, James (1986): Urubu-Kaapor, in: Derbyshire, Desmond & Pullum, Geoffrey (eds.), *Handbook of Amazonian languages*, Vol. I. Berlin: Mouton de Gruyter, 326–403.
- Koehn, Edward & Koehn, Sally (1986): Apalaí, in: Derbyshire, Desmond & Pullum, Geoffrey (eds.), *Handbook of Amazonian languages*, Vol. I. Berlin: Mouton de Gruyter, 33–127.
- Magalhães, Marina (2007): Sobre a morfologia e a sintaxe da língua Guajá (família Tupí-Guaraní). PhD Dissertation. *Departamento de Lingüística, Português e Línguas Clássicas*. Universidade de Brasília.
- Matras, Yaron (2007): The borrowability of structural categories, in: Matras, Yaron & Sakel, Jeanette (eds.), *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter, 31–73.
- Matras, Yaron & Sakel, Jeanette (2007): Investigating the mechanisms of pattern replication in language convergence, in: *Studies in Language* 31(4), 829–865.
- Maurel, Didier (2000): Taila. 1596–1789: deux siècles de relations inter-communautaires teko-kaliña, Master thesis, Department of ethnology, Université M.Bloch, Strasbourg II.
- Mazières, Stéphane; Guitard, Evelyne; Sevin, André; Joly, Nicole; Dugoujon, Jean-Michel; Salzano, Francisco; Larrouy, Georges & Crubézy, Eric (2006): Structure génétique et histoire biologique de trois populations amérindiennes de Guyane française, in: *Antropo* 11, 51–59.
- Mazières, Stéphane; Sevin, André; Bonnet, Françoise; Crubézy, Eric; Salzano, Francisco & Larrouy, Georges (2007): Genetic studies in French Guiana populations: a synthesis, in: *American Journal of Physical Anthropology* 132(2), 292–300.
- Meira, Sergio (1999): *A grammar of Tiriýó*. Houston: Rice University.

- Métraux, Alfred (1927): Migrations historiques des Tupi-Guarani, in: *Journal de la Société des Américanistes de Paris* 19(1), 1–45.
- Moravcsik, Edith (1978): Language contact, in: Greenberg, Joseph (ed.), *Universals of human language*, Vol. 1: Method & Theory. Stanford: Stanford University Press, 99–122.
- Praça, Walkiria (2007): *Morfossintaxe da língua Tapirapé*. Universidade de Brasília.
- Queixalós, Francesc (2001): Le suffixe référentiel en émerillon, in: Queixalós, Francesc (ed.), *Noms et des verbes en tupi-guarani: état de la question*. München: LINCOM Europa, 115–132.
- Queixalós, Francesc (2006): The primacy and fate of predicativity in Tupi-Guarani, in: Lois, Ximena & Vapnarsky, Valentina (eds.), *Lexical categories and root classes in Amerindian languages*. Bern: Peter Lang, 249–287.
- Queixalós, Francesc & Renault-Lescure, Odile (eds.) (2000): *As línguas amazônicas hoje*. São Paulo: IRD, ISA & MPEG.
- Renault-Lescure, Odile (1984): A propos des premières descriptions d'une langue caribe, le galibi, in: *Amerindia* numéro spécial 6, 183–208.
- Rodrigues, Aryon (1985): Evidence for Tupi-Carib relationships, in: Klein, Harriet E. & Stark, Louisa (eds.), *South American Indian languages. Retrospect and prospect*. Austin: University of Texas Press, 371–404.
- Rodrigues, Aryon (2000): Hipótese sobre as migrações dos três subconjuntos meridionais da família tupí-guaraní, in: *II Congresso Nacional da Abralín e XIV Instituto Lingüístico*, 1596–1605.
- Rodrigues, Aryon & Cabral, Ana Suely (2002): Revendo a classificação interna da família Tupí-Guaraní, in: Cabral, Ana Suely & Rodrigues, Aryon (eds.), *Línguas Indígenas Brasileiras. Fonologia, Gramática e História. Atas do I Encontro Internacional do GTLI da ANPOLL*. Belem: Editora Universitária U.F.P.A., 327–337.
- Rose, Françoise (2008): A typological overview of Emerillon, a Tupi-Guarani language from French Guiana, in: *Linguistic typology* 12(3), 431–460.
- Rose, Françoise (2011): *Grammaire de l'émerillon teko. Une langue tupi-guarani de Guyane française*. Leuven: Peeters.
- Rose, Françoise & Renault-Lescure, Odile (2008): Contact-induced changes in Amerindian languages of French Guiana, in: Stolz, Thomas; Salas Palomo, Rosa & Bakker, Dik (eds.), *Aspects of language contact. New theoretical, methodological and empirical findings with special focus on Romancisation processes*. Berlin: Mouton de Gruyter, 349–376.
- Schleicher, Charles (1998): *Comparative and internal reconstruction of the Tupi-Guarani language family*. Ann Arbor: University of Wisconsin – Madison.
- Seki, Lucy (2000): *Gramática do Kamaiurá*. Campinas: Editora da Unicamp.
- Stolz, Thomas (2008): Romancisation worldwide, in: Stolz, Thomas; Salas Palomo, Rosa & Bakker, Dik (eds.), *Aspects of language contact. New theoretical, methodological and empirical findings with special focus on Romancisation processes*. Berlin: Mouton de Gruyter, 1–42.
- Taylor, Douglas (1954): Diachronic note on the Carib contribution to Island Carib, in: *International Journal of American Linguistics* 20(1), 28–33.
- Taylor, Douglas & Hoff, Berend (1980): The linguistic repertory of the Island-Carib in the seventeenth century: the men's language – A Carib pidgin?, in: *International Journal of American Linguistics* 46(4), 301–312.
- Thomason, Sandra & Kaufman, Terrence (1988): *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Weinreich, Uriel (1953 [1970]): *Languages in contact. Findings and problems*. The Hague/Paris: Mouton.
- Winford, Donald (2005): Contact-induced changes. Classification and processes, in: *Diachronica* 22(2), 373–427.

